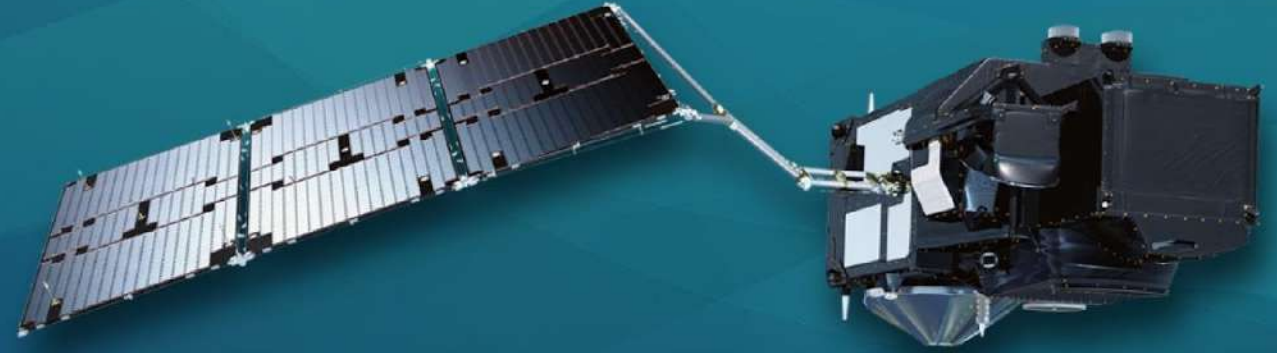




PROGRAMME OF THE
EUROPEAN UNION



co-funded with



9th Sentinel-3 Validation Team meeting 2026

30 March–01 April 2026 | ESA–ESRIN | Frascati (Rome), Italy

Evaluation of the collection4 Sentinel-3 SLSTR IST retrieval

Dybkjær, G. (DMI), Eastwood, S. (MET), Killie, M. (MET), Ribere C. (Noveltis), Lalloue, N. (Noveltis), Igor Tomazic (EUMETSAT)



Project context and presentation layout

Project context:

- Copernicus initiated the Science for Marine Surface Temperature (Sci4MaST) project at EUMETSAT to develop and operate and improve Sentinel 3 (S3) Marine Surface Temperature retrievals, to provide monitoring tools and to coordinate GHRSSST project office.
- This work is Sci4Mast WP1-5 concerning retrieval of IST.

Presentation outline

- Retrieval algorithms, and other algorithms
- Reference data sets
- General performance of the S3 IST retrieval.
- Evaluation of uncertainties and new developments (work in progress)
- Cloud mask validation



Sentinel-3 SLSTR IST retrieval algorithms

Two algorithms selected for operation - from 16 tested algorithms

Split Window (NadirView) and Single Channel
(DualView)

$$IST2/NV = a_0 + a_1 T b_{11_{nadir}} + a_2 T b_{12_{nadir}} + a_3 ((T b_{11_{nadir}} - T b_{12_{nadir}})(\sec \theta - 1))$$

$$IST12/DV = a_0 + a_1 T b_{11_{nadir}} + a_2 T b_{11_{oblique}}$$

Reference data sets

– Searching high quality reference data from polar sea ice



PROGRAMME OF THE EUROPEAN UNION

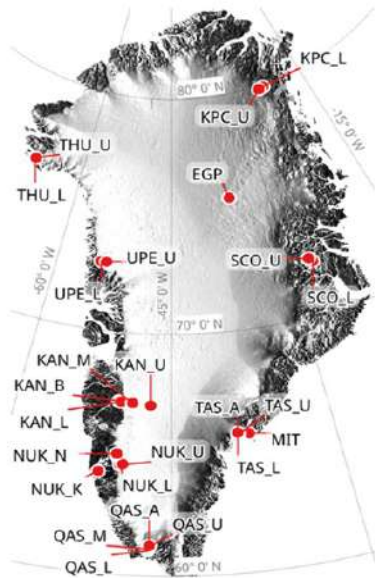


EUMETSAT

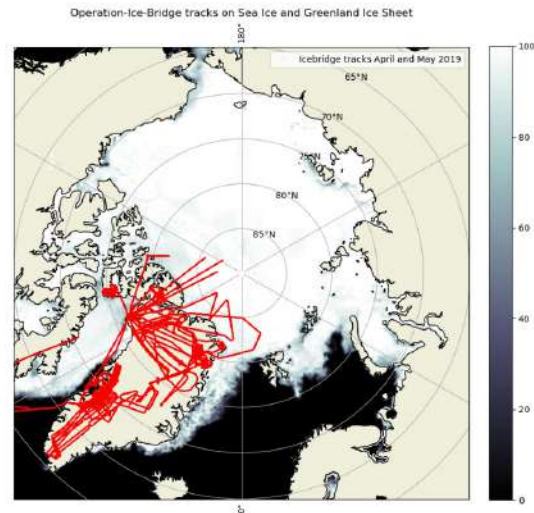
co-funded with



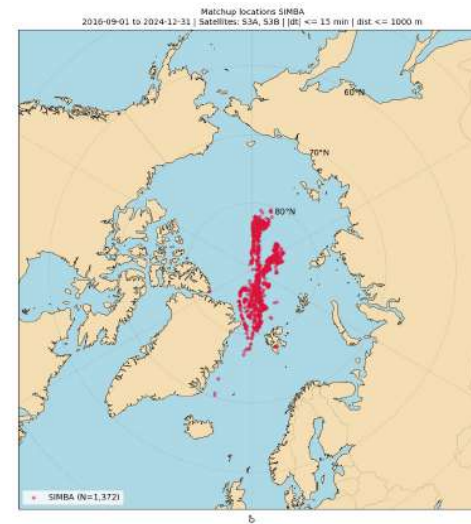
See also talk (Marc Lucas) and poster (Gorm Dybkjaer), about Copernicus reference buoy development program for Sentinel-3 SLSTR SST and IST



Promice, Automatic Weather Stations, Greenland Ice Sheet



Operation Ice-Bridge, Aerial data Greenland Ice Sheet and Sea Ice



Simba, Ice Massbalance Buoys (IMB) Sea Ice



MDB also contains Sea Ice data from IABP buoy program
Data not presented here, because of low reliability

Sentinel-3 Collection4, IST Performance - PROMICE

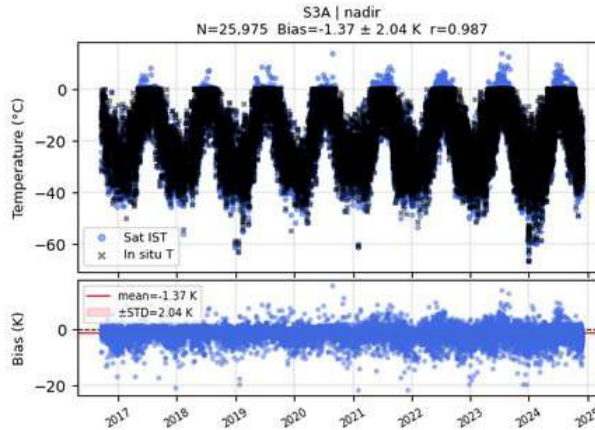


PROGRAMME OF THE
EUROPEAN UNION



EUMETSAT

co-funded with



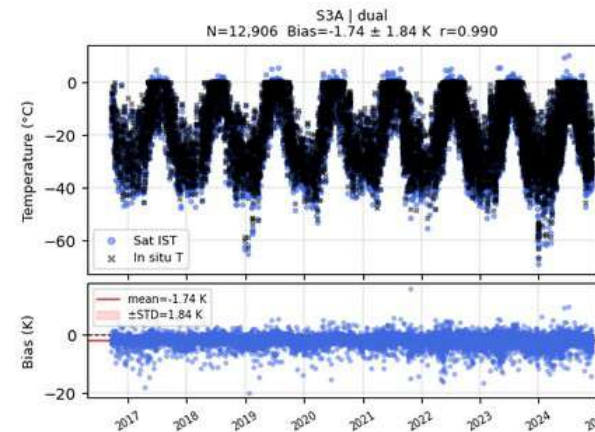
The EUMETSAT team issued MDB data from the full reprocessed Collection4 SST/IST data set (2016-2025).

REFERENCE:

PROMICE Skin Temperatures, GIS QL 4+5, dt < 15 min.; dx < 1000 m

Generally fine performance:

- Relatively small STD of differences to ref.
- Stable bias across timeseries
- DV performs best on ice sheet
- DV has less warm outliers



First glance at full timeseries of collection 4 data (From Igor T, EUMETSAT)

Sentinel-3 Collection4, IST Performance – Operation IceBridge



PROGRAMME OF THE
EUROPEAN UNION



EUMETSAT

co-funded with

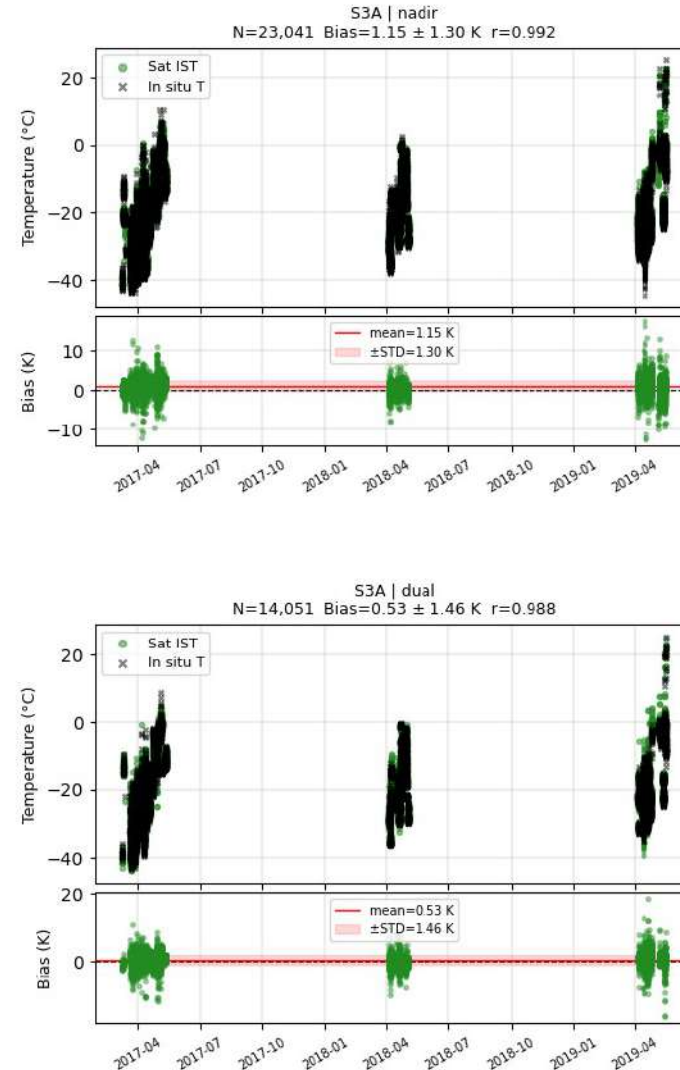


REFERENCE:

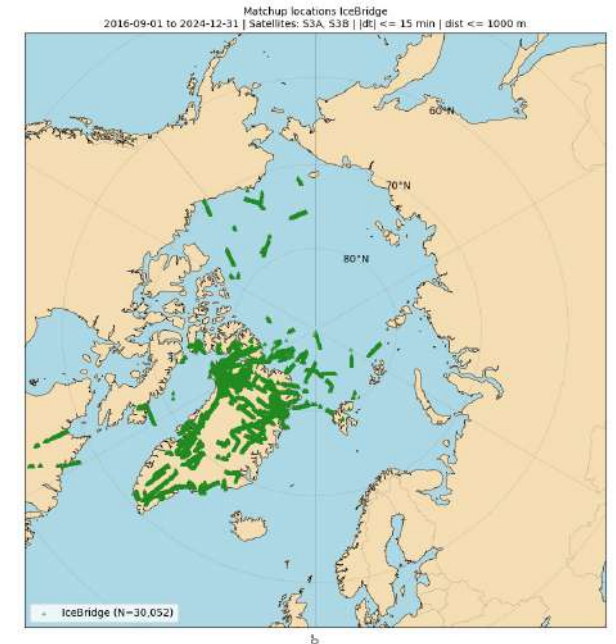
IceBridge Skin Temperatures (averaged over 1km)

QL 4+5, dt < 15 min.; dx < 1000 m

- Very high performance of both algorithms
-
- STD of "errors" less than 1.5 K
- High stability



(From Igor T, EUMETSAT)



Sentinel-3 Collection4, IST Performance - SIMBA IMB



PROGRAMME OF THE
EUROPEAN UNION



EUMETSAT

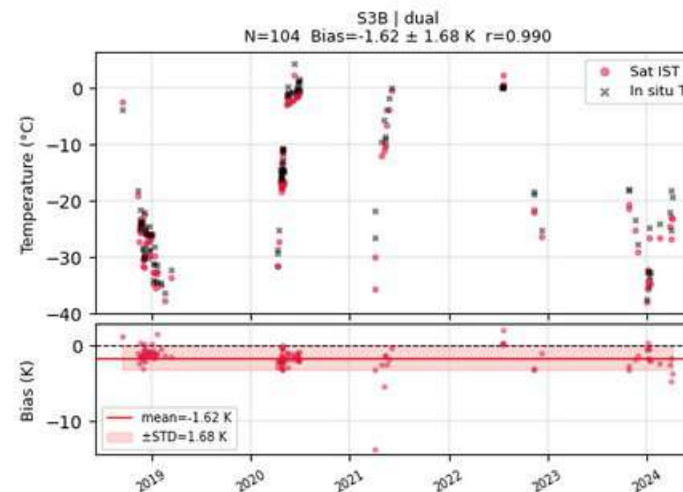
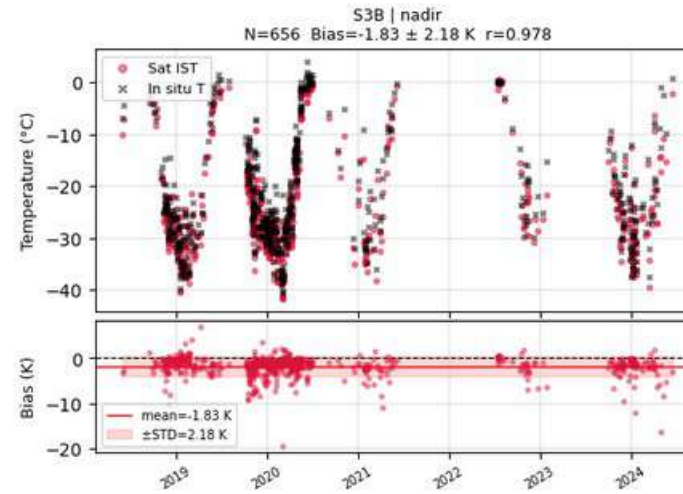
co-funded with



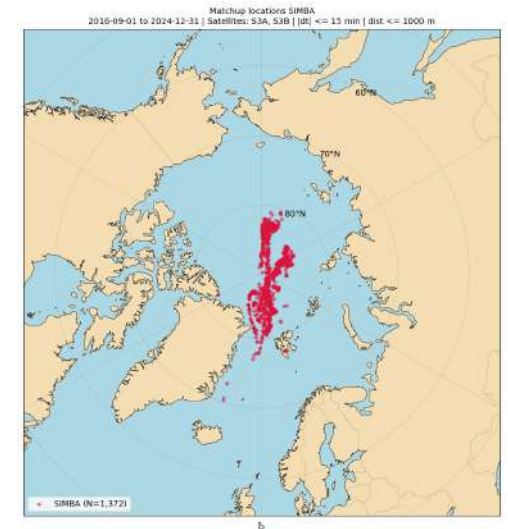
REFERENCE:

Simba IMBs from AWI
QL 4+5, dt < 15 min.; dx < 1000 m

- High performance of both algorithms - DV in particular
- Slightly larger 'errors' than OIB and Promice reference - most likely due to Tskin retrieval from IMB.
- High stability



(From Igor T, EUMETSAT)

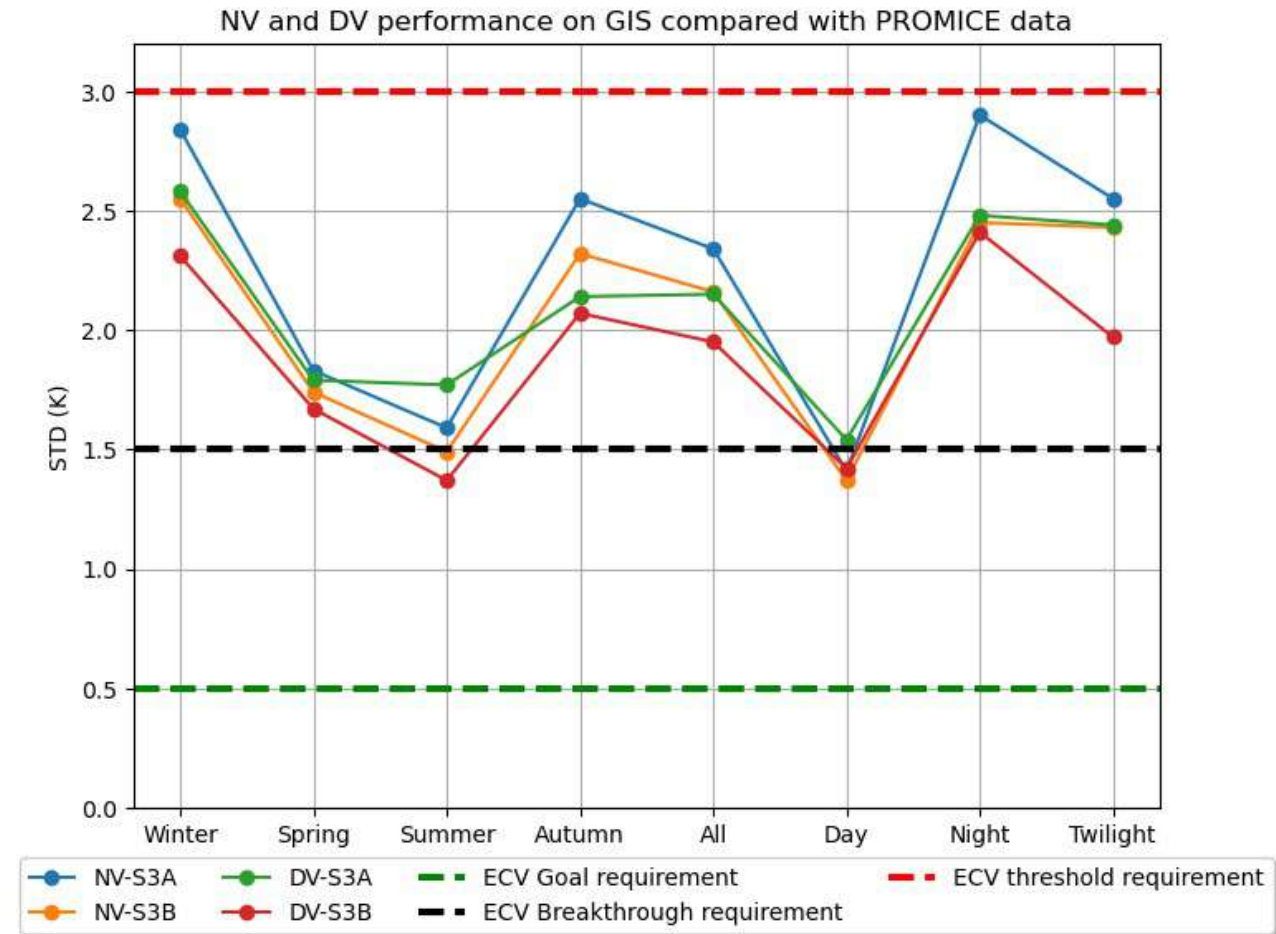


Generally

- S3B performs better than S3A
- DV performs better than NV

Validation of S3A/B NV and DV algorithms on Greenland Ice Sheet. Reference is Promice.

Graph shows STD of differences between the respective algorithm and reference. Data with QL 3+4+5 are evaluated.



S3-IST, Collection4 2019 Performance by Quality Level



PROGRAMME OF THE
EUROPEAN UNION



co-funded with



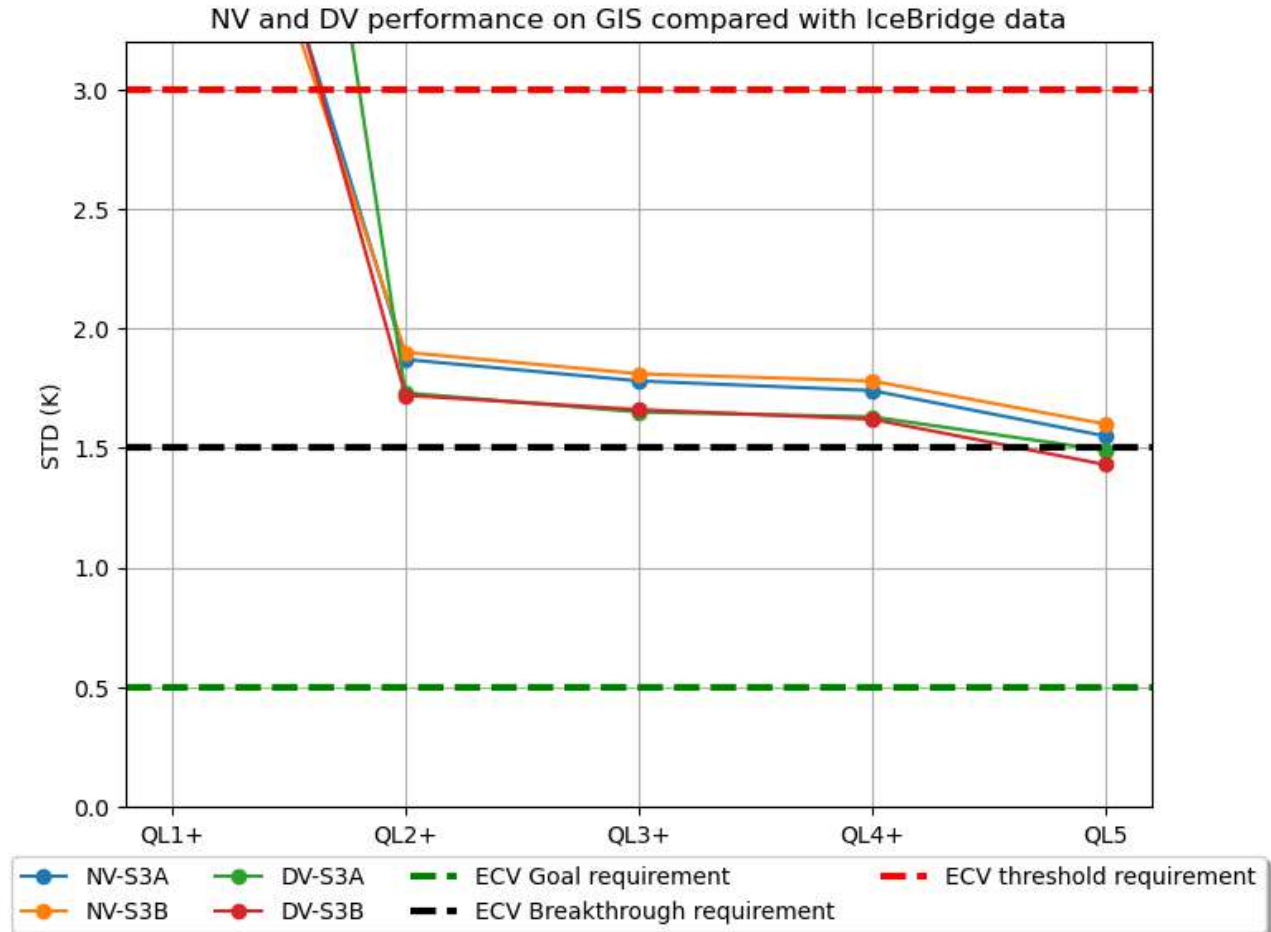
Generally

- S3B performs better than S3A
- DV performs best on the GIS

Validation of S3A/B NV and DV algorithms over the Greenland Ice Sheet. Reference is IceBridge.

All reference data are recorded in daylight April and May 2019.

Graph shows STD for the respective algorithms and sensors. Performance values are shown a function of QL, where QLX+ represent all data with QLX and larger.



Sentinel-3 IST Uncertainty algorithm

$$\text{UNC}_{\text{total}}^2 = \text{UNC}_{\text{random}}^2 + \text{UNC}_{\text{correlated}}^2$$

- $\text{UNC}_{\text{random}}^2 = \text{UNC}_{\text{geo}}^2 + \text{UNC}_{\text{NEdT}}^2$
- $\text{UNC}_{\text{correlated}}^2 = \text{UNC}_{\text{synoptic}}^2 + \text{UNC}_{\text{global}}^2$

$$\underbrace{\hspace{10em}}_{\text{UNC}_{\text{algo}}^2 + \text{UNC}_{\text{emissivity}}^2}$$

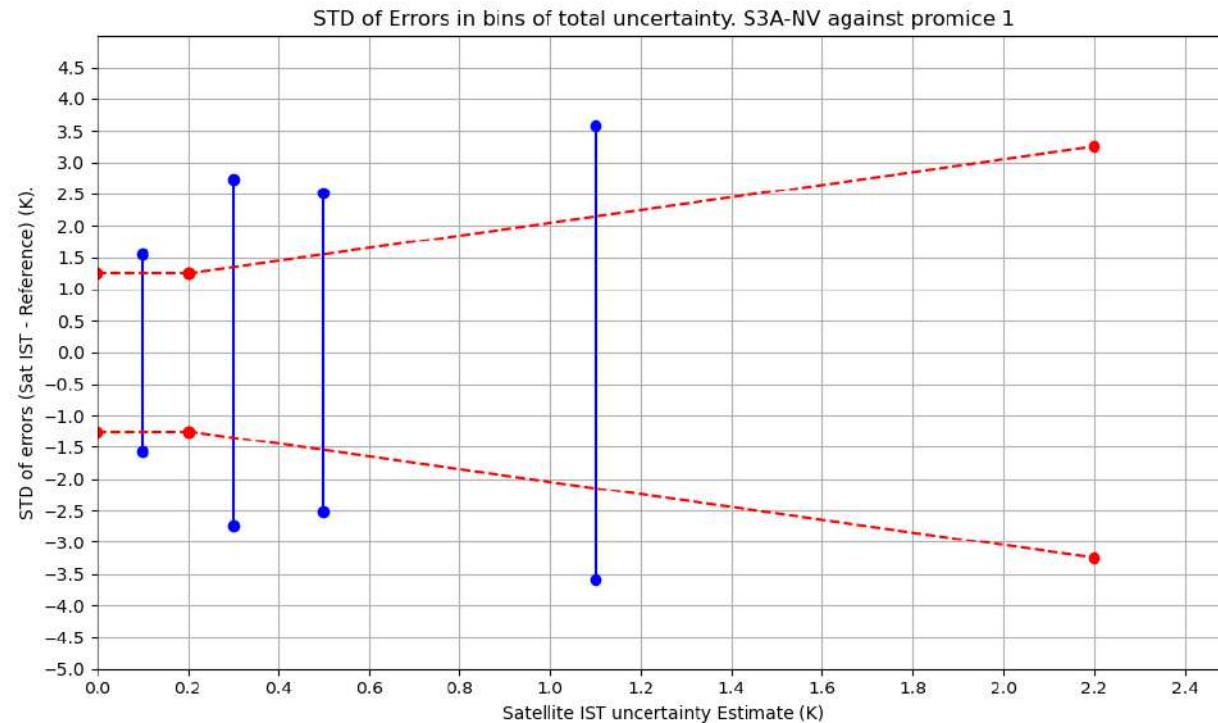
The UNC components:

- **Total** uncertainty for a given IST pixel value
- **Random** uncertainty is composed by NEdT induced level-2 uncertainty and a geoposition error induced uncertainty
- **Correlated** uncertainty is composed by a synoptic component and Global uncertainty that is associated with each quality level.
 - The **synoptic** uncertainty is composed by the algorithm uncertainty (from calibration) and uncertainty from snow emissivity variations.

Sentinel-3 IST Uncertainties - performance

Baseline uncertainty, $\pm 1\text{-}\sigma$

Uncertainty associated with point to area temperature comparison



Example of uncertainty – Promice data, NV algorithm.

Blue vertical lines is ± 1 sigma of observed "errors" in bins of total uncertainty

Red dotted lines indicate ideal 'error' distribution for uncertainty bins.

Evaluation reveal that the uncertainty algorithm is not working well



Estimating the *baseline uncertainty*

Baseline Uncertainty for obstype ['promice', 'icebridge', 'simba'];

$$\text{unc_baseline}^2 = \text{unc_ref}^2 + \text{unc_dx}^2 + \text{unc_dt}^2 + \text{unc_dz}^2$$

unc_ref_simba	= 0.1K	; Thermistor uncertainty, measured
unc_ref_promice	= 0.4 K	; Broadband radiometric measurement / 2 mT air, estimate
unc_ref_icebridge	= 0.2 K	; Narrowband TIR radiometer, estimate
unc_dx_icesheet	= 1.0 K	; level ice and level ice with snow piles, guess from drone footage
unc_dx_seaice	= 1.5 K	; 50 % added to STD of level ice, guess from drone footage
unc_dt_all	= 0.5 K	; winter+spring (december-may) 15 minutes, observations
unc_dz_icebridge	= 0.0 K	; Measurements are from surfaces
unc_dz_promice	= 0.3 K	; Modelled from surfaces and 2mt air, guess
unc_dz_simba	= 1.2 K	; selection of non-surface thermistor, observation

SIMBA:	2.0 K
PROMICE:	1.2 K
ICEBRIDGEicesheet:	1.2 K
ICEBRIDGEseaice:	1.6 K

Adjustments to Collection4 uncertainty algorithm → DEV uncertainty algorithm



PROGRAMME OF THE
EUROPEAN UNION



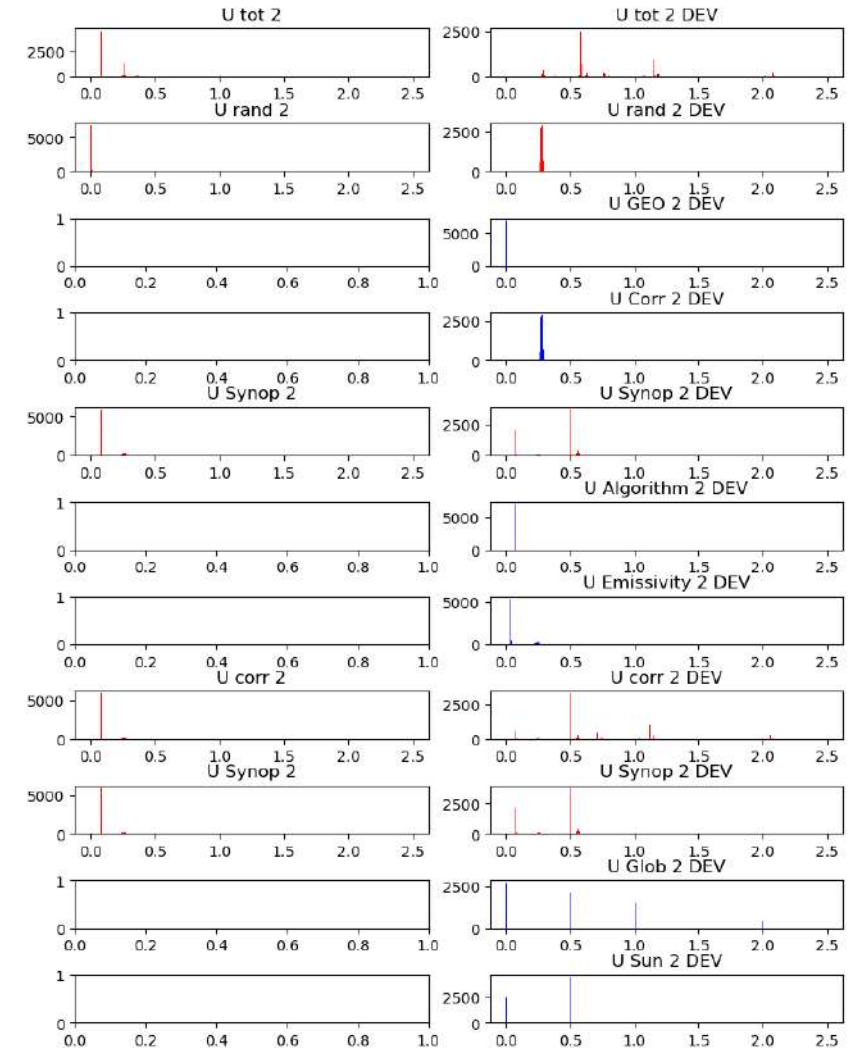
co-funded with



Dev-uncertainties is adjustments to collection4 uncertainties is work in progress

- New synoptic uncertainty component: 0.5 K for sun-zen angle > 80 deg
- Global uncertainty bug-fix
- NEdT set to 0.1 K for all TIR Tb's, and pertubeted (100x) through algorithms
- Geo Position error increased. Only effect for sea ice (85 % > conc > 15 %)

sun-zen angle



DEV uncertainty evaluation – ice sheet

Work in progress...



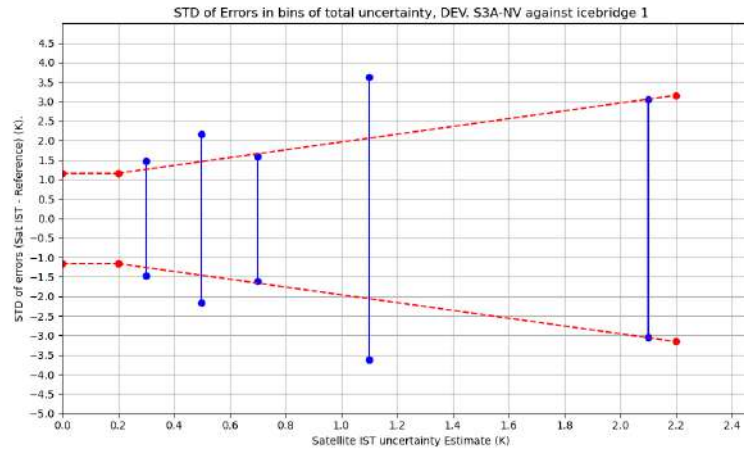
PROGRAMME OF THE EUROPEAN UNION



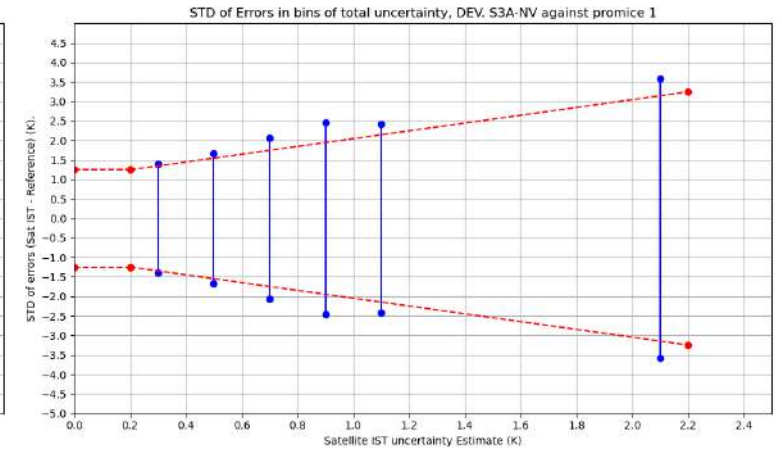
co-funded with



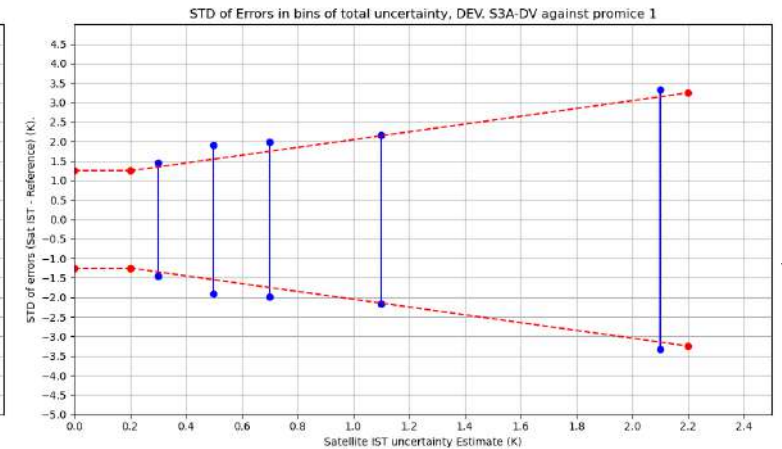
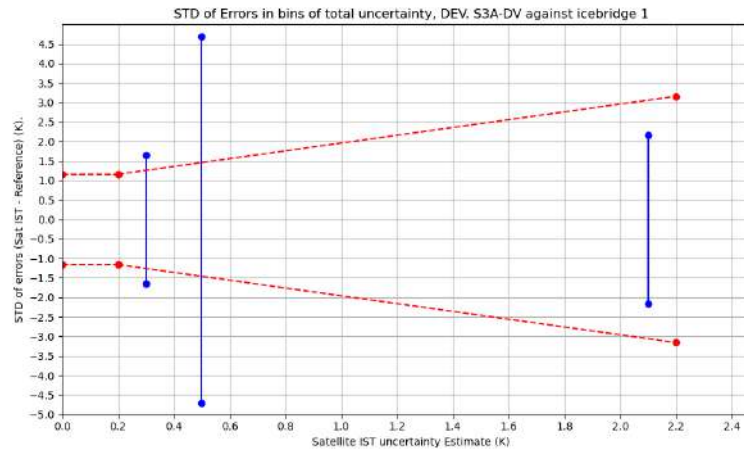
IceBridge



Promice



N
V



D
V

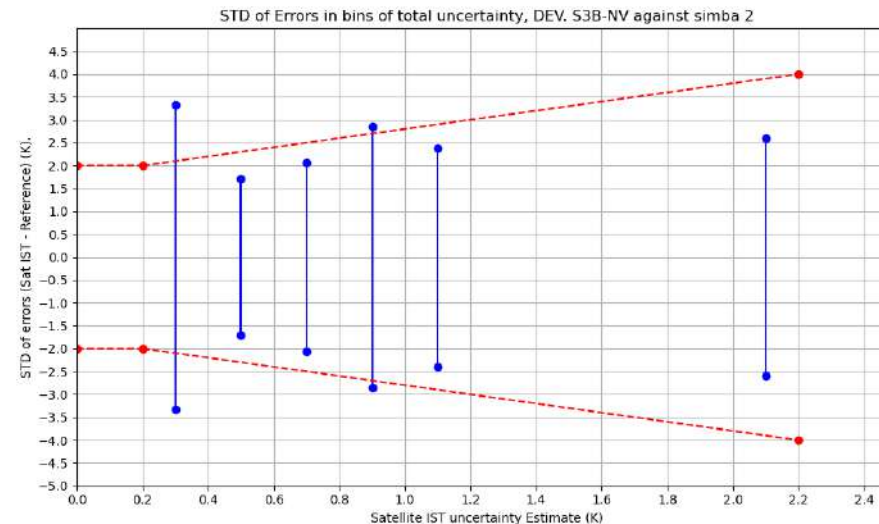
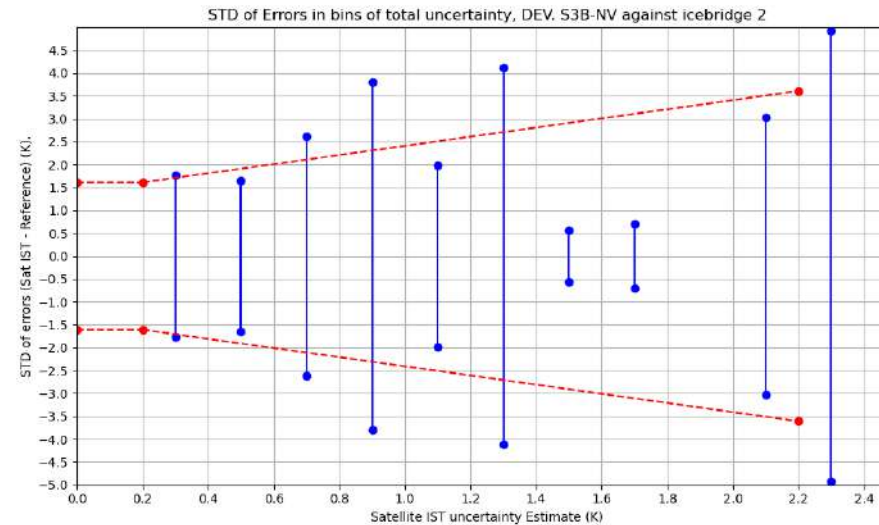
DEV uncertainty algorithm seems to handle uncertainties on ice sheet fair.

With issues with icebridge data...

DEV uncertainty evaluation – sea ice

DEV uncertainty algorithm has troubles handling uncertainties for **sea ice**

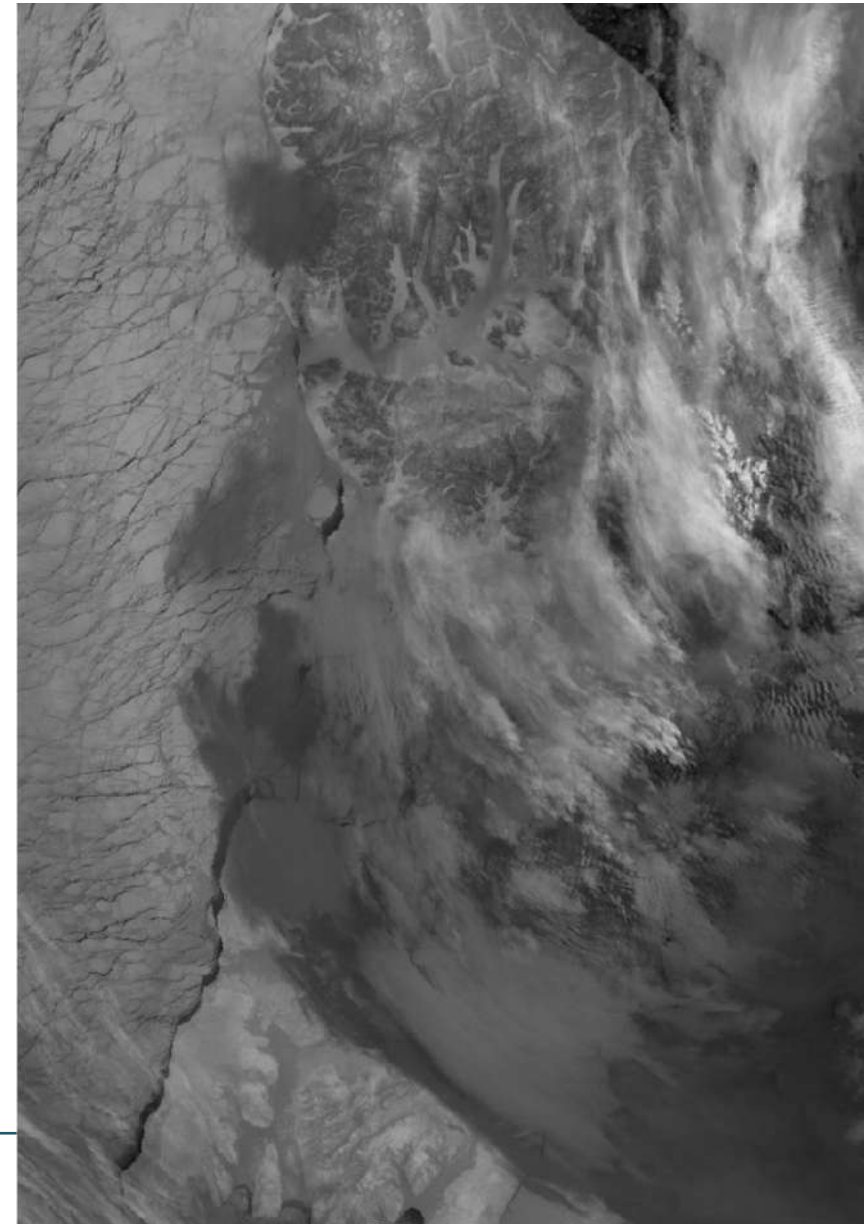
Top: NV vs icebridge.
Bottom: NV vs Simba





Cloud mask validation

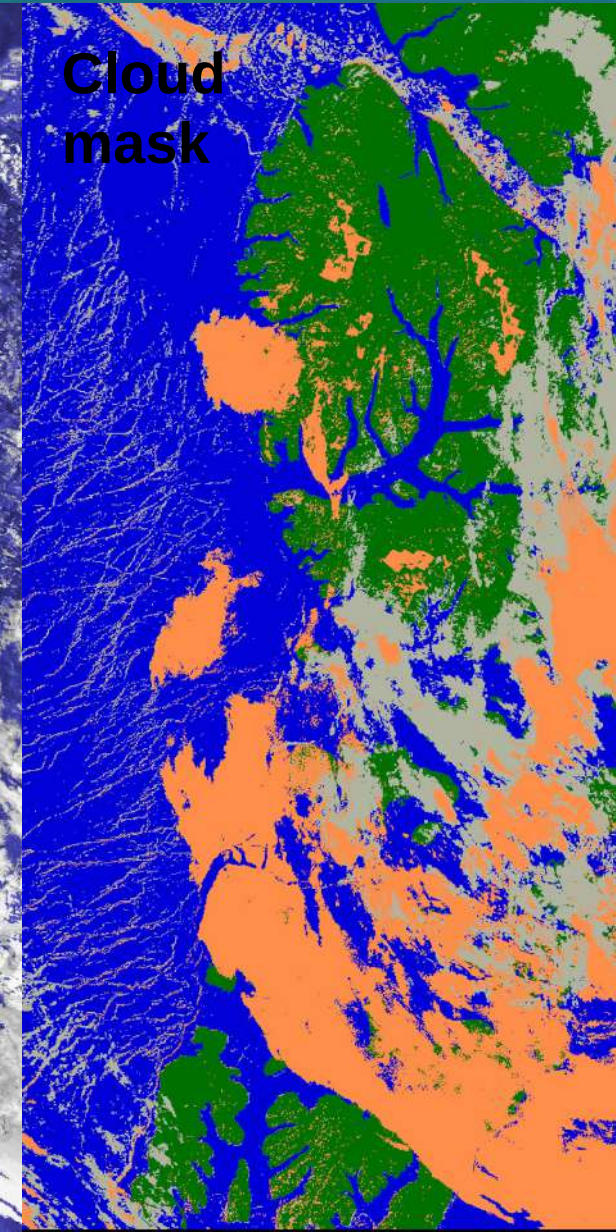
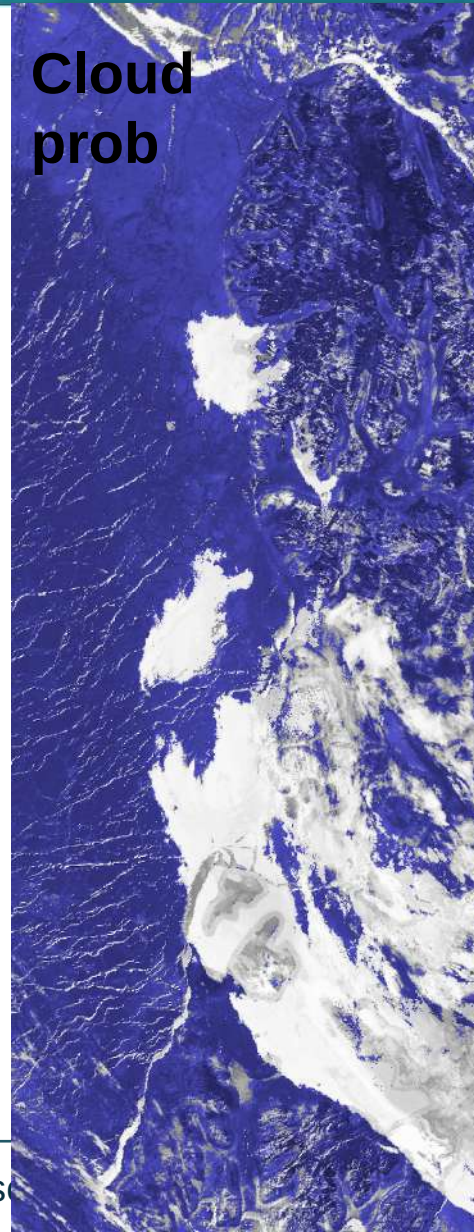
- Validation of the cloud mask used in the SLSTR IST processing



S-3B 2023-11-30 01:45UTC

Cloud mask validation

- Validation of the cloud mask used in the SLSTR IST processing
- PPS v2021 from EUMETSAT Nowcasting Satellite Application Facility (NWC SAF)
- Provides two independent cloud products:
 - CMA: Classical threshold-based cloud mask
 - CMAProb: Naive Bayesian cloud probability

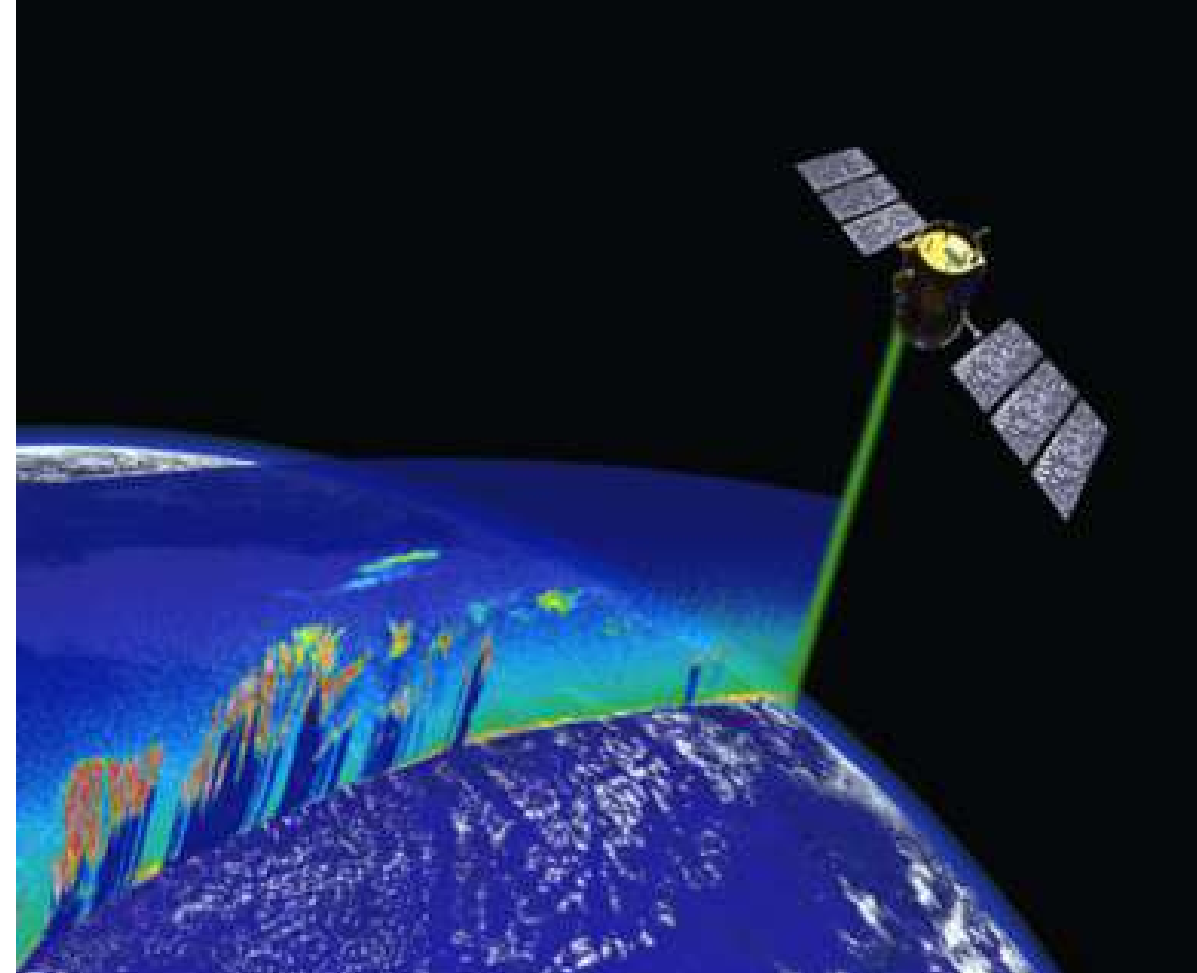


S-3B 2023-11-30 01:45UTC



Validation data

- Using Calipso lidar from the Calipso satellite
- Detects presence of cloud and aerosols in multiple levels
- Sampled at 70m at 330m intervals in nadir
- Use data resampled to 1km



Definition of validation scores

- Contingency table with binary cloud info from PPS and Caliop
- Probability of Detection and False Alarm Ratio scores are used to evaluate the cloud mask
- FAR-clear important for temperature retrieval: how often the cloud mask falsely classifies cloudy pixels as cloud free

		Validation truth	
		Cloudy	Cloud-free
PPS	Cloudy	A	B
	Cloud-free	C	D

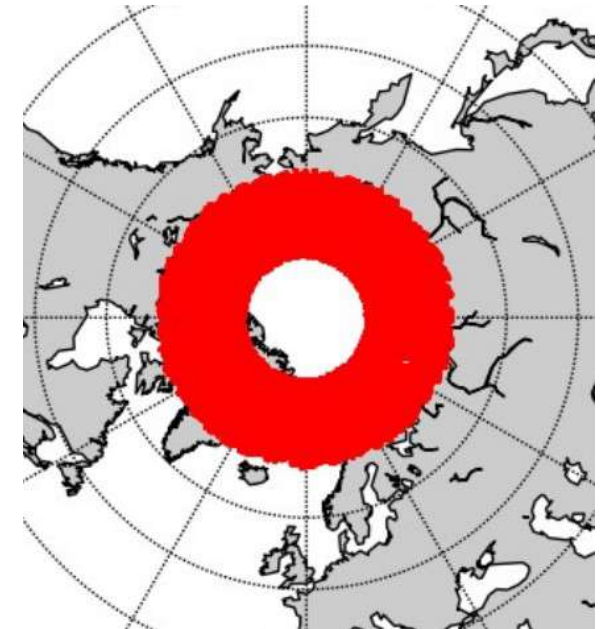
Hit rate (HR) / PC (percent correct)	$(A+D)/N$
POD-cloudy	$A/(A+C)$
FAR-cloudy	$B/(A+B)$
POD-clear	$D/(B+D)$
FAR-clear	$C/(C+D)$
Hanssen-Kuiper Skill Score	$(AD - BC) / ((A+C) (B+D))$

$$Bias = 1/N \sum_N (y_k - o_k)$$

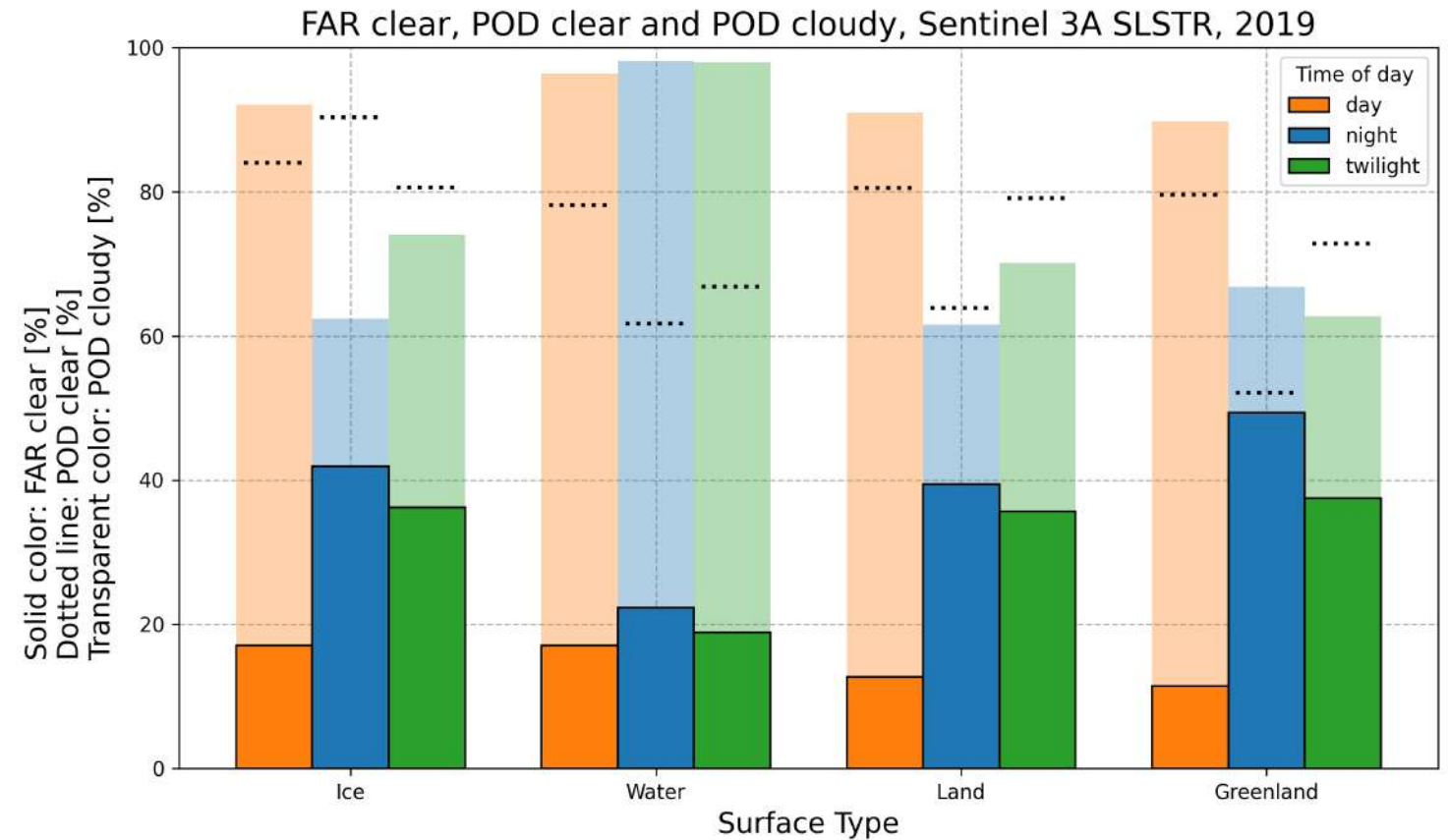


Validation results – preliminary results

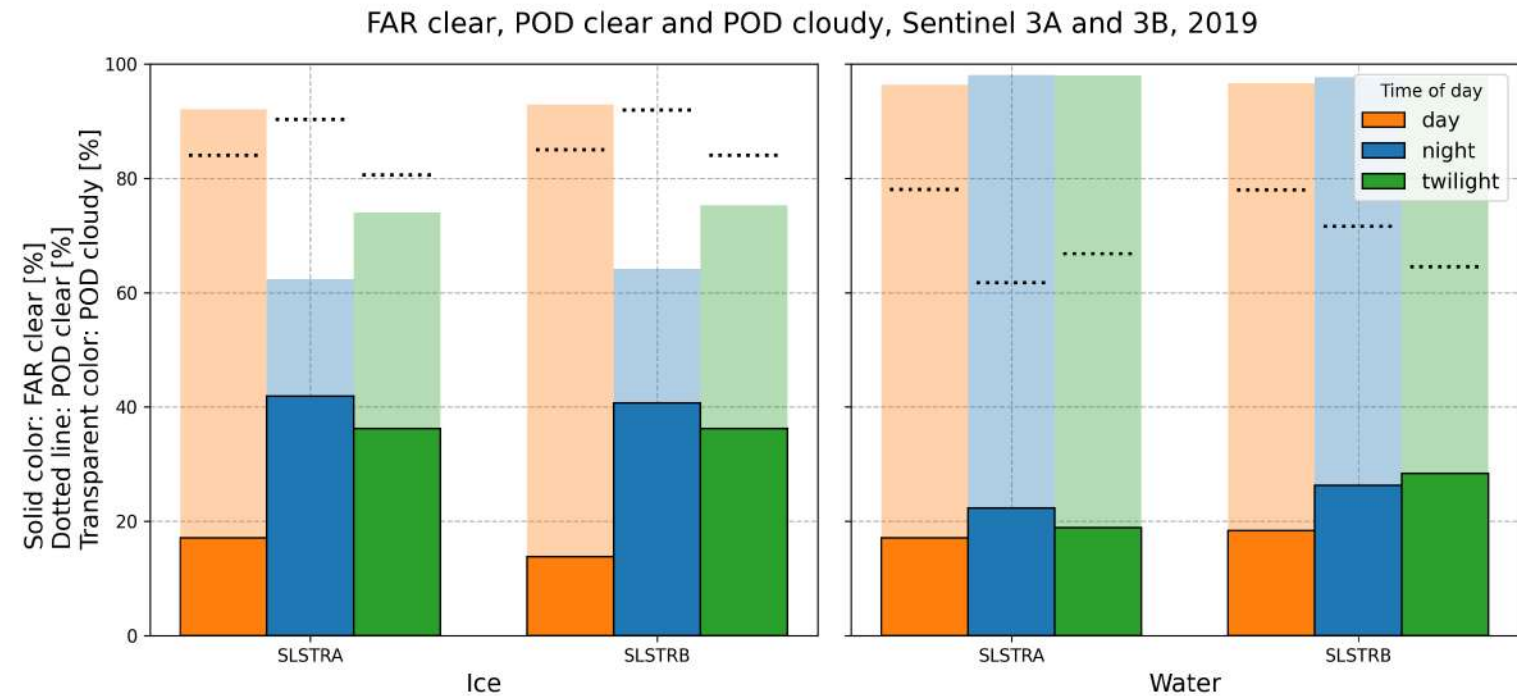
- Preliminary cloud masking results
- Based on 2019 data from the most recent Multi-mission Matchup Data Set (see poster by Igor Tomazic et al)
- Only Northern Hemisphere so far
- Will extend to 2016-2022 and Southern Hemisphere when MMDS is ready



Validation of CMA for different surface types

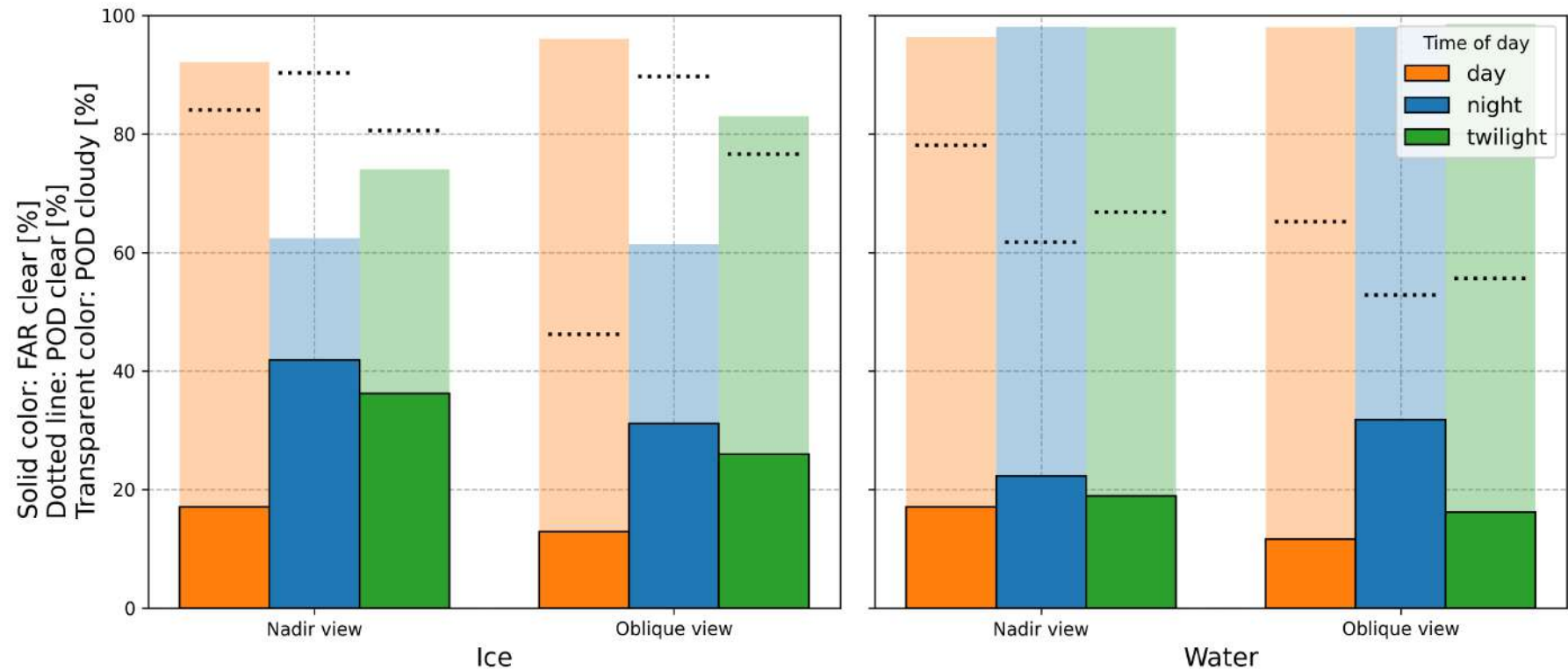


Validation of CMA for S3-A vs S3-B



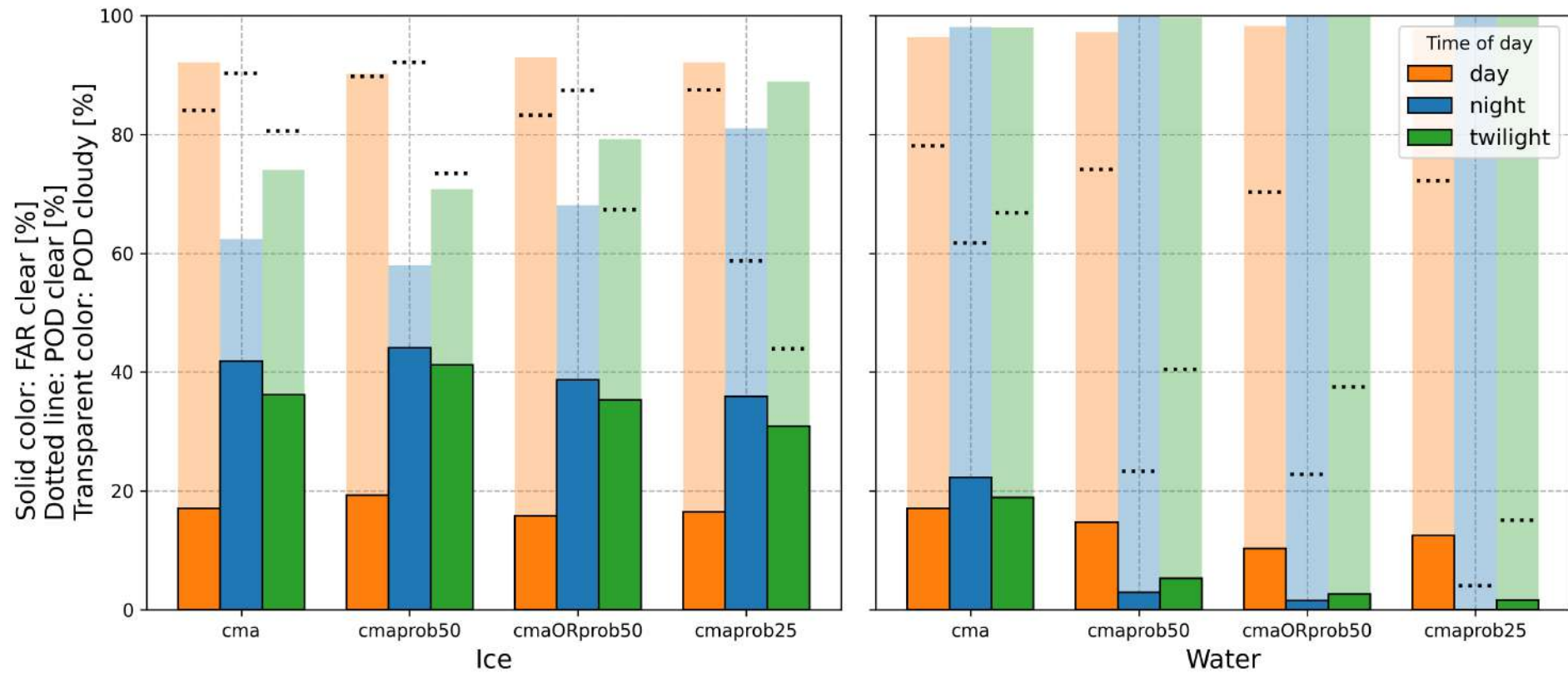
Validation nadir vs oblique views

FAR clear, POD clear and POD cloudy, Sentinel 3A, 2019

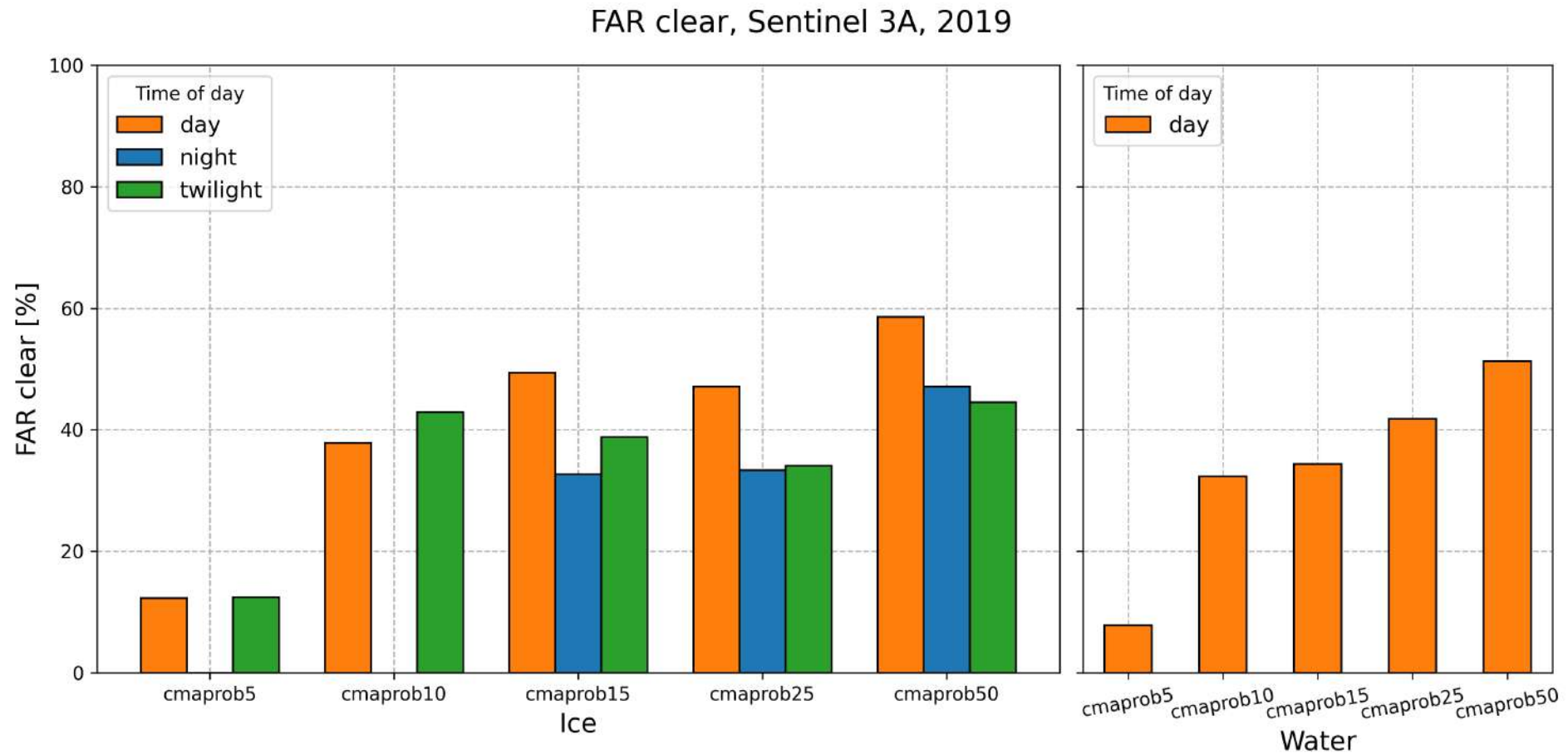


Validation of cloud mask vs cloud probability

FAR clear, POD clear and POD cloudy, Sentinel 3A SLSTR, 2019

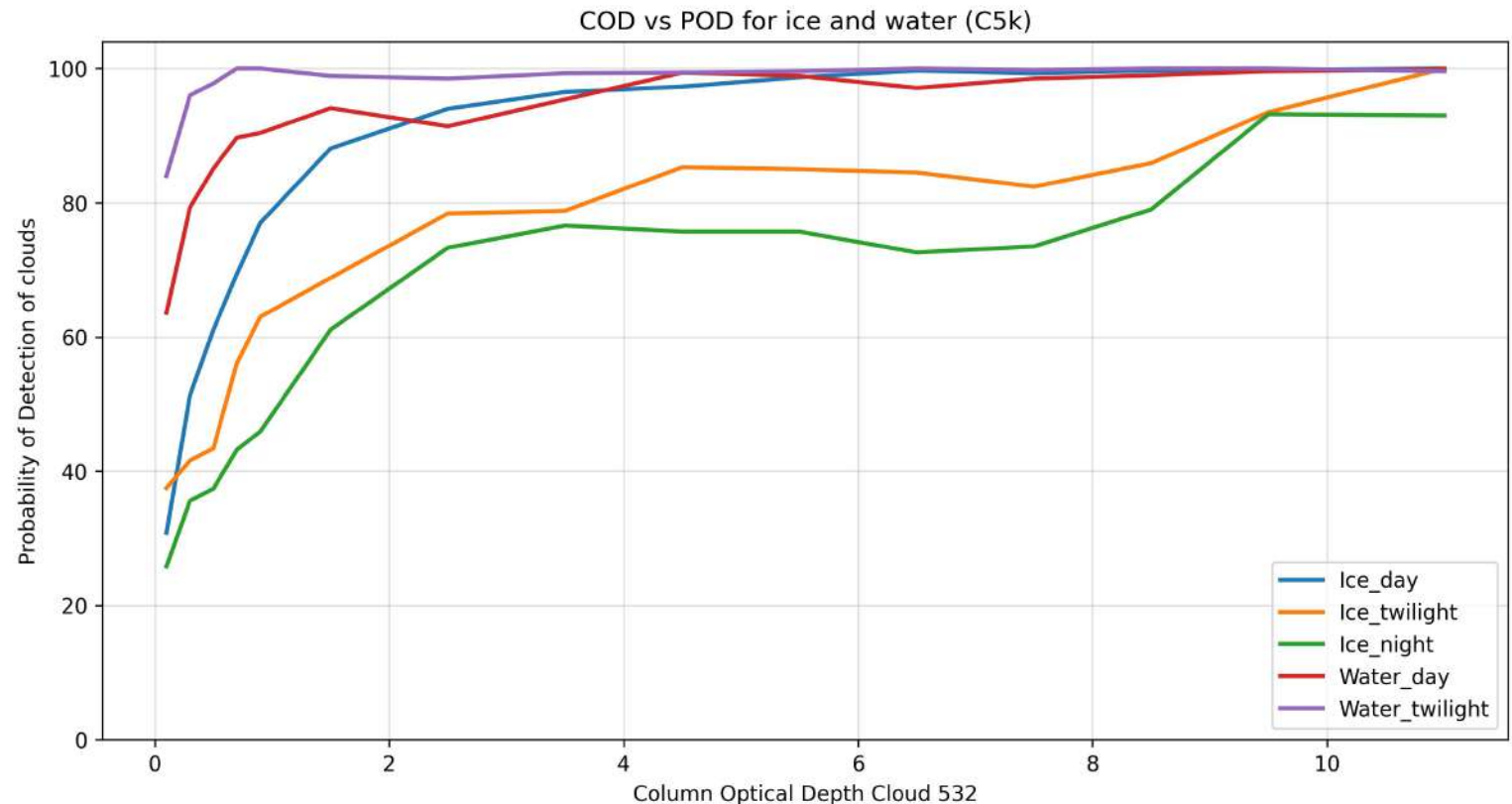


Sensitivity of different cloud probability levels



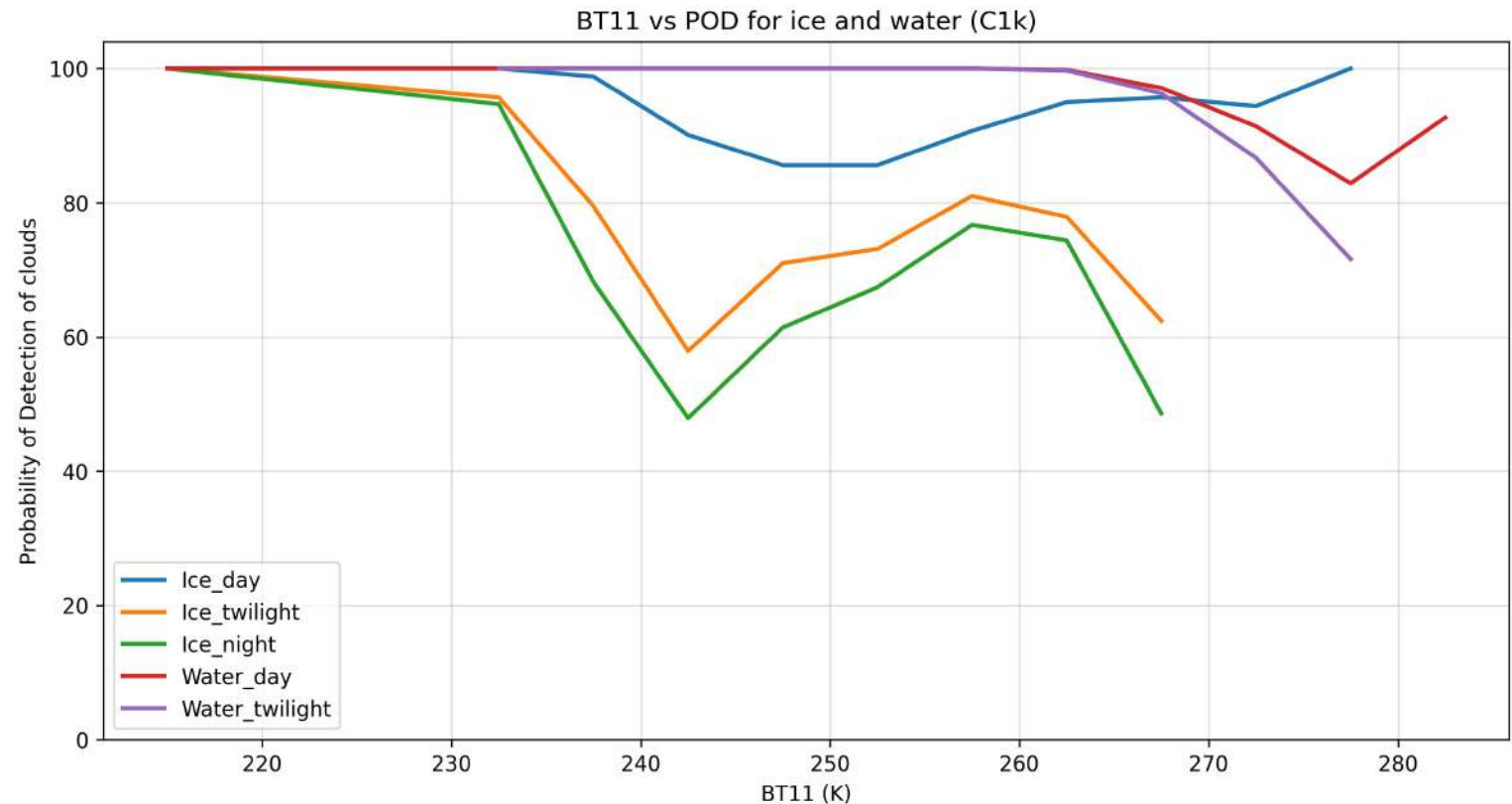
Which clouds are not detected ?

- Comparing the PODcloud with cloud optical depth from Caliop



Which clouds are not detected ?

- Comparing the PODcloud with SLSTR 11 μ m brightness temperature





Thank you for your
attention.

Any questions?