BROCKMANN CONSULT

Ground based sky camera images as a validation source for satellite cloud masking algorithms

Jan Wevers[1]

C. Brockmann[1], J. Scholze[1], F. Niro[4,], A. Santamaria-Artigas[2], S. Skakun[2], E. Vermote[3], S. Casadio[4,], A. M. Iannarelli[5]

[1] Brockmann Consult GmbH, [2] University of Maryland College Park, [3] NASA Goddard Space Flight Center, [4] SERCO c/o ESA-ESRIN, [5]SERCO

ESA Living Planet Symposium 2022, Bonn, Germany

Overview

- Background
- Objective/Overview
- Validation site and methods preparation
- Validation results
- Limitations
- Conclusion

Background

- Validation of satellite-based cloud masks is commonly done by the algorithm developers themselves.
  - Non-independent validation
- A few attempts have been made to objectively inter-compare performances of satellite-based cloud masking algorithms (e.g., Skakun et al. 2021, Zekoll et al. 2021, Tarrio et al. 2020, Hammersson Sanchez et al. 2020, Chi & Zhang 2020).
- All these validations/inter-comparisons are based on different datasets, leading to variable results even if the same algorithm is analyzed.
  - This was shown during the Cloud Mask Intercomparison eXercise (CMIX)
- Most validation datasets are sensor dependent and don't allow cross-sensor validation of multi-sensor cloud detection algorithms.

Goal: An independent validation source for cloud masking algorithms is needed, which is sensor independent

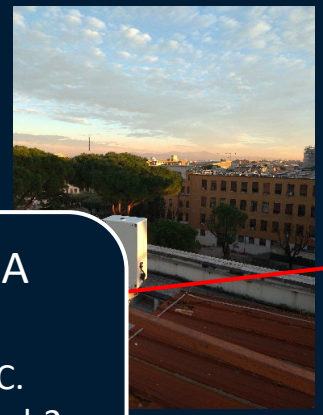
Objective / Overview

- The project was conducted in the frame of ESA's Quality assurance framework for earth observation (QA4EO).
- The objective of the project was to analyse the usage of ground-based sky cameras, as an independent validation source for satellite cloud masking algorithms.
- The scope of this work was to prototype algorithms and methods to process sky camera data and compare them with satellite-based cloud masks.
- There are two instruments for validation that have been compared:
  1. stereo sky camera (SC)
  2. Ceilometer - Raymetrics Aerosol Profiler (RAP).
- The work included 4 tasks:
  1. Requirements and state of the art analysis
  2. Validation sites and methods preparation
  3. Experimental operations
  4. Evaluation and conclusion

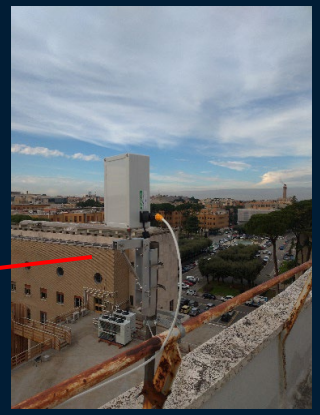
Validation site and methods preparation

Instrumentation setup

- A set of two cameras (stereo pair) was setup at La Sapienza University in Rome.
- The cameras use a
  the Omnivision C
  field of view is 194
  (vertical). Distance
  around 260 met
  cameras are co
  minute between 08
- Sky camera two
  approx. 20m apart from the ceilometer (RAP)
  - comparisons between the RAP and SC based cloud detection
  - validate the SC based cloud height estimation with RAP measurements.



Sky Camera 1: Marconi



Sky Camera 2: Fermi

Sky cameras are developed by University of Maryland & NASA

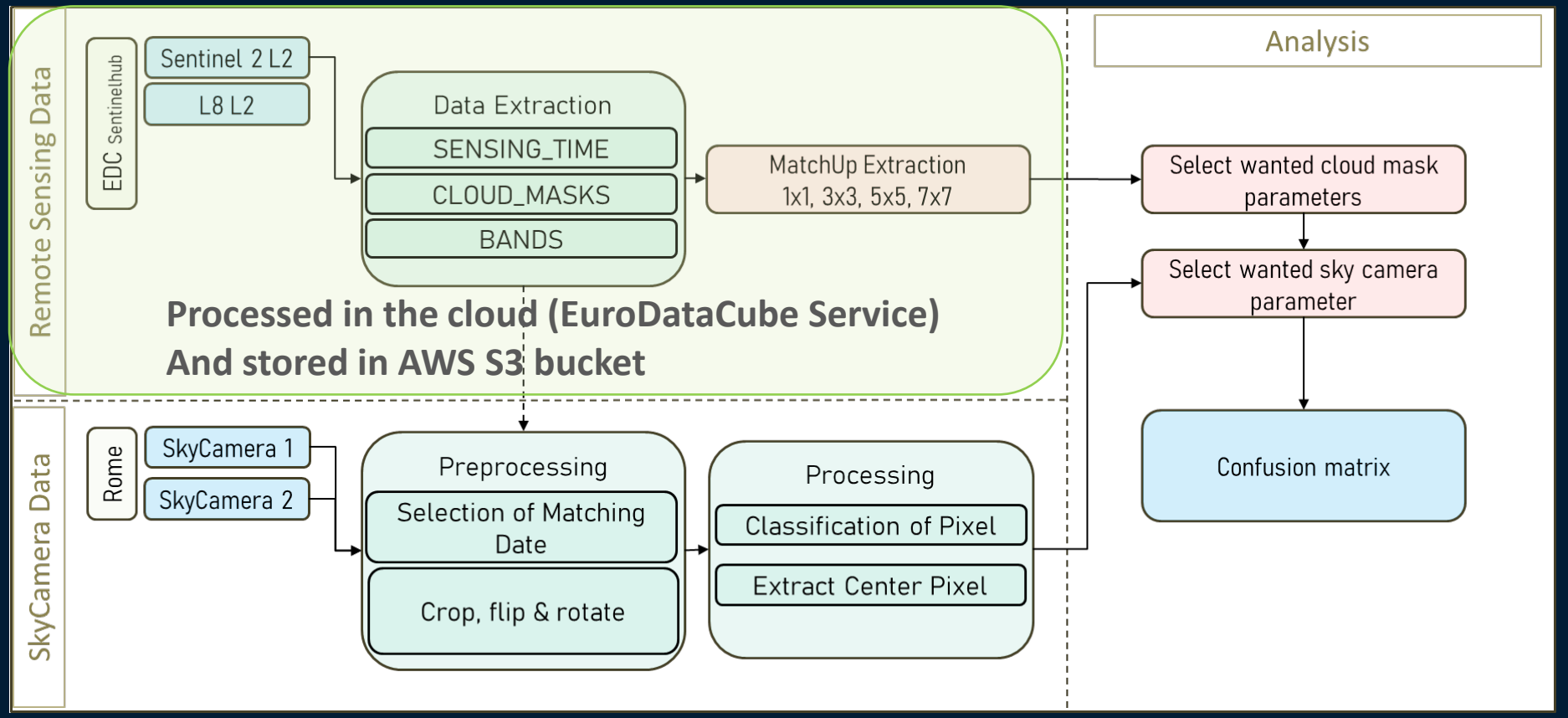
Skakun, S., Vermote, E. F., Santamaria-Artigas, A., Rountree, W. H., & Roger, J. C. (2021). An experimental sky-image-derived cloud validation dataset for Sentinel-2 and Landsat 8 satellites over NASA GSFC. International Journal of Applied Earth Observation and Geoinformation, 95, 102253.

Raymetrics Aerosol Profiler (RAP)

BROCKMANN CONSULT

Validation site and methods preparation

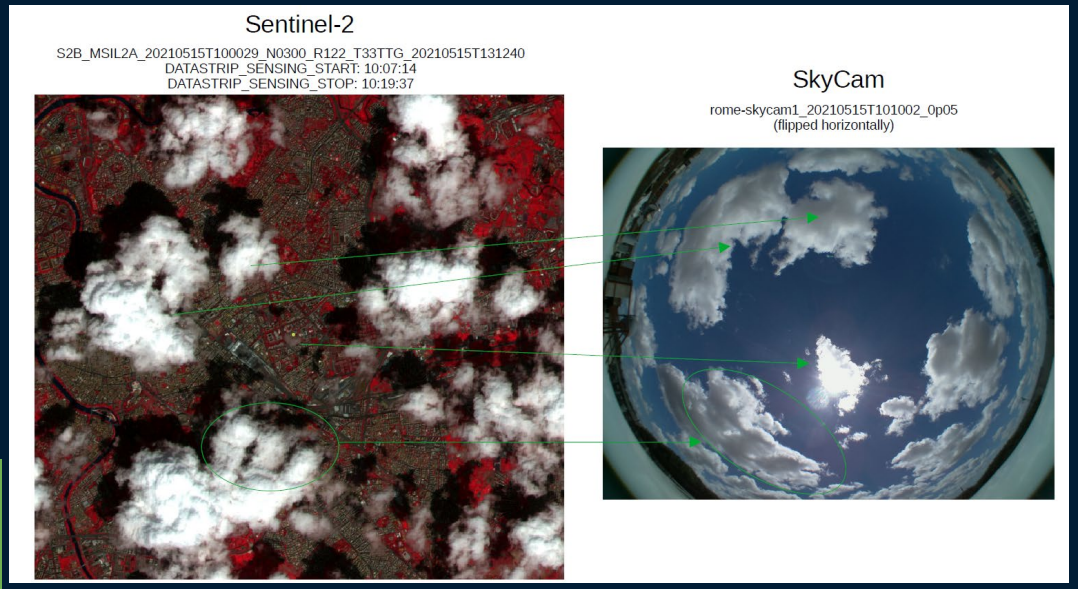
- Processing overview

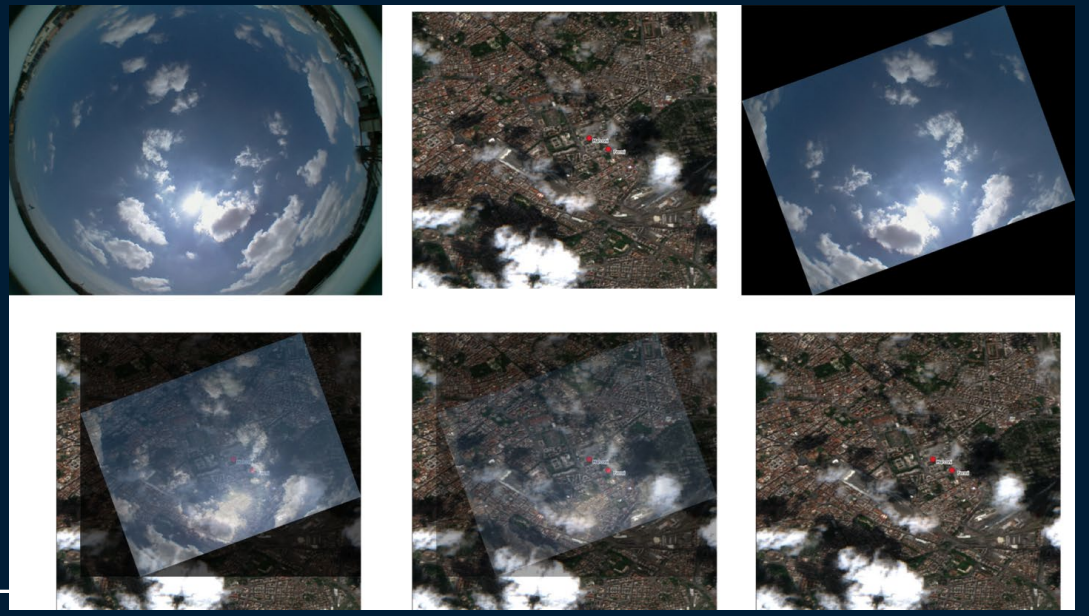


Validation site and methods preparation

Pre-processing of sky camera data to better match the satellite observations

- Crop: Reduce geometric distortion (increasing outside of the center).
- Rotate: The SCs are installed looking a bit northwest.
- Flip: The SC is looking from the ground up and the satellite does the opposite.





BROCKMANN CONSULT

Validation site and methods preparation

Finding an appropriate classification method

- A few methods hav uired accuracies.
  - Simple threshold on
  - Otsu thresholding
  - Otsu thresholding a
  - Implementing a line ges to improve the results of the pr
  - Brightness index (BI
- Training of a rando
  - 12 to 15 SC images
  - Polygons representi ages.
  - Inside these polygo
  - Overall, 11,100 sam en collected.

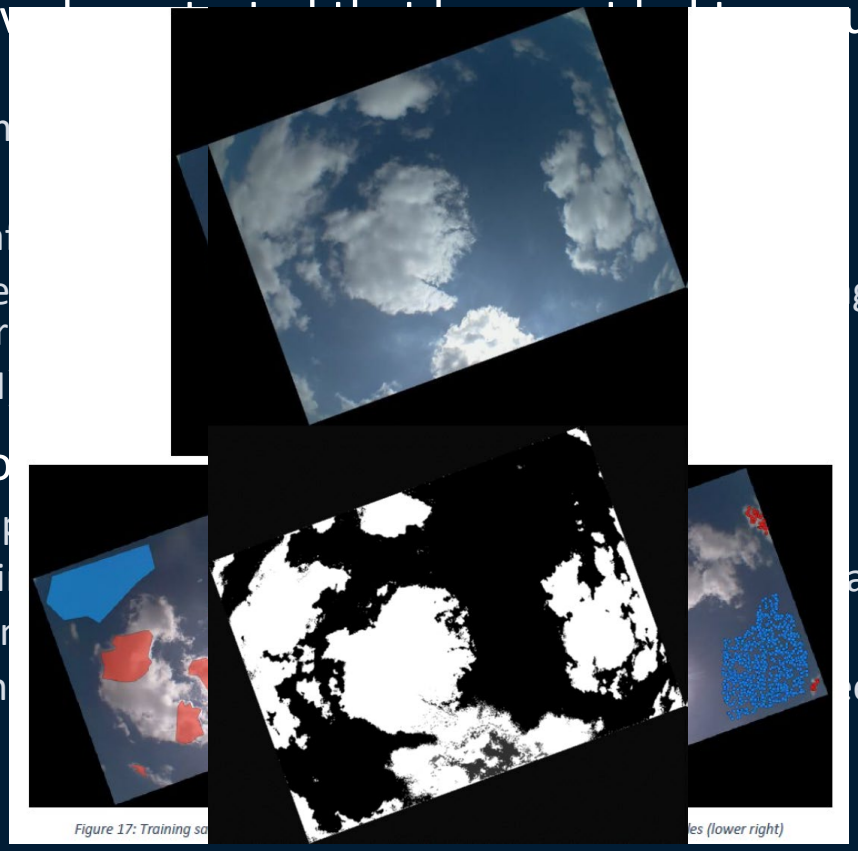


Figure 17: Training sa les (lower right)

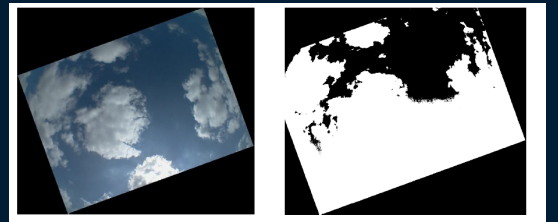


Figure 14: SC 1 image 26.03.2021

Figure 15: Classification using Otsu threshold after Gaussian filtering

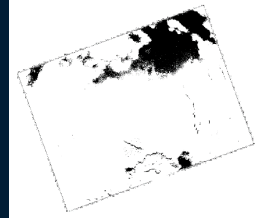


Figure 16: Classification using BI SI method from Letu et al. 2014

Validation results

- Validation of the RF classifier shows high accuracy (93-96% OA)

SkyCam 1 manual classification vs. SkyCam 1 auto classification

Sky Camera 1 manual classification

Sky Camera 1 automatic classification	Clear	Cloud	Sum	U A	E
CLEAR	30	2	32	93.8	6.2
CLOUD	2	27	29	93.1	6.9
Sum	32	29	61		
P A	93.8	93.1		OA:	93.44
E	6.2	6.9		BOA:	93.45

Scotts Pi: 0.868
Krippendorfs alpha: 0.869
Cohens kappa: 0.868

SkyCam 2 manual classification vs. SkyCam 2 auto classification

Sky Camera 2 manual classification

Sky Camera 2 automatic classification	Clear	Cloud	Sum	U A	E
CLEAR	38	1	39	97.4	2.6
CLOUD	1	26	27	96.3	3.7
Sum	39	27	66		
P A	97.4	96.3		OA:	96.97
E	2.6	3.7		BOA:	96.85

Scotts Pi: 0.937
Krippendorfs alpha: 0.937
Cohens kappa: 0.937

S2 Validation results – automatic SC classification

- Sentinel-2 results between 12.02.2021 and 12.02.2022
- OA is between 86% and 88%.
- These numbers are quite comparable with the validation results of sen2cor during the CMIX exercise

Sky Camera 1 automatic classification vs. Sentinel-2 L2A SCL (8', 9, '10)

Sky Camera 1

Sentinel-2 L2A	Class	Clear	Cloud	Sum	U A	E
	CLEAR	35	7	42	83.3	16.7
	CLOUD	2	24	26	92.3	7.7
	Sum	37	31	68		
	P A	94.6	77.4		OA:	86.76
	E	5.4	22.6		BOA:	86.0

Scotts Pi: 0.728
Krippendorfs alpha: 0.73
Cohens kappa: 0.729

'Sky Camera 2 automatic classification vs. Sentinel-2 L2A SCL (8', 9, '10)'

Sky Camera 2

Sentinel-2 L2A	Class	Clear	Cloud	Sum	U A	E
	CLEAR	36	5	41	87.8	12.2
	CLOUD	3	23	26	88.5	11.5
	Sum	39	28	67		
	P A	92.3	82.1		OA:	88.06
	E	7.7	17.9		BOA:	87.2

Scotts Pi: 0.751
Krippendorfs alpha: 0.753
Cohens kappa: 0.752

Label	Classification
0	NO_DATA
1	SATURATED_OR_DEFECTIVE
2	DARK_AREA_PIXELS
3	CLOUD_SHADOWS
4	VEGETATION
5	NOT_VEGETATED
6	WATER
7	UNCLASSIFIED
8	CLOUD_MEDIUM_PROBABILITY
9	CLOUD_HIGH_PROBABILITY
10	THIN_CIRRUS
11	SNOW

S2 Validation results – manual SC classification

- Sentinel-2 results between 12.02.2021 and 12.02.2022
- OA is between 86% and 88%.
- The results for SC1 completely match those of the automatic classification, while the results for SC2 differ a tiny bit.

Sky Camera 1 manual classification vs. Sentinel-2 L2A SCL (8', 9, '10)'

Sky Camera 1 manual

Sentinel-2 L2A	Class	Clear	Cloud	Sum	U A	E
	CLEAR	35	7	42	83.3	16.7
	CLOUD	2	24	26	92.3	7.7
	Sum	37	31	68		
	P A	94.6	77.4		OA:	86.76
	E	5.4	22.6		BOA:	86.0

Scotts Pi: 0.728
Krippendorfs alpha: 0.73
Cohens kappa: 0.729

Sky Camera 2 manual classification vs. Sentinel-2 L2A SCL (8', 9, '10)

Sky Camera 2 manual

Sentinel-2 L2A	Class	Clear	Cloud	Sum	U A	E
	CLEAR	37	5	42	88.1	11.9
	CLOUD	3	24	27	88.9	11.1
	Sum	40	29	69		
	P A	92.5	82.8		OA:	88.41
	E	7.5	17.2		BOA:	87.65

Scotts Pi: 0.759
Krippendorfs alpha: 0.761
Cohens kappa: 0.759

Label	Classification
0	NO_DATA
1	SATURATED_OR_DEFECTIVE
2	DARK_AREA_PIXELS
3	CLOUD_SHADOWS
4	VEGETATION
5	NOT_VEGETATED
6	WATER
7	UNCLASSIFIED
8	CLOUD_MEDIUM_PROBABILITY
9	CLOUD_HIGH_PROBABILITY
10	THIN_CIRRUS
11	SNOW

Comparison between RAP and SC2 (Fermi) automatic classification

- The result shows a comparably low agreement (below 80%).
- This result was a bit surprising.
- Comparison with manual classification needed

Sky Camera 2 automatic classification vs. RAP cloud top

Sky Camera 2

RAP \ Sky Camera 2	Clear	Cloud	Sum	U A	E
CLEAR	19	5	24	79.2	20.8
CLOUD	3	11	14	78.6	21.4
Sum	22	16	38		
P A	86.4	68.8		OA:	78.95
E	13.6	31.2		BOA:	77.6

Scotts Pi: 0.559
Krippendorfs alpha: 0.565
Cohens kappa: 0.56

Comparison between RAP and SC2 (Fermi) manual classification

- Agreement increased to above 84% OA
- Nevertheless, the agreement was lower than expected.
- Further analysis was required

SkyCam 2 manual classification vs. RAP

Sky Camera 2 manual classification for RAP position

RAP \ Sky Camera 2	Clear	Cloud	Sum	U A	E
CLEAR	20	3	23	87.0	13.0
CLOUD	2	8	10	80.0	20.0
Sum	22	11	33		
P A	90.9	72.7		OA:	84.85
E	9.1	27.3		BOA:	81.8

Scotts Pi: 0.65
Krippendorfs alpha: 0.656
Cohens kappa: 0.651

	date	time_x	skycam_class	RAP_QF
0	20210316	101002	255	11
1	20210321	101002	255	11
2	20210326	101002	0	0
3	20210405	101002	255	11
4	20210410	101002	255	11
5	20210415	101002	100	11
6	20210420	101003	0	0
7	20210425	101002	0	0
9	20210430	101002	0	0
11	20210505	101002	0	0
12	20210510	101002	0	0
13	20210515	101002	255	0
14	20210520	101002	0	0
15	20210525	101002	0	0
16	20210604	100902	0	0
17	20210609	100902	0	11
18	20210614	100902	0	0
19	20210619	100902	255	0
20	20210624	100903	255	10
21	20210629	100902	255	10
22	20210704	100902	100	10
24	20210709	100902	0	0
26	20210714	100902	255	0
28	20210719	100902	0	0
29	20210724	100902	0	0
30	20210729	100902	0	0
31	20210803	100902	0	11
32	20210808	100902	0	0
33	20210813	100902	100	10
34	20210818	100902	0	0
35	20210823	100902	100	0
36	20210828	100902	0	11
37	20210902	100902	0	0
38	20210907	100902	0	0
39	20210917	100902	255	11
40	20210922	100902	0	0
42	20210927	100902	255	0

Comparison between RAP and SC2 (Fermi) manual classification

- Tables shows matchup between RAP QF flag (RAP_QF) and classification of SC 2 (skycam_class)
- The red marked entries show disagreements in the classification
- The sky camera data for those dates have been analyzed.

SkyCam 2 manual classification vs. RAP

Sky Camera 2 manual adjusted classification for RAP position

RAP \ Class	Clear	Cloud	Sum	U A	E
CLEAR	21	2	23	91.3	8.7
CLOUD	0	9	9	100.0	0.0
Sum	21	11	32		
P A	100.0	81.8		OA:	93.75
E	0.0	18.2		BOA:	90.9

Scotts Pi: 0.854
Krippendorfs alpha: 0.856
Cohens kappa: 0.855

Comparison between RAP and SC2 (Fermi) manual classification

- The most likely explanation is the location difference of 22m between the two instruments and RAP observation is a bit tilted.
- A red/green cross marks the potential location of the RAP acquisition within the SC image
- The potential location of the RAP acquisition has been manually classified for all SC2 data, to ensure a “true” comparison between the two instruments.

Sentinel-2 L2A cloud mask over SkyCam 1 manual L1C classification

In-Situ Database

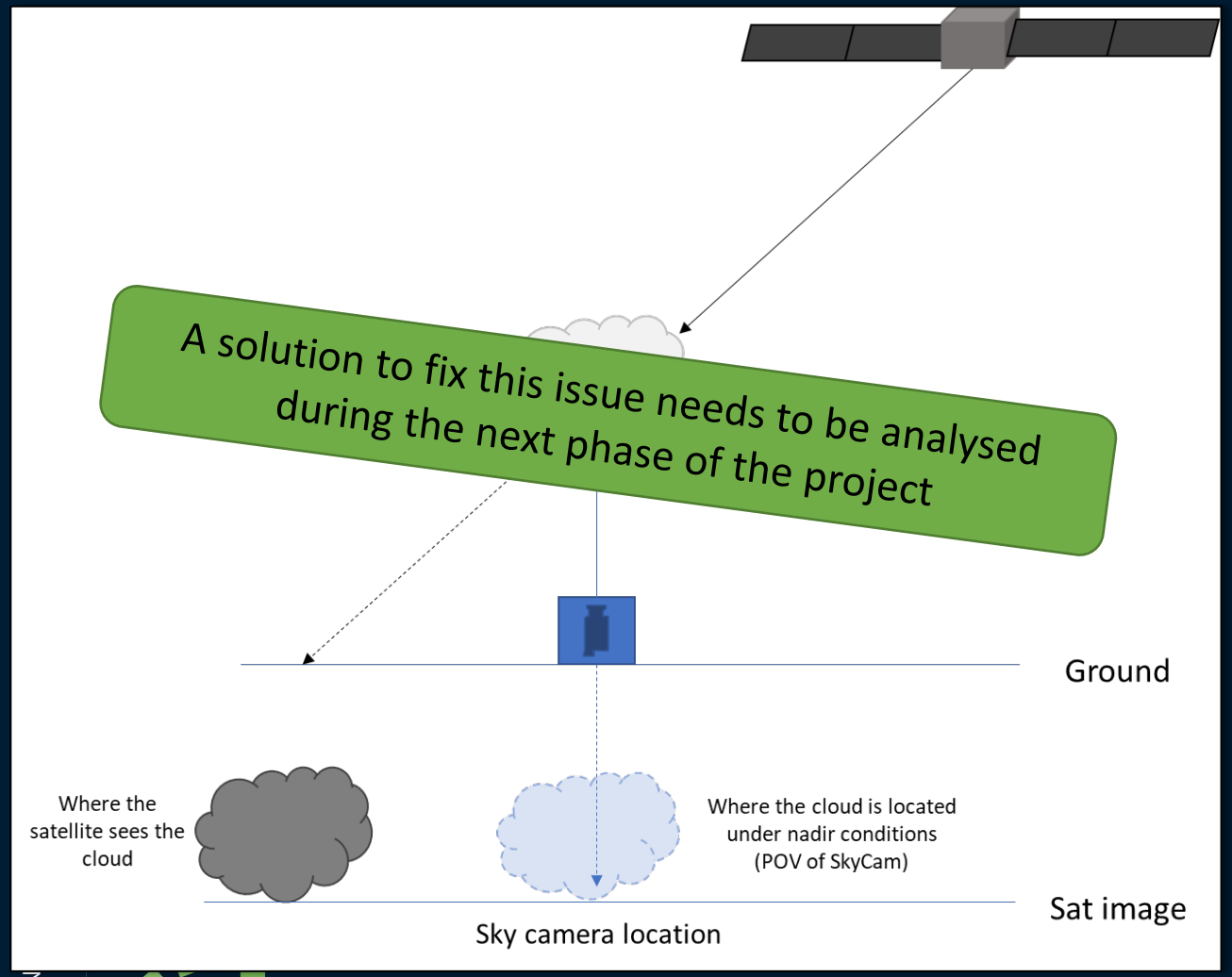
Sentinel-2 L2A	Class	Clear	Cloud	Sum	U A	E
	CLEAR	24	5	29	82.8	17.2
	CLOUD	0	14	14	100.0	0.0
	Sum	24	19	43		
	P A	100.0	73.7		OA:	88.37
	E	0.0	26.3		BOA:	86.85

Scotts Pi: 0.754
Krippendorfs alpha: 0.757
Cohens kappa: 0.757

Limitations

- To eliminate the bias from the S2 L2A scene classification and to compare clouds visible in the satellite image and the sky camera, a subset of the above used S2 data was manually classified for the SC1 location.
- The OA is still below 90%.
- Therefore, the question arose why there is no better agreement.
- S2 products and SC1 (as well as SC2) data for cases without matching classifications have been compared.

Limitations

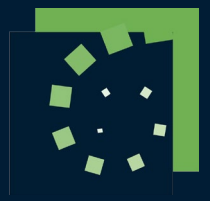


- The images show that the cloud in the center of SC2 (Fermi) is located northeast of SC2 in the S2 L2A image.
- While the same cloud is located southwest of the center of SC1 (Marconi) and south/over SC1 in the S2 L2A product.
- The cause for this mismatch can be explained by the viewing differences of the three instruments and the location of the cloud above ground.
- The S2 L2A data have been acquired off-nadir with a VAA mean of 130.28053 and a VZA mean of 3.3807745 (purple arrow viewing direction of S2 MSI).
- The parallax between true nadir and the actual S2 location cause the cloud to be projected in north-western direction onto the ground
- Reducing the S2 and L8 observation close to nadir might help circumventing this issue

Conclusion from experimental operations

- Sky camera data provide an interesting and valuable reference source for comparison
- The strength of the data is
  - the constant acquisition (leading to a dataset with a high temporal resolution),
  - quite high classification accuracy that could be achieved by the RF classifier,
  - the comparable low costs for the instrument
- While the validation or better intercomparison results had shown a quite good agreement between the SC classification and the satellite (S2 & L8) cloud masks, the study had also revealed geometric issues that can lead to incomparability between SC and satellite data.
- Further studies are needed to analyse if these issues/disagreements can be circumvented/corrected.

Thank you for the attention!



BROCKMANN CONSULT