QA4EO

Support to cloud mask validation for CMIX

CAL/VAL WORKSHOP 01.04.2022

JAN WEVERS

BROCKMANN CONSULT GMBH

Overview

- Objective / WP Overview
- Requirements & state of the art analysis
- <u>Validation site and methods preparation</u>
- Experimental operation
- Validation results
- Identified issues
- Roadmap

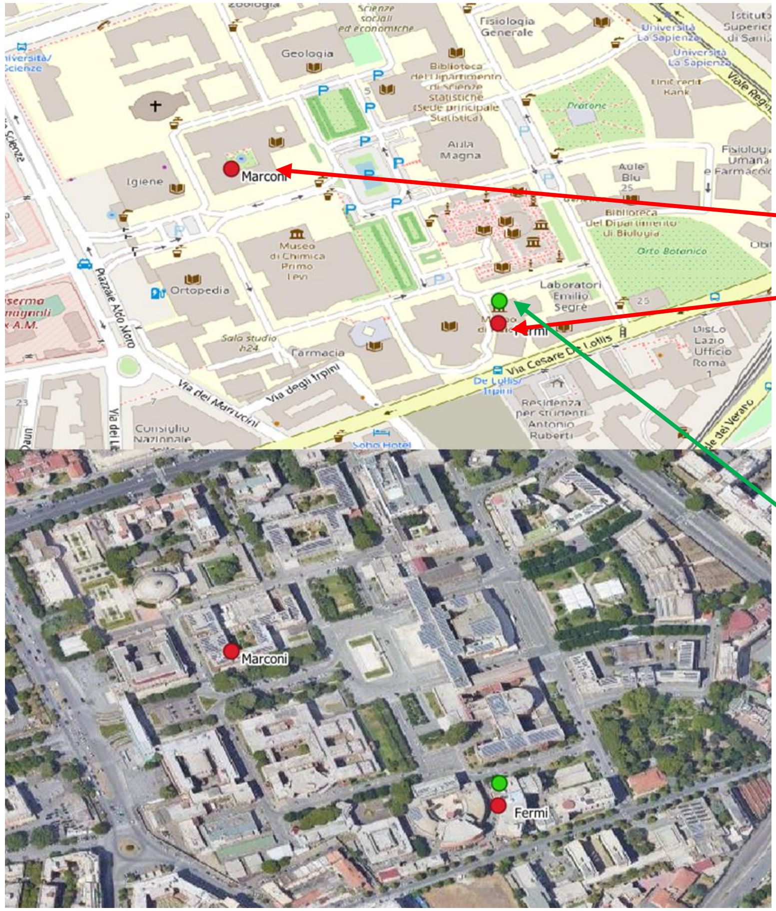
Objective / Overview

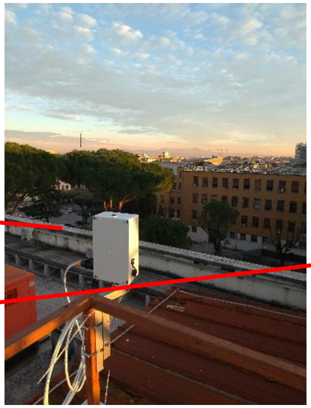
- The objective was to prepare for establishing and for operating a network of cloud mask validation sites.
- The goal of this work package was to prototype algorithms and methods to process sky camera data and compare them with satellite algorithms for cloud masking.
- There are two approaches for validation that have been planned to be compared:
  1. using stereo sky camera (SC) data and
  2. a ceilometer (RAP).
- The work package 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 Raspberry Pi 4 and the Omnivision OV5647 sensor. The field of view is 194 (horizontal) and 142 (vertical). Distance between cameras is around 260 meters. Currently, the cameras are collecting data every minute between 08:00 and 14:00 UTC.
- Sky camera two (Fermi) is located 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

Raymetrics Aerosol Profiler (RAP)

Validation site and methods preparation

- Setup phase activities
  - Setup data transfer of Rome SC data form UoM to BC
  - Setup of server and archiving system to store the SC data
  - Develop scripts to find matching S2/L8 data with SC site location
  - Develop scripts to find matching SC data to the S2/L8 acquisitions
  - Develop preprocessing methods for SC data (crop/flip/rotate)
  - Finding an approriate classification method for SC data
  - Classifier training
  - Development of scripts for SC classification & sample extraction
  - Development of scripts to create previews and Geotiffs
  - Development of scripts to create confusion matrices

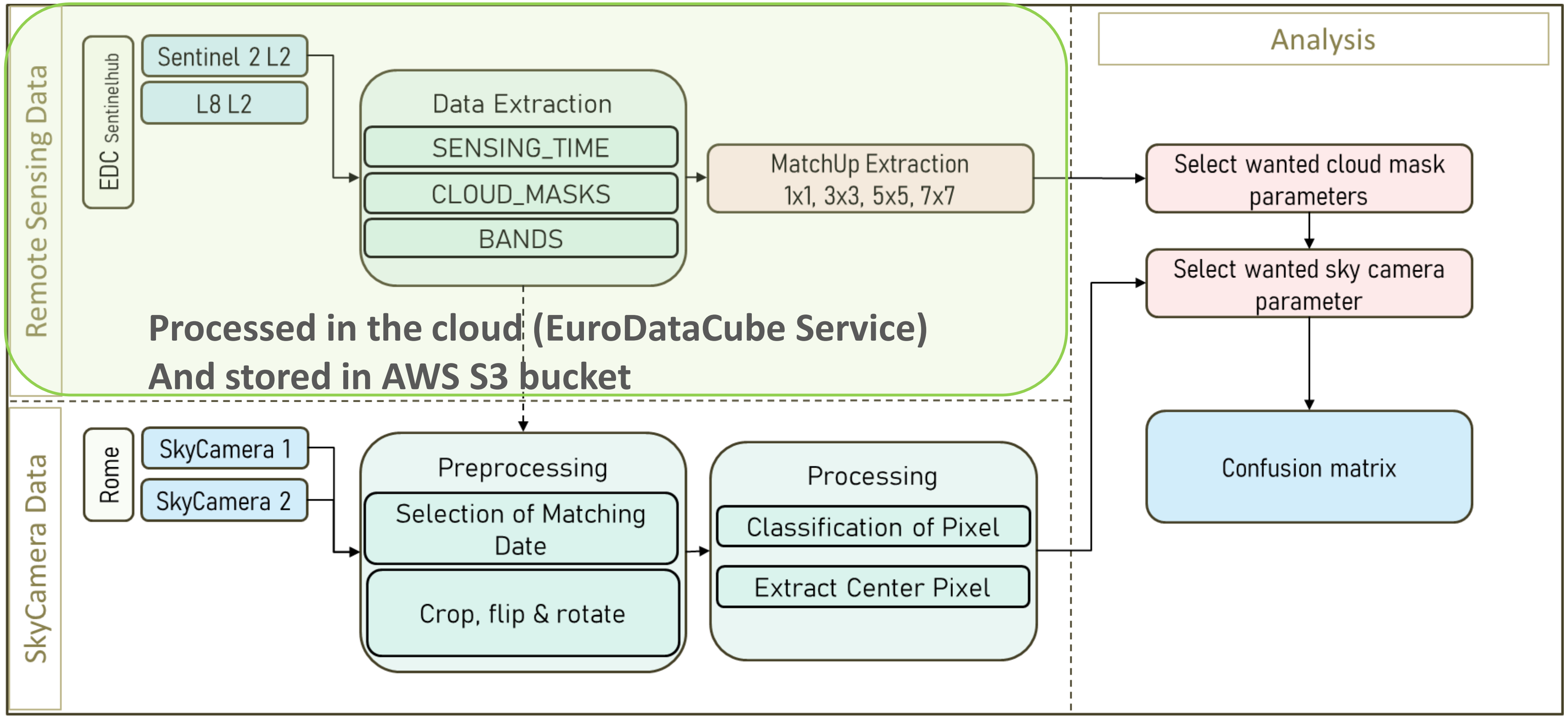
Validation site and methods preparation

- Data transfer and archiving server
  - A dedicated FTP server and archiving system was implemented at Brockmann Consult (BC) premises.
  - The sky camera data are collected by UoM from La Sapienza University (LSU) using rsync
  - Afterwards the data is again transferred via rsync from UoM to a BC server.
  - Due to the time difference between the US and Europe and rsync being executed only once a day, there is a delay of one day in the data availability.
  - Since a direct data transfer between the LSU and BC cannot be implemented (data are property of UoM) this delay cannot be circumvented at this stage.

  - *Note: The data amount of the complete SC archive is quite big*
    - *29 acquisitions between 10:00h and 10:29h (rough S2/L8 overpass window) for both SCs between 12.02.2021 and 16.03.2022 correspond to 135 GB of data.*
    - *The complete archive already exceeds 1TB at the moment*

Validation site and methods preparation

- Processing overview



Remote Sensing Data

EDC Sentinelhub

Sentinel 2 L2

L8 L2

Data Extraction

SENSING_TIME

CLOUD_MASKS

BANDS

MatchUp Extraction
1x1, 3x3, 5x5, 7x7

Processed in the cloud (EuroDataCube Service)
And stored in AWS S3 bucket

SkyCamera Data

Rome

SkyCamera 1

SkyCamera 2

Preprocessing

Selection of Matching Date

Crop, flip & rotate

Processing

Classification of Pixel

Extract Center Pixel

Analysis

Select wanted cloud mask parameters

Select wanted sky camera parameter

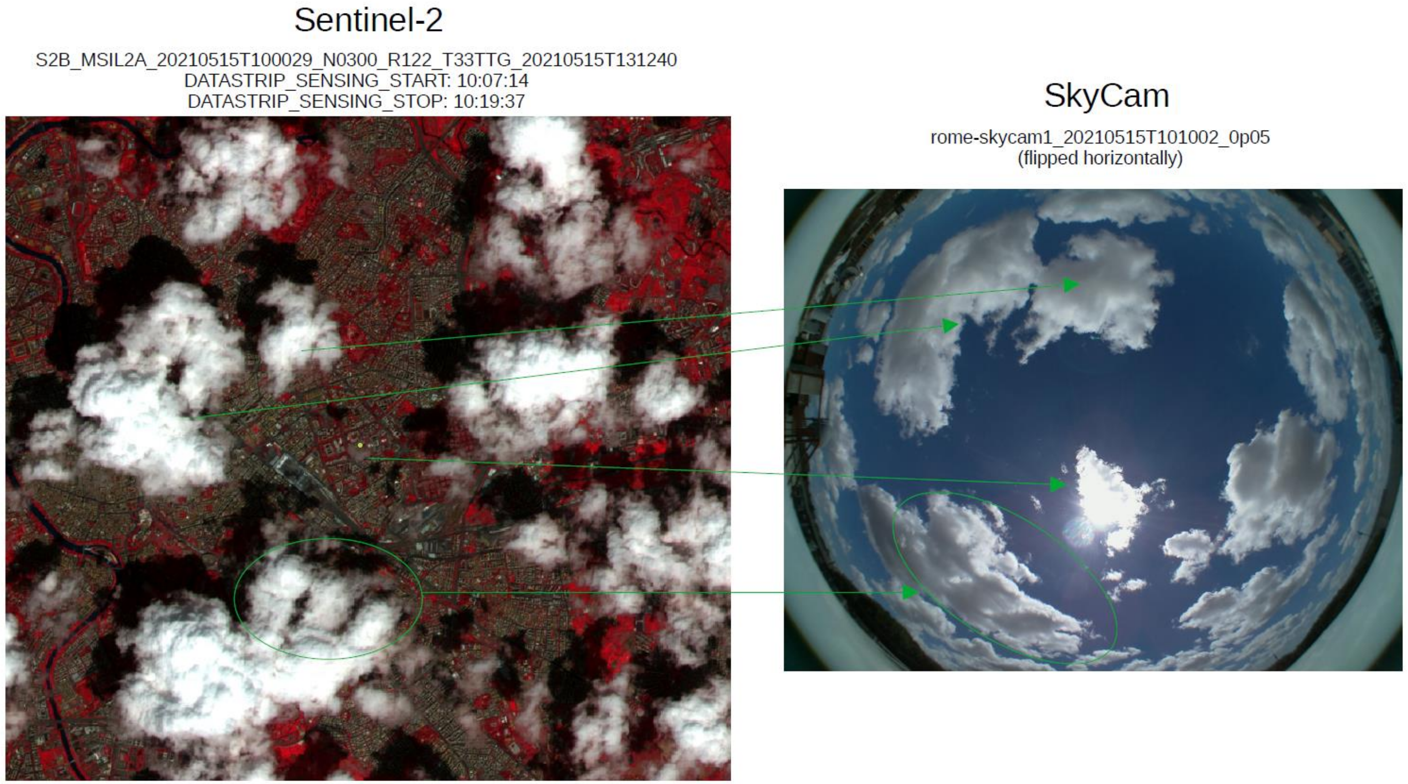
Confusion matrix

Validation site and methods preparation

Pre-processing of sky camera data (crop, flip, rotate)

- The matching SC images show the complete FOV of the camera, which is quite a lot geometrically distorted outside of the center.
- The upper part of the image does not represent north, since the SCs are installed looking a bit northwest.
- Compared to the satellite acquisitions, the images are flipped left to right, since the camera is looking from the ground upwards and the sensor das the opposite

Validation site and methods preparation



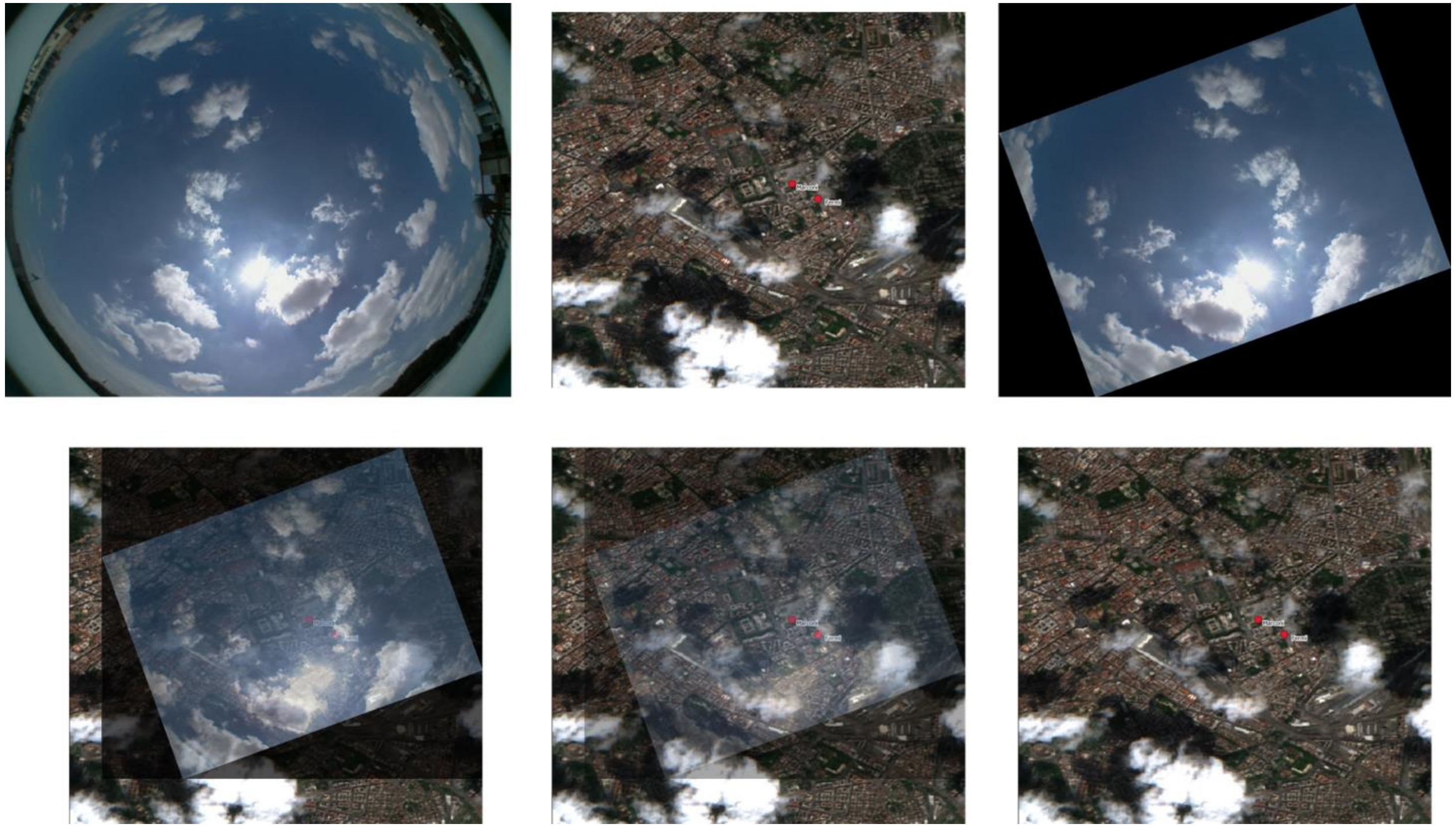
Pre-processing of sky camera data (crop, flip, rotate)

- The image shows a comparison between a Sentinel-2 acquisition and the corresponding complete SC image flipped left to right
- The example shows the drastic distortion caused by the fisheye lens and the difference in orientation

Brockmann
Consult GmbH

QA4EO

Validation site and methods preparation



Pre-processing of sky camera data (crop, flip, rotate)

- To compensate for all this, methods have been developed and implemented in Python, to
    - crop the image to the center part, which is less affected by distortion,
    - flipped horizontally and then
    - rotated to match the cardinal directions properly and to allow direct comparison with the satellite data

Validation site and methods preparation

Figure 14: SC 1 image 26.03.2021

Figure 15: Classification using Otsu threshold after Gaussian filtering

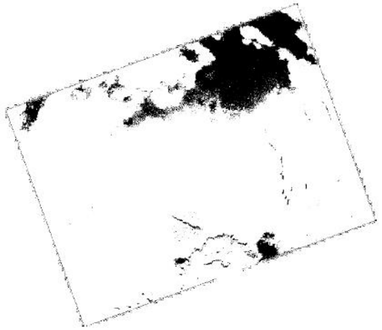


Figure 16: Classification using BI SI method from Letu et al. 2014

<u>Finding an appropriate classification method</u>

- It was not intended to develop a classification method for the SC data within the scope of the project. But due to the prior described pandemic induced delays, a solution needed to be found.
- A few methods have been tested that have not led to required accuracies.
  - Simple threshold on a greyscale representation of the RGB image
  - Otsu thresholding
  - Otsu thresholding after Gaussian filtering
  - Implementing a linear light filter, to enhance the contrasts in the images to improve the results of the prior three methods
  - Brightness index (BI), Sky index (SI) method by Letu et al. 2014

Validation site and methods preparation

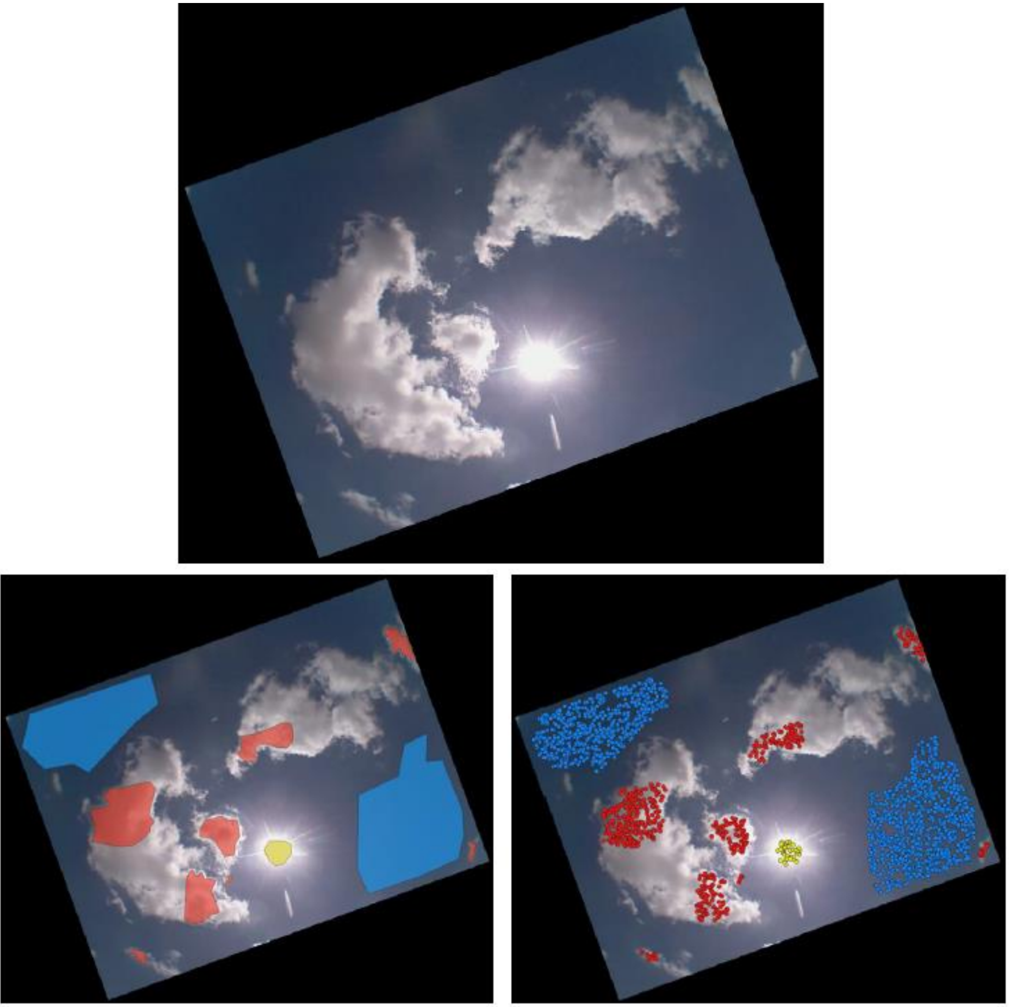
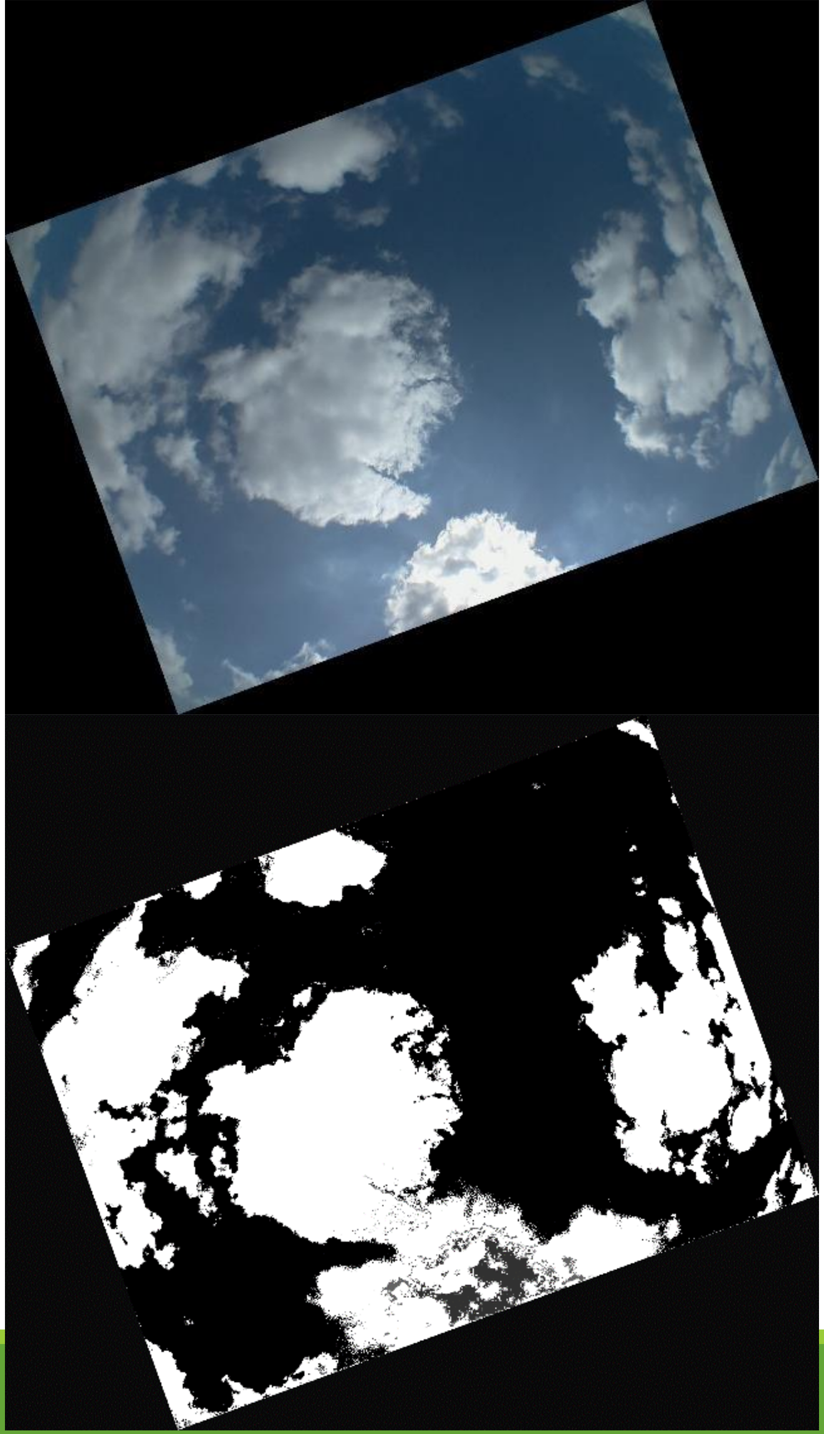


Figure 17: Training sample generation. SC image (top), training polygons (lower left), training samples (lower right)

Finding an appropriate classification method

- Random forest classifier
  - Since none of these methods had reached the necessary accuracies, it was tested if a random forest (RF) classifier can be trained for each camera, to achieve the necessary quality in classification.
- Generation of training samples:
  - 12 to 15 SC images per SC have been selected
  - Polygons representing the same class have been drawn on the SC images.
  - Inside these polygons random samples have been generated.
  - Overall, 11,100 samples for SC1 and 27,300 samples for SC 2 have been collected.



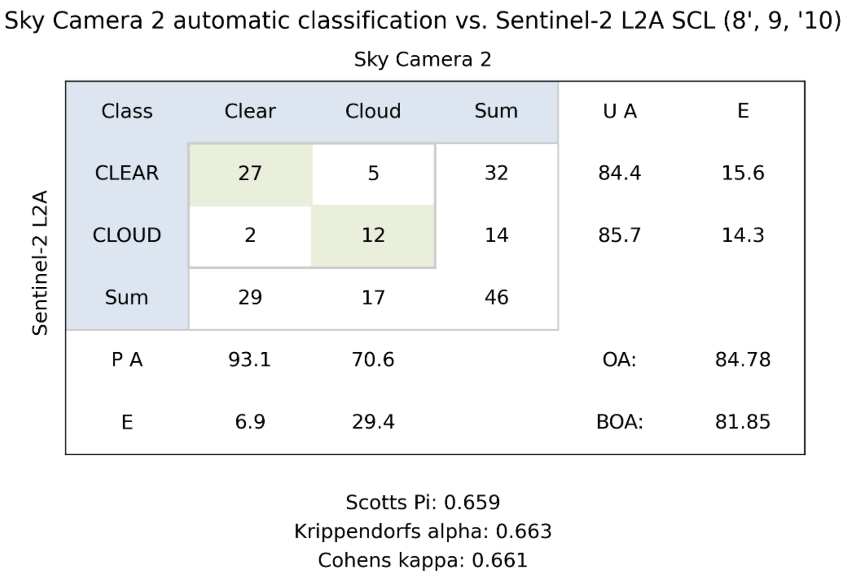
Validation site and methods preparation

Finding an appropriate classification method

- The following classes have been trained
  - 0 = Clear
  - 10 = NoData
  - 50 = Sun
  - 100 = Thin clouds (cirrus)
  - 255 = Opaque clouds
- Even though there are some smaller omissions and commissions, the overall accuracy is quite high and much better compared to the previously tested methods.

Validation site and methods preparation

Confusion matrix example



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	27	5	32	84.4	15.6
CLOUD	2	12	14	85.7	14.3
Sum	29	17	46		
P A	93.1	70.6		OA:	84.78
E	6.9	29.4		BOA:	81.85

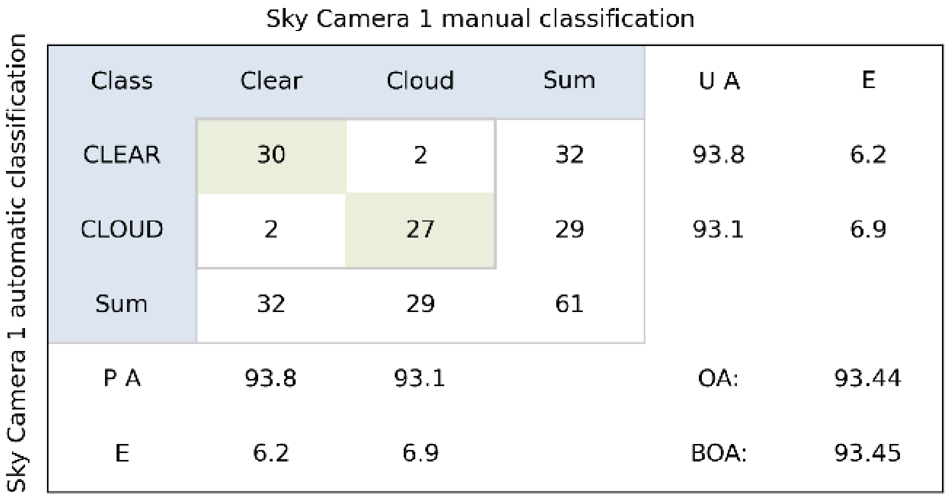
Scotts Pi: 0.659
Krippendorfs alpha: 0.663
Cohens kappa: 0.661

Validation

- A tool has been developed to generate confusion matrices between the sky camera classifications (used as reference) and the satellite cloud mask (product to be validated).
- This tool
  - harnesses the satellite data extractions and sky camera classification extractions stored in a csv file,
  - joints the data based on dates, and
  - automatically plots confusion matrices

Validation results

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



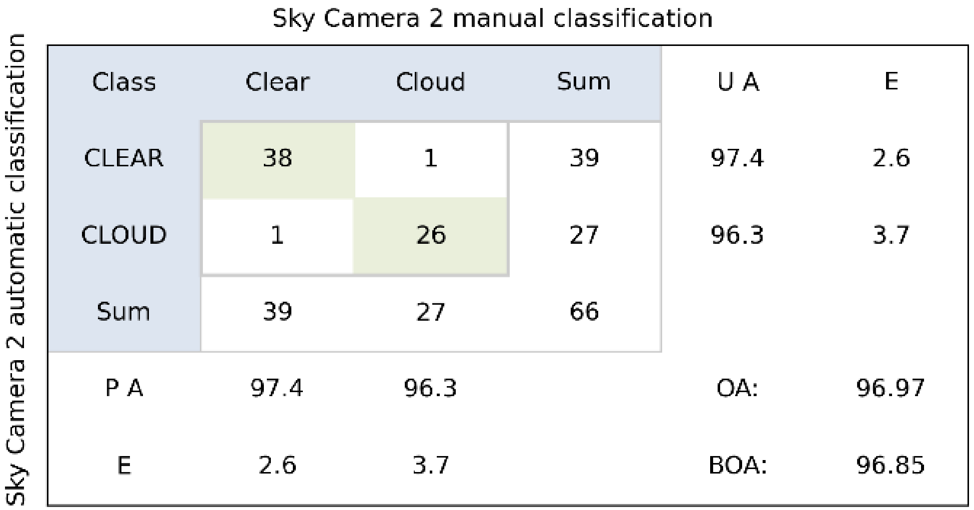
SkyCam 1 manual classification vs. SkyCam 1 auto classification

Sky Camera 1 manual classification

Sky Camera 1 automatic classification

Class	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

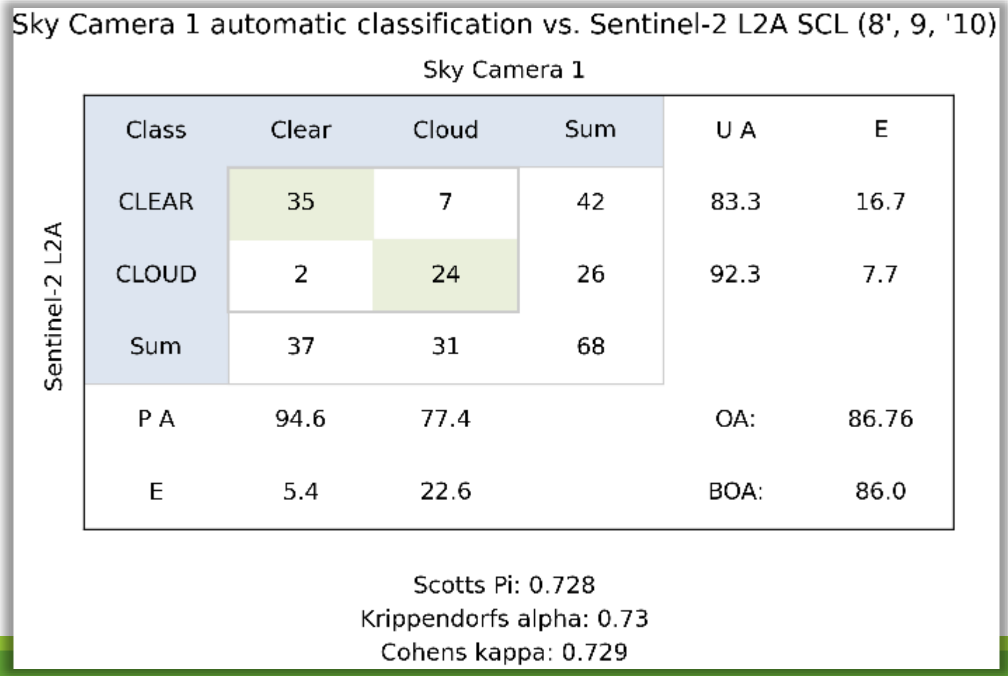
Sky Camera 2 automatic classification

Class	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



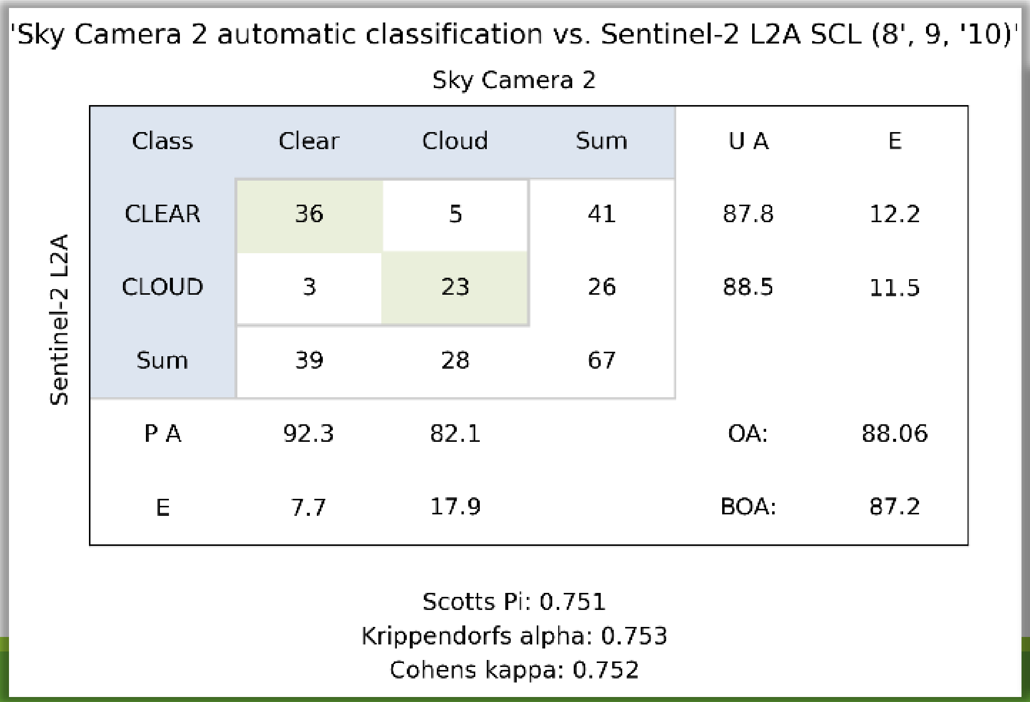
Sky Camera 1 automatic classification vs. Sentinel-2 L2A SCL (8', 9, '10)

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



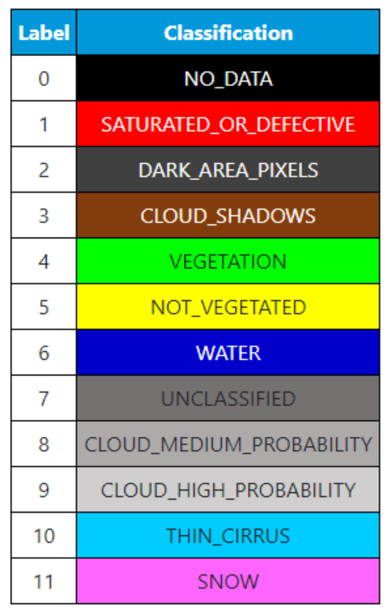
'Sky Camera 2 automatic classification vs. Sentinel-2 L2A SCL (8', 9, '10)'

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

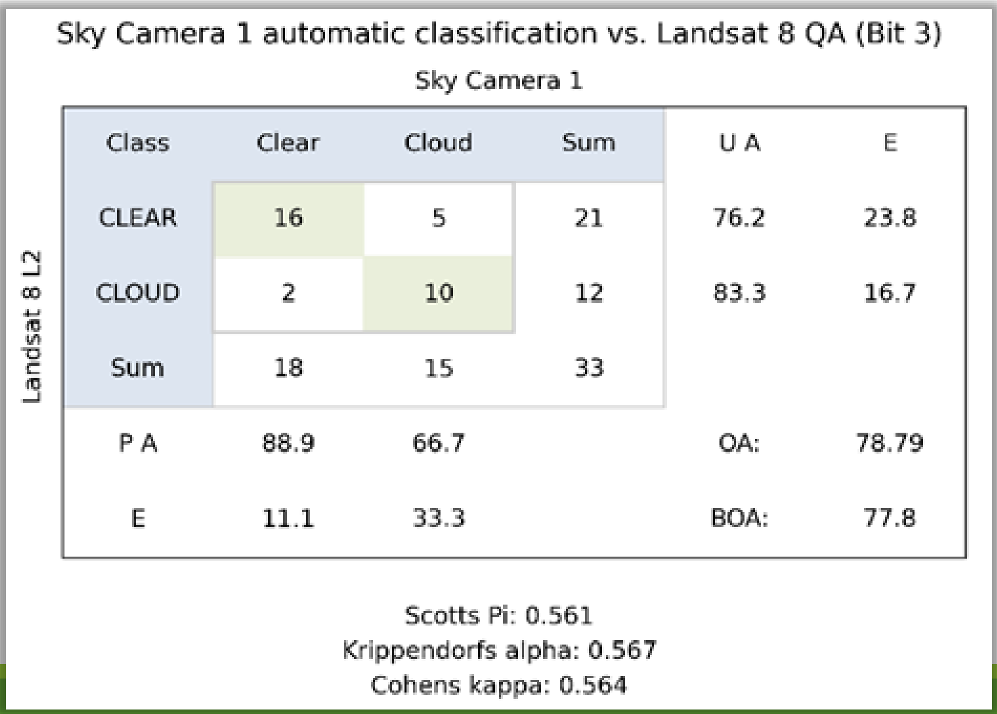
Scotts Pi: 0.751
Krippendorfs alpha: 0.753
Cohens kappa: 0.752



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

L8 Validation results – automatic SC classification

- Landsat 8 L2 BQA cloud mask (Bit 3) results between 12.02.2021 and 12.02.2022
- OA is between 78% and 80%.
- Again, these numbers are quite comparable with the validation results of LaSRC during the CMIX exercise for the PixBox dataset.



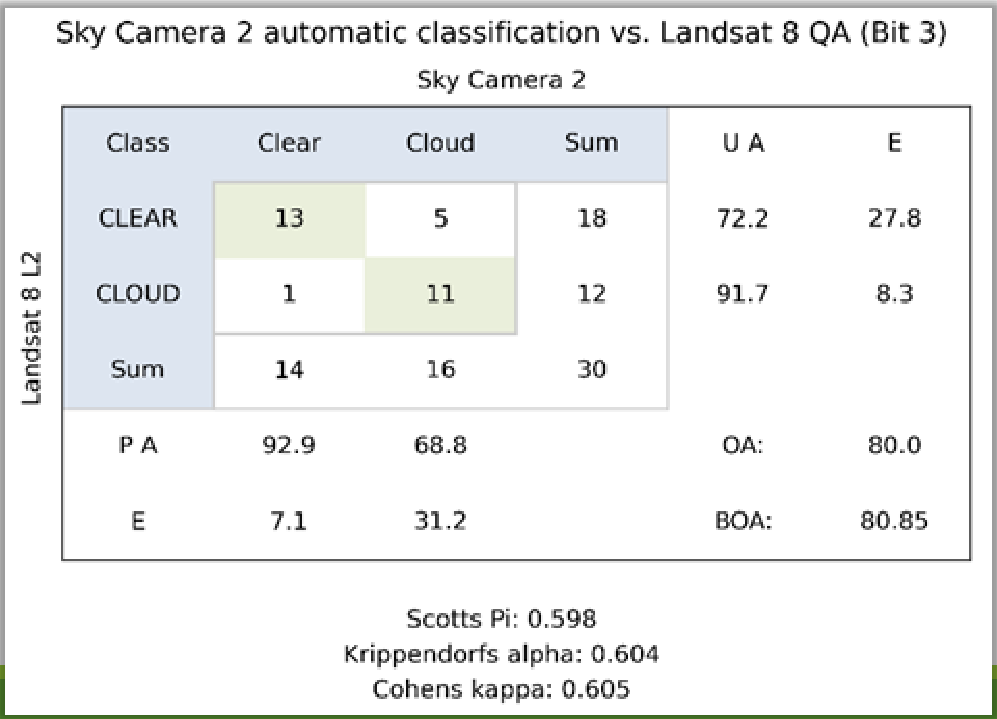
Sky Camera 1 automatic classification vs. Landsat 8 QA (Bit 3)

Sky Camera 1

Landsat 8 L2

Class	Clear	Cloud	Sum	U A	E
CLEAR	16	5	21	76.2	23.8
CLOUD	2	10	12	83.3	16.7
Sum	18	15	33		
P A	88.9	66.7		OA:	78.79
E	11.1	33.3		BOA:	77.8

Scotts Pi: 0.561
Krippendorfs alpha: 0.567
Cohens kappa: 0.564



Sky Camera 2 automatic classification vs. Landsat 8 QA (Bit 3)

Sky Camera 2

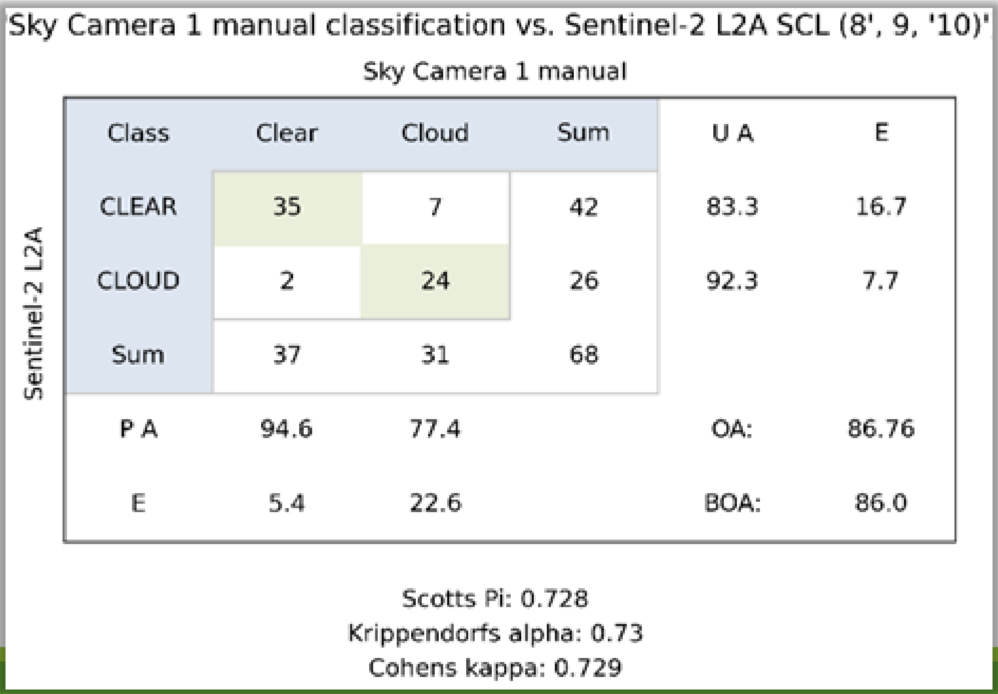
Landsat 8 L2

Class	Clear	Cloud	Sum	U A	E
CLEAR	13	5	18	72.2	27.8
CLOUD	1	11	12	91.7	8.3
Sum	14	16	30		
P A	92.9	68.8		OA:	80.0
E	7.1	31.2		BOA:	80.85

Scotts Pi: 0.598
Krippendorfs alpha: 0.604
Cohens kappa: 0.605

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.



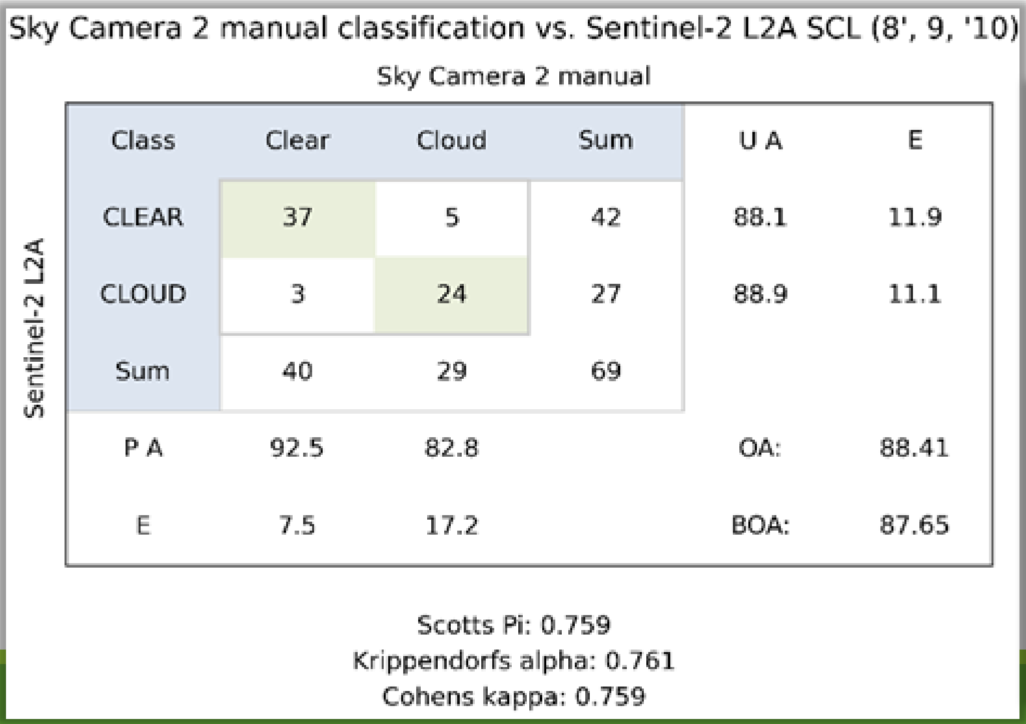
Sky Camera 1 manual classification vs. Sentinel-2 L2A SCL (8', 9, '10)'

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



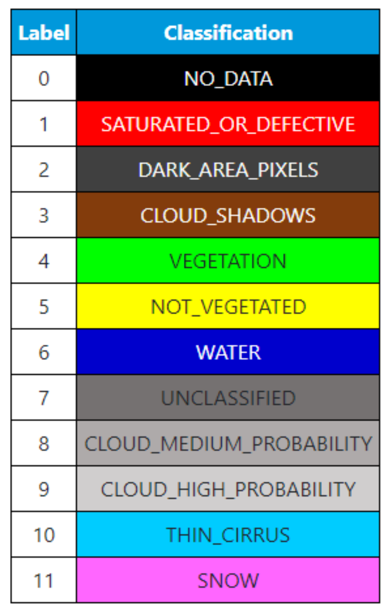
Sky Camera 2 manual classification vs. Sentinel-2 L2A SCL (8', 9, '10)

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

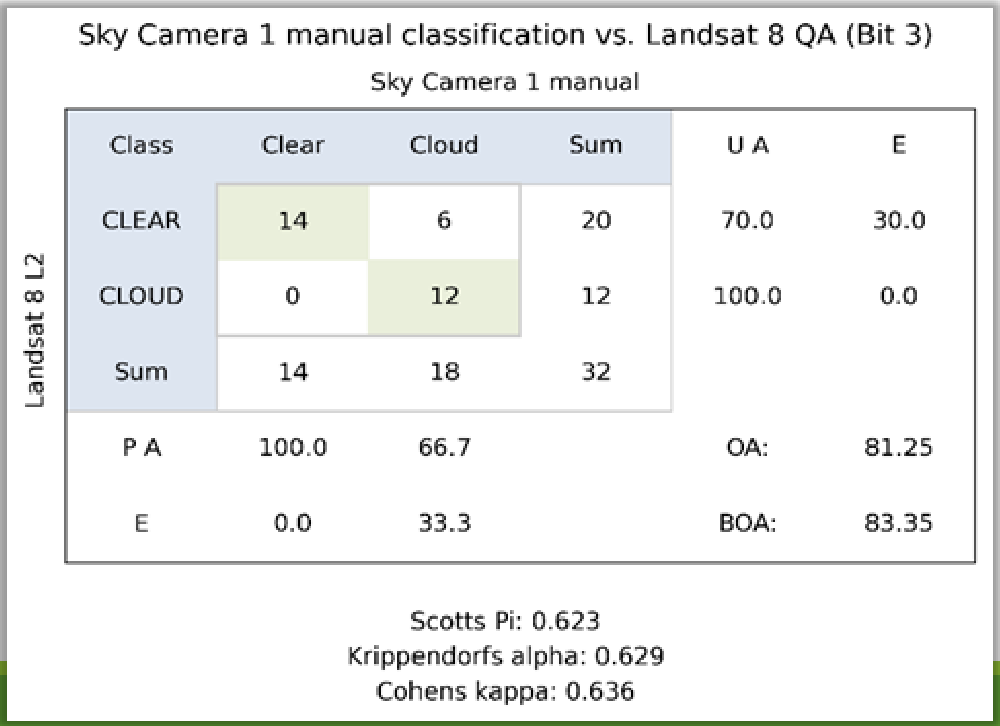
Scotts Pi: 0.759
Krippendorfs alpha: 0.761
Cohens kappa: 0.759



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

L8 Validation results – manual SC classification

- Landsat 8 L2 BQA cloud mask (Bit 3) results between 12.02.2021 and 12.02.2022
- OA is between 81% and 84%.
- The numbers for the manually classified SC data are a bit higher compared to the automatic classified results (78% and 80%).



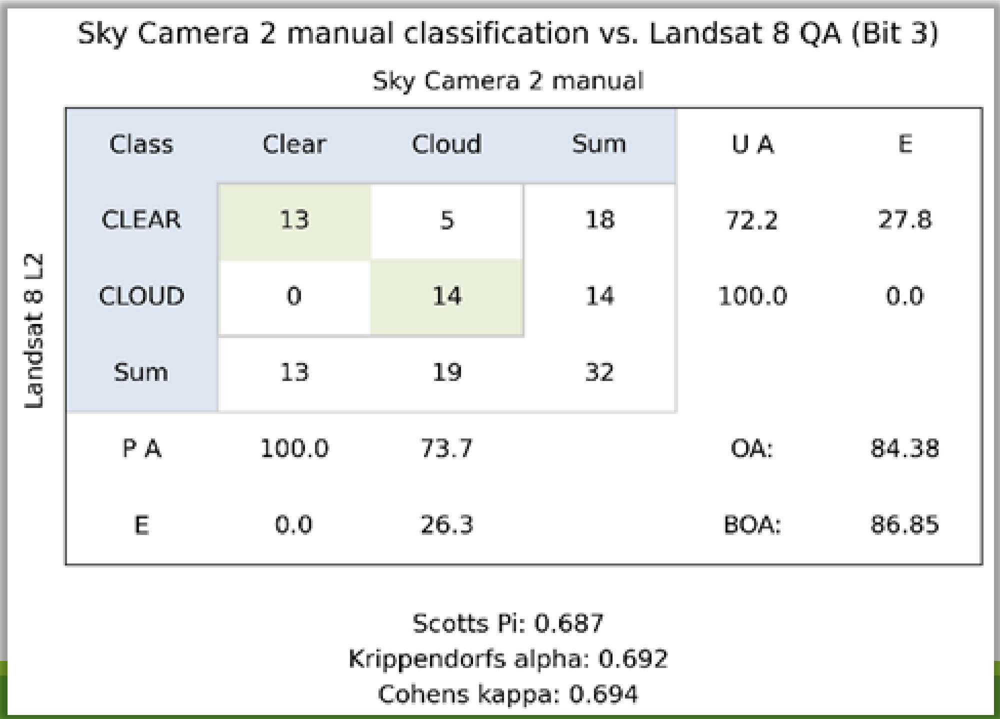
Sky Camera 1 manual classification vs. Landsat 8 QA (Bit 3)

Sky Camera 1 manual

Landsat 8 L2

Class	Clear	Cloud	Sum	U A	E
CLEAR	14	6	20	70.0	30.0
CLOUD	0	12	12	100.0	0.0
Sum	14	18	32		
P A	100.0	66.7		OA:	81.25
E	0.0	33.3		BOA:	83.35

Scotts Pi: 0.623
Krippendorfs alpha: 0.629
Cohens kappa: 0.636



Sky Camera 2 manual classification vs. Landsat 8 QA (Bit 3)

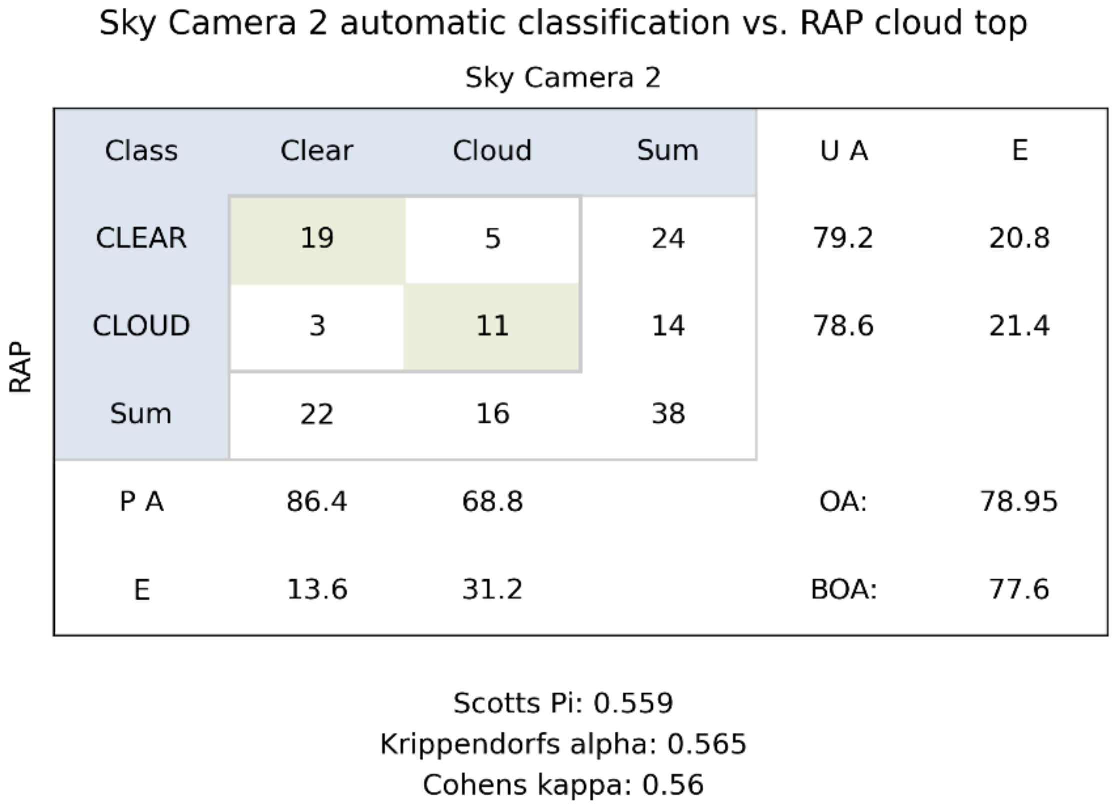
Sky Camera 2 manual

Landsat 8 L2

Class	Clear	Cloud	Sum	U A	E
CLEAR	13	5	18	72.2	27.8
CLOUD	0	14	14	100.0	0.0
Sum	13	19	32		
P A	100.0	73.7		OA:	84.38
E	0.0	26.3		BOA:	86.85

Scotts Pi: 0.687
Krippendorfs alpha: 0.692
Cohens kappa: 0.694

Comparison between RAP and SC2 (Fermi) automatic classification



Sky Camera 2 automatic classification vs. RAP cloud top

Sky Camera 2

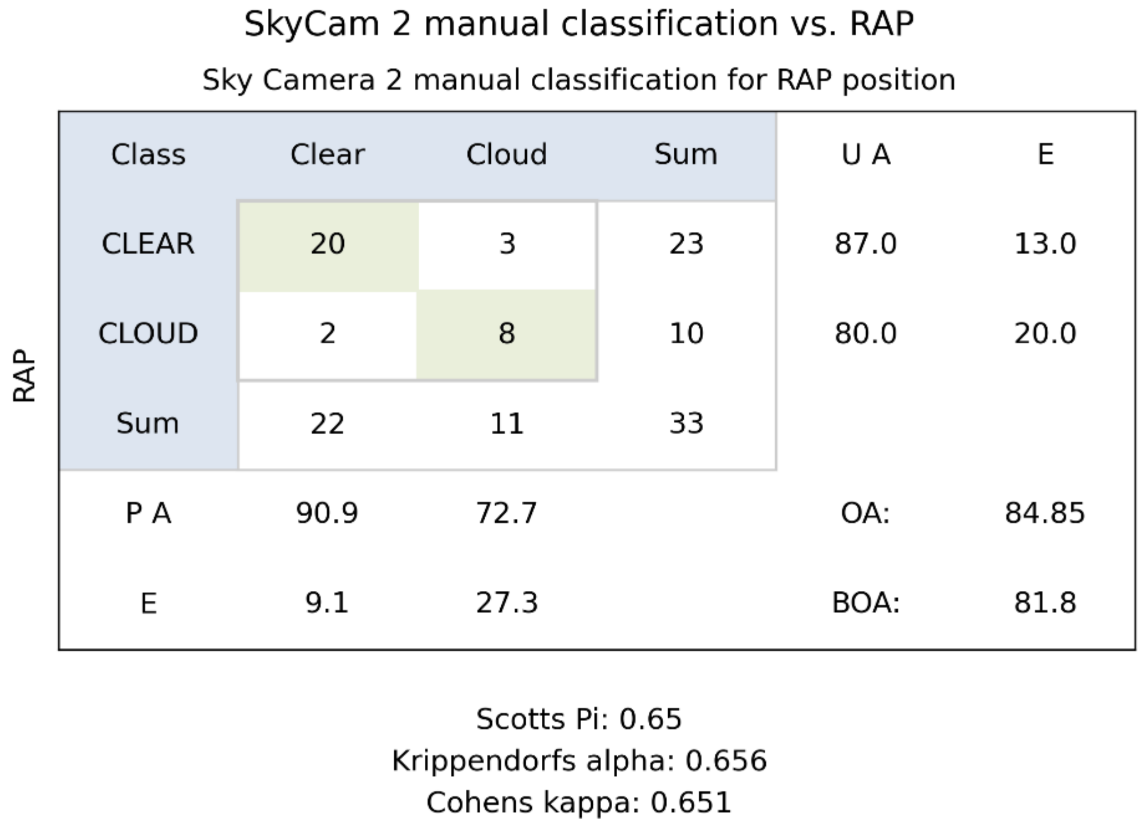
RAP

Class	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

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

Comparison between RAP and SC2 (Fermi) manual classification



SkyCam 2 manual classification vs. RAP

Sky Camera 2 manual classification for RAP position

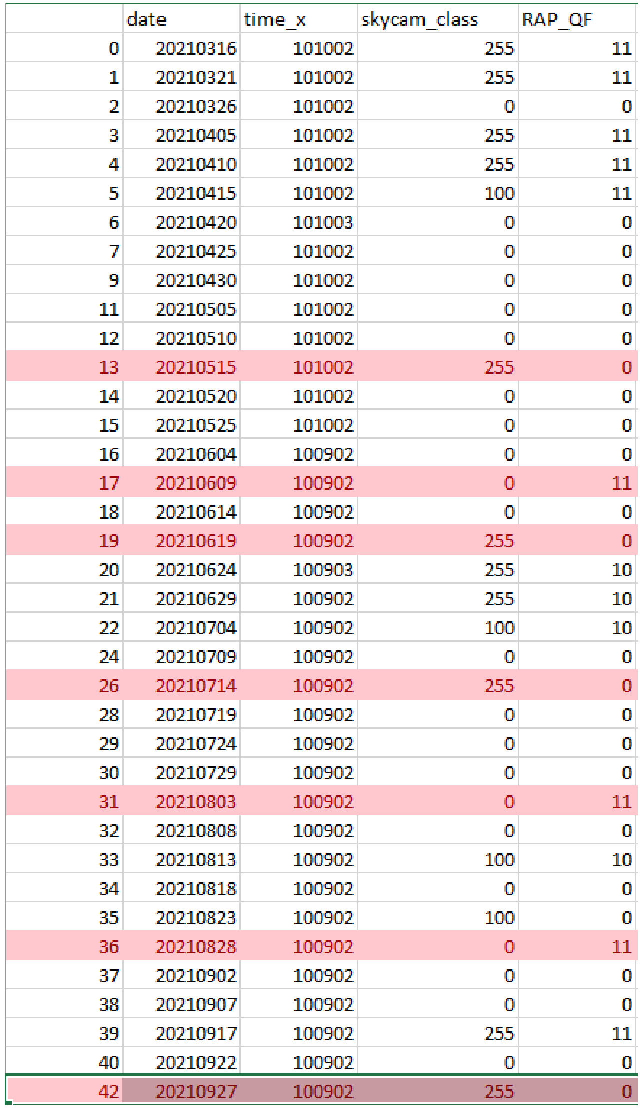
RAP

Class	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

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

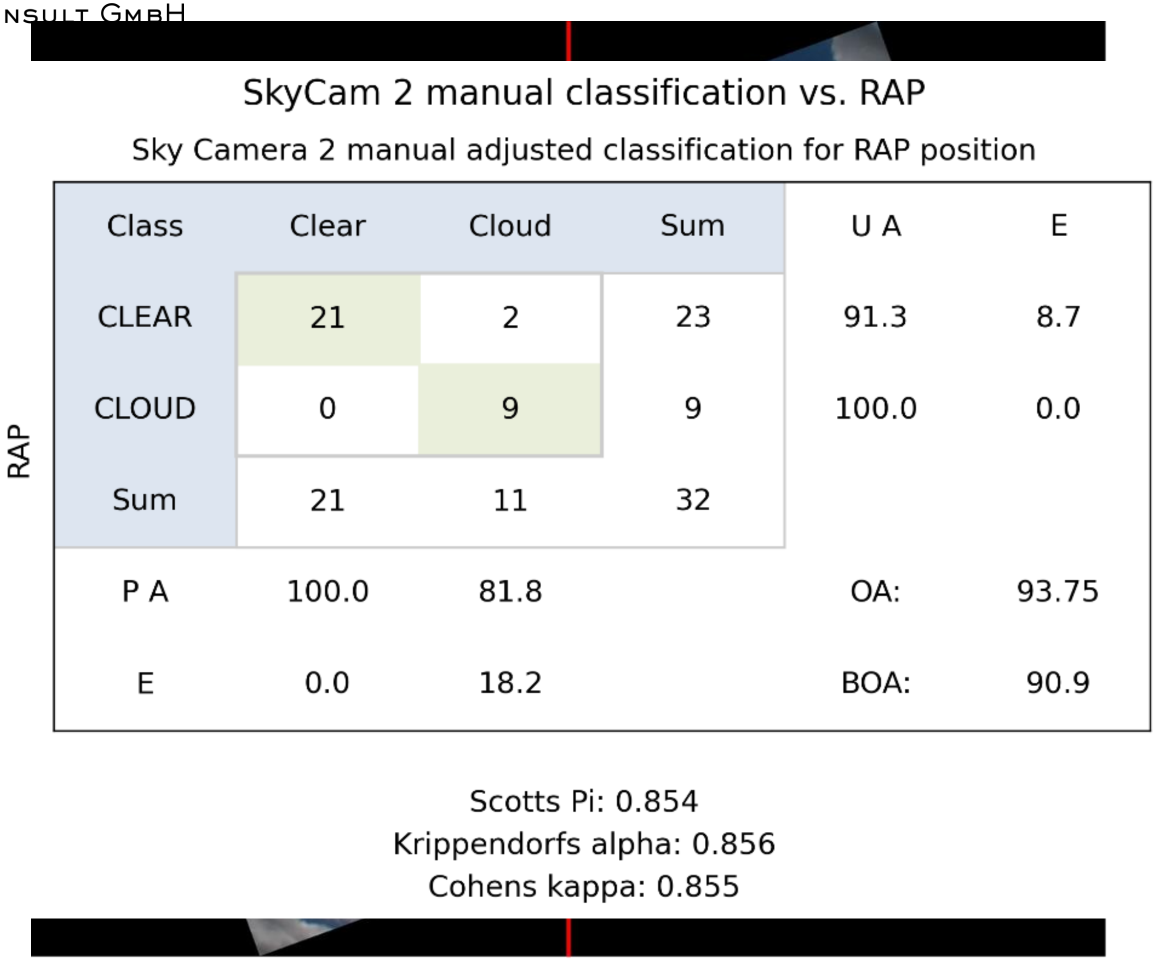
Comparison between RAP and SC2 (Fermi) manual classification



	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

- 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.

Comparison between RAP and SC2 (Fermi) manual classification



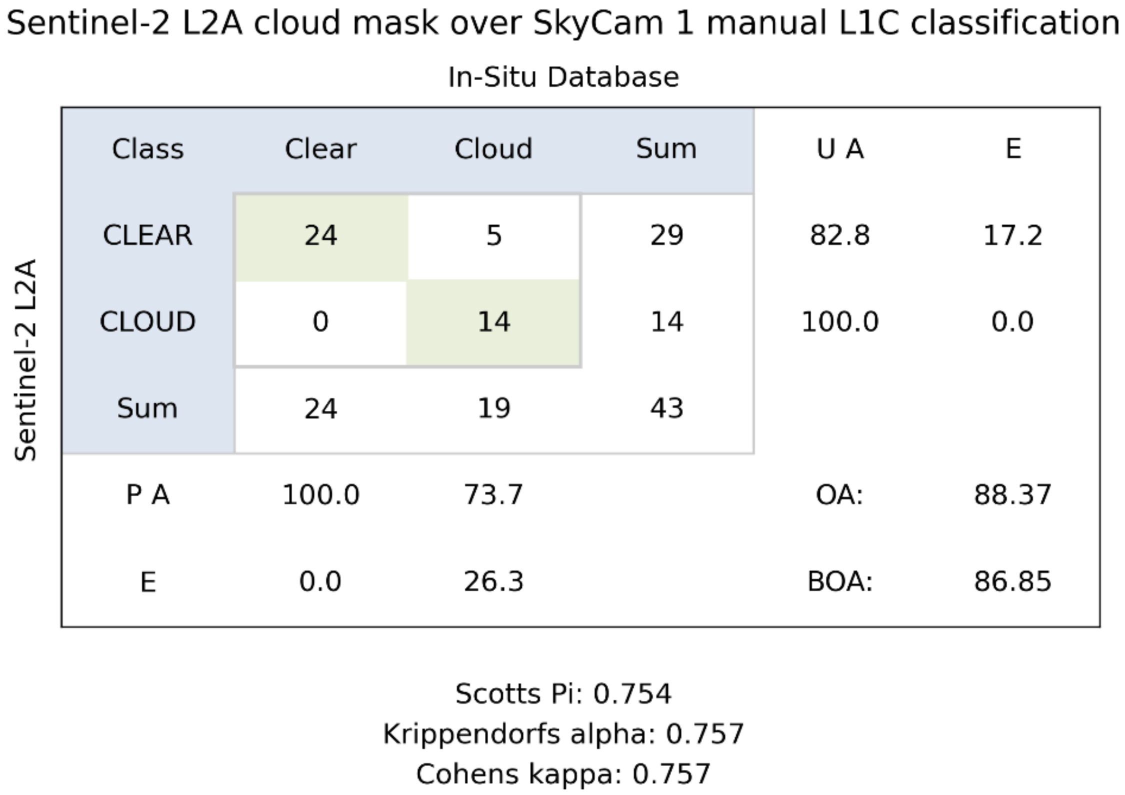
SkyCam 2 manual classification vs. RAP

Sky Camera 2 manual adjusted classification for RAP position

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

- The most likely explanation is the location difference of 22m between the two instruments.
- 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

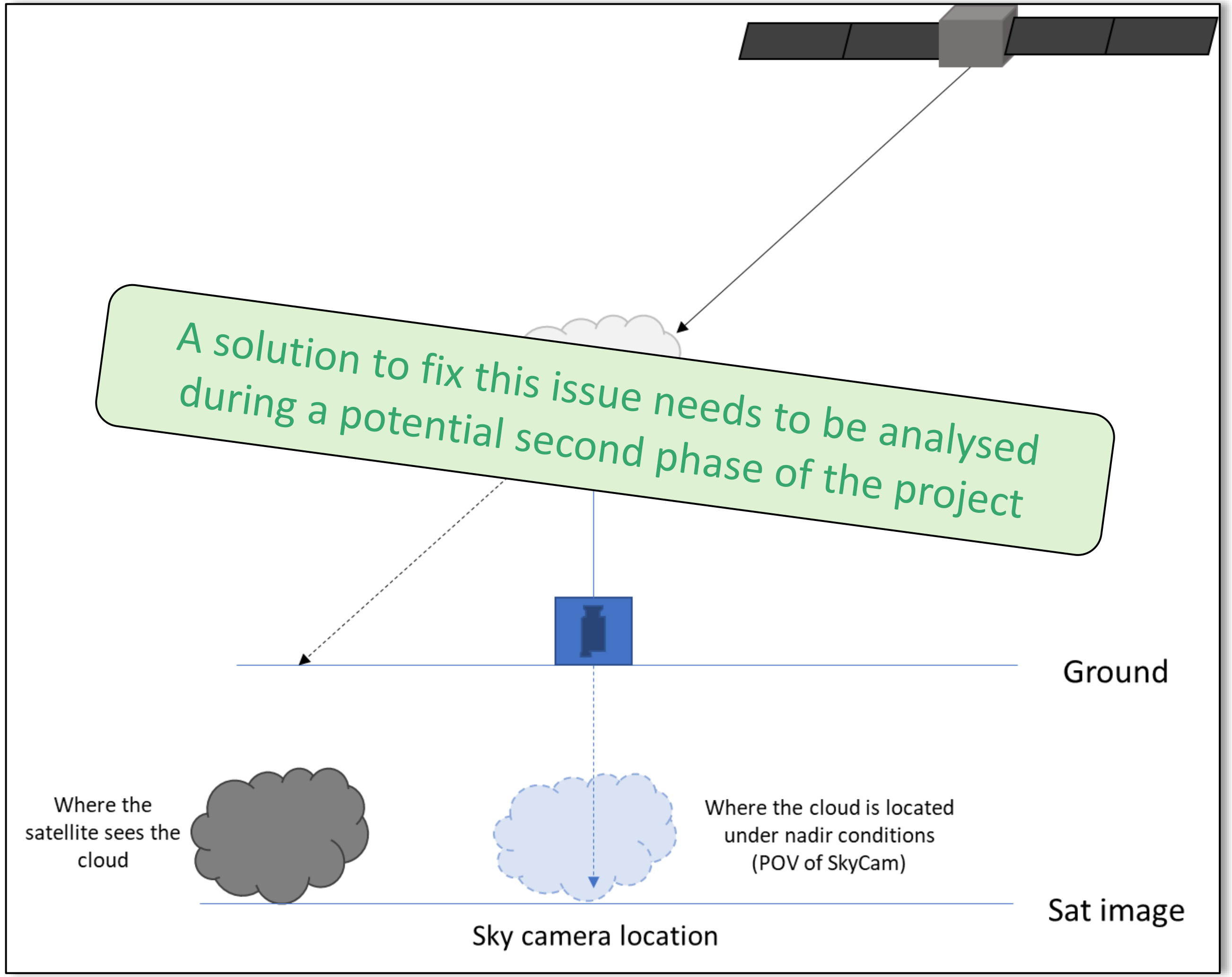
Class	Clear	Cloud	Sum	U A	E
CLEAR	24	5	29	82.8	17.2
CLOUD	0	14	14	100.0	0.0
Sum	24	19	43		
P A	100.0	73.7		OA:	88.37
E	0.0	26.3		BOA:	86.85

Sentinel-2 L2A

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.



A solution to fix this issue needs to be analysed during a potential second phase of the project

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

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.

Tasks that could not be executed

- Due to the pandemic induced delays at UoM no algorithm for cloud height estimation from SC data was available
  - Planned comparison between SC and RAP cloud base height could not be done
- Due to the missing algorithm for SC classification, a classification algorithm had to be designed by BC
  - Implementation of OLCI validation had to be skipped to make time for classification research

- Both task can be executed during a potential next phase of the project

BROCKMANN
CONSULT GMBH

QA4EO

Thank you for the attention!