PROGRAMME OF THE EUROPEAN UNION

co-funded with esa

CMIX-II: Cloud Mask Intercomparison eXercise – second edition

Jan Wevers[a], *Sergii Skakun*[b,c], *Carsten Brockmann*[a], *Eric Vermote*[c]

[a] Brockmann Consult GmbH, 21029 Hamburg, Germany
[b] Department of Geographical Science, University of Maryland, College Park, MD 20742, USA
[c] NASA Goddard Space Flight Center Code 619, Greenbelt, MD 20771, USA

7th Sentinel-2 Validation Team Meeting
13 – 15 October 2025 | ESA – ESRIN | Frascati (RM), Italy

Objective

- The **Cloud Masking Inter-comparison eXercise** (CMIX) is an international collaborative effort aimed at intercomparing cloud detection algorithms for moderate-spatial resolution (10-30 m) spaceborne optical sensors.
- CMIX is a joint activity by **ESA** and **NASA** in the frame of the **CEOS working group on calibration and validation** (WGCV)
- The first CMIX was started in 2018, the final meeting was held in Dec. 2019. Internal report was circulated between 2020 and 2021 leading to a final publication of the results in 2022 (Skakun et al. 2022).
- Report and paper provided recommendations and ways forward to advance inter-comparison and validation of cloud detection algorithms. Those were related to providing a quantitative definition of clouds, generating new cloud reference/validation datasets, and expanding the analysis framework.
  - These necessitated carrying out the second CMIX (**CMIX-II**).
- Like CMIX, **CMIX-II** utilizes open and free **multi-spectral** data
  - **Sentinel-2** and
  - **Landsat 8/9**
- The Focus of CMIX-II is not only on clouds but also on **cloud shadow** detection.

Design – course of events

- Definition of the inter-comparison protocol and reference datasets
  - June 2022
- **Phase 1 – Test Dataset (TDS) exercise**
  - Provision of reference datasets samples
  - Application of the CM processors
  - Initial analysis of the results and provision to participants
  - Feedback from participants
- **Phase 2 – Main exercise**
  - Preparation of complete datasets
  - Provision of datasets to the participants
  - Application of the CM processors
  - Analysis of the results and preparation of report
- Provision of Analysis Report (Internally): Early 2026
- 2nd Workshop of CEOS-WGCV CMIX II
  - Short after release of report @ ESA or NASA
- Preparation of a publication: Mid 2026

CMIX-II was split into two phases to make sure all participating parties agree on the design of the datasets.

Participants

- 25 participating algorithms from 20 participating parties

➔ **largest intercomparison of cloud masking algorithms**

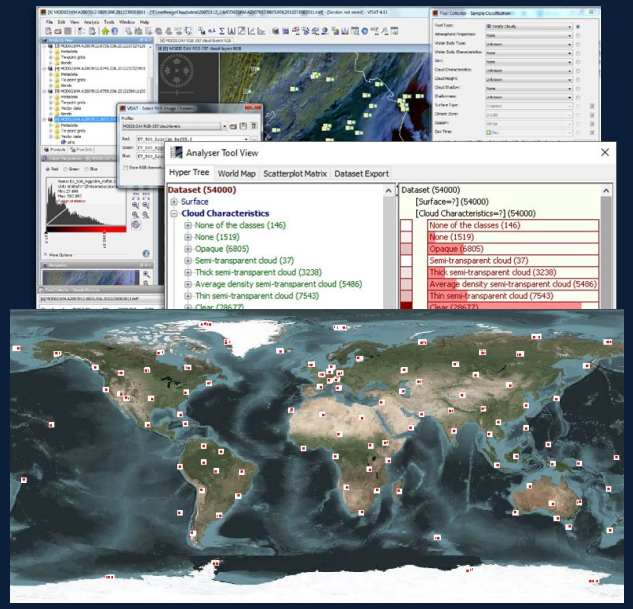
Participant	Processor/Model	Affiliation
Sergii Skakun Eric Vermote Jean-Claude Roger	LaSRC	NASA Goddard Space Flight Center / University of Maryland
Jan Wevers Jorrit Scholze	IdePix	Brockmann Consult GmbH
Alistair Francis	SEnSeI v2	ESA/ESRIN
Feng Yin	SIAC	University College London
David Frantz	FORCE	Trier University
Béatrice Berthelot	MAGELLIUM	MAGCMA
Luis Gómez-Chova	UVDeepCloud	University of Valencia
Kaupo Voormansik Tetiana Shtym	KappaMask	KappaZeta Ltd.
Jerome Louis	Sen2cor & prototype	Telespazio France
Christopher Brown Valerie Pasquarella	CloudScore+ & CloudScore CDF	Google, LLC
Hervé Poilvé	Overland & OneCloudDetector	Airbus Geo-Intelligence Toulouse (Airbus DS)

Participant	Processor/Model	Affiliation
Bringfried Pflug, Avi Avi Pertiwi Raquel de los Reyes	PACO	DLR
Hankui Zhang	LANA	South Dakota State University
Zhe Zhu Shi Qiu	Fmask 4.7 & Fmask 5	University of Conneticut
Ute Gangkofner	noClouds	eoConsultancy
Christian Köppl Andreas Brunn Hannah Kofler	ConstellR	ConstellR
Matthieu Molinier Jean-Eudes Gbodjo	MoCo & DeepCluster	VTT
Ruben van de Kerchove Carolien Toté Yannis Kalfas	LOS	Vito
Jakub Nalepa Bartosz Grabowski	KPLabs_nnUNet	KP Labs
Chris Rampersad Rick Chern		EarthDaily

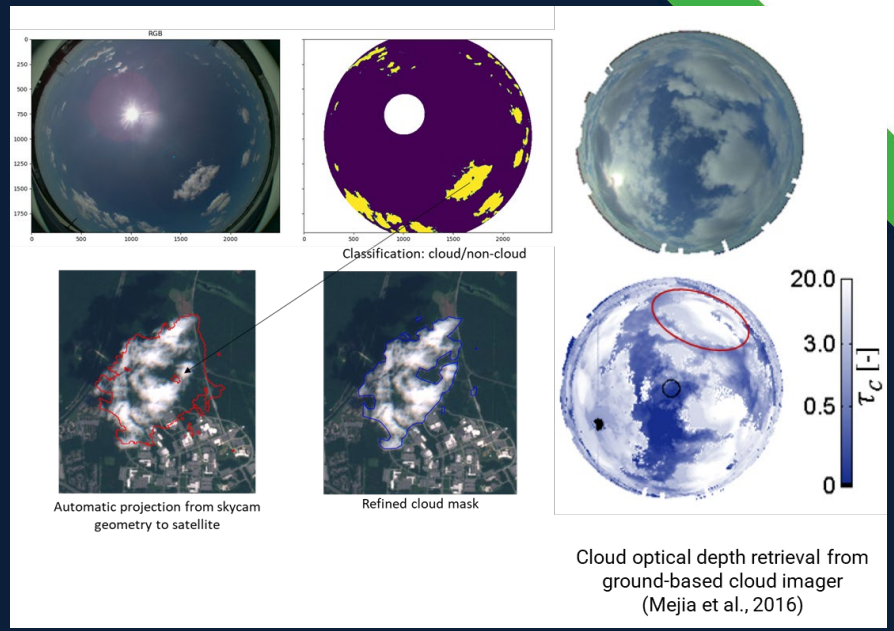
Datasets

- Sky Camera Network
- PixBox (expert pixel collection) dataset (incl. COD)
- Multi-temporal (time series) - identifying potential systematic errors
- Collaborative dataset using IRIS (active learning): classification of subsets by the participants

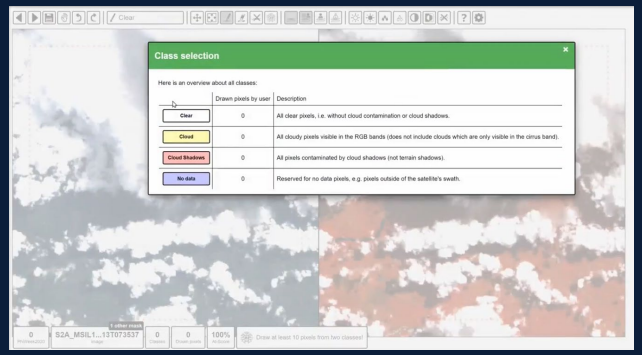
PixBox

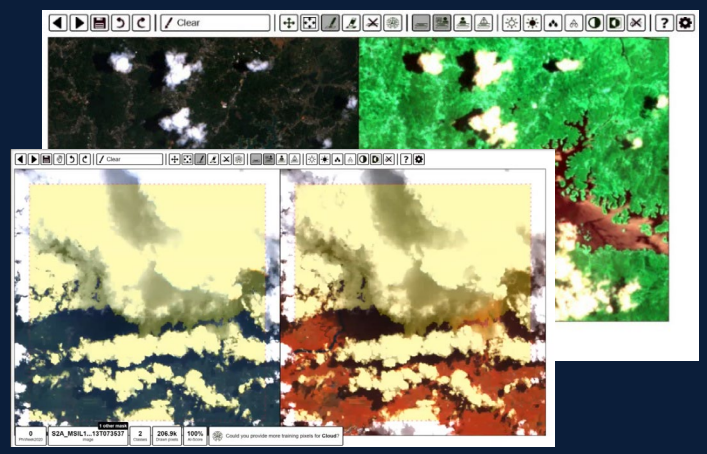


Sky Camera

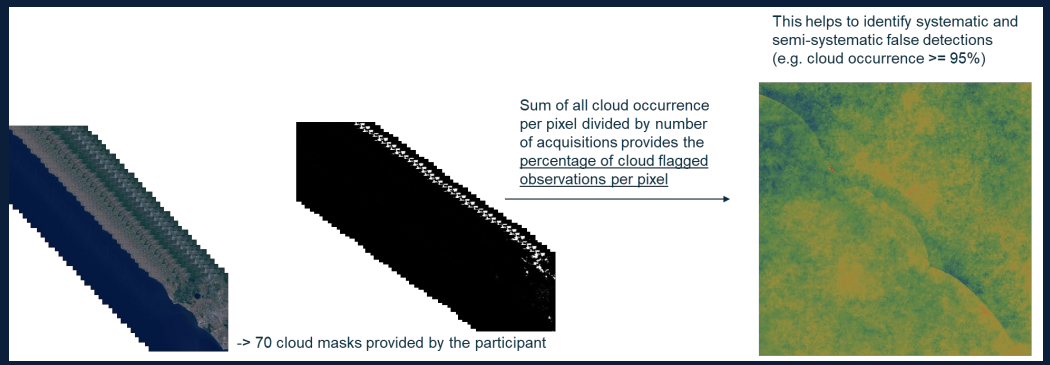


Collaborative dataset





Multi-temporal



Metrics

- Per pixel: confusion matrix and estimated OA, PA, UA
- Visual inspection (potentially study an impact on SR, especially with transparent/cirrus clouds)
- Comparisons to sky imagery outputs:
  - Per-pixel comparison, when a precise co-registration between sky and satellite images will be feasible.
  - Cloud fraction: cloud fractions derived from sky and satellite imagery in a specified ROI will be estimated and will be compared. A separate analysis of completely clear and fully cloudy scenes will be performed. Analysis will be performed using a varying threshold for defining clouds (fuzzy logic analysis).
- Time series analysis:
  - Temporal analysis on cloud (mask) occurrence per pixel per algorithm, to identify potential systematic and semi-systematic errors.
  - Statistic analysis between number of systematic errors and confusion matrix performance.

Results – PixBox dataset

- CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.
- On the next slides the performances of the algorithms for Cloud/Clear and Cloud/Clear/Cloud shadow will be shown.
  - 4 algorithms do not provide a cloud shadow flag
- Questions to be answered:
  - Machine learning is superior to classic spectral test based approaches?
  - High accuracies of cloud detection automatically leads to good cloud shadow detections?

Results – Cloud vs. Clear

CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.

PixBox test validation - Sentinel-2

Algo A (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	9824	462	10286	95.5	4.5
CLOUD	3429	15328	18757	81.7	18.3
Sum	13253	15790	29043		
P A	74.1	97.1		OA:	86.6
E	25.9	2.9		BOA:	85.6

Scotts Pi: 0.722
Krippendorfs alpha: 0.722
Cohens kappa: 0.725

PixBox test validation - Sentinel-2

Algo B (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	12236	1721	13957	87.7	12.3
CLOUD	1017	14069	15086	93.3	6.7
Sum	13253	15790	29043		
P A	92.3	89.1		OA:	90.57
E	7.7	10.9		BOA:	90.7

Scotts Pi: 0.81
Krippendorfs alpha: 0.81
Cohens kappa: 0.81

PixBox test validation - Sentinel-2

Algo C (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	12726	370	13096	97.2	2.8
CLOUD	527	15420	15947	96.7	3.3
Sum	13253	15790	29043		
P A	96.0	97.7		OA:	96.91
E	4.0	2.3		BOA:	96.85

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

PixBox test validation - Sentinel-2

Algo D (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	6279	337	6616	94.9	5.1
CLOUD	6974	15453	22427	68.9	31.1
Sum	13253	15790	29043		
P A	47.4	97.9		OA:	74.83
E	52.6	2.1		BOA:	72.65

Scotts Pi: 0.44
Krippendorfs alpha: 0.44
Cohens kappa: 0.471

PixBox test validation - Sentinel-2

Algo E (N=64 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	7798	1612	9410	82.9	17.1
CLOUD	567	8638	9205	93.8	6.2
Sum	8365	10250	18615		
P A	93.2	84.3		OA:	88.29
E	6.8	15.7		BOA:	88.75

Scotts Pi: 0.765
Krippendorfs alpha: 0.765
Cohens kappa: 0.766

PixBox test validation - Sentinel-2

Algo F (N=64 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	7817	719	8536	91.6	8.4
CLOUD	548	9531	10079	94.6	5.4
Sum	8365	10250	18615		
P A	93.4	93.0		OA:	93.19
E	6.6	7.0		BOA:	93.2

Scotts Pi: 0.862
Krippendorfs alpha: 0.862
Cohens kappa: 0.862

PixBox test validation - Sentinel-2

Algo G (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	10815	1154	11969	90.4	9.6
CLOUD	2438	14636	17074	85.7	14.3
Sum	13253	15790	29043		
P A	81.6	92.7		OA:	87.63
E	18.4	7.3		BOA:	87.15

Scotts Pi: 0.748
Krippendorfs alpha: 0.748
Cohens kappa: 0.748

PixBox test validation - Sentinel-2

Algo H (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	11449	1757	13206	86.7	13.3
CLOUD	1804	14033	15837	88.6	11.4
Sum	13253	15790	29043		
P A	86.4	88.9		OA:	87.74
E	13.6	11.1		BOA:	87.65

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

Results – Cloud vs. Clear

CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.

PixBox test validation - Sentinel-2

PixBox

Algo I (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	12340	329	12669	97.4	2.6
CLOUD	913	15461	16374	94.4	5.6
Sum	13253	15790	29043		
P A	93.1	97.9		OA:	95.72
E	6.9	2.1		BOA:	95.5

Scotts Pi: 0.913
Krippendorfs alpha: 0.913
Cohens kappa: 0.913

PixBox test validation - Sentinel-2

PixBox

Algo J (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	13034	590	13624	95.7	4.3
CLOUD	219	15200	15419	98.6	1.4
Sum	13253	15790	29043		
P A	98.3	96.3		OA:	97.21
E	1.7	3.7		BOA:	97.3

Scotts Pi: 0.943
Krippendorfs alpha: 0.943
Cohens kappa: 0.943

PixBox test validation - Sentinel-2

PixBox

Algo K (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	11243	388	11631	96.7	3.3
CLOUD	2010	15402	17412	88.5	11.5
Sum	13253	15790	29043		
P A	84.8	97.5		OA:	91.74
E	15.2	2.5		BOA:	

Presentation last modified: Fr at 13:

Scotts Pi: 0.831
Krippendorfs alpha: 0.831
Cohens kappa: 0.831

PixBox test validation - Sentinel-2

PixBox

Algo L (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	12731	1102	13833	92.0	8.0
CLOUD	522	14688	15210	96.6	3.4
Sum	13253	15790	29043		
P A	96.1	93.0		OA:	94.41
E	3.9	7.0		BOA:	94.55

Scotts Pi: 0.887
Krippendorfs alpha: 0.887
Cohens kappa: 0.887

PixBox test validation - Sentinel-2

PixBox

Algo M (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	12982	1469	14451	89.8	10.2
CLOUD	271	14321	14592	98.1	1.9
Sum	13253	15790	29043		
P A	98.0	90.7		OA:	94.01
E	2.0	9.3		BOA:	94.35

Scotts Pi: 0.879
Krippendorfs alpha: 0.879
Cohens kappa: 0.88

PixBox test validation - Sentinel-2

PixBox

Algo N (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	6659	172	6831	97.5	2.5
CLOUD	6594	15618	22212	70.3	29.7
Sum	13253	15790	29043		
P A	50.2	98.9		OA:	76.7
E	49.8	1.1		BOA:	74.55

Scotts Pi: 0.485
Krippendorfs alpha: 0.485
Cohens kappa: 0.511

PixBox test validation - Sentinel-2

PixBox

Algo O (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	11801	1067	12868	91.7	8.3
CLOUD	1452	14723	16175	91.0	9.0
Sum	13253	15790	29043		
P A	89.0	93.2		OA:	91.33
E	11.0	6.8		BOA:	91.1

Scotts Pi: 0.824
Krippendorfs alpha: 0.824
Cohens kappa: 0.824

PixBox test validation - Sentinel-2

PixBox

Algo P (N=100 products)

Class	Clear	Cloud	Sum	U A	E
CLEAR	13043	672	13715	95.1	4.9
CLOUD	210	15118	15328	98.6	1.4
Sum	13253	15790	29043		
P A	98.4	95.7		OA:	96.96
E	1.6	4.3		BOA:	97.05

Scotts Pi: 0.938
Krippendorfs alpha: 0.938
Cohens kappa: 0.938

Results – Cloud vs. Clear

CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.

PixBox test validation - Sentinel-2

Algo Q (N=100 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	12409	5183	17592	70.5	29.5
CLOUD	844	10607	11451	92.6	7.4
Sum	13253	15790	29043		
P A	93.6	67.2		OA:	79.25
E	6.4	32.8		BOA:	80.4

Scotts Pi: 0.583
Krippendorfs alpha: 0.583
Cohens kappa: 0.592

PixBox test validation - Sentinel-2

Algo R (N=89 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	4678	10103	14781	31.6	68.4
CLOUD	6807	4384	11191	39.2	60.8
Sum	11485	14487	25972		
P A	40.7	30.3		OA:	34.89
E	59.3	69.7		BOA:	35.5

Scotts Pi: -0.302
Krippendorfs alpha: -0.302
Cohens kappa: -0.281

PixBox test validation - Sentinel-2

Algo S (N=100 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	12021	2260	14281	84.2	15.8
CLOUD	1232	13530	14762	91.7	8.3
Sum	13253	15790	29043		
P A	90.7	85.7		OA:	87.98
E	9.3	14.3		BOA:	88.2

Scotts Pi: 0.758
Krippendorfs alpha: 0.758
Cohens kappa: 0.759

PixBox test validation - Sentinel-2

Algo T (N=100 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	11830	1912	13742	86.1	13.9
CLOUD	1423	13878	15301	90.7	9.3
Sum	13253	15790	29043		
P A	89.3	87.9		OA:	88.52
E	10.7	12.1		BOA:	88.6

Scotts Pi: 0.769
Krippendorfs alpha: 0.769
Cohens kappa: 0.769

PixBox test validation - Sentinel-2

Algo U (N=100 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	13003	552	13555	95.9	4.1
CLOUD	250	15238	15488	98.4	1.6
Sum	13253	15790	29043		
P A	98.1	96.5		OA:	97.24
E	1.9	3.5		BOA:	97.3

Scotts Pi: 0.944
Krippendorfs alpha: 0.944
Cohens kappa: 0.944

PixBox test validation - Sentinel-2

Algo V (N=100 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	12702	208	12910	98.4	1.6
CLOUD	551	15582	16133	96.6	3.4
Sum	13253	15790	29043		
P A	95.8	98.7		OA:	97.39
E	4.2	1.3		BOA:	97.25

Scotts Pi: 0.947
Krippendorfs alpha: 0.947
Cohens kappa: 0.947

PixBox test validation - Sentinel-2

Algo W (N=100 products)

PixBox

Class	Clear	Cloud	Sum	U A	E
CLEAR	12955	442	13397	96.7	3.3
CLOUD	298	15348	15646	98.1	1.9
Sum	13253	15790	29043		
P A	97.8	97.2		OA:	97.45
E	2.2	2.8		BOA:	97.5

Scotts Pi: 0.948
Krippendorfs alpha: 0.948
Cohens kappa: 0.948

Results – Cloud vs. Clear

- Does Machine Learning (ML) outperform classical approaches?
  - Yes
  - But: ML method as such is not a guarantee for good performance -> Strongly depending on training data and know-how

Anynomys_Name	PixBox BOA	ML Algorithm	ML type	Use of time series
Algo A	85.6	Yes	Deep learning model	Yes
Algo B	90.7	Yes	Deep learning model	Yes
Algo C	**96.85**	**Yes**	**Deep convolution (U-Net)**	**No**
Algo D	72.65	Yes	Self-supervised learning algorithms (DL)	No
Algo E	88.75	No	-	No
Algo F	93.2	Yes	Physics-Informed Machine Learning (PIML)	
Algo G	87.15	No	-	No
Algo H	87.65	No	-	No
Algo I	**95.5**	**Yes**	**Deep learning segmentation methods**	**No**
Algo J	**97.3**	**Yes**	**U-Net**	**No**
Algo K	91.15	Yes	U-Net	No
Algo L	94.55	Yes	EvoNet	?
Algo M	94.35	Yes	U-Net	No
Algo N	74.55	Yes	Self-supervised learning algorithms (DL)	No
Algo O	91.1	No	-	No
Algo P	**97.05**	**Yes**	**U-Net CNN with a regnety006 encoder**	**No**
Algo Q	80.4	?	?	No
Algo R	35.5	No	-	No
Algo S	88.2	No	-	No
Algo T	88.6	?	-	No
Algo U	**97.3**	**Yes**	**SegFormer**	**No**
Algo V	**97.25**	**Yes**	**Deep learning for semantic segmentation**	**No**
Algo W	**97.5**	**Yes**	**U-Net CNN with a regnety006 encoder**	**No**

Results – Cloud vs. Clear vs. Cloud Shadow

BROCKMANN CONSULT

CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.

Algo A

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9094	462	730	10286	88.4	11.6
CLOUD	1050	15328	2379	18757	81.7	18.3
SHADOW	0	0	0	0	0	100
Sum	10144	15790	3109	29043		
P A	89.6	97.1	0.0		OA:	84.09
E	10.4	2.9	100.0		BOA:	82.23

Scotts Pi: 0.693
Krippendorfs alpha: 0.693
Cohens kappa: 0.697

Algo B

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9964	1721	2272	13957	71.4	28.6
CLOUD	180	14069	837	15086	93.3	6.7
SHADOW	0	0	0	0	0	100
Sum	10144	15790	3109	29043		
P A	98.2	89.1	0.0		OA:	82.75
E	1.8	10.9	100.0		BOA:	82.43

Scotts Pi: 0.681
Krippendorfs alpha: 0.681
Cohens kappa: 0.686

Algo C

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9674	304	735	10713	90.3	9.7
CLOUD	372	15420	155	15947	96.7	3.3
SHADOW	98	66	2219	2383	93.1	6.9
Sum	10144	15790	3109	29043		
P A	95.4	97.7	71.4		OA:	94.04
E	4.6	2.3	28.6		BOA:	88.17

Scotts Pi: 0.894
Krippendorfs alpha: 0.894
Cohens kappa: 0.894

Algo D

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	3461	236	340	4037	85.7	14.3
CLOUD	6192	15453	782	22427	68.9	31.1
SHADOW	491	101	1987	2579	77.0	23.0
Sum	10144	15790	3109	29043		
P A	34.1	97.9	63.9		OA:	71.97
E	65.9	2.1	36.1		BOA:	65.3

Scotts Pi: 0.436
Krippendorfs alpha: 0.436
Cohens kappa: 0.463

Algo E

PixBox test validation - Sentinel-2

PixBox

(N=64 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	5691	1331	936	7958	71.5	28.5
CLOUD	543	8638	24	9205	93.8	6.2
SHADOW	233	281	938	1452	64.6	35.4
Sum	6467	10250	1898	18615		
P A	88.0	84.3	49.4		OA:	82.01
E	12.0	15.7	50.6		BOA:	73.9

Scotts Pi: 0.683
Krippendorfs alpha: 0.683
Cohens kappa: 0.685

Algo F

PixBox test validation - Sentinel-2

PixBox

(N=64 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	5709	450	553	6712	85.1	14.9
CLOUD	472	9531	76	10079	94.6	5.4
SHADOW	286	269	1269	1824	69.6	30.4
Sum	6467	10250	1898	18615		
P A	88.3	93.0	66.9		OA:	88.69
E	11.7	7.0	33.1		BOA:	82.73

Scotts Pi: 0.8
Krippendorfs alpha: 0.8
Cohens kappa: 0.8

Algo G

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	7753	924	713	9390	82.6	17.4
CLOUD	1477	14636	961	17074	85.7	14.3
SHADOW	714	151	1414	2279	62.0	38.0
Sum	9944	15711	3088	28743		
P A	78.0	93.2	45.8		OA:	82.81
E	22.0	6.8	54.2		BOA:	72.33

Scotts Pi: 0.689
Krippendorfs alpha: 0.689
Cohens kappa: 0.689

Algo H

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	8117	1535	865	10517	77.2	22.8
CLOUD	1614	14033	190	15837	88.6	11.4
SHADOW	413	222	2054	2689	76.4	23.6
Sum	10144	15790	3109	29043		
P A	80.0	88.9	66.1		OA:	83.34
E	20.0	11.1	33.9		BOA:	78.33

Scotts Pi: 0.706
Krippendorfs alpha: 0.706
Cohens kappa: 0.706

Results – Cloud vs. Clear vs. Cloud Shadow

CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.

Algo I

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	8651	238	378	9267	93.4	6.6
CLOUD	797	15461	116	16374	94.4	5.6
SHADOW	694	90	2615	3399	76.9	23.1
Sum	10142	15789	3109	29040		
P A	85.3	97.9	84.1		OA:	92.04
E	14.7	2.1	15.9		BOA:	89.1

Scotts Pi: 0.86
Krippendorfs alpha: 0.86
Cohens kappa: 0.86

Algo J

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9849	326	367	10542	93.4	6.6
CLOUD	177	15200	42	15419	98.6	1.4
SHADOW	118	264	2700	3082	87.6	12.4
Sum	10144	15790	3109	29043		
P A	97.1	96.3	86.8		OA:	95.54
E	2.9	3.7	13.2		BOA:	93.4

Scotts Pi: 0.922
Krippendorfs alpha: 0.922
Cohens kappa: 0.922

Algo K

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	8589	388	2654	11631	73.8	26.2
CLOUD	1555	15402	455	17412	88.5	11.5
SHADOW	0	0	0	0	0	100
Sum	10144	15790	3109	29043		
P A	84.7	97.5	0.0		OA:	82.61
E	15.3	2.5	100.0		BOA:	60.73

Scotts Pi: 0.671
Krippendorfs alpha: 0.671
Cohens kappa: 0.674

Algo L

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9471	654	783	10908	86.8	13.2
CLOUD	460	14688	62	15210	96.6	3.4
SHADOW	149	234	2262	2645	85.5	14.5
Sum	10080	15576	3107	28763		
P A	94.0	94.3	72.8		OA:	91.86
E	6.0	5.7	27.2		BOA:	87.03

Scotts Pi: 0.857
Krippendorfs alpha: 0.857
Cohens kappa: 0.857

Algo M

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9884	1188	1231	12303	80.3	19.7
CLOUD	198	14321	73	14592	98.1	1.9
SHADOW	62	281	1805	2148	84.0	16.0
Sum	10144	15790	3109	29043		
P A	97.4	90.7	58.1		OA:	89.56
E	2.6	9.3	41.9		BOA:	82.07

Scotts Pi: 0.816
Krippendorfs alpha: 0.816
Cohens kappa: 0.817

Algo N

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	3691	75	83	3849	95.9	4.1
CLOUD	5752	15618	842	22212	70.3	29.7
SHADOW	701	97	2184	2982	73.2	26.8
Sum	10144	15790	3109	29043		
P A	36.4	98.9	70.2		OA:	74.0
E	63.6	1.1	29.8		BOA:	68.5

Scotts Pi: 0.483
Krippendorfs alpha: 0.483
Cohens kappa: 0.506

Algo O

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	8462	897	907	10266	82.4	17.6
CLOUD	1238	14723	214	16175	91.0	9.0
SHADOW	444	170	1988	2602	76.4	23.6
Sum	10144	15790	3109	29043		
P A	83.4	93.2	63.9		OA:	86.67
E	16.6	6.8	36.1		BOA:	80.17

Scotts Pi: 0.763
Krippendorfs alpha: 0.763
Cohens kappa: 0.763

Algo P

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9825	442	385	10652	92.2	7.8
CLOUD	176	15118	34	15328	98.6	1.4
SHADOW	143	230	2690	3063	87.8	12.2
Sum	10144	15790	3109	29043		
P A	96.9	95.7	86.5		OA:	95.15
E	3.1	4.3	13.5		BOA:	93.03

Scotts Pi: 0.915
Krippendorfs alpha: 0.915
Cohens kappa: 0.915

Results – Cloud vs. Clear vs. Cloud Shadow

CMIX-II is actively running in Phase 2. Therefore, only anonymized results will be shown.

Algo Q

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9314	5183	3095	17592	52.9	47.1
CLOUD	830	10607	14	11451	92.6	7.4
SHADOW	0	0		0	0	100
Sum	10144	15790	3109	29043		
P A	91.8	67.2	0.0		OA:	68.59
E	8.2	32.8	100.0		BOA:	53.0

Scotts Pi: 0.428
Krippendorfs alpha: 0.428
Cohens kappa: 0.452

Algo R

PixBox test validation - Sentinel-2

PixBox

(N=89 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	3770	9304	697	13771	27.4	72.6
CLOUD	4796	4384	2011	11191	39.2	60.8
SHADOW	179	784	32	995	3.2	96.8
Sum	8745	14472	2740	25957		
P A	43.1	30.3	1.2		OA:	31.54
E	56.9	69.7	98.8		BOA:	24.87

Scotts Pi: -0.217
Krippendorfs alpha: -0.217
Cohens kappa: -0.186

Algo S

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	8948	1974	1488	12410	72.1	27.9
CLOUD	987	13530	245	14762	91.7	8.3
SHADOW	145	72	1374	1591	86.4	13.6
Sum	10080	15576	3107	28763		
P A	88.8	86.9	44.2		OA:	82.93
E	11.2	13.1	55.8		BOA:	73.3

Scotts Pi: 0.696
Krippendorfs alpha: 0.696
Cohens kappa: 0.697

Algo T

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	8814	1859	1469	12142	72.6	27.4
CLOUD	1174	13878	249	15301	90.7	9.3
SHADOW	156	53	1391	1600	86.9	13.1
Sum	10144	15790	3109	29043		
P A	86.9	87.9	44.7		OA:	82.92
E	13.1	12.1	55.3		BOA:	73.17

Scotts Pi: 0.694
Krippendorfs alpha: 0.694
Cohens kappa: 0.695

Algo U

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9855	287	377	10519	93.7	6.3
CLOUD	203	15238	47	15488	98.4	1.6
SHADOW	86	265	2685	3036	88.4	11.6
Sum	10144	15790	3109	29043		
P A	97.2	96.5	86.4		OA:	95.64
E	2.8	3.5	13.6		BOA:	93.37

Scotts Pi: 0.923
Krippendorfs alpha: 0.923
Cohens kappa: 0.923

Algo V

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9579	153	272	10004	95.8	4.2
CLOUD	452	15582	99	16133	96.6	3.4
SHADOW	113	55	2738	2906	94.2	5.8
Sum	10144	15790	3109	29043		
P A	94.4	98.7	88.1		OA:	96.06
E	5.6	1.3	11.9		BOA:	93.73

Scotts Pi: 0.93
Krippendorfs alpha: 0.93
Cohens kappa: 0.93

Algo W

PixBox test validation - Sentinel-2

PixBox

(N=100 products)

Class	Clear	Cloud	Shadow	Sum	U A	E
CLEAR	9612	226	251	10089	95.3	4.7
CLOUD	266	15348	32	15646	98.1	1.9
SHADOW	266	216	2826	3308	85.4	14.6
Sum	10144	15790	3109	29043		
P A	94.8	97.2	90.9		OA:	95.67
E	5.2	2.8	9.1		BOA:	94.3

Scotts Pi: 0.924
Krippendorfs alpha: 0.924
Cohens kappa: 0.924

Results – Cloud vs. Clear vs. Cloud Shadow

- Does an overall good cloud detection (BOA), or a high Producer Accuracy (PA) of cloud, guarantee a good cloud shadow detection?
  - No!
  - But it is the essential first step, since most cloud shadow detection algorithms rely on geometric approaches, using the cloud mask as a reference.

BOA for Cloud/Clear ⇩ ⇩ PA Coud for Cloud/Clear/Shadow

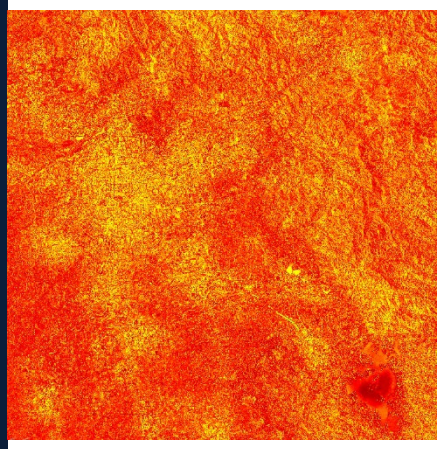
Anynomys_Name	PixBox BOA	PA Cloud	PA Cloud Shadow
Algo A	85.6	0	0
Algo B	90.7	0	0
Algo C	96.85	97.7	71.4
Algo D	72.65	97.9	63.9
Algo E	88.75	84.3	49.4
Algo F	93.2	93	66.9
Algo G	87.15	93.2	45.8
Algo H	87.65	88.9	66.1
Algo I	**95.5**	**97.9**	**84.1**
Algo J	**97.3**	**96.3**	**86.8**
Algo K	91.15	0	0
Algo L	94.55	94.3	72.8
Algo M	94.35	90.7	58.1
Algo N	74.55	98.9	70.2
Algo O	91.1	93.2	63.9
Algo P	**97.05**	**95.7**	**86.5**
Algo Q	80.4	0	0
Algo R	35.5	30.3	1.2
Algo S	88.2	86.9	44.2
Algo T	88.6	87.9	44.7
Algo U	**97.3**	**96.5**	**86.4**
Algo V	**97.25**	**98.7**	**88.1**
Algo W	**97.5**	**97.2**	**90.9**

Results – Multi-temporal dataset

Multi-temporal analysis:

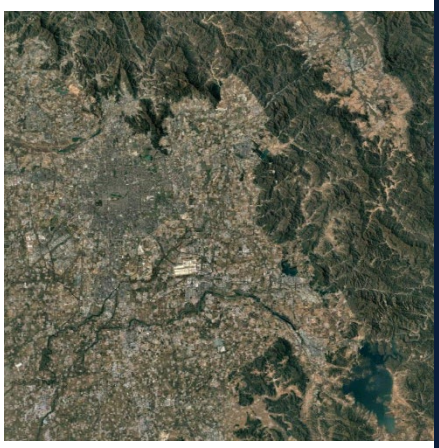
- A pixel-based analysis (like PixBox) does not give any insight into the spatial behavior of the algorithms, and it is not capable of identifying systematic and semi-systematic false detections.
- By processing an annual time series over a given location and calculating the percentage of the cloud flagged observations in relation to all observations, those systematic issues can be identified.
- 13 participants provided the results for the multi-temporal dataset
  - Again, only anonymized results will be shown.
  - The BOA for Cloud/Clear from the PixBox dataset is noted next to the Algorithm ID
- The main question: Does a good performance for the PixBox dataset mean the absence of systematic errors?

Algo N – BOA 74.55

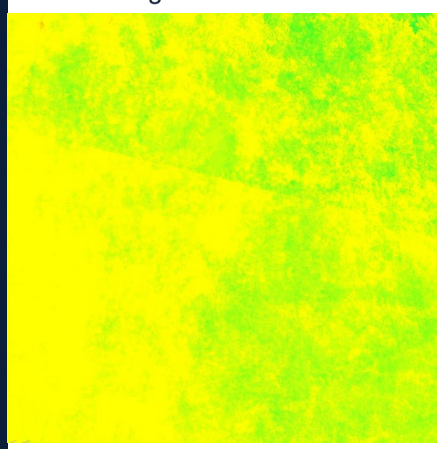


Percentage of cloud masked pixels

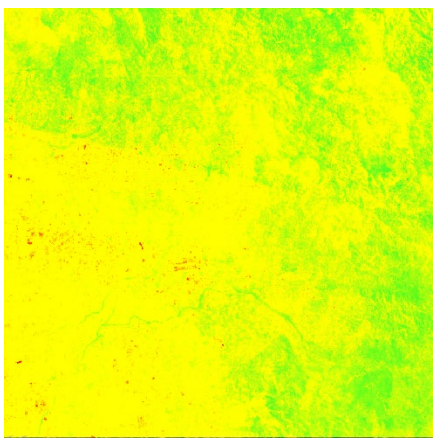
China, Beijing



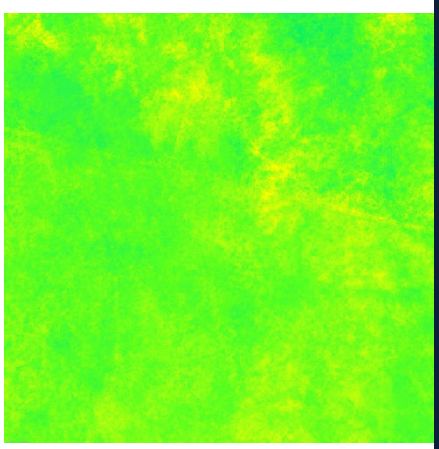
Algo S – BOA 88.2



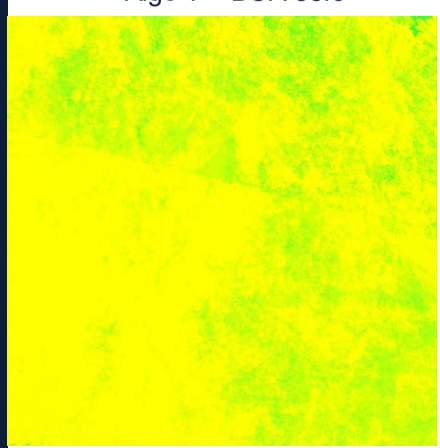
Algo H – BOA 87.65



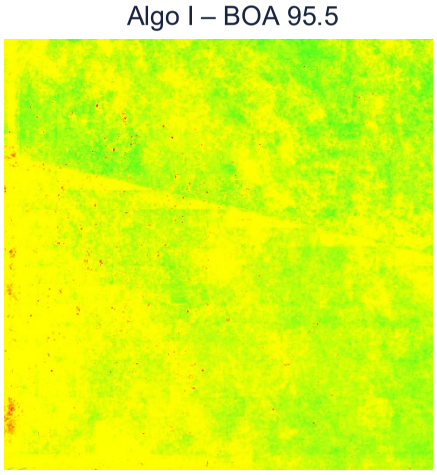
Algo A – BOA 85.6



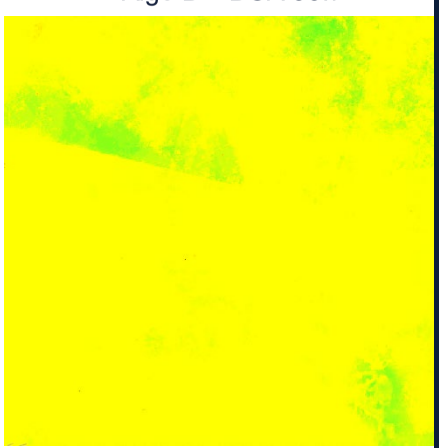
Algo T – BOA 88.6



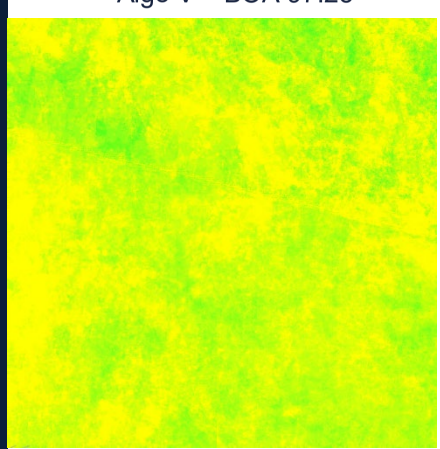
Algo I – BOA 95.5



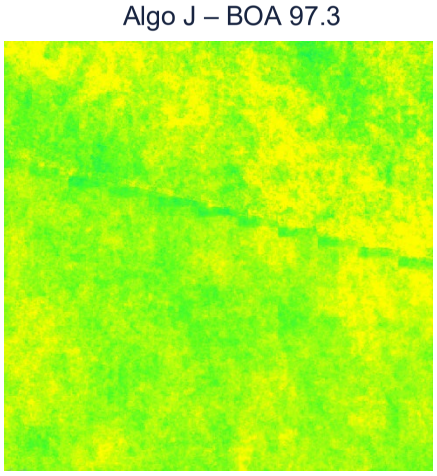
Algo B – BOA 90.7



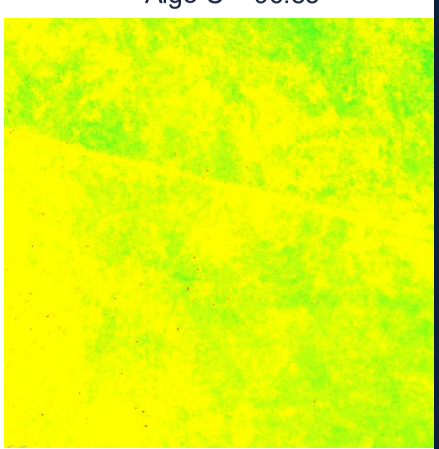
Algo V – BOA 97.25



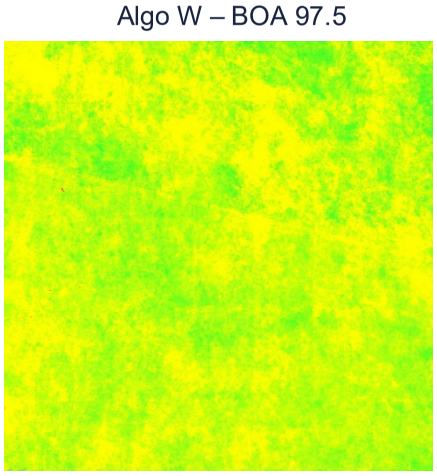
Algo J – BOA 97.3



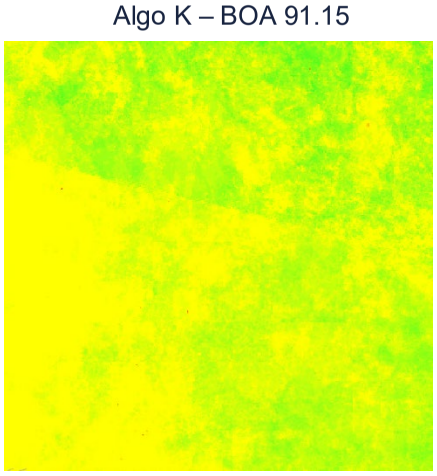
Algo C – 96.85



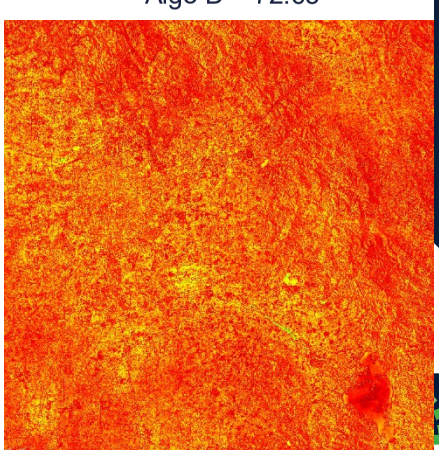
Algo W – BOA 97.5



Algo K – BOA 91.15

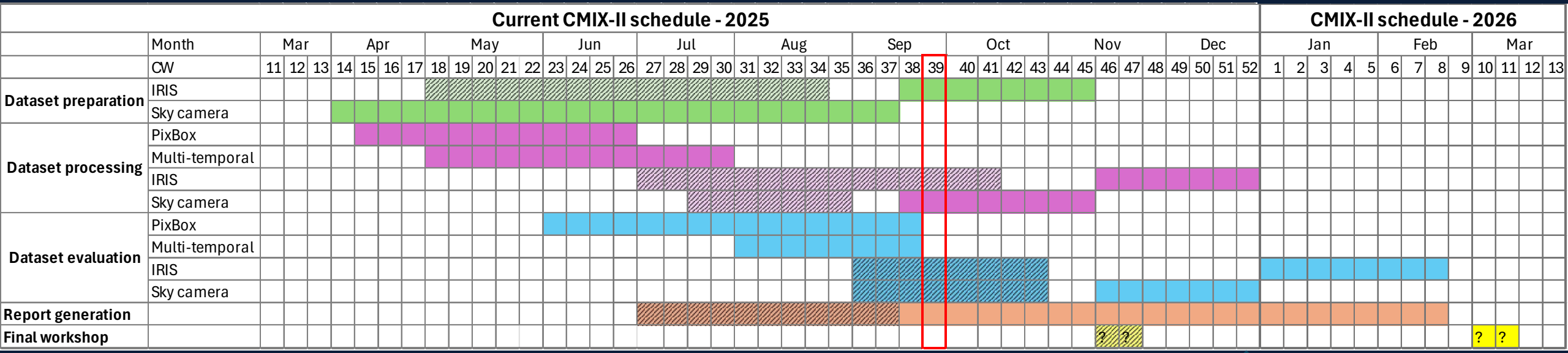


Algo D – 72.65



Schedule

- Phase 1 – Test Dataset (TDS) exercise: **Completed**
- Phase 2 – Main exercise: **Running**
- Current schedule shifted due to very large number of participations



End of WP 2490

Conclusion

- The exercise is actively running with strong engagement from the participant community.
- Its outcomes are of great interest, both to the participants and to the wider cloud detection community, as highlighted by feedback at the Living Planet Symposium.
- The activity delivers valuable insights into the performance of diverse algorithms and methodological approaches.
- It also represents the largest intercomparison of cloud detection algorithms carried out to date.

Thank you for the attention!

BROCKMANN CONSULT

Turning Environmental Data Into Knowledge

brockmann-consult.de
info@brockmann-consult.de
+49 40 696 389 300