RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

cnes
CENTRE NATIONAL D'ÉTUDES SPATIALES

SWOT HR
LAKE PRODUCT VALIDATION

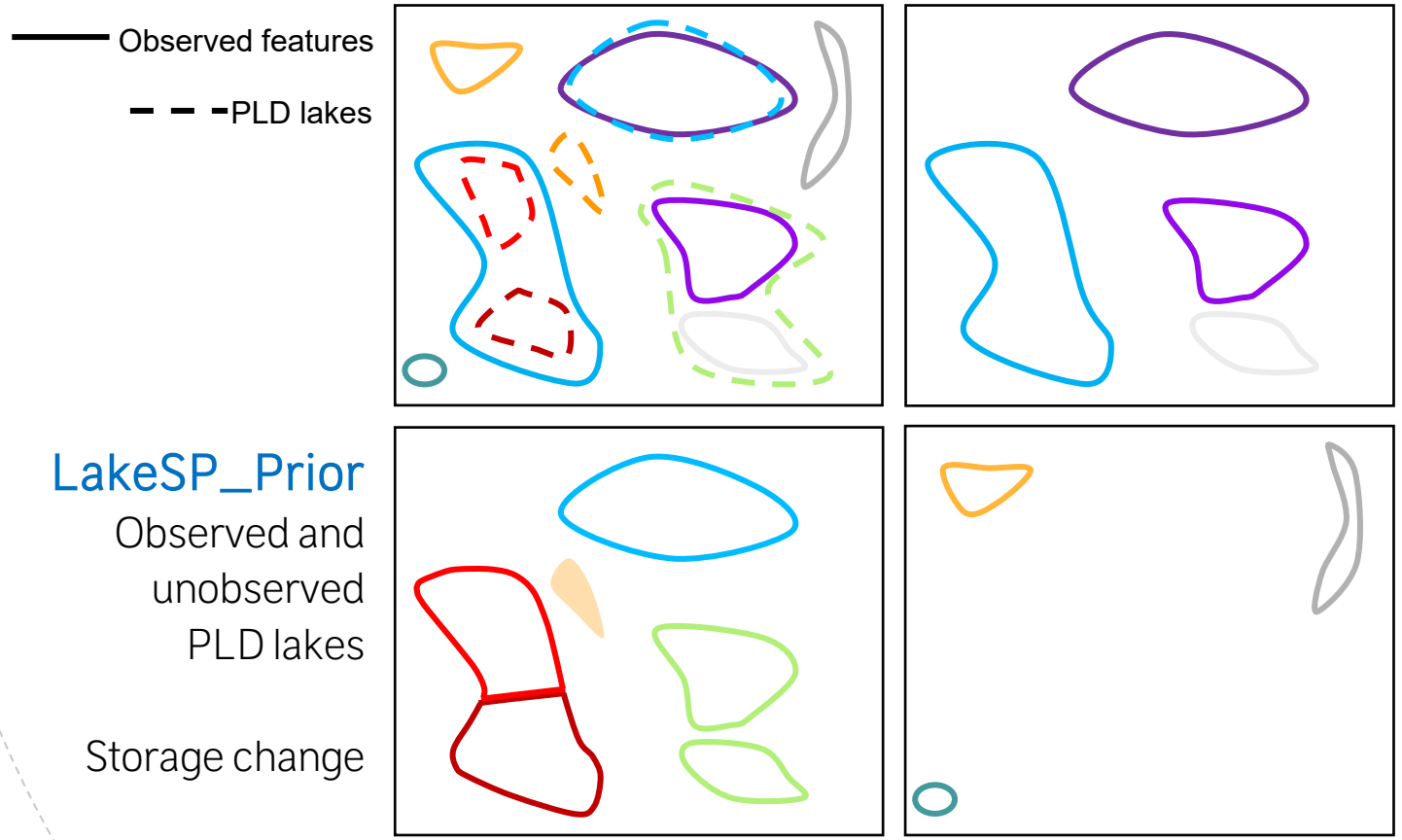
Claire Pottier, Roger Fjørtoft,
on behalf of the HR Cal/Val Team

SWOT VALIDATION MEETING, CHAPEL HILL, NC
19 JUNE 2024

OUTLINE

LakeSP PRODUCTS

3 shapefiles based on the intersection with the polygons of the Prior Lake Database (PLD)



LakeSP_Obs

Observed features intersecting one or more PLD lakes

LakeSP_Prior

Observed and unobserved PLD lakes

Storage change

LakeSP_Unassigned

Unassigned features (neither assigned to SWORD rivers nor PLD lakes)

LAKE VALIDATION APPROACH

General features

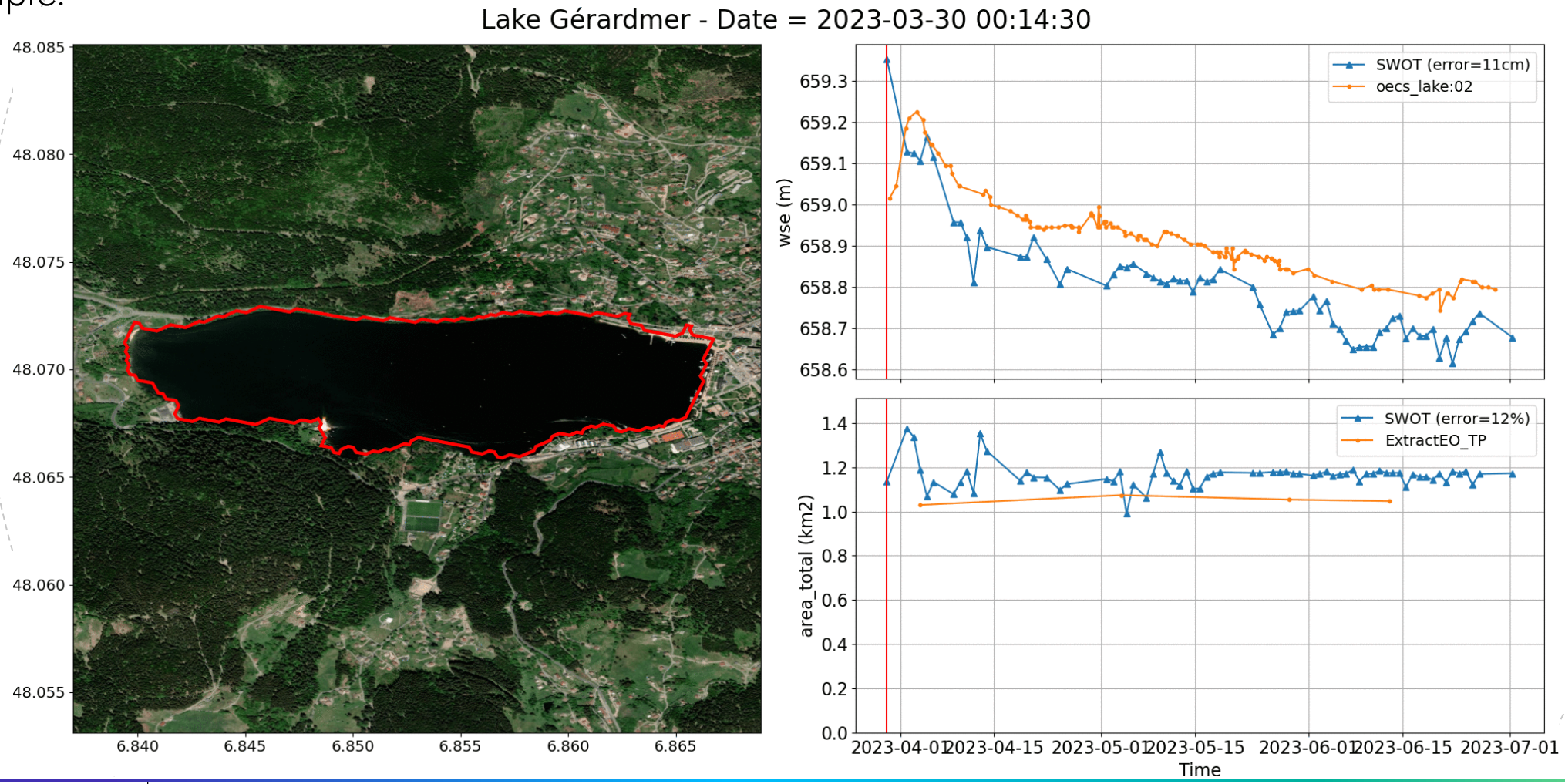
- Validation on LakeSP_Prior products (version “C”, i.e. PIC0/PGC0)
- Limited to the principal variables: WSE and area

- Dedicated in situ measurements and acquisition of satellite images
  - Mainly during Cal/Val period (1-day orbit, March 30 – July 10, 2023)
- Extensive use of existing gauges and publicly available satellite data
  - Additional levelling activities and preprocessing
- Computation of global statistics on a large number of PLD lakes worldwide
  - No distinction made between Tier 1 and Tier 2 sites in what follows (similar accuracy)

- The reference data are targeted to be more accurate than SWOT requirements, but they are not perfect
  - The measured errors may partially stem from inaccuracies in the reference data

LAKE VALIDATION APPROACH

Example:



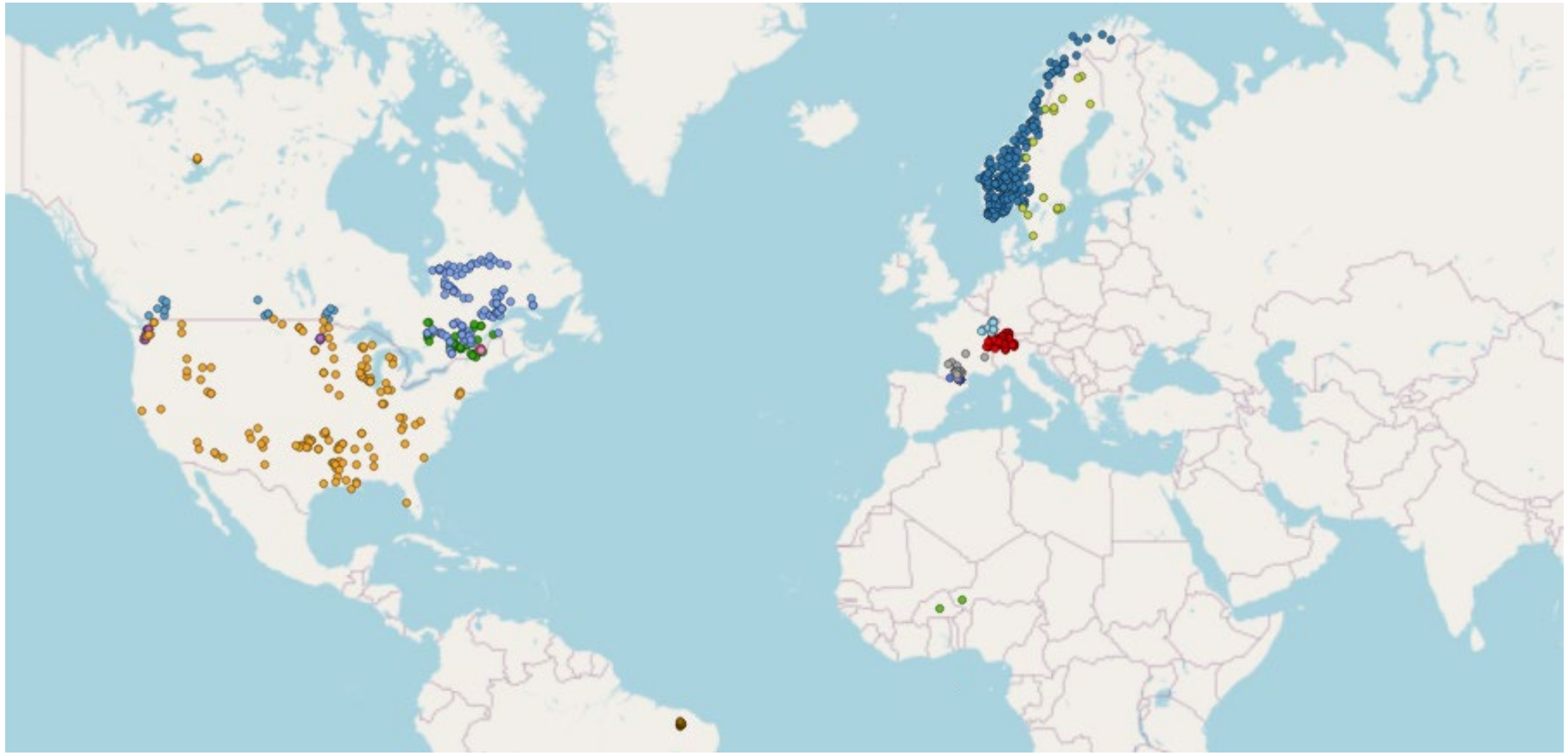
LAKE WSE VALIDATION

Principles

- LakeSP_Prior products exclude:
  - Lakes outside the nominal swath (10-60 km)
  - Lakes whose size is below 100x100 m^2
- **WSE science requirements and goals apply to lakes > 250x250 m^2**
- In situ / reference data:
  - Clearly erroneous ones are discarded
  - **Focus on absolute (not relative) WSE → only leveled in situ data are analyzed here**
- Identification of matching LakeSP and in situ / reference data:
  - Spatially: intersection with the PLD lake polygon
  - Temporally: interpolation between in situ WSE measurement before/after SWOT acquisition (generally <1 h, maximum **3** days)

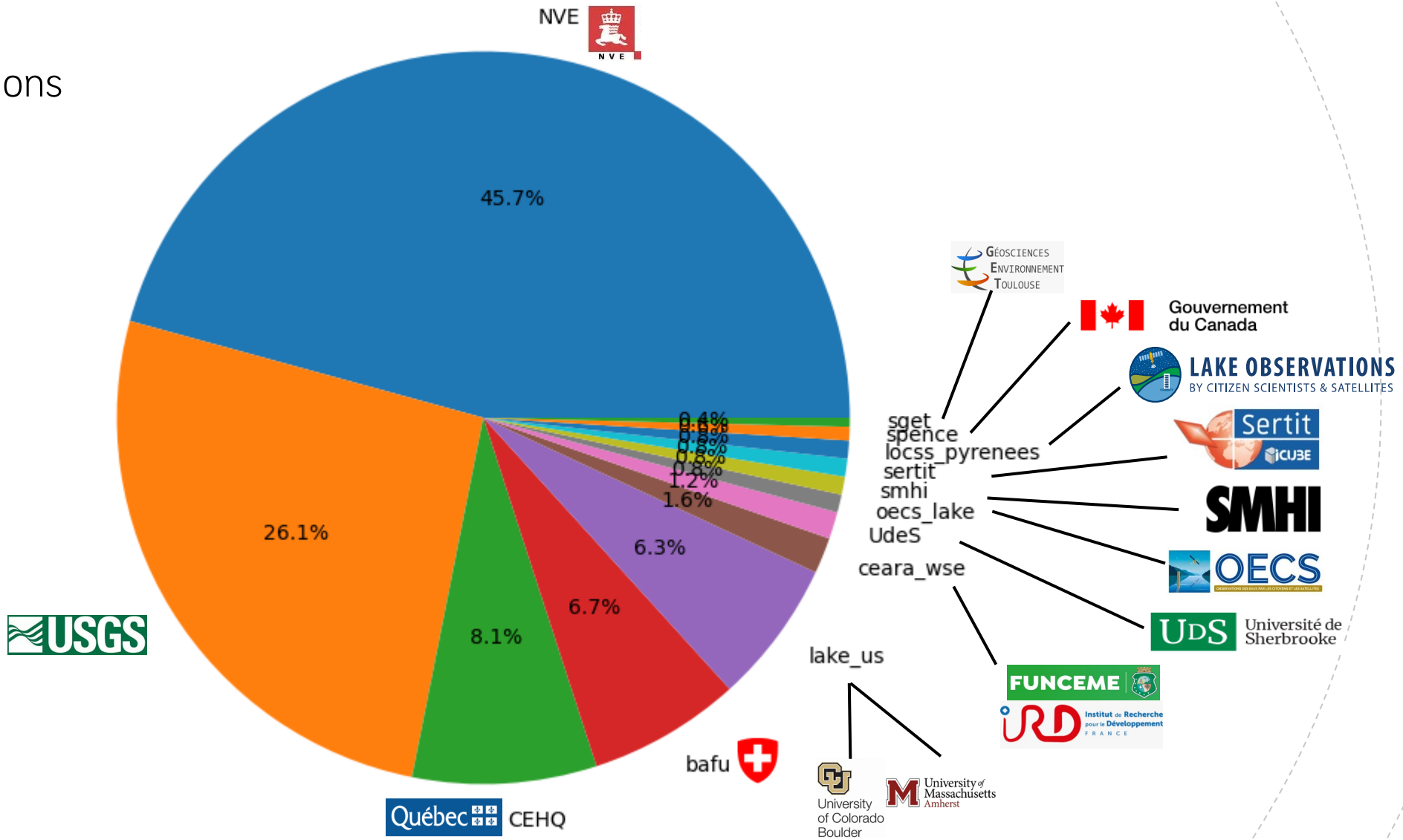
GEOGRAPHICAL DISTRIBUTION

PLD lakes with ground truth available for WSE validation (leveled gauges only)



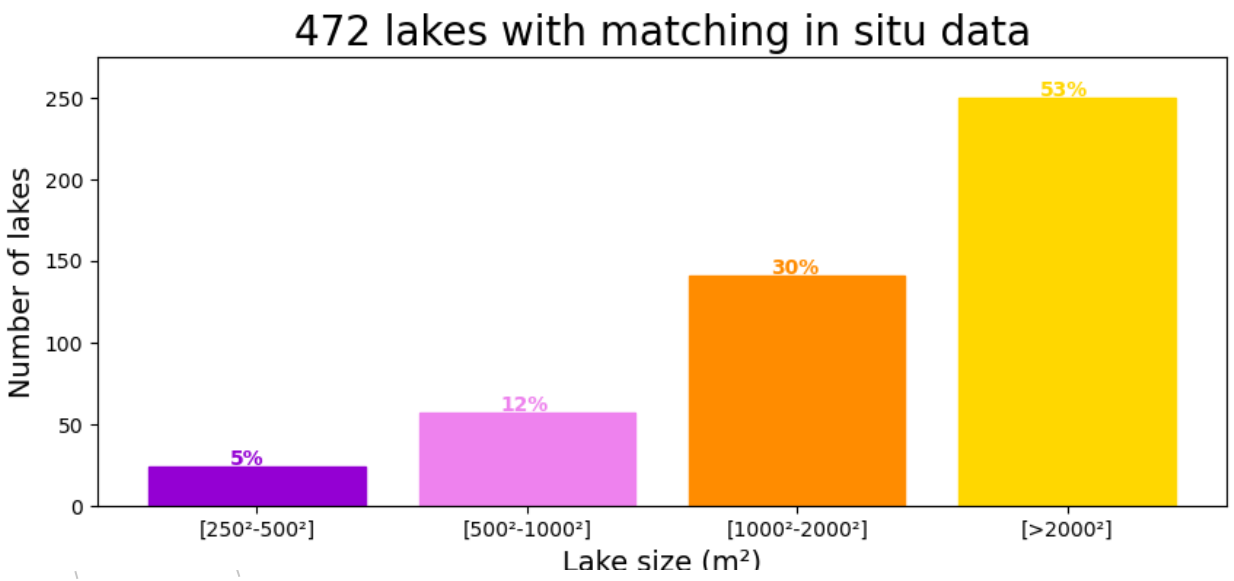
DATA SOURCE DISTRIBUTION

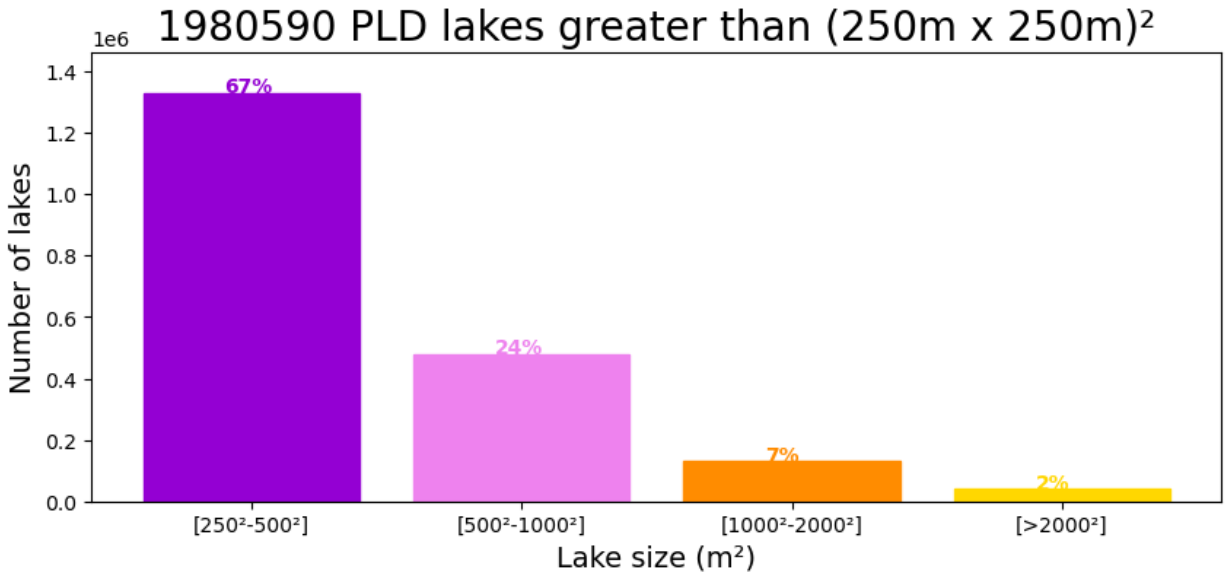
Leveled in situ stations



SIZE DISTRIBUTION

Number of PLD lakes with ground truth available for WSE validation vs. total number of PLD lakes



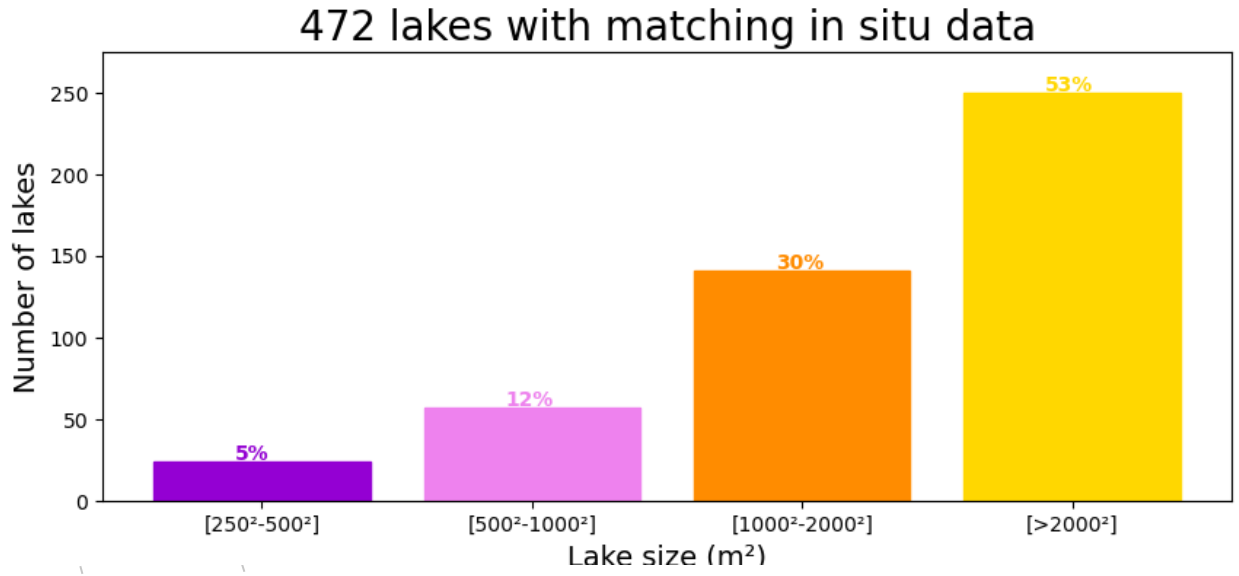


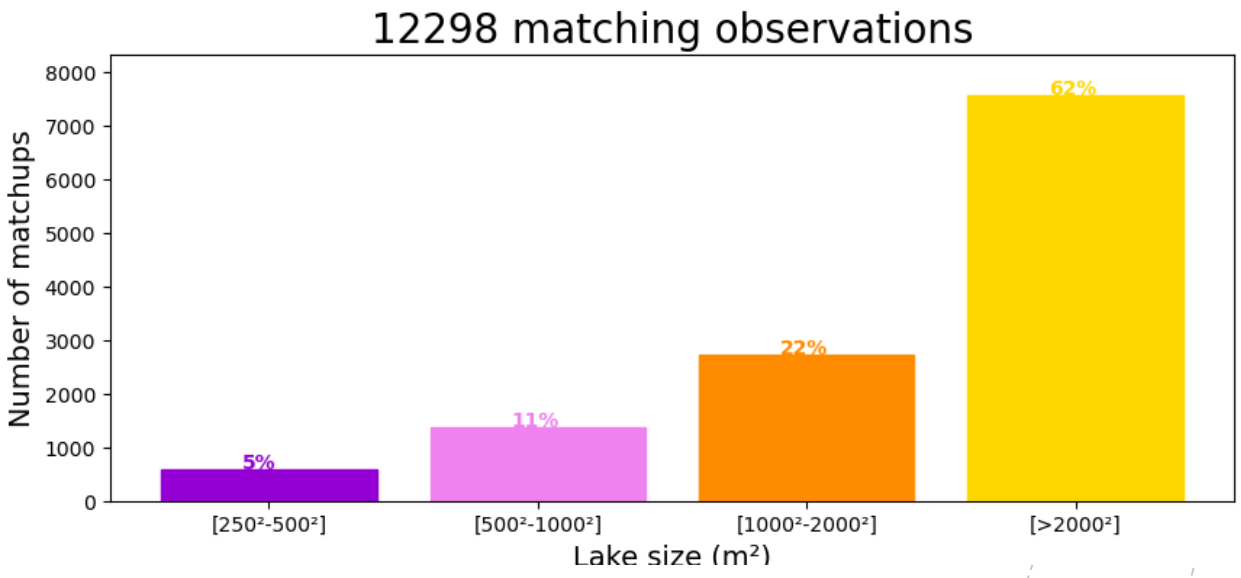
~4 millions PLD lakes < (250m x 250m)²

- Inverse of expected global distribution: increasing number of lakes with increasing size

SIZE DISTRIBUTION

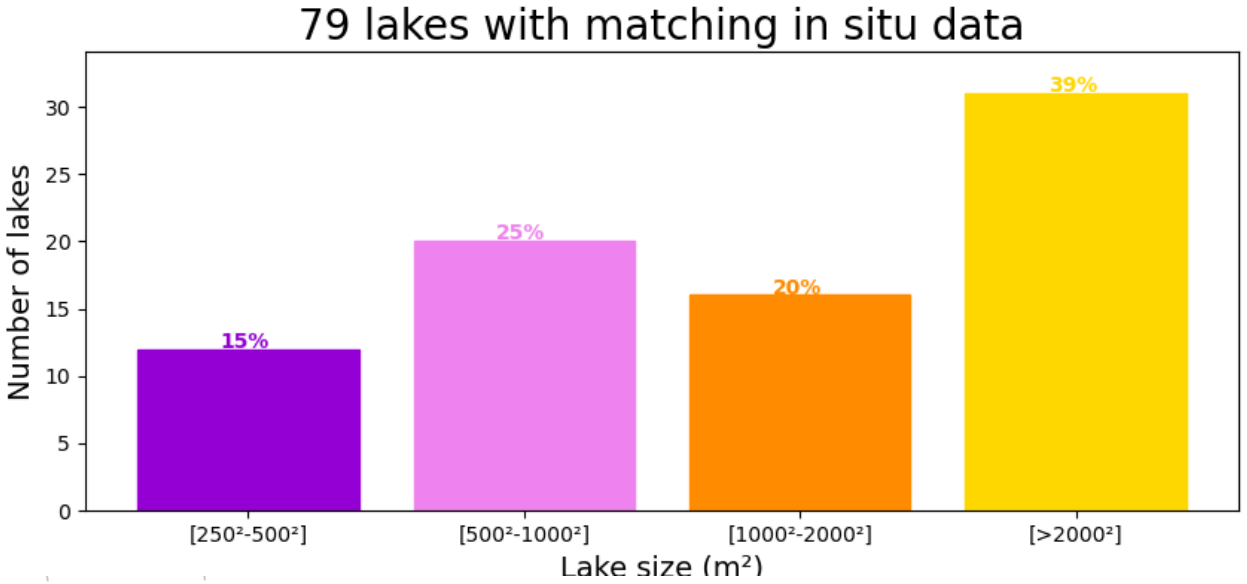
Number of PLD lakes with ground truth available for WSE validation

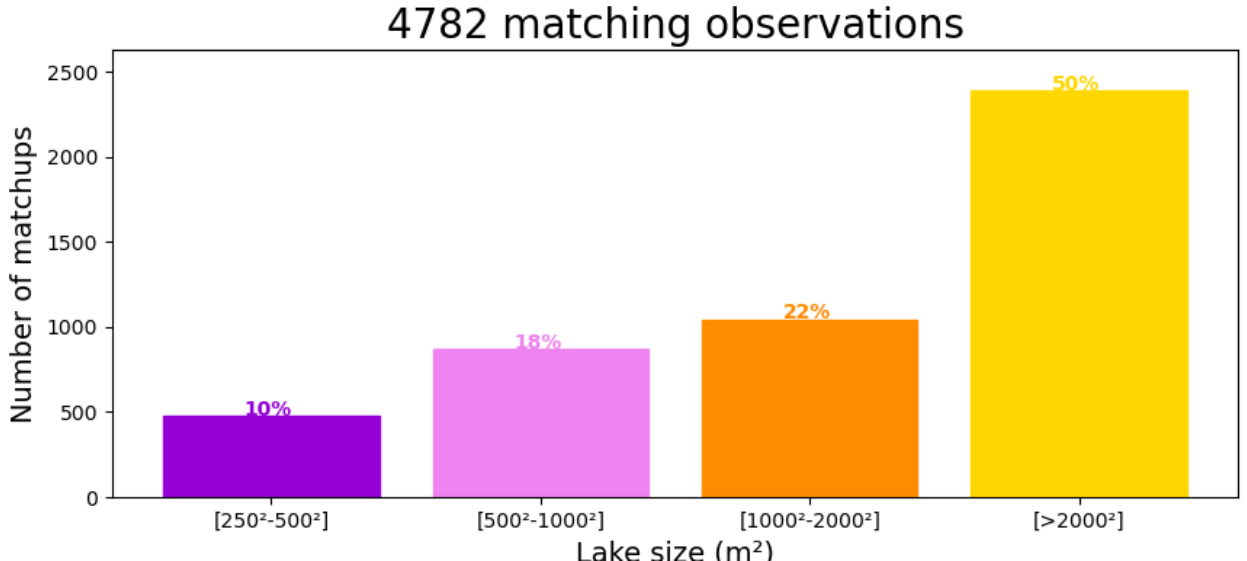




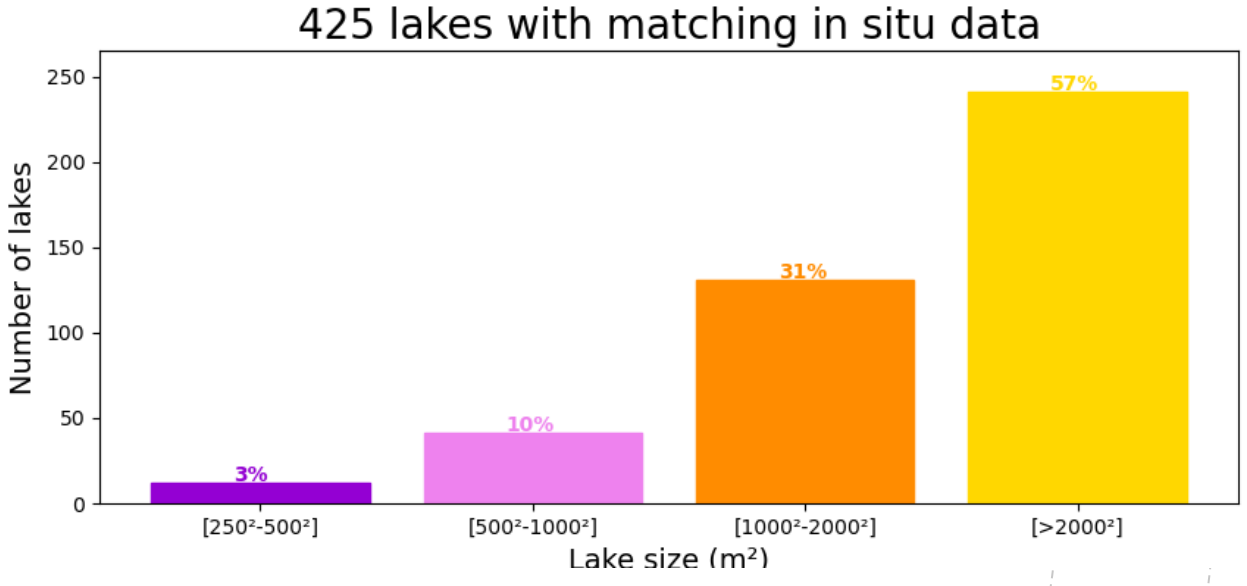
SIZE DISTRIBUTION

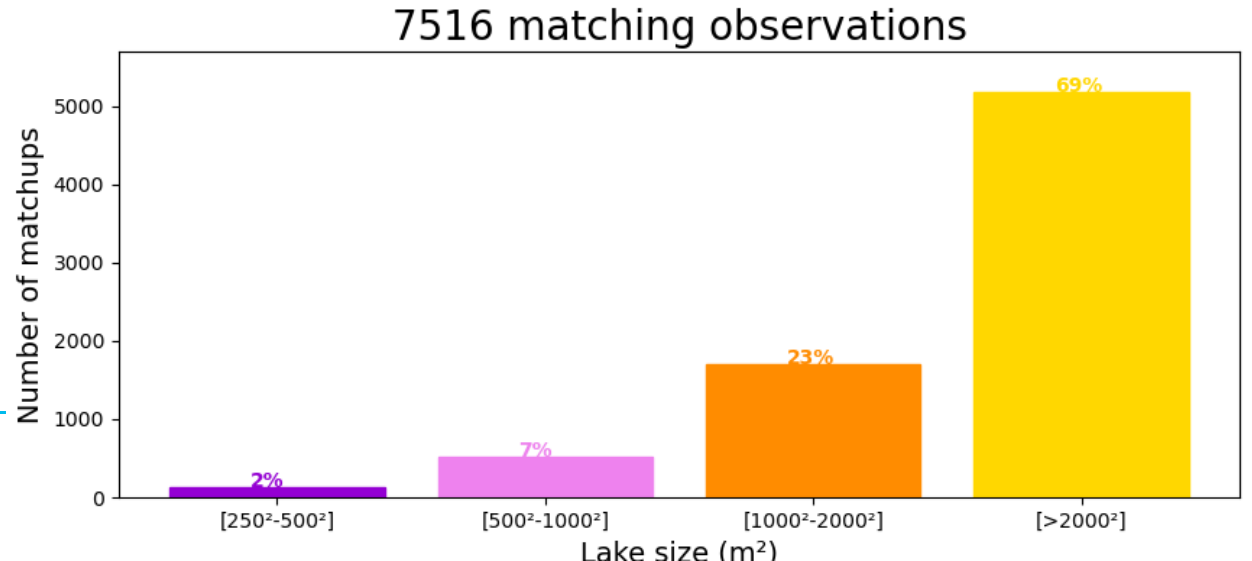
Cal/Val phase





Science phase *(from Nov 22th 2023)*





FILTERING OF RESULTS

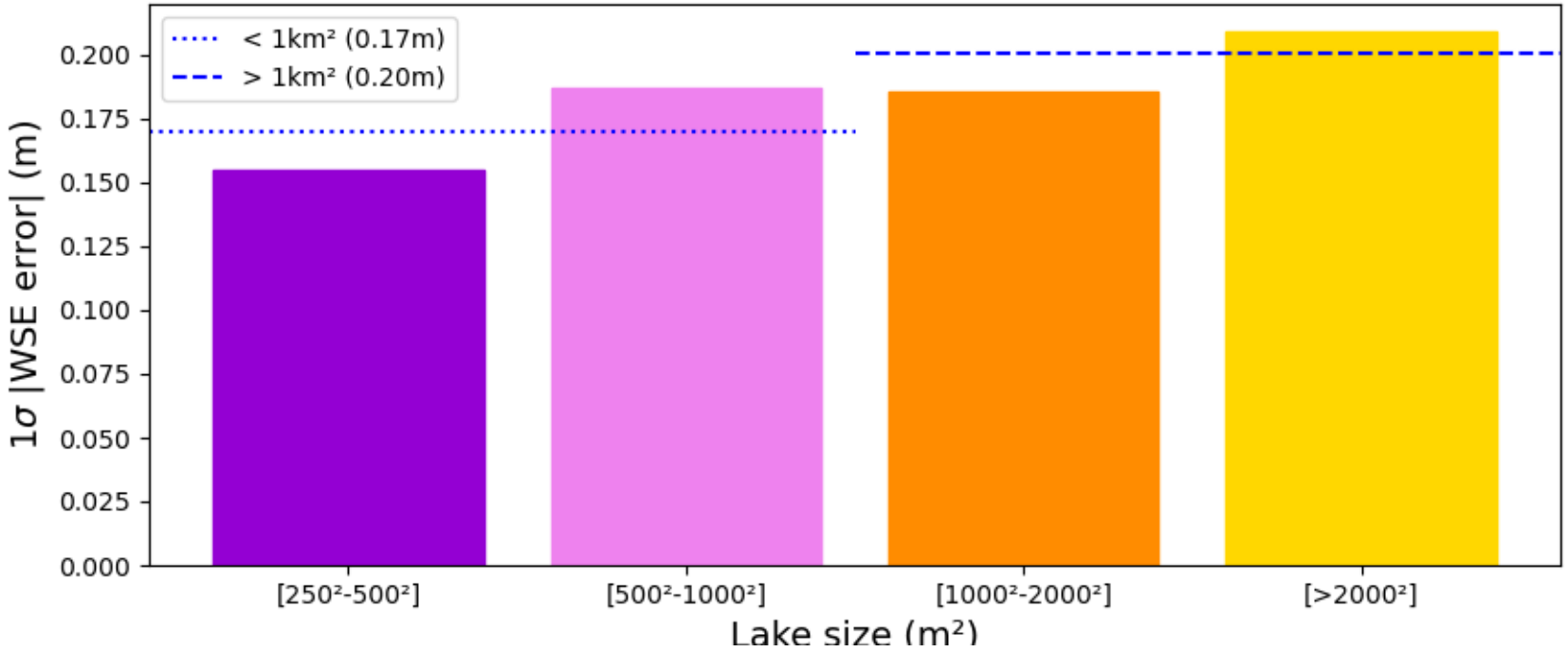
Based on various combinations of quality flags/indicators

Flag name	Flag definition
quality_f	Summary quality indicator for the lake measurement • 0 if nb_good_pixels / total_nb_pixels > 70% • 1 otherwise
xovr_qual_q	Quality of the cross-over calibration: 0=good; 1=suspect; 2=bad
ice_f	Ice cover flag, from in situ data, or climatological flag given in LakeSP products [Yang et al. 2020] • 0=no ice cover • 1=uncertain ice cover • 2=full ice cover
partial_f	Flag that indicates only partial lake coverage: 0=covered; 1=partially covered
dark_frac	Fraction of lake area_total covered by dark water

LAKE WSE ERROR – BASIC FILTERING

quality_f = 0

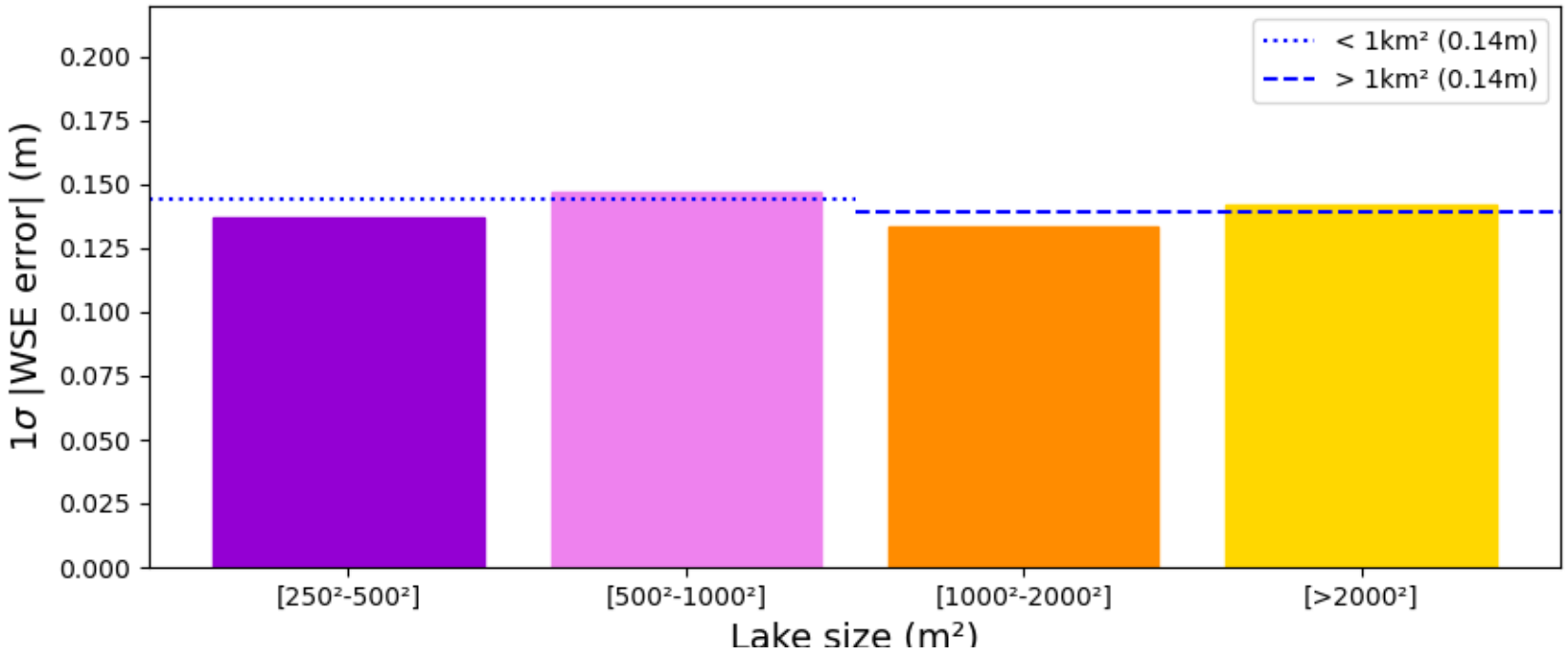
461 lakes – 8073 matchups



LAKE WSE ERROR – IMPACT OF ICE FLAG

quality_f = 0 & ice_flag =0 OR 1

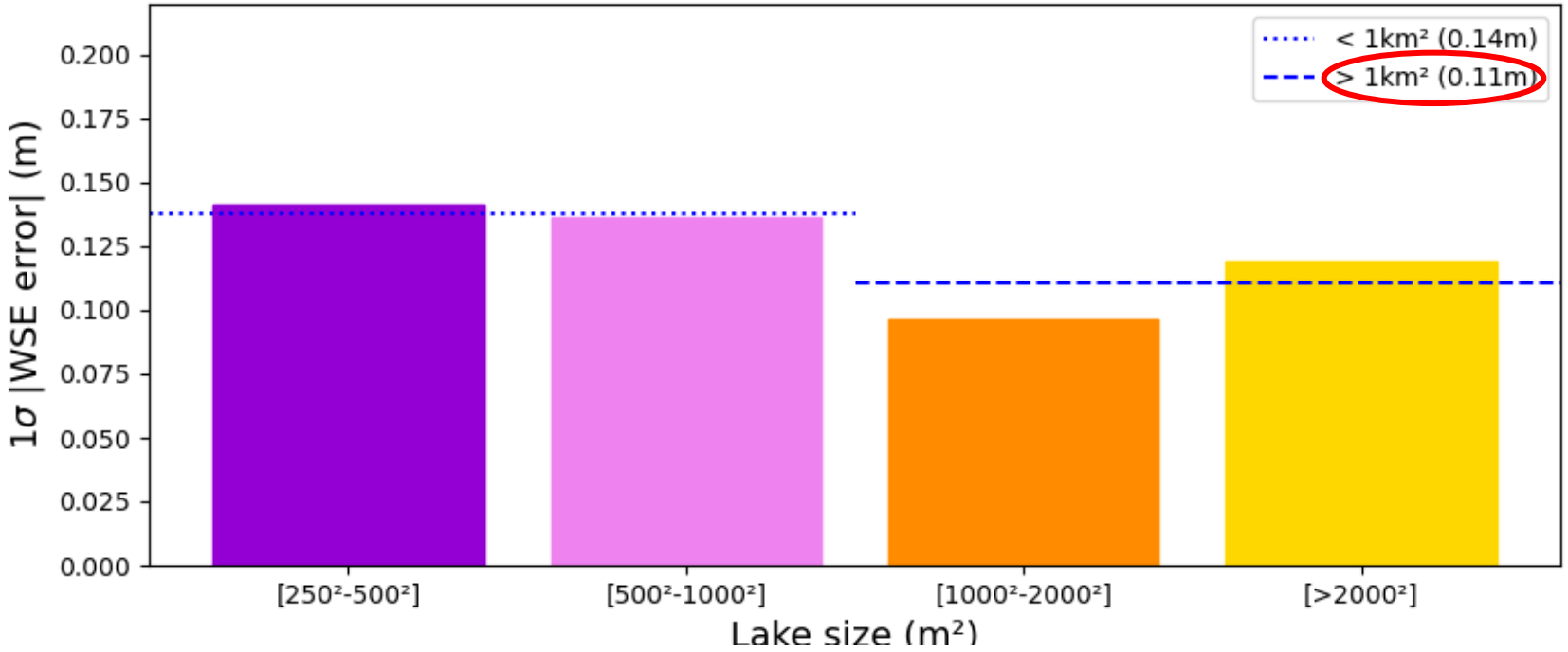
386 lakes – 6054 matchups



LAKE WSE ERROR – IMPACT OF ICE FLAG

quality_f = 0 & ice_flag =0

257 lakes – 3587 matchups



- Impact mainly on large lakes, for which ice coverage is less uniform

LAKE WSE ERROR

Impact of filtering of quality flags/indicators on WSE error

Flag	Value	1σ|error| for lakes < 1km²	1σ|error| for lakes > 1km²
quality_f	0	17cm	20cm
+ ice_f	0&1 0	14cm 14cm	14cm 11cm
+ xovr_qual_q	0&1 0	14cm 14cm	10cm 10cm
Baseline + partial_f	0	14cm	9cm
Baseline + dark_frac	< 50%	13cm	10cm

Baseline

Not significant: too few matchups with dark water > 50% (less than 100)

Lakes in the Pyrenees (France) – Comparison of LakeSP and in situ WSE

Blue: PLD
Yellow circle: PIXCVec
Green line: Prior

- Part of the lakes in the PLD never observed
- Some lakes not in the PLD observed
- Overdetection of water oberved
- Some unexplained outliers in the time series
- For large lakes like Fontargente 1: 2/3 of the cycles were observed
- Smaller like Estanyols: only half of the cycles observed

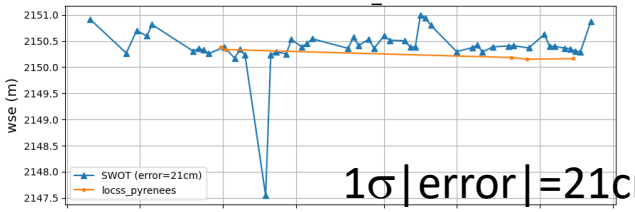
Fontargente 2

Fontargente 1

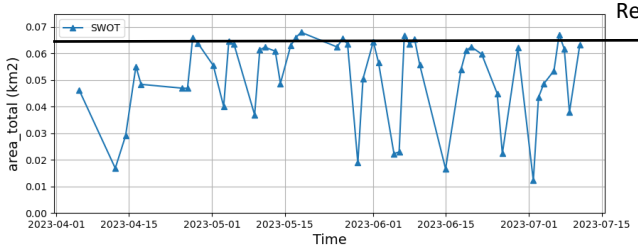
Estanyols

PLD not observed

< 0.01 km²

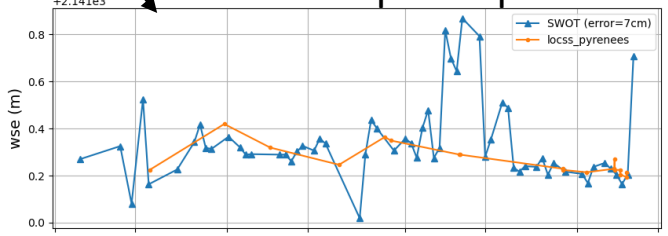


1σ|error|=21cm

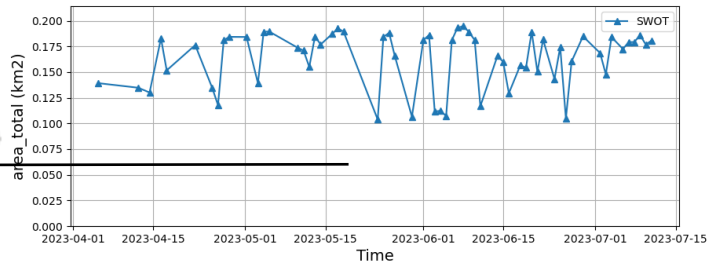


Req (250*250m)

1σ|error|=7cm

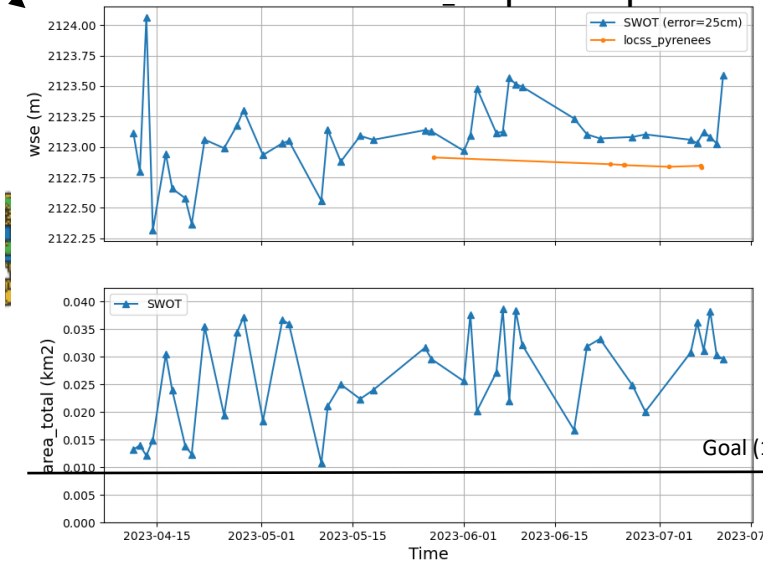


Over-detection: no lake here



Req (250*250m)

1σ|error|=25cm



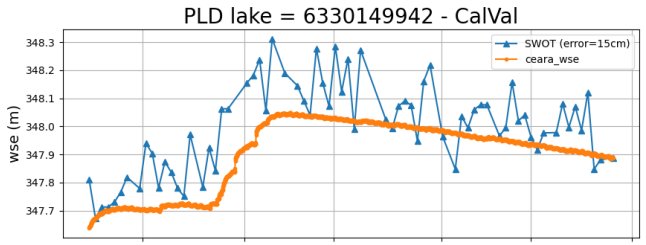
Goal (100*100m)

Courtesy: Jean François Cretaux

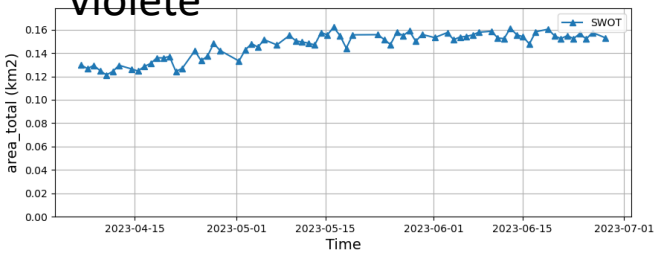
FUNCEME | Courtesy: Marielle Gosset, Raphael Reis, and many others.

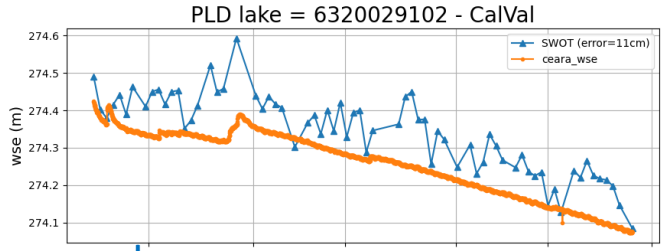
Institut de Recherche pour le Développement FRANCE

Ceara (Brazil) – WSE dynamics seen by LakeSP products (Cal/Val phase)

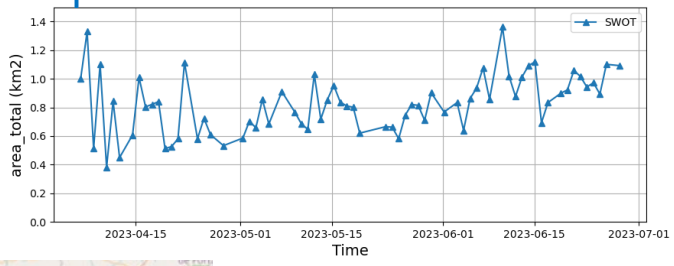


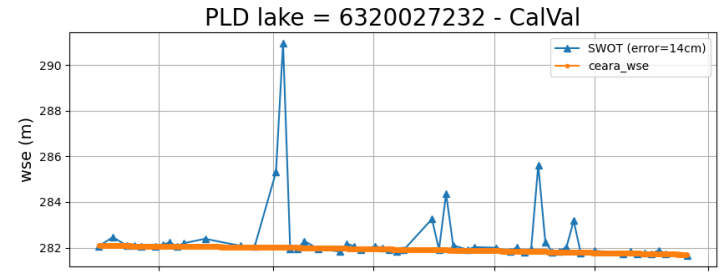
violete



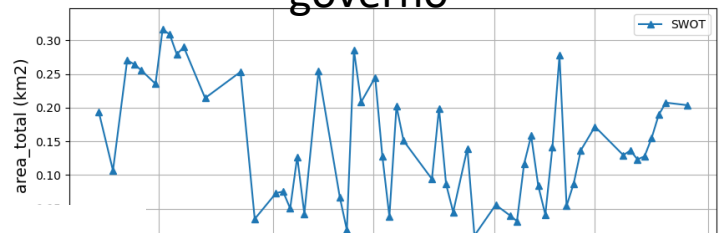


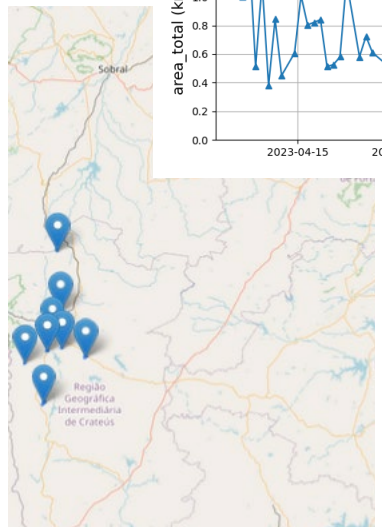
palmares

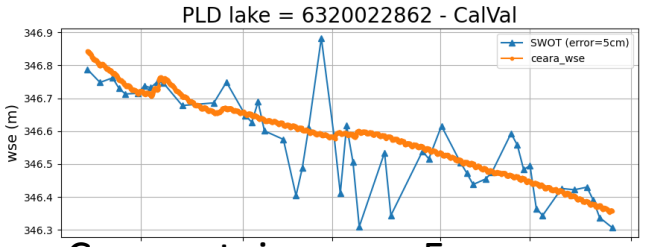




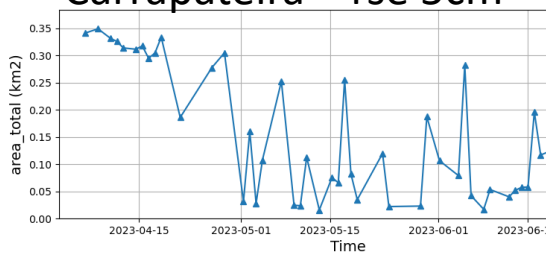
governo

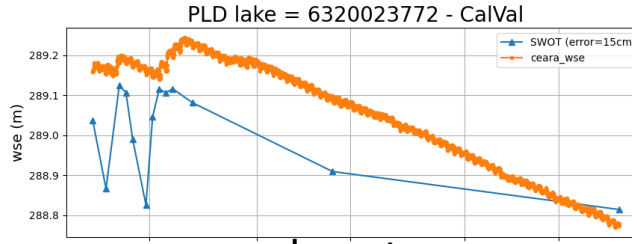




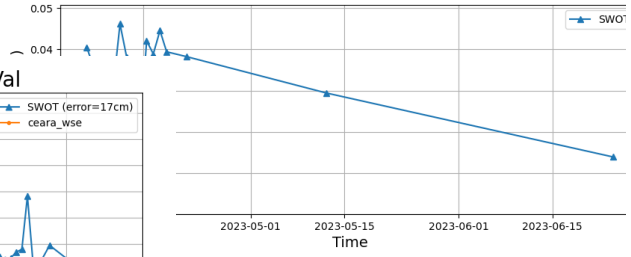


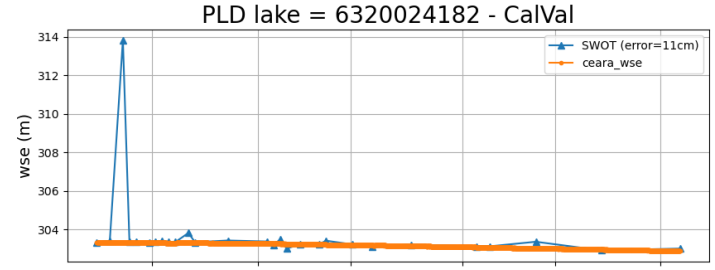
Carrapateira – rse 5cm



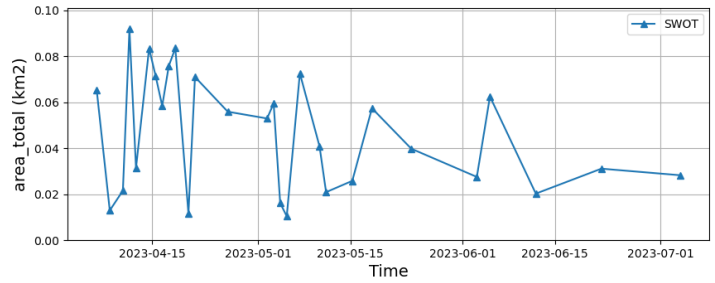


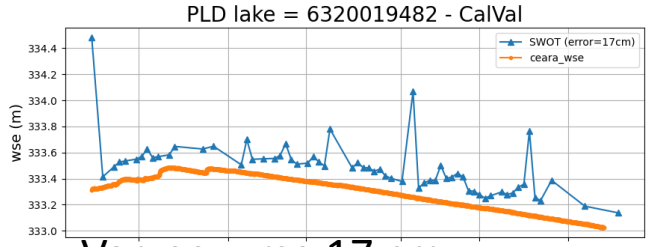
deserto



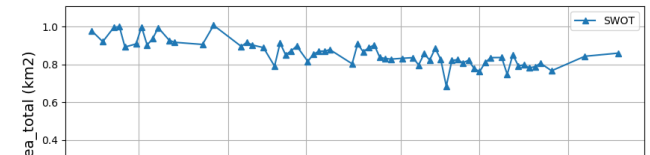


convento





Varzea – rse 17 cm



- **WSE well represents dynamics (outliers need to be understood and filtered out)**
- Surface area more unstable…

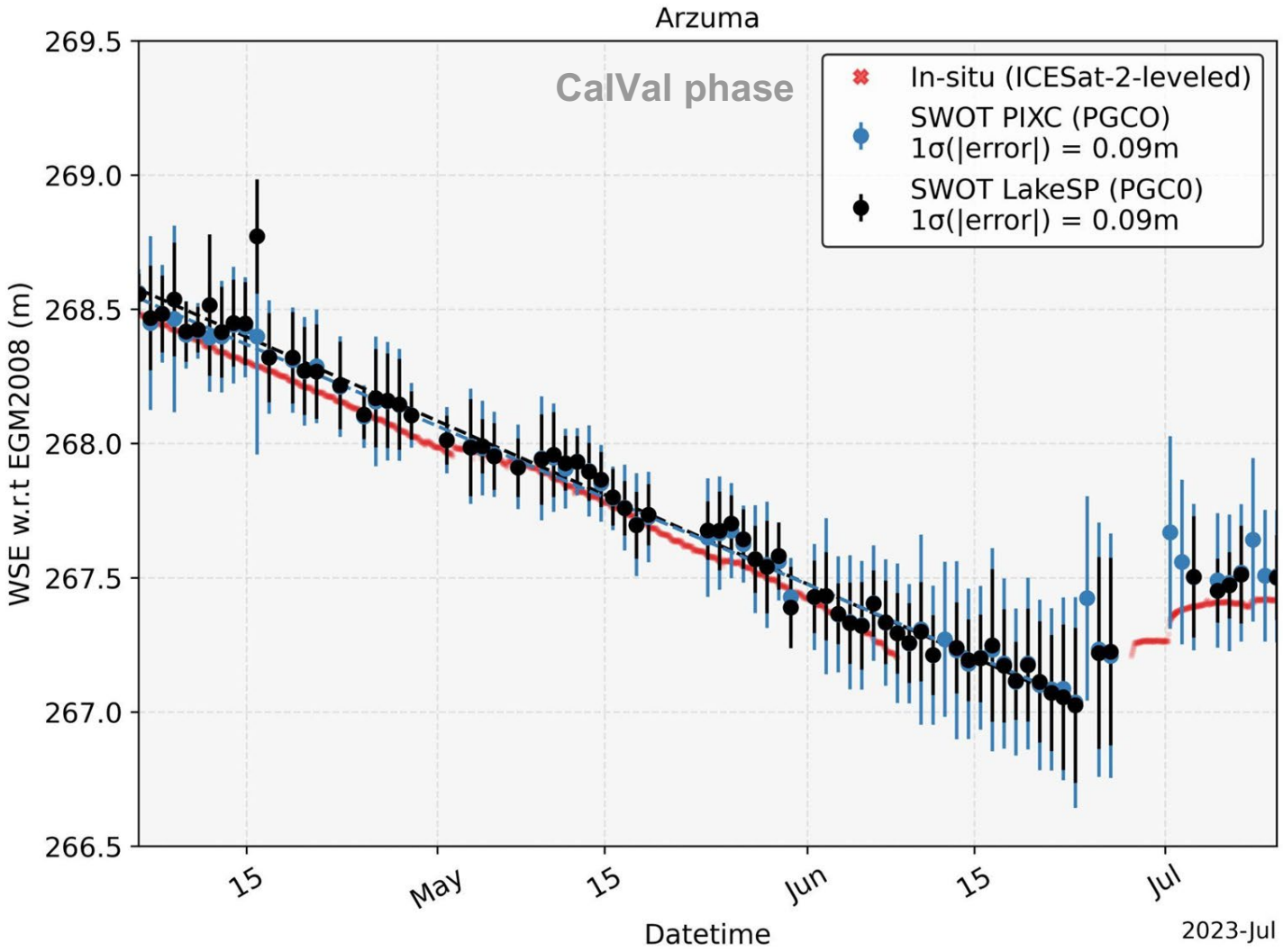
SWOT DOWNSTREAM

Early SWOT validation results on Arzuma shallow reservoir in West Africa

PIXC and LakeSP WSE assessment

F. Girard[1,2], M. Grippa[1], L. Kergoat[1], M. Vayre[2], J. Renou[2] and N. Taburet[2]

[1]GET, Toulouse (France), [2]CLS, Ramonville-Saint-Agne (France)



Comparison of SWOT PIXC and LakeSP data with in-situ (ICESat-2-leveled).

PIXC WSE of a lake is computed as median value of WSE of open water (class 4) pixels within PLD polygon.

SWOT PIXC and LakeSP data version is PIC0 or PGC0.

CalVal phase WSE assessment:

- **1σ (|error|):** 0.09 m (LakeSP and PIXC aggregated to lake level)
- **PIXC pixel-level standard deviation:** 0.10 - 0.40 m (mostly below 0.25 m)
- **Water level decrease during the dry season:** 1 mm/day error

SUMMARY AND OUTLOOK

- It is essential to use quality flags to filter out bad (and suspect) data:
  - quality_f
  - ice_f, impacting mainly large lakes
  - Lower impact of xover_cal_q (if already filtering quality_f), partial_f and dark_frac

- Ongoing work:
  - Analyze the entire reprocessed SWOT dataset (version “C”)
  - Refine quality_f in the LakeSP product
  - Enlarge in situ dataset with other leveled data (analysis in progress)
  - Look at unleveled gauge data (relative WSE errors rather than absolute errors)

LAKE AREA VALIDATION

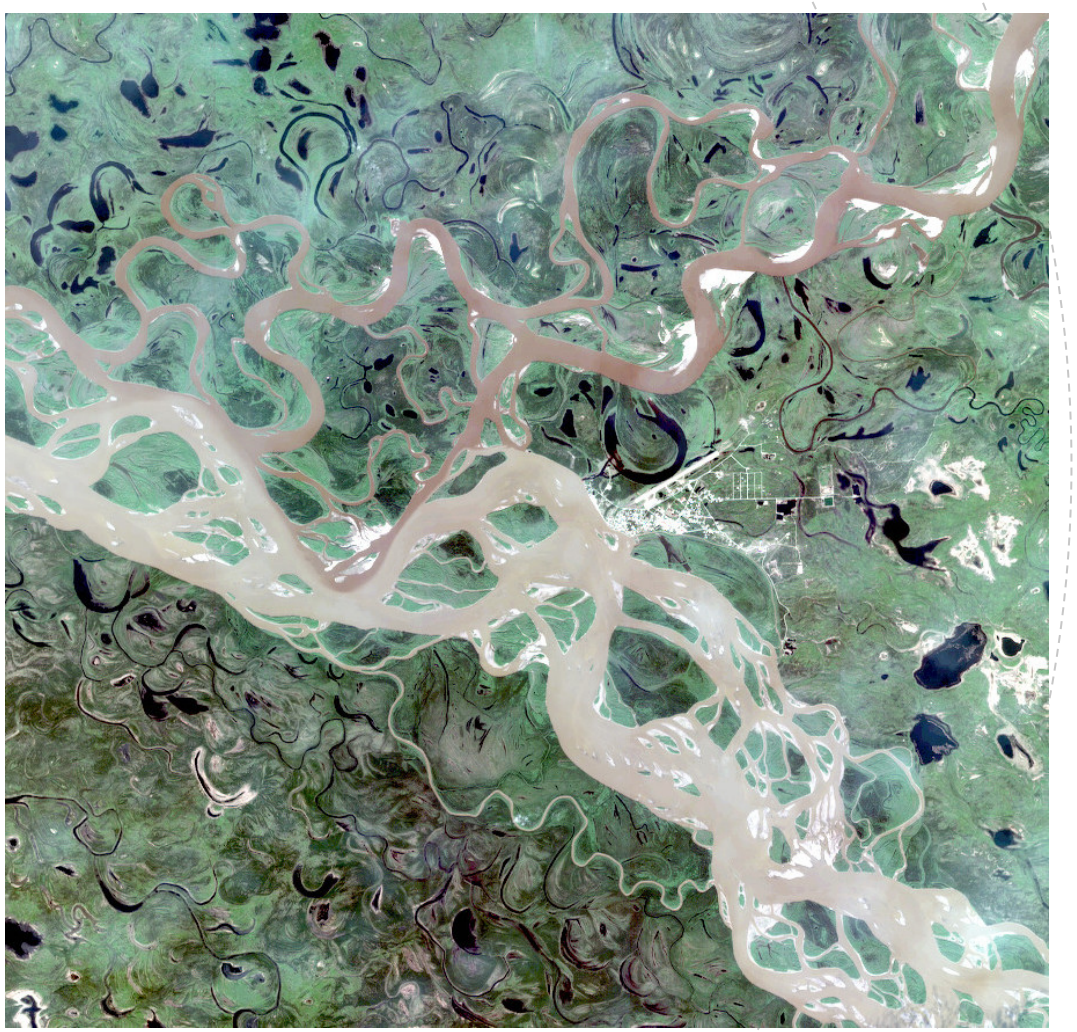
Lake area reference data are based on water masks derived from high-resolution optical and radar satellite images.

- 34 x Pleiades [0.5 m]
- 204 x Radarsat Constellation Mission (RCM) [5 m]
- 283 x Sentinel-2 (S2) [10 m]

- Mainly over Cal/Val sites and during Cal/Val phase

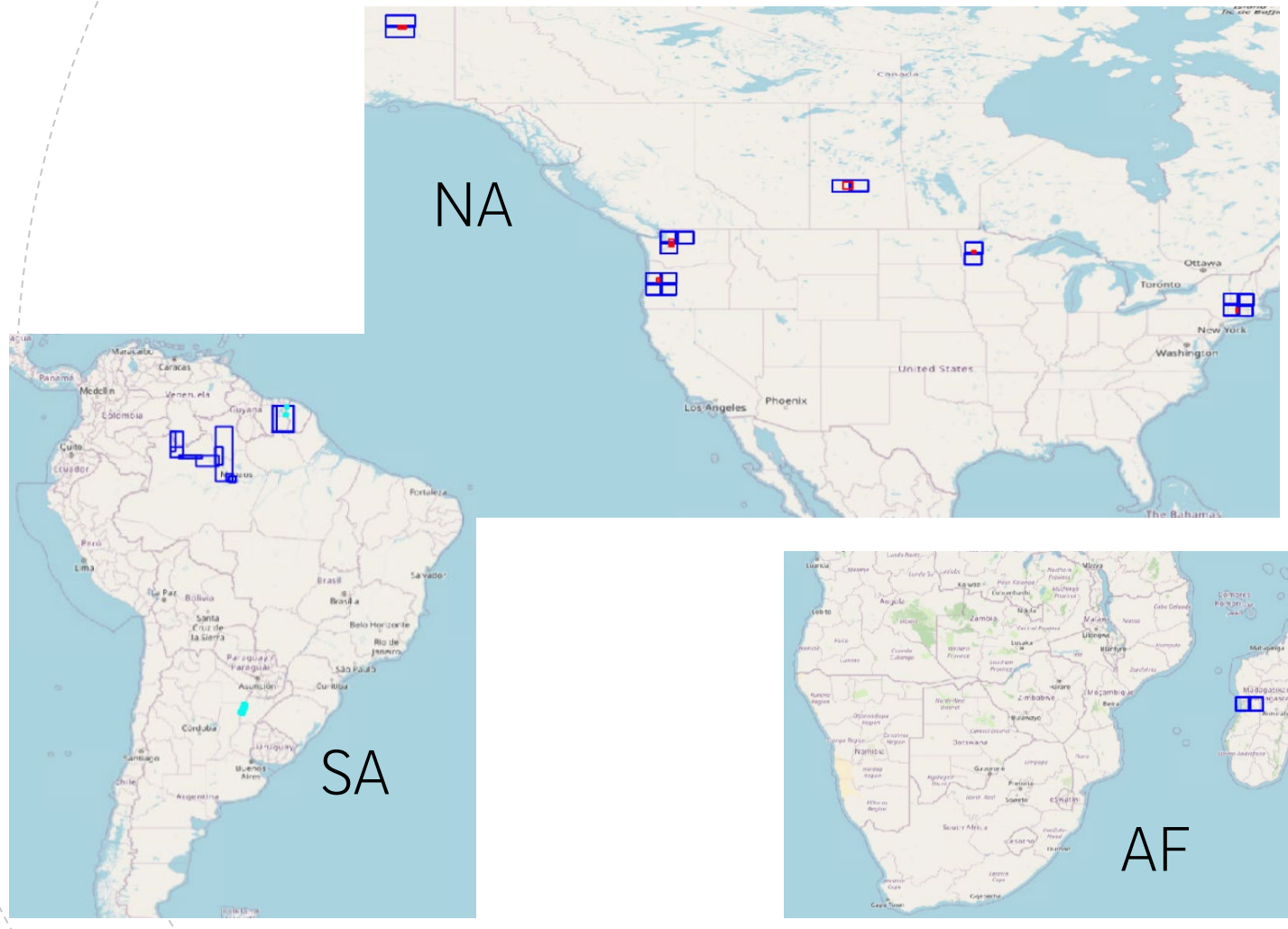
- Pre-processed to obtain reference areas [m^2] for all PLD lakes covered by the images
  - >10 000 lakes with matching SWOT LakeSP data

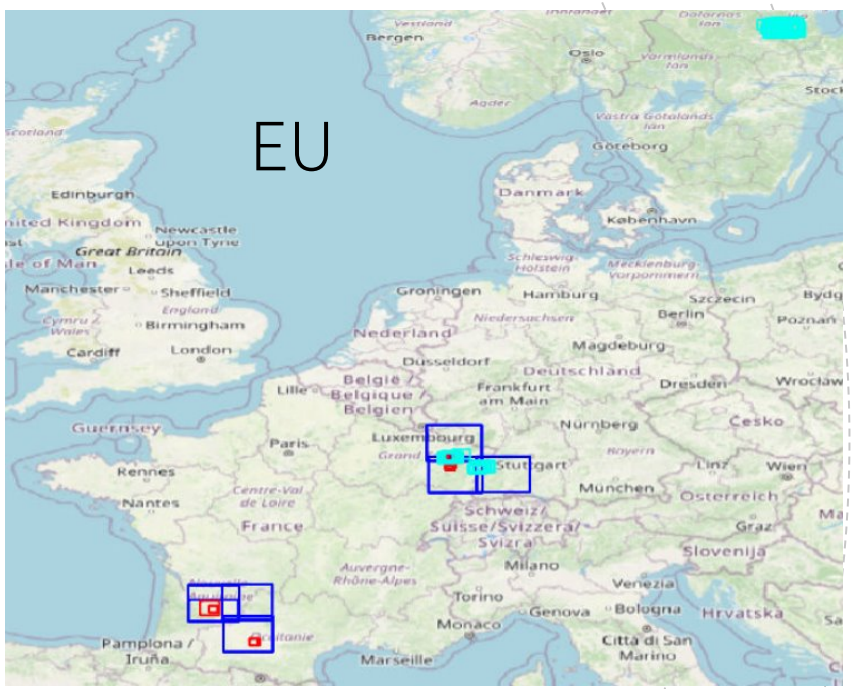
- Lake area error metric: |relative area error| (1σ) [%]
  - |relative area error| = |area_total-area_truth|/area_truth



Pleiades image, Yukon Flats, Alaska, June 6, 2023

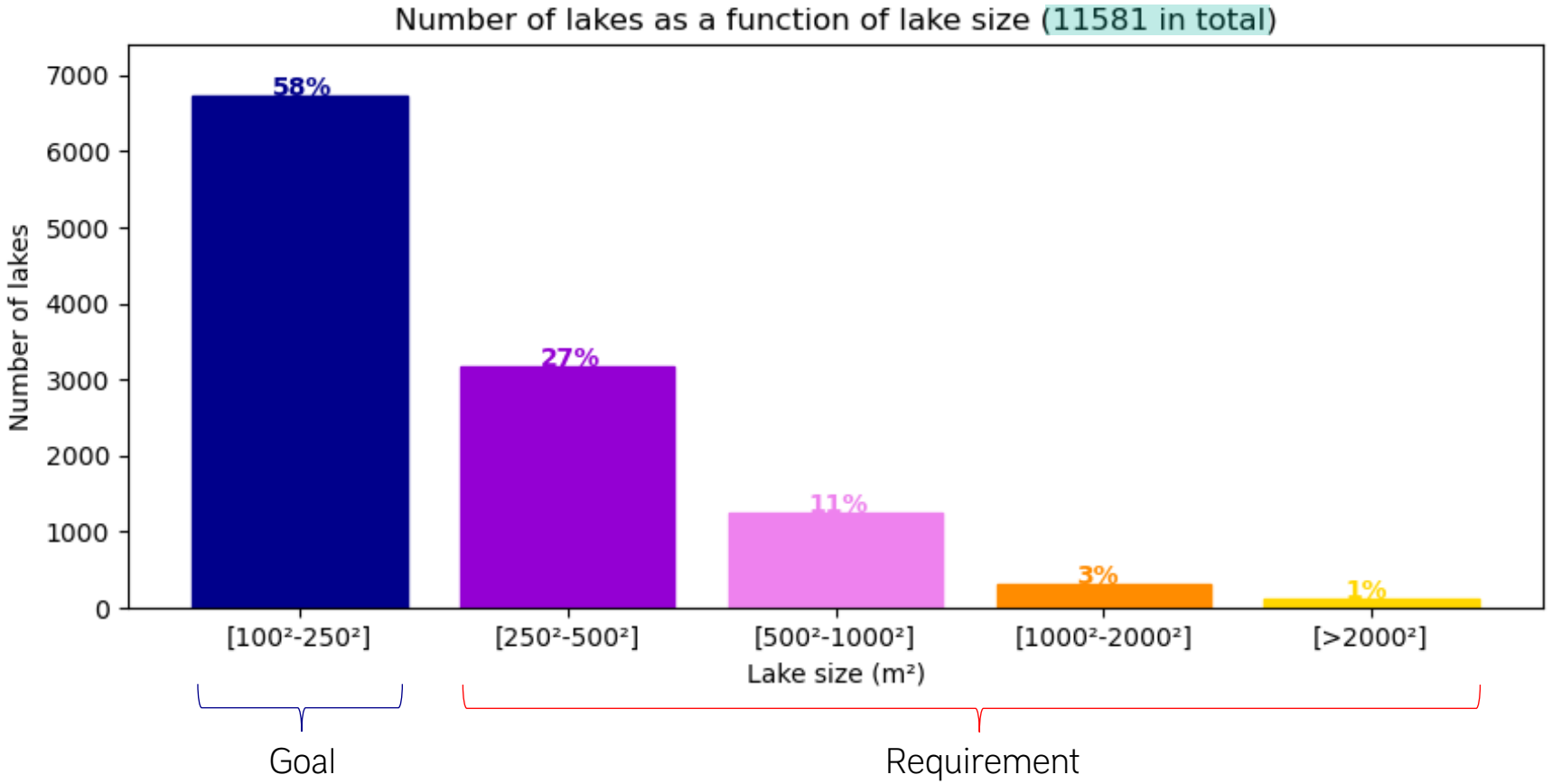
GEOGRAPHICAL DISTRIBUTION OF AREA REFERENCE DATA





283 x S2 [10 m]
204 x RCM [5 m]
34 x Pleiades [0.5 m]

SIZE DISTRIBUTION OF PLD LAKES WITH MATCHING AREA REFERENCE DATA



- Corresponds roughly to expected global distribution: decreasing number of lakes with increasing size

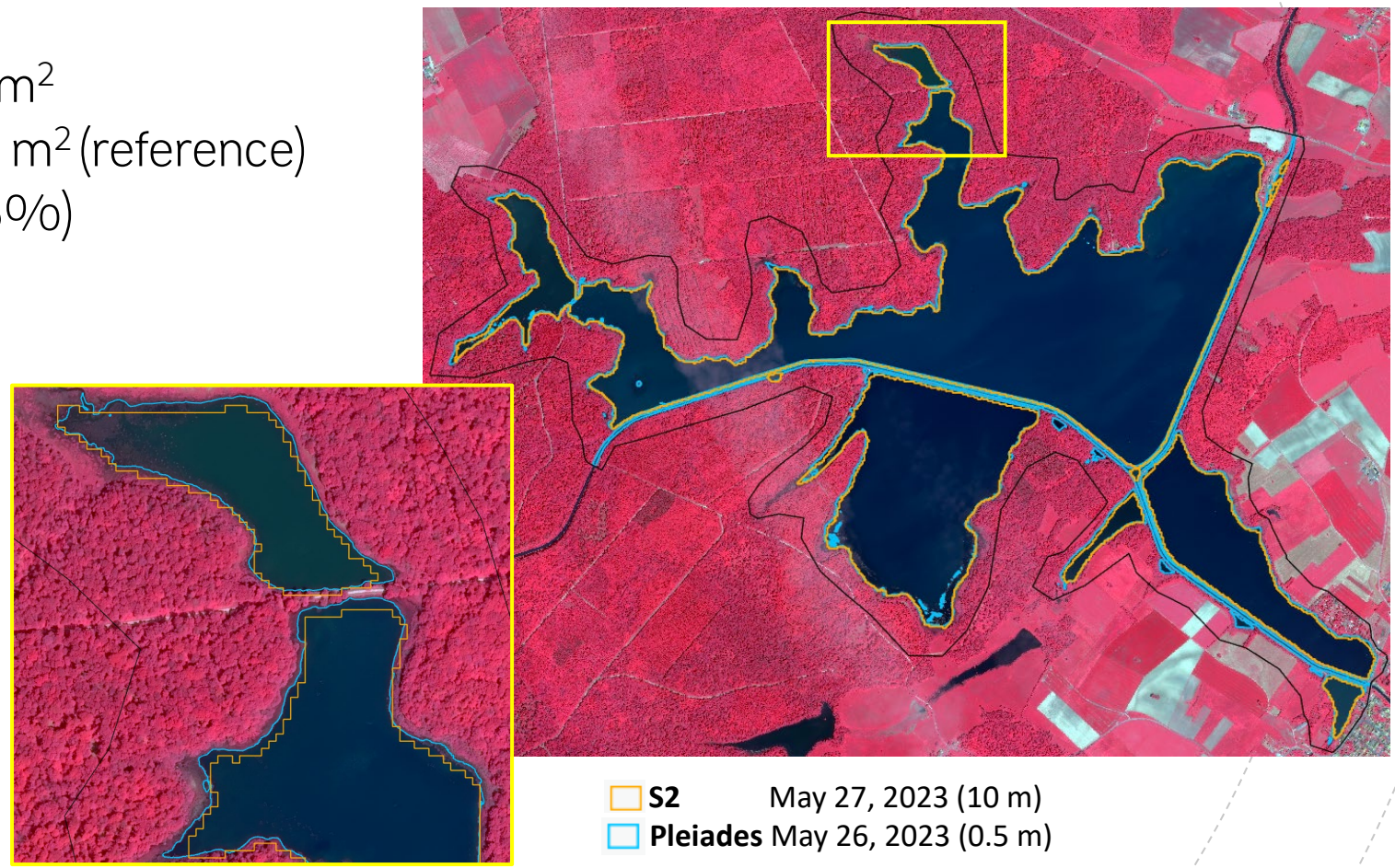
ACCURACY OF REFERENCE AREAS

Example: S2 vs. Pleiades water mask

- Gondrexange Lake, France, ~5.5 km^2
  - Pleiades water mask: 5460560 m^2 (reference)
  - S2 water mask: 5161300 m^2 (-5%)

Indicator	Score
Precision	0.98
Recall	0.93
Fscore	0.96
CSI	0.92

- Reference water areas are not perfect.
  - S2 water masks tend to underestimate area.



Courtesy: Sabrine Amzil, Thomas Ledauphin, Jérôme Maxant, Hervé Yésou

FILTERING OF REFERENCE AND SWOT DATA

- LakeSP_Prior products (and Science Requirements/Goals for area) exclude:
  - Lakes outside the nominal swath (10-60 km)
  - Lakes whose size is below 100x100 m^2
- Reference data that are clearly erroneous are discarded (but we may have missed some)
- Identification of matching LakeSP and in situ / reference data:
  - Spatially: intersection with the PLD lake polygon (truth processing for water masks)
  - Temporally: reference mask nearest to SWOT acquisition (maximum 3 days by default)
- Filtering (inclusion criteria) based on various combinations of LakeSP quality flags/indicators:
  - partial_f=0
  - quality_f =0
  - xovr_cal_q =0
  - ice_f=0

  } We call this our baseline filtering in what follows

  The area of a partially covered lake should not be used
  Significant impact of allowing quality_f =0 OR 1
  Limited impact of allowing xovr_cal_q=0 OR 1 OR 2
  Limited impact of allowing ice_f=0 OR 1 OR 2

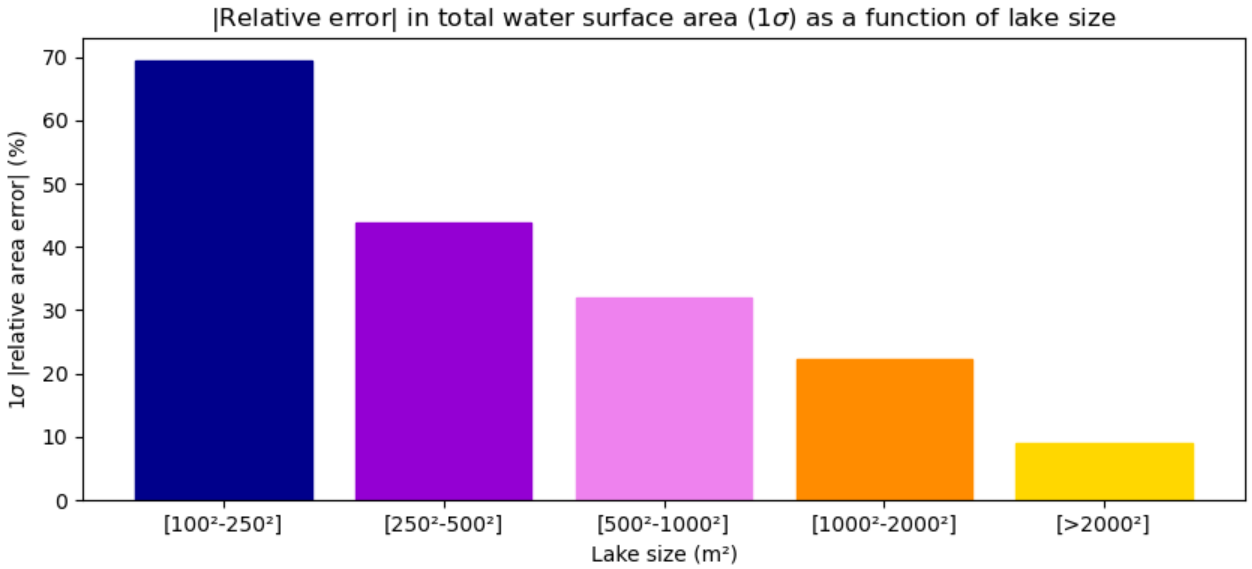
  } Indvidual results not shown
- Additional filtering tested w.r.t. the above:
  - dark_frac <50%
  - |area_truth −area_PLD| / area_PLD <50%
  - time − time_truth < 1 OR 2 days (3 by default)

  Very limited impact – Results not shown

RELATIVE LAKE AREA ERROR

Baseline filtering = partial_f=0, quality_f =0, xovr_cal_q =0, ice_f=0

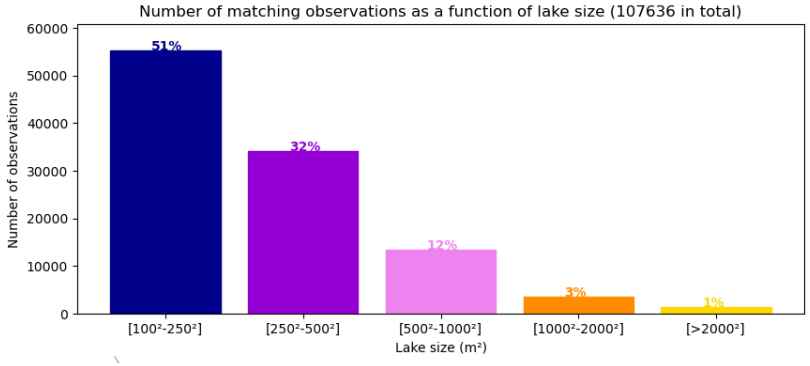
All water mask types

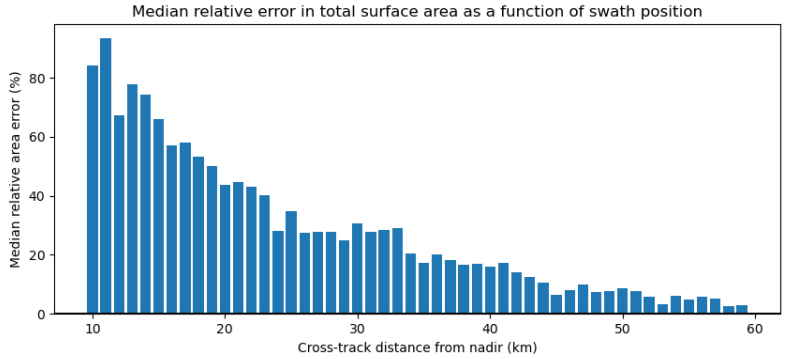


> 250x250 m^2: 38% (1σ)

> 1 km^2: 17% (1σ)

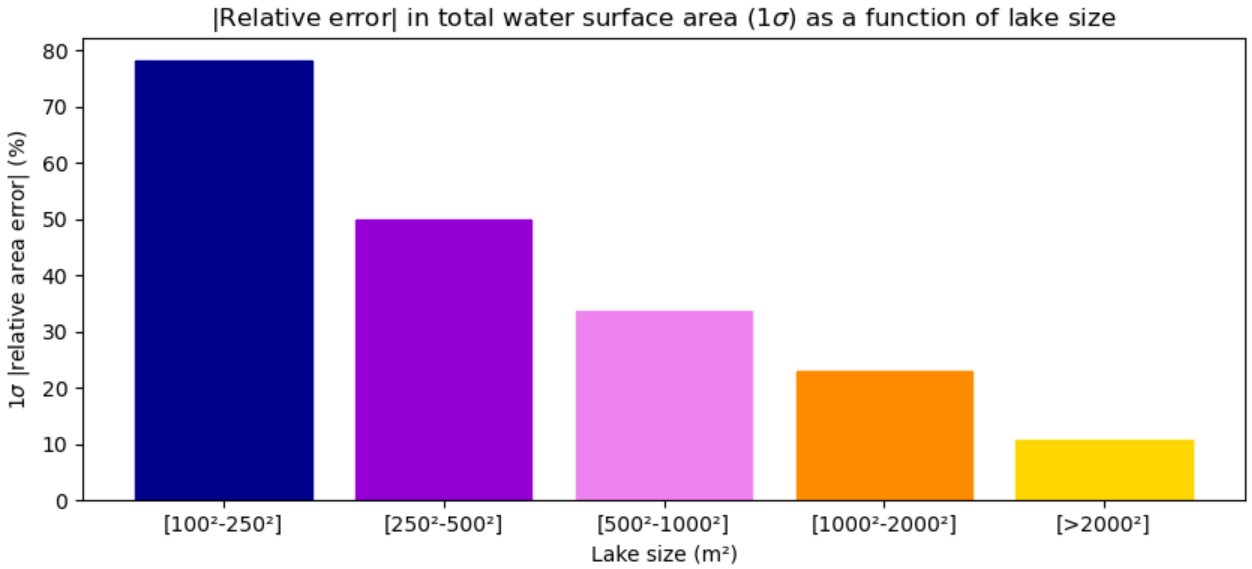
10191 PLD lakes





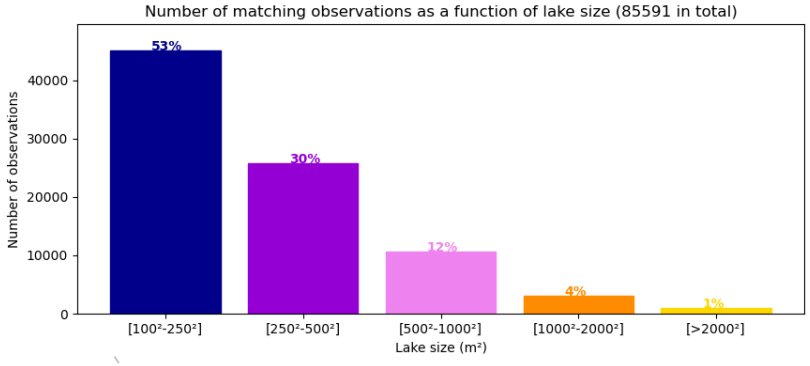
RELATIVE LAKE AREA ERROR

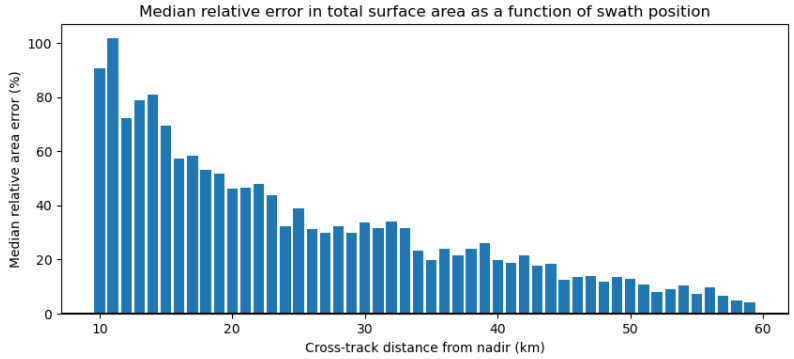
Baseline filtering = partial_f=0, quality_f =0, xovr_cal_q =0, ice_f=0 S2



> 250x250 m^2: 42% (1σ)
> 1 km^2: 19% (1σ)

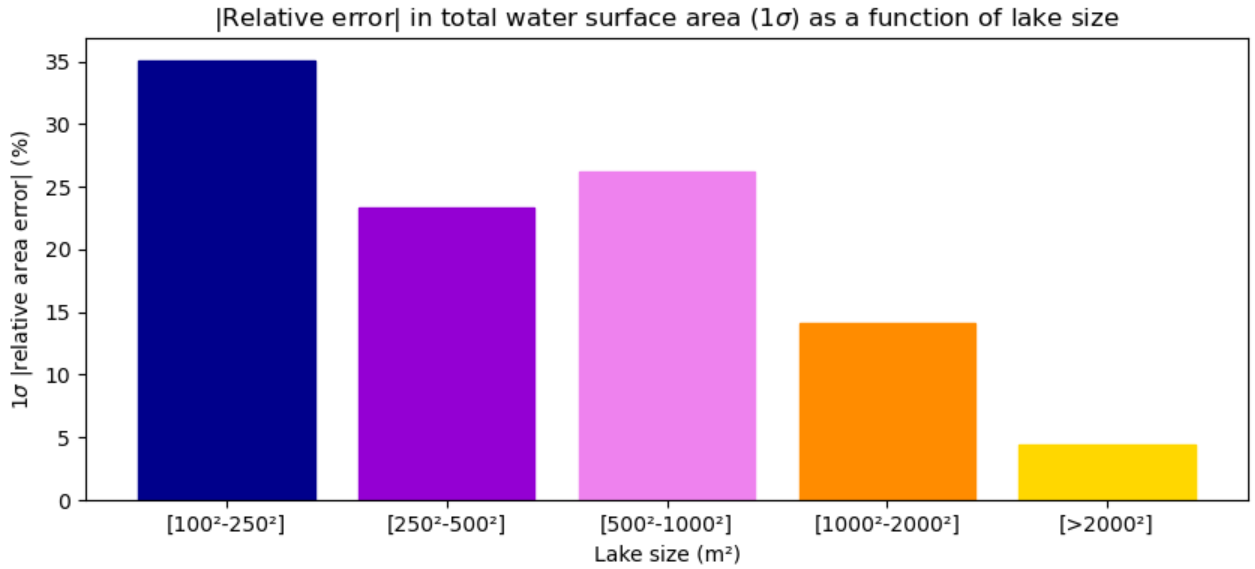
8973 PLD lakes





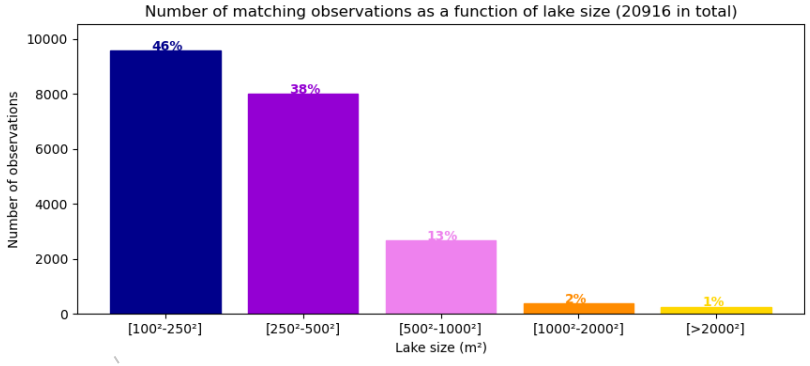
RELATIVE LAKE AREA ERROR

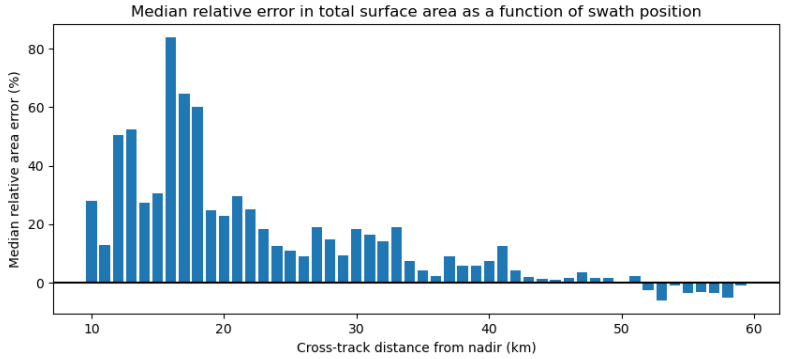
Baseline filtering = partial_f=0, quality_f =0, xovr_cal_q =0, ice_f=0 RCM



> 250x250 m^2: 23% (1σ)
> 1 km^2: 9% (1σ)

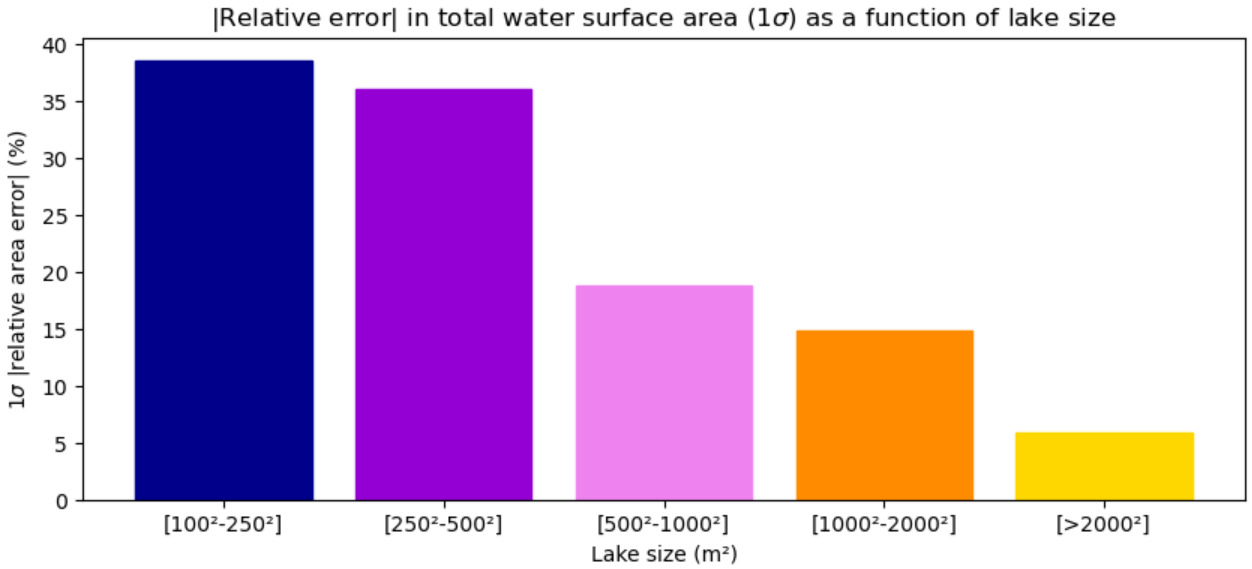
981 PLD lakes





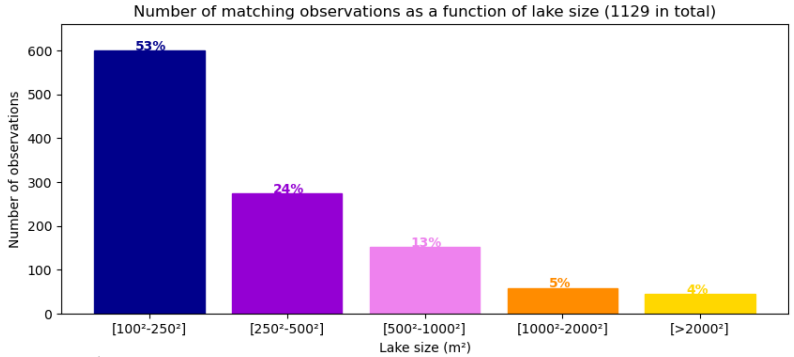
RELATIVE LAKE AREA ERROR

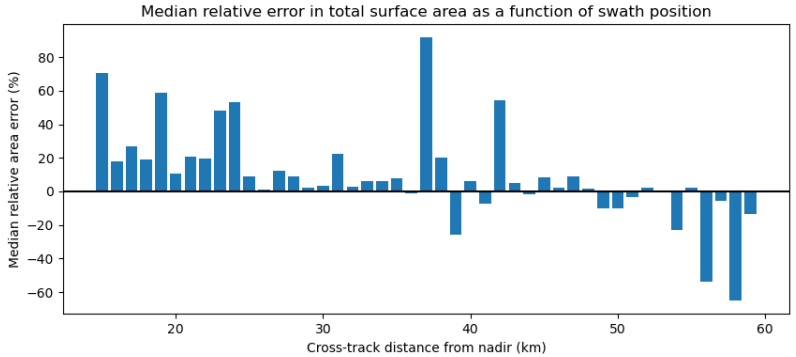
Baseline filtering = partial_f=0, quality_f =0, xovr_cal_q =0, ice_f=0 Pleiades



> 250x250 m^2: 24% (1σ)
> 1 km^2: 10% (1σ)

237 PLD lakes

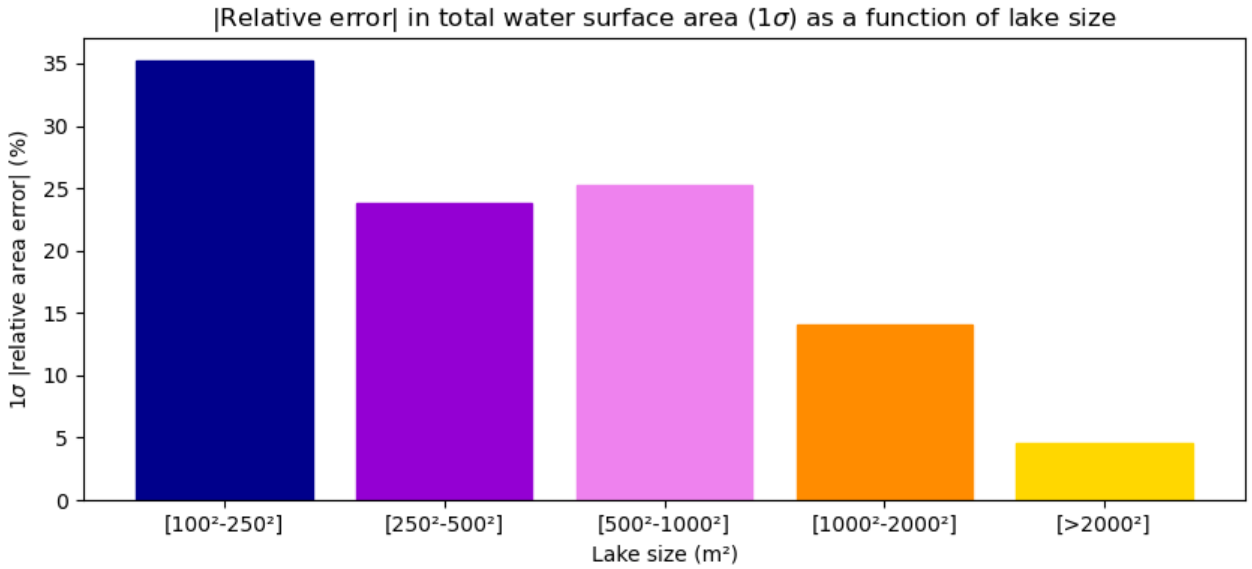




RELATIVE LAKE AREA ERROR

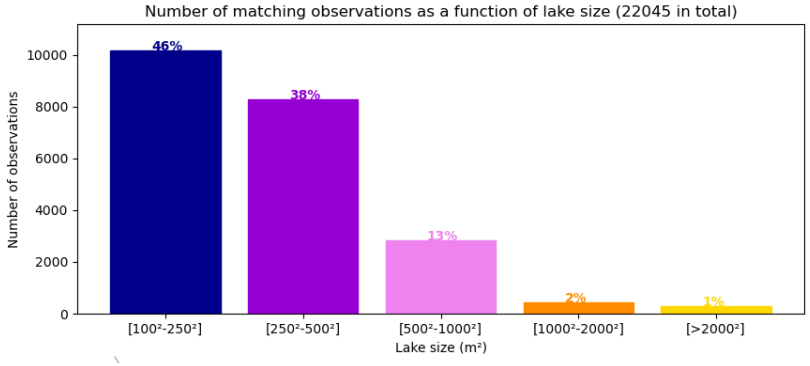
Baseline filtering = partial_f=0, quality_f =0, xovr_cal_q =0, ice_f=0

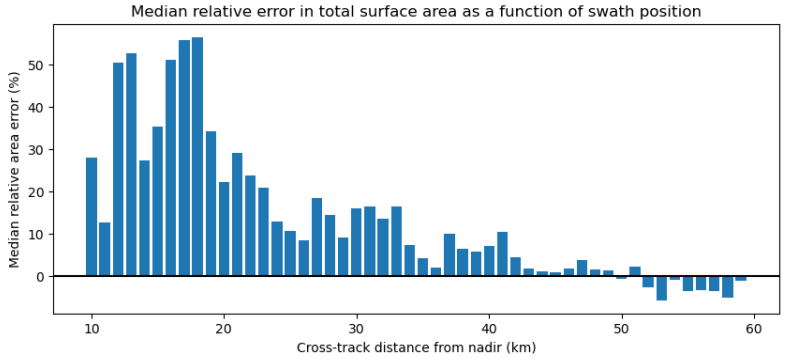
Pleiades + RCM



> 250x250 m^2: 23% (1σ)
> 1 km^2: 9% (1σ)

1218 PLD lakes

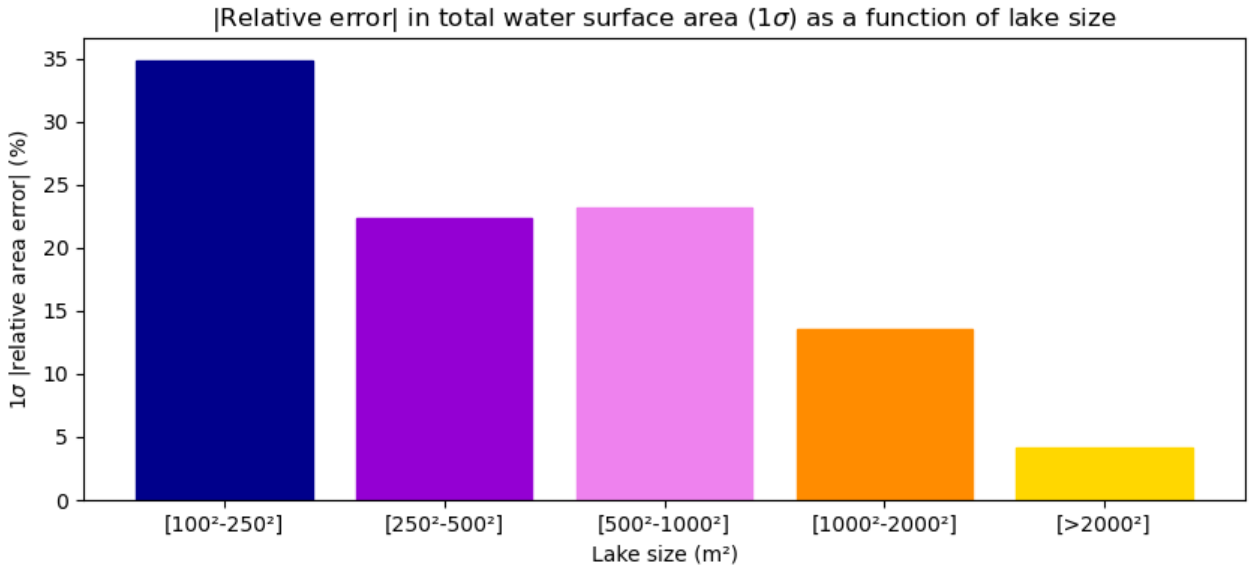




RELATIVE LAKE AREA ERROR

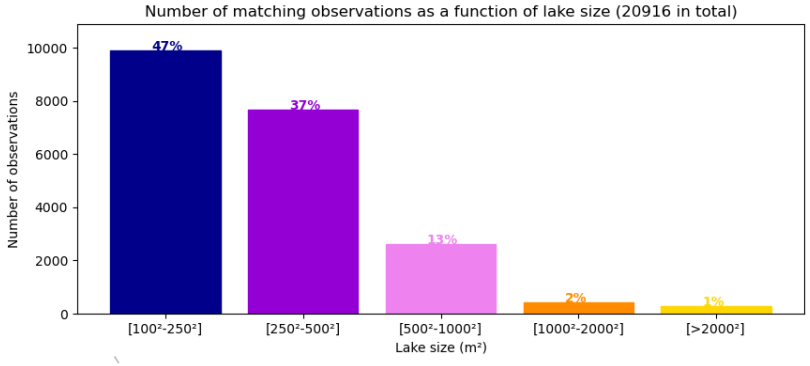
Baseline filtering + dark_frac <50%

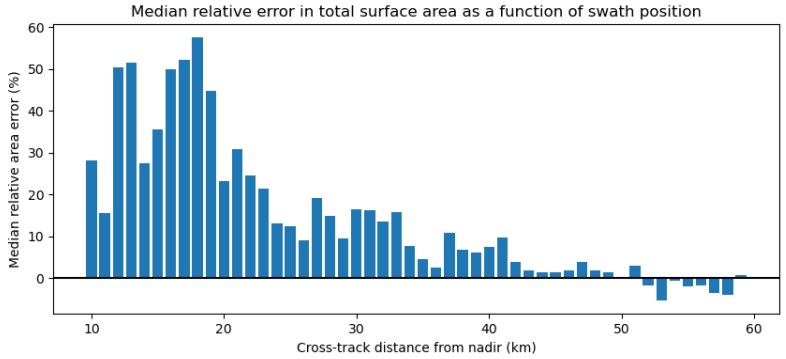
Pleiades + RCM



> 250x250 m^2: 22% (1σ)
> 1 km^2: 9% (1σ)

1218 PLD lakes

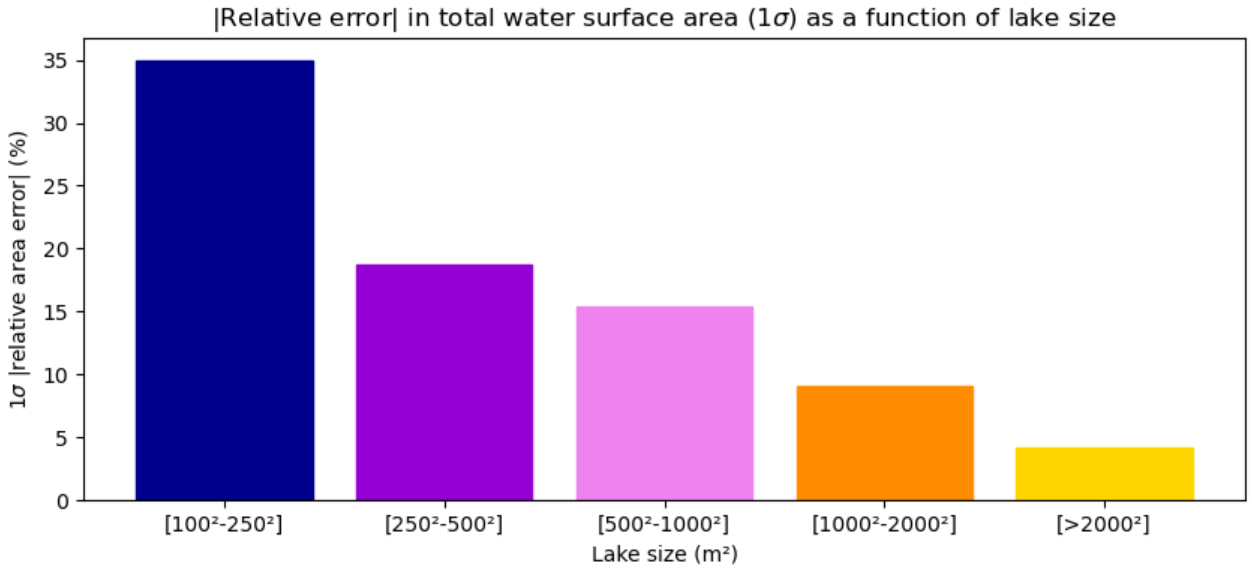




RELATIVE LAKE AREA ERROR

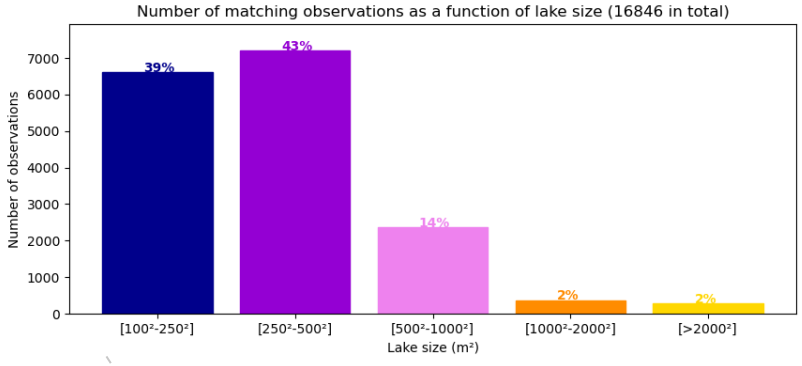
Baseline filtering + |area_truth – area_PLD|/area_PLD <50%

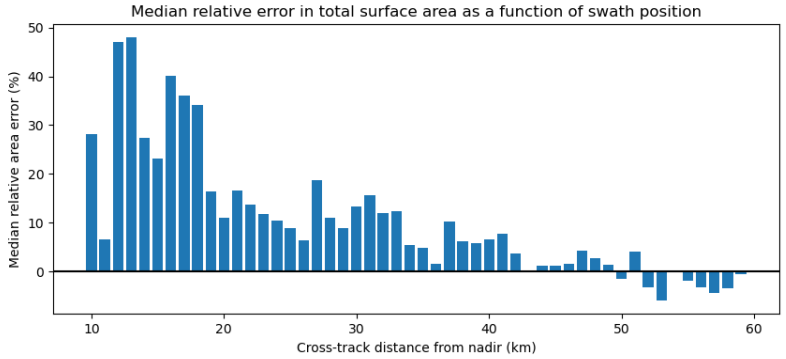
Pleiades + RCM



> 250x250 m^2: 17% (1σ)
> 1 km^2: 7% (1σ)

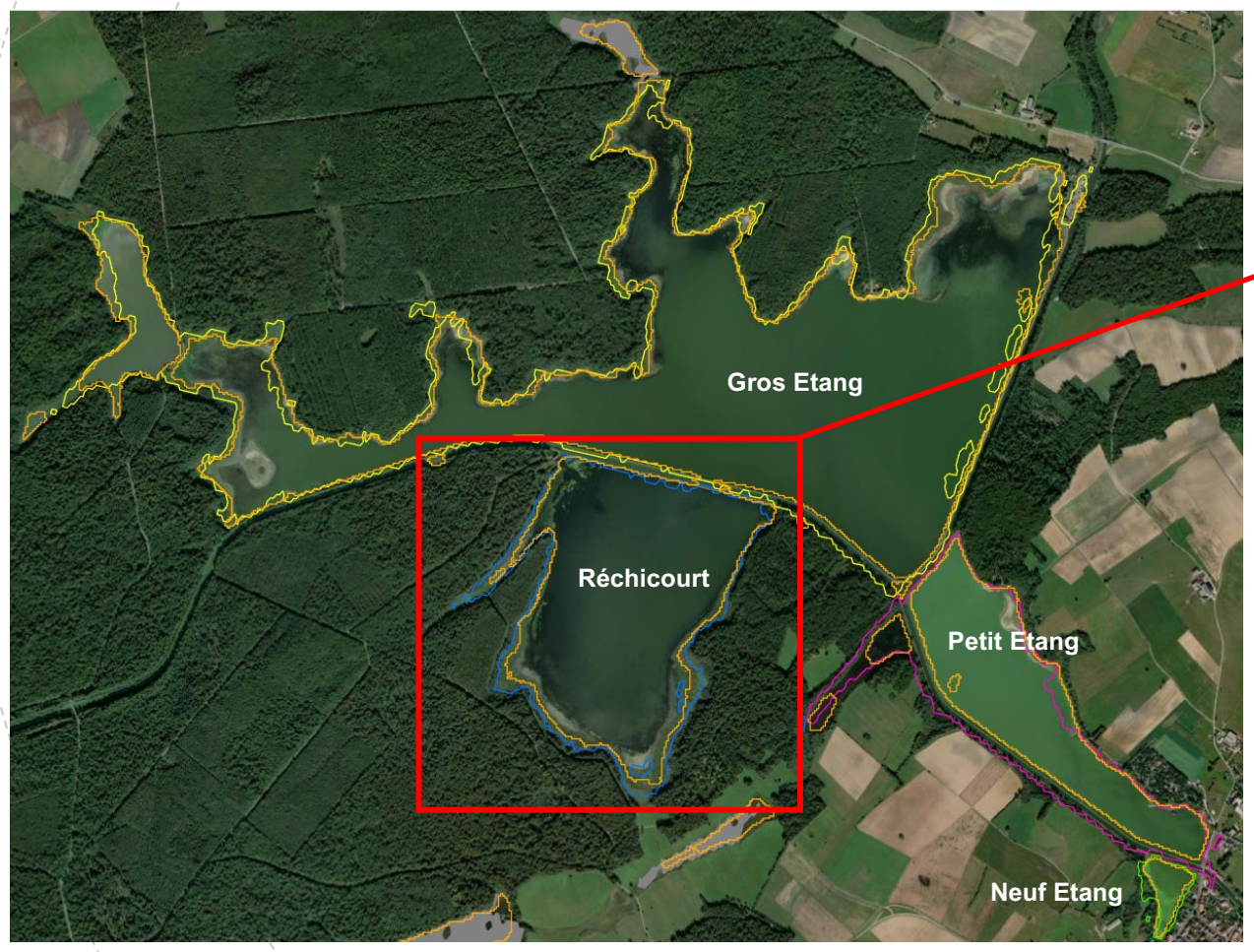
921 PLD lakes





EXAMPLE

Gondrexange – Réchicourt, France (0.93 km^2 in PLD)



S2 (June 8, 2023)

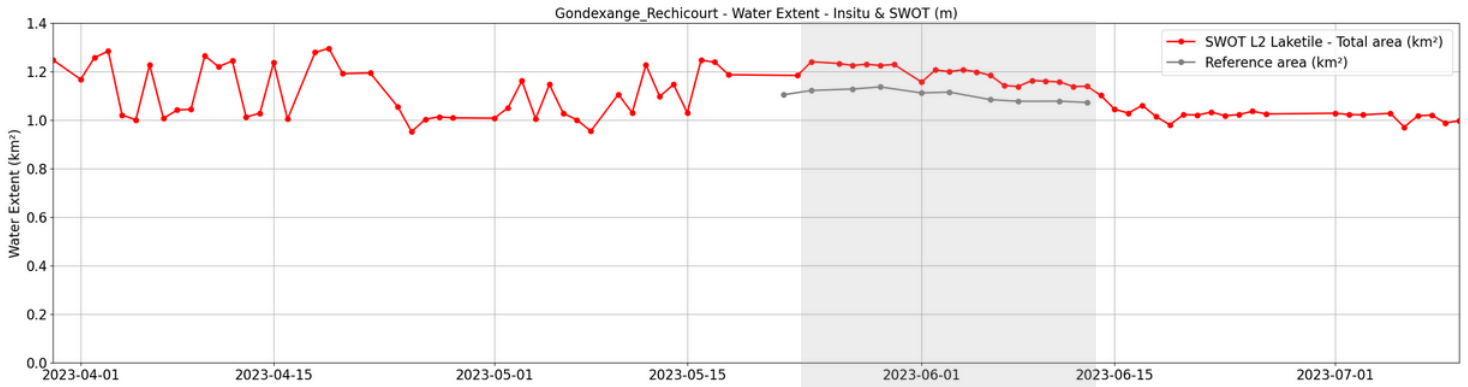
lake_id
- 2320162172 Neuf Etang
- 2320166602 Gros Etang
- 2320166622 **Réchicourt**
- 2320166632 Petit Etang

SWOT LakeSP (June 8, 2023)

Courtesy: Sabrine Amzil, Thomas Ledauphin, Jérôme Maxant, Hervé Yésou

EXAMPLE

Gondrexange – Réchicourt, France (0.93 km^2 in PLD)



	date_dt	area_total	area	diff	diff_percent
0	2023-05-22	NaN	1.1038	NaN	NaN
1	2023-05-24	1.239906	1.1215	-0.118406	-10.557824
2	2023-05-27	1.224501	1.1273	-0.097201	-8.622461
3	2023-05-29	1.224348	1.1360	-0.088348	-7.777113
4	2023-06-01	1.155810	1.1112	-0.044610	-4.014579
5	2023-06-03	1.198661	1.1146	-0.084061	-7.541809
6	2023-06-06	1.183346	1.0835	-0.099846	-9.215136
7	2023-06-08	1.137814	1.0768	-0.061014	-5.666233
8	2023-06-11	1.156486	1.0771	-0.079386	-7.370346
9	2023-06-13	1.138650	1.0722	-0.066450	-6.197538

|relative area error| [%]

count	9.000000
mean	7.440338
std	1.962858
min	4.014579
25%	6.197538
50%	7.541809
68.2% (1σ)	8.162591
75%	8.622461
max	10.557824

Recall that S2 mask (May 27, 2023) was 5% smaller than Pleiades mask (May 26, 2023) for the entire Gondrexange Lake.

S2 (June 8, 2023)

lake_id
- 2320162172 Neuf Etang
- 2320166602 Gros Etang
- 2320166622 **Réchicourt**
- 2320166632 Petit Etang

SWOT LakeSP (June 8, 2023)

SUMMARY AND OUTLOOK

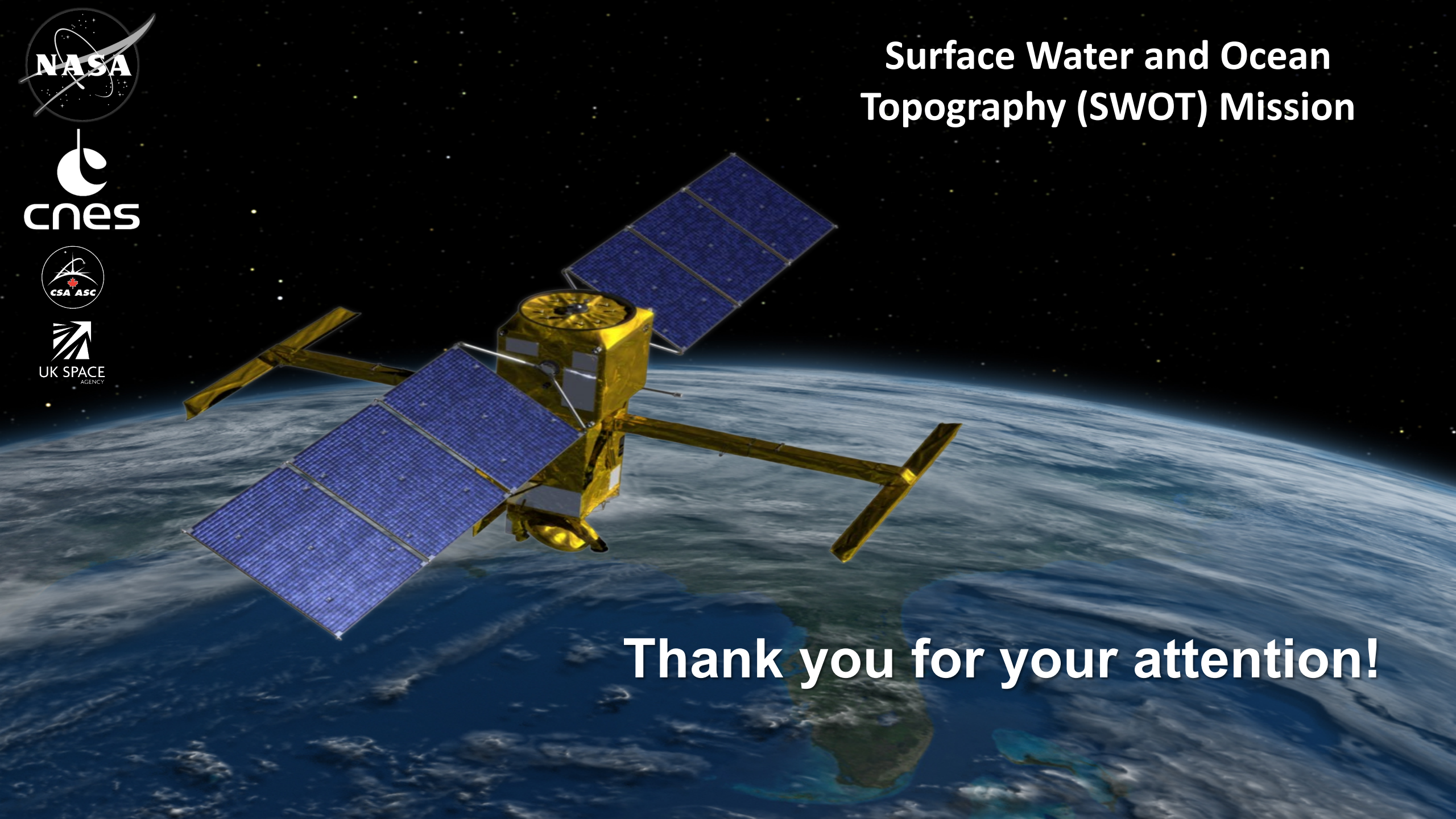
Lake area performance of LakeSP products assessed on a large number of lakes worldwide

- Lake area performance varies a lot
- Much better for lakes > 1 km^2 than for smaller lakes
- General over-estimation due to smearing
- Error figures depend strongly on the accuracy of the reference data.
- It is essential to use quality flags to filter out bad (and suspect) data.
- Main error causes identified
- Performance will become better through improvements in LakeSP and upstream algorithms, and in prior data (PLD, SWORD, water occurrence mask, bright land mask).

|relative area error| (1σ)

Filtering	S2+RCM+Pleiades		RCM+Pleiades	
	> (250 m)²	> 1 km²	> (250 m)²	> 1 km²
Baseline*	38%	17%	23%	9%
+ ~dark_frac**	39%	16%	22%	9%
+ ~area_truth***	31%	14%	17%	7%

*) include data with quality_f, ice_f, xovr_qual_q, partial_f and ice_f = 0
**) include data with dark_frac < 50%
***) include data with |area_truth – area_PLD|/area_PLD < 50%


NASA
cnes
CSA ASC
UK SPACE AGENCY
Surface Water and Ocean Topography (SWOT) Mission
Thank you for your attention!

BACK-UP

© CNES/DERAMECOURT Arnaud, 2007

PROPORTION OF RAISED FLAGS IN AREA VALIDATION LAKESP PRODUCTS

Flag	Percentage
quality_f > 0	19.7%
ice_f > 0	7.7%
xovr_qual_q > 0	3.0%
partial_f > 0	12,4%
dark_frac > 50%	7.9%

- Flags here considered independently
- However, there is some overlap
  - For example, quality_f >0 covers most xovr_qual_q > 0

MAIN LAKE AREA ERROR SOURCES

- Azimuth smearing
  - Lake extent systematically over-estimated
  - Larger impact on relative area error of small lakes
    - Can be improved through algorithm modifications (better handling of edge pixels, water fraction estimates…)
- Dark water
  - Area errors because of imperfect dark water flagging (estimation of extent or projection)
    - Can be improved through improved prior water occurrence masks , reference DEM and projection algorithm
- Bright land (humid soil, urban areas…)
  - Bright land detected as water adjacent to PLD lakes may cause important overestimation of lake area
    - Can be partially mitigated through active use of bright land flag
- Specular ringing
  - Specular ringing may seriously deteriorate lake polygon and degrade lake area and wse
    - Handling of specular ringing will be improved in future versions
- Assignment errors
  - Missing connected rivers in SWORD and missing nearby lakes in PLD may cause assignment and area errors
    - Improved versions of SWORD and PLD will reduce the assignments errors, likewise improved assignment algorithms

MAIN SCIENCE REQUIREMENTS FOR LAKES

- Global inventory of lakes, reservoirs and wetlands > 250x250 m^2 (Goal: >100x100 m^2)

- Water surface elevation (WSE) error
  - Requirement: < 10 cm (1σ) for lakes > 1 km^2
  - Goal: < 25 cm (1σ) for lakes > 250x250 m^2 and < 1 km^2
  - Threshold requirement: < 11 cm (1σ) for lakes > 1 km^2

- Relative surface area error
  - Requirement: < 15% (1σ) for lakes > 250x250 m^2
  - Goal: < 25% (1σ) for lakes > 100x100 m^2 and < 250x250 m^2
  - Threshold requirement: < 15% (1σ) for lakes > 1 km^2

(1σ) means that

68%

of the errors are smaller, and

32%

are bigger.