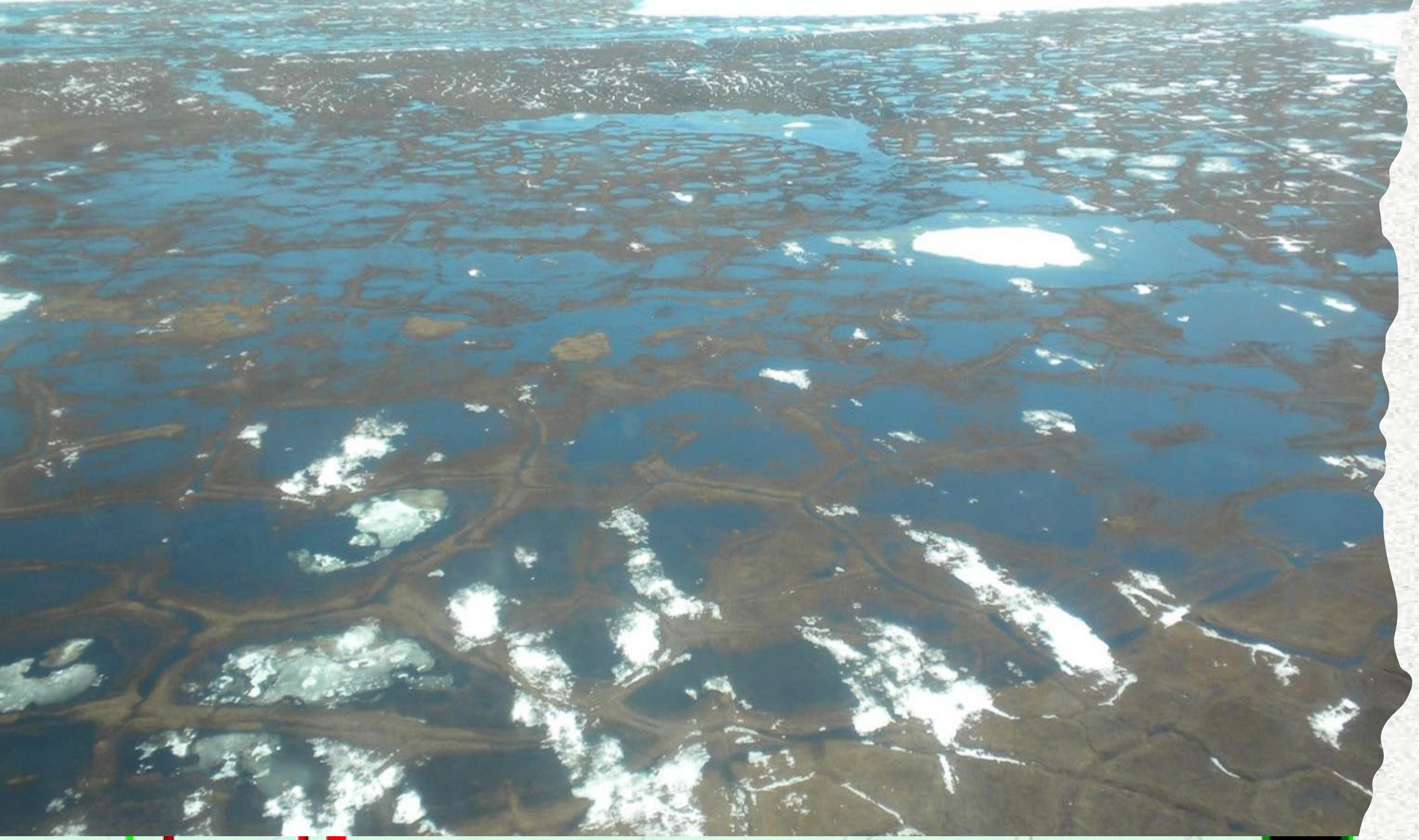
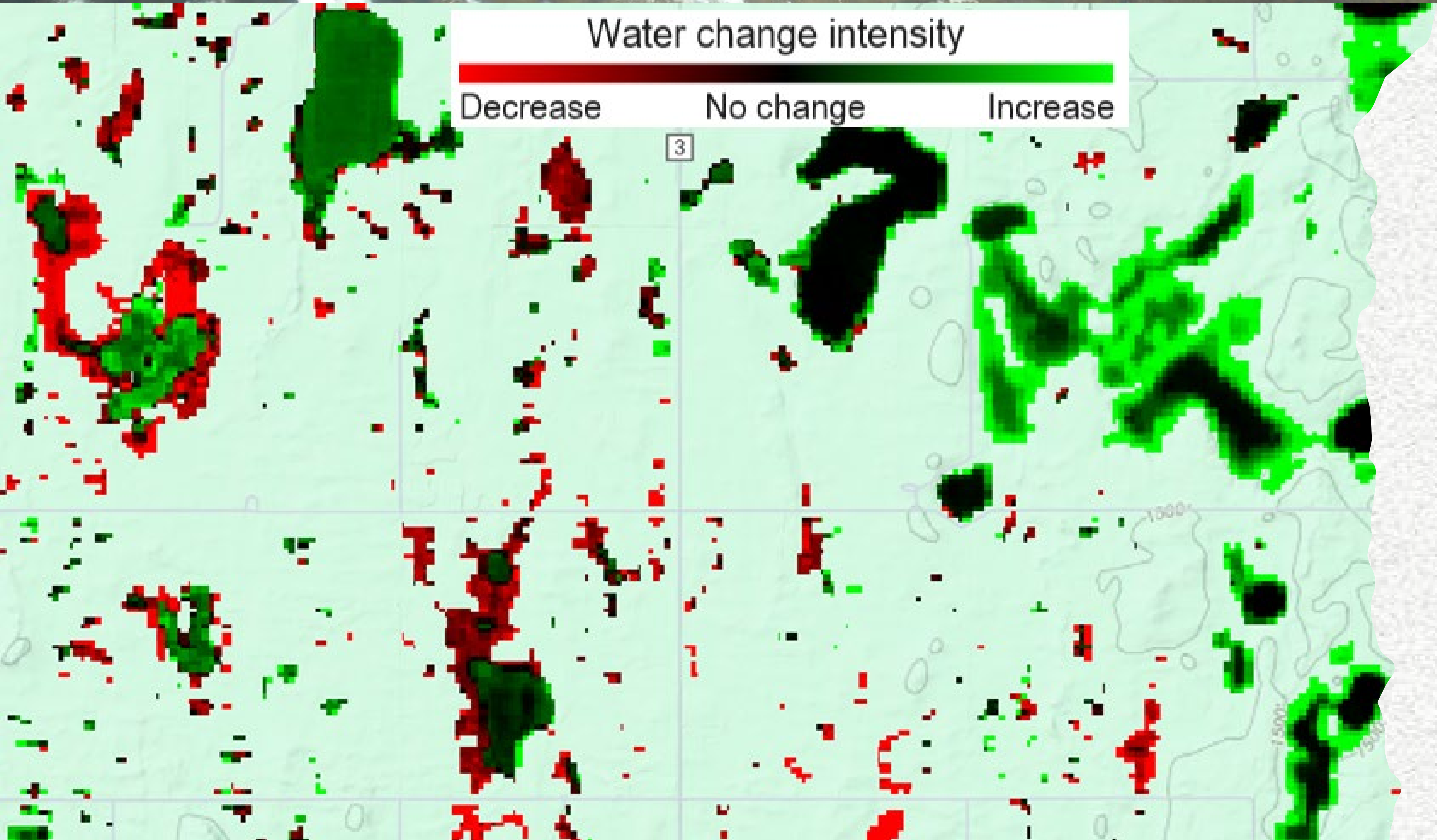
SURFACE WATER & OCEAN TOPOGRAPHY

Update on the SWOT Prior Lake Database (PLD)

- Review of the latest version (v1.05)
- Updates:
  - Plan for key operational deliveries
  - Scientific attributes
  - Harmonization with SWORD

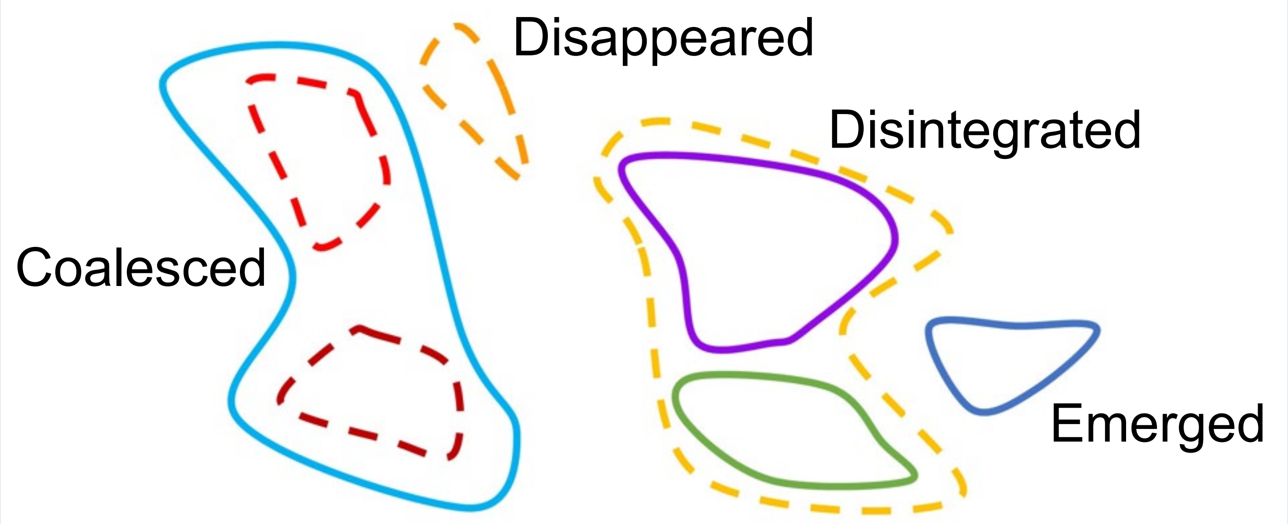
Jida Wang (UIUC) and Claire Pottier (CNES),
on behalf of the PLD team

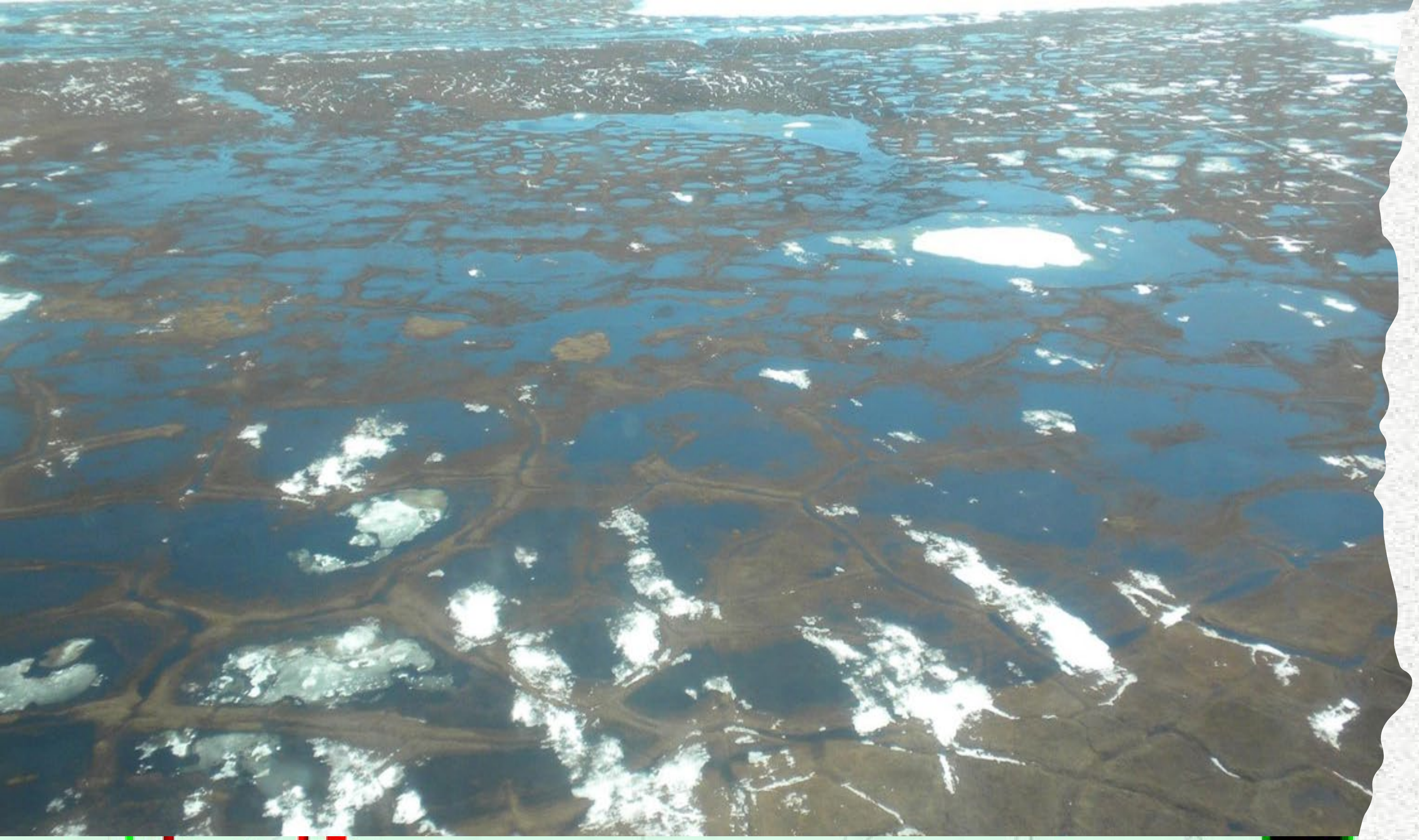


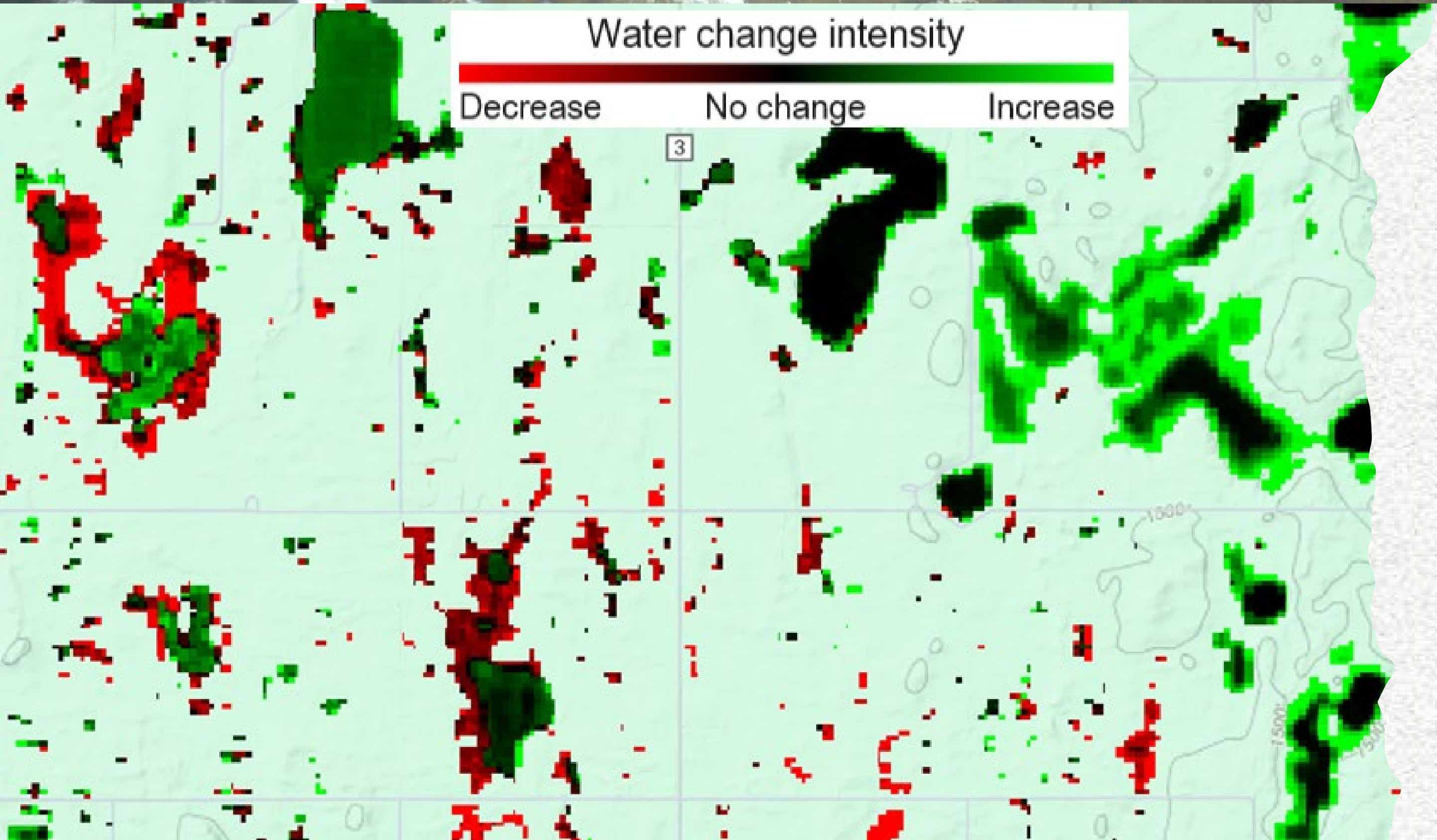


Purposes of the PLD

- Provide **benchmark** locations and extents for water bodies known as lakes (≥ 1 ha), i.e., prior lakes.
- Allow SWOT observations over time to be linked to the prior lakes so lake dynamics can be **consistently** computed.
- Identify and track the "**surprising**" water features that are observed by SWOT but cannot be linked to any prior lake.

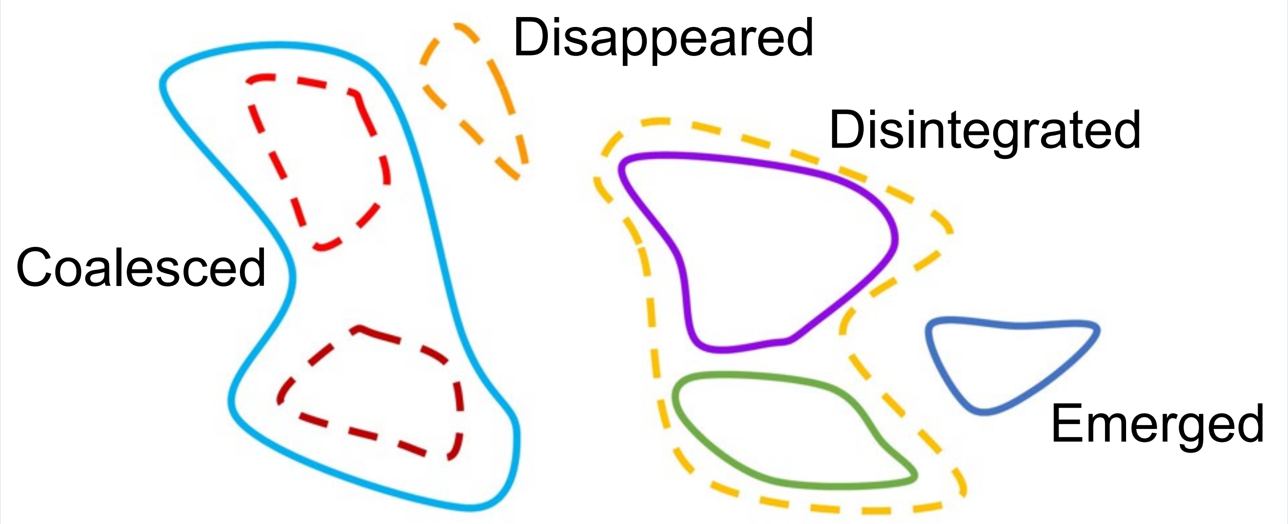




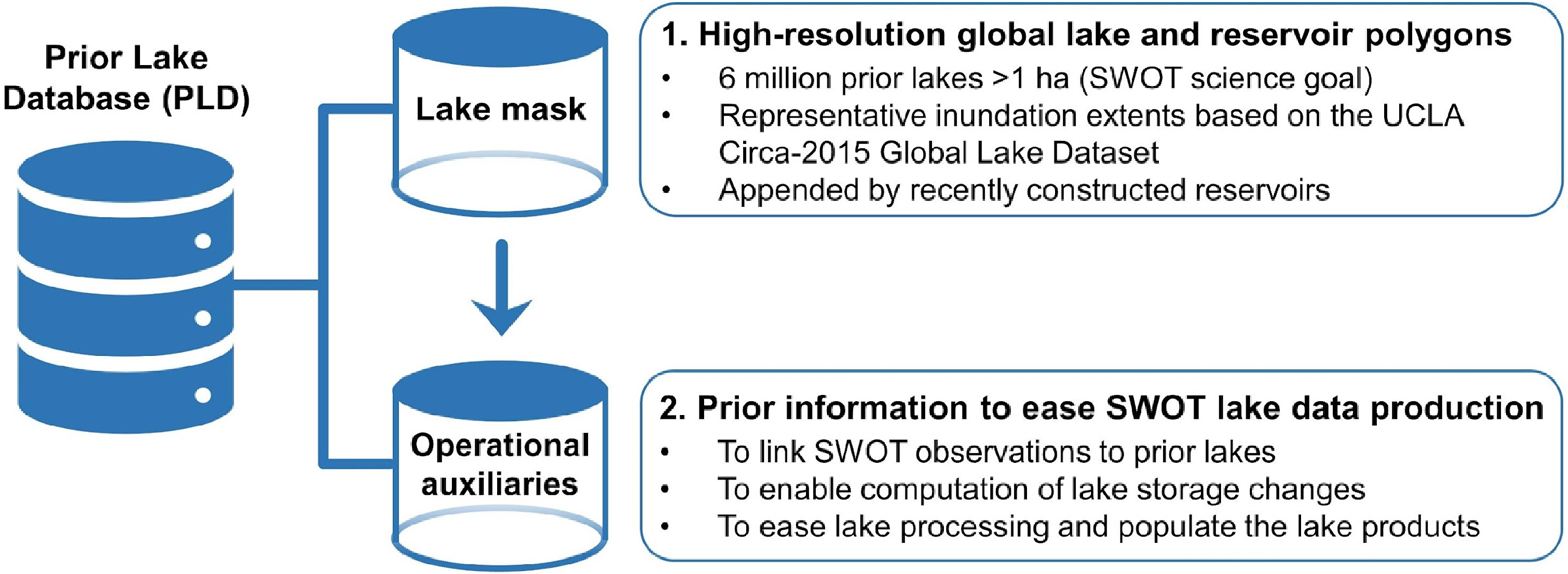


What PLD doesn't do…

- The PLD does not provide **dynamic** lake boundaries over time.
- At least for now, the PLD does not always offer the **maximum** lake extent.
- The PLD does not define **all** SWOT can observe.
- The PLD should not prevent users from processing PIXC based on their **own** lake definitions.



Conceptual structure of the PLD (v103)

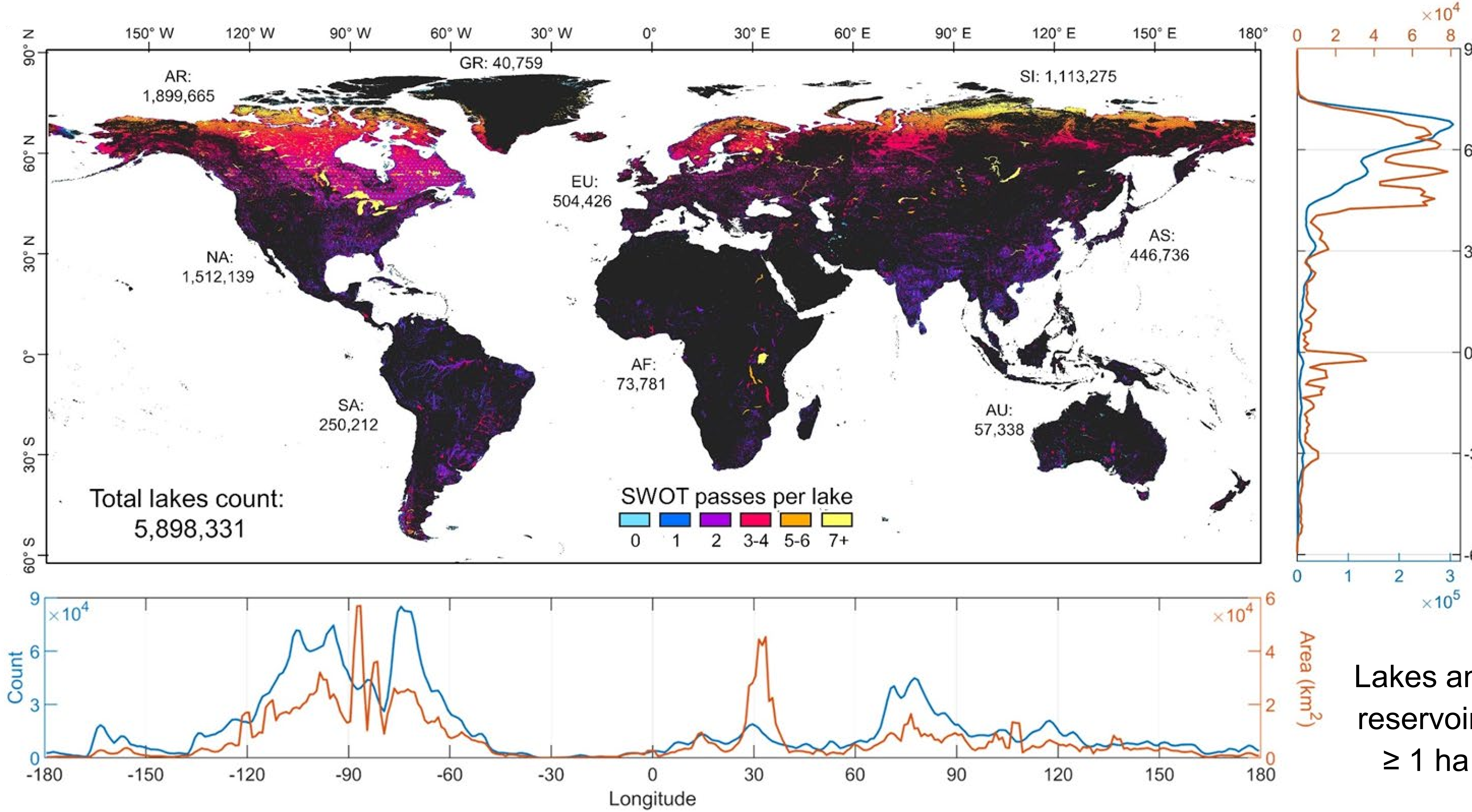


- **Accessible on the Hydroweb-next website (https://hydroweb.next.theia-land.fr)**
- *Wang, Pottier, et al.* WRR, in revision, doi:10.22541/au.170258987.72387777/v1

Data sources for harmonizing the PLD global lake mask

Source	Component	Contribution
UCLA Circa-2015 Global Lake Dataset (*Sheng et al., 2016.* doi:10.1016/j.rse.2015.12.041)	~6 million lakes representing **intermediate inundation extents** during 2013-2015 mapped from Landsat-8 images	Provides **the main source** for the prior lake mask
GeoDAR (Georeferenced global Dams And Reservoirs) dataset v1.1 (*Wang et al., 2022.* doi:10.5194/essd-14-1869-2022)	22,560 large reservoirs based on the International Commission on Large Dams (ICOLD)	Supplements the UCLA Circa-2015 Global Lake Dataset with **additional reservoirs**
GREI-p2k reservoir dataset (*Fan et al., 2024.* doi:10.1016/j.scib.2024.04.043)	~6,700 new reservoirs (>0.5 km^2) that emerged after 2000	
Other miscellaneous regional reservoirs (*Collected by coauthors*)	~7,000 reservoirs in parts of South America and Africa	

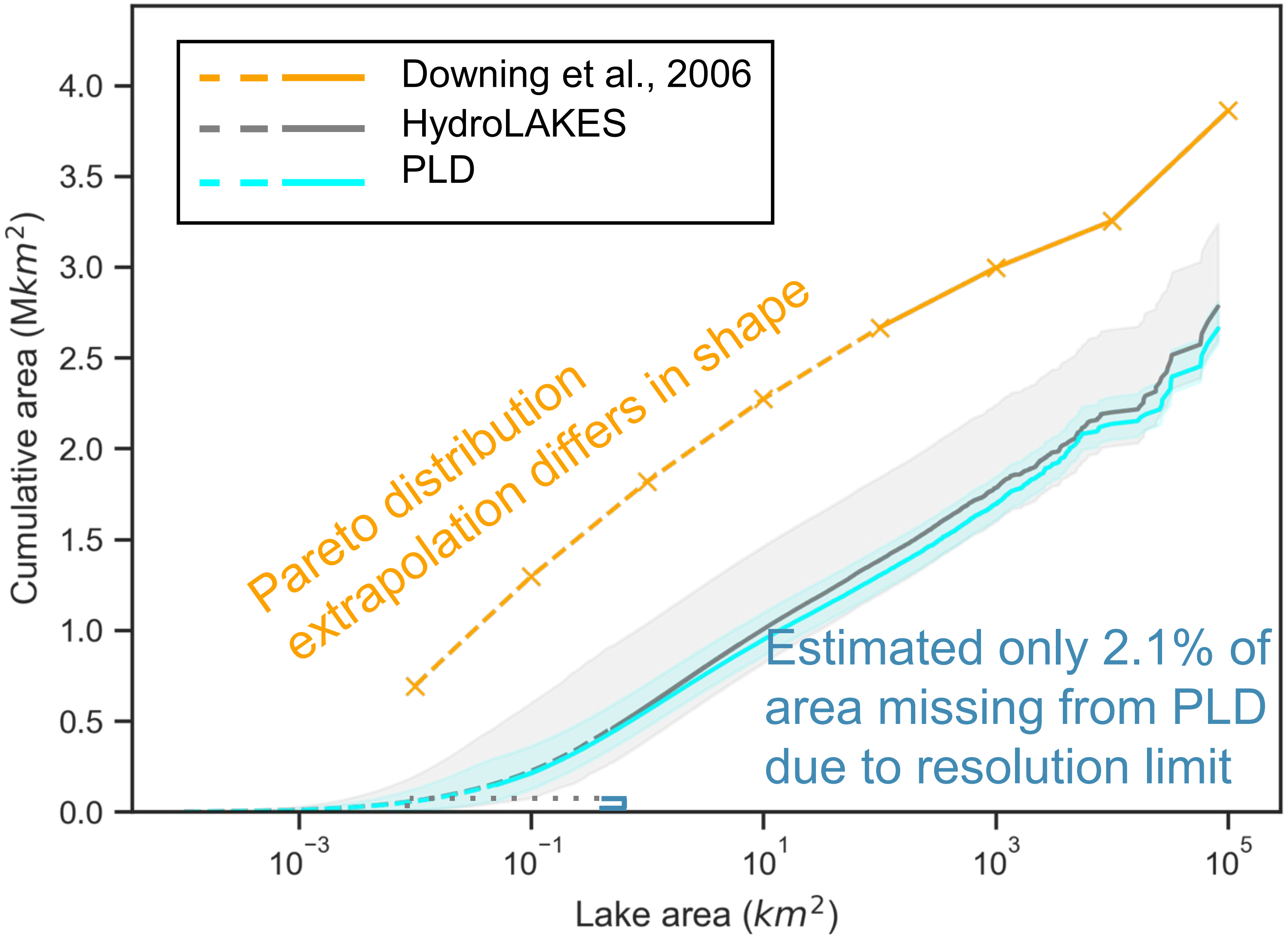
SWOT Prior Lake Mask



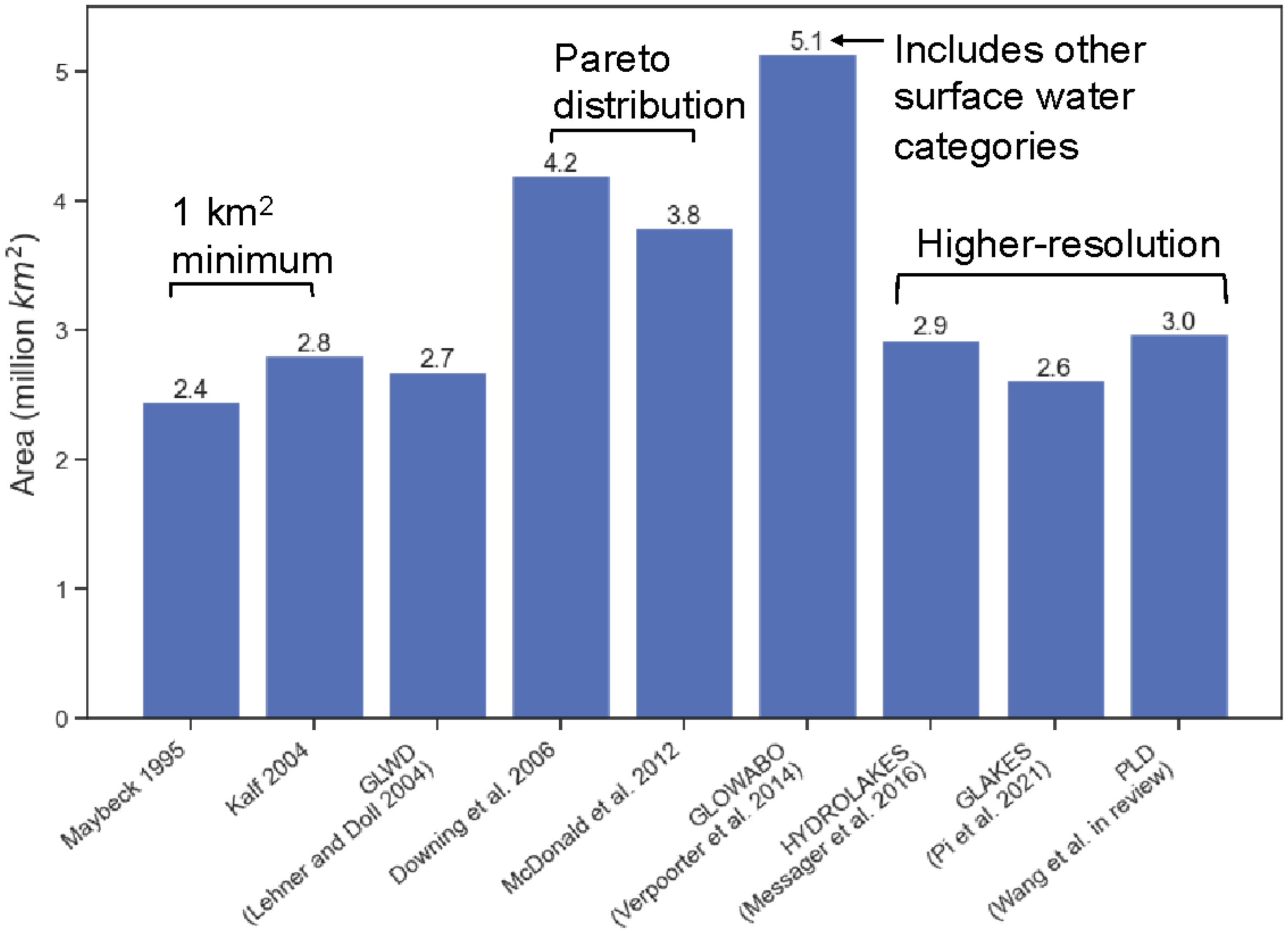
Lakes and reservoirs ≥ 1 ha

The SWOT PLD provides the most comprehensive inventory of global lake area to date

Kyzivat & Smith 2023, *GRL*



Global Lake Area Estimates



Total lake area is much smaller than previous extrapolated estimates, with implications for greenhouse gas emissions

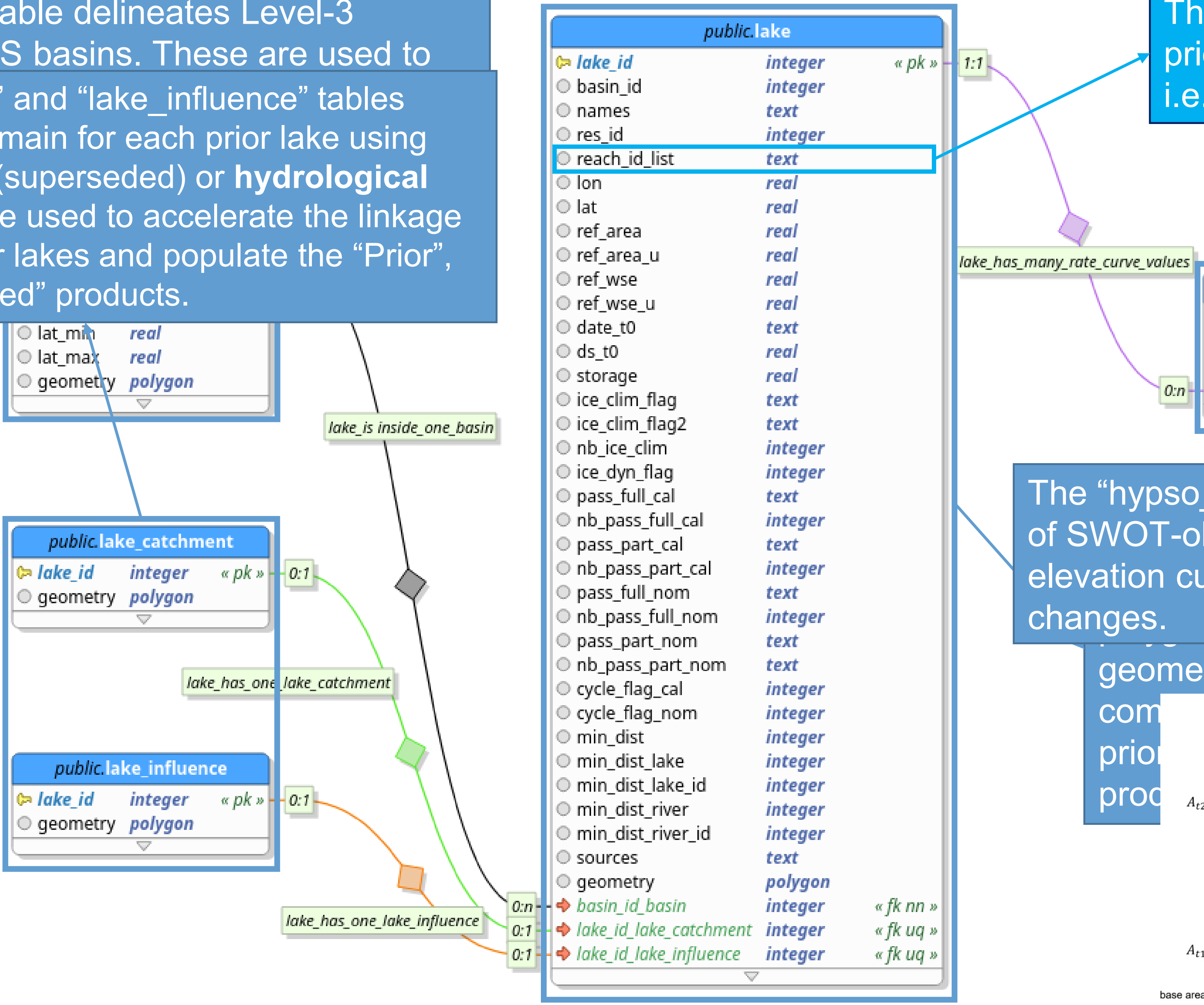
Ethan Kyzivat

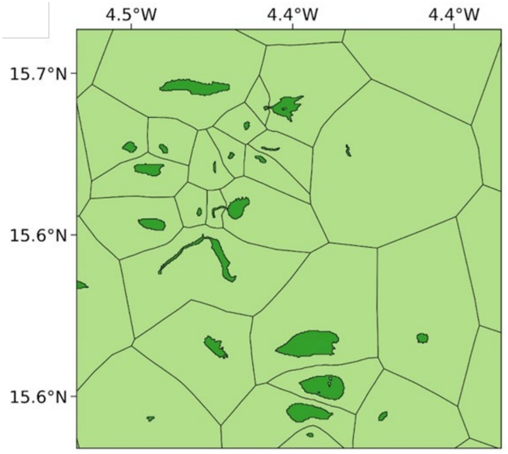
Structure of operational auxiliaries

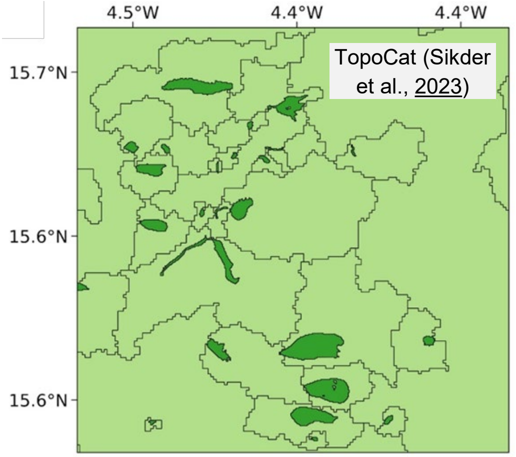
The “basin” table delineates Level-3 HydroBASINS basins. These are used to

The “lake_catchment” and “lake_influence” tables delineate a spatial domain for each prior lake using **Thiessen polygons** (superseded) or **hydrological catchments.** They are used to accelerate the linkage of PIXC pixels to prior lakes and populate the “Prior”, “Obs”, and “Unassigned” products.

The “reach_id_list” attribute identifies prior lakes that intersected SWORD, i.e., so called “river-connected lakes”

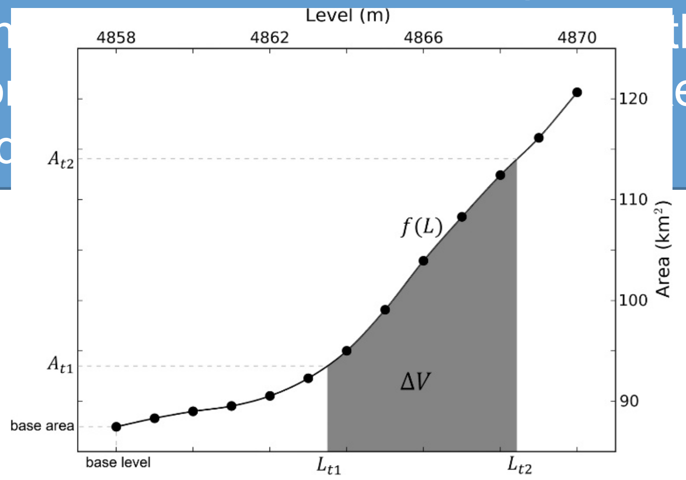






The “hypso_curve” table stores the information of SWOT-observed lake hypsometry (i.e., area-elevation curve) for computing water storage changes.

geometries. It also contains parameters to com… ther prior… e proc…

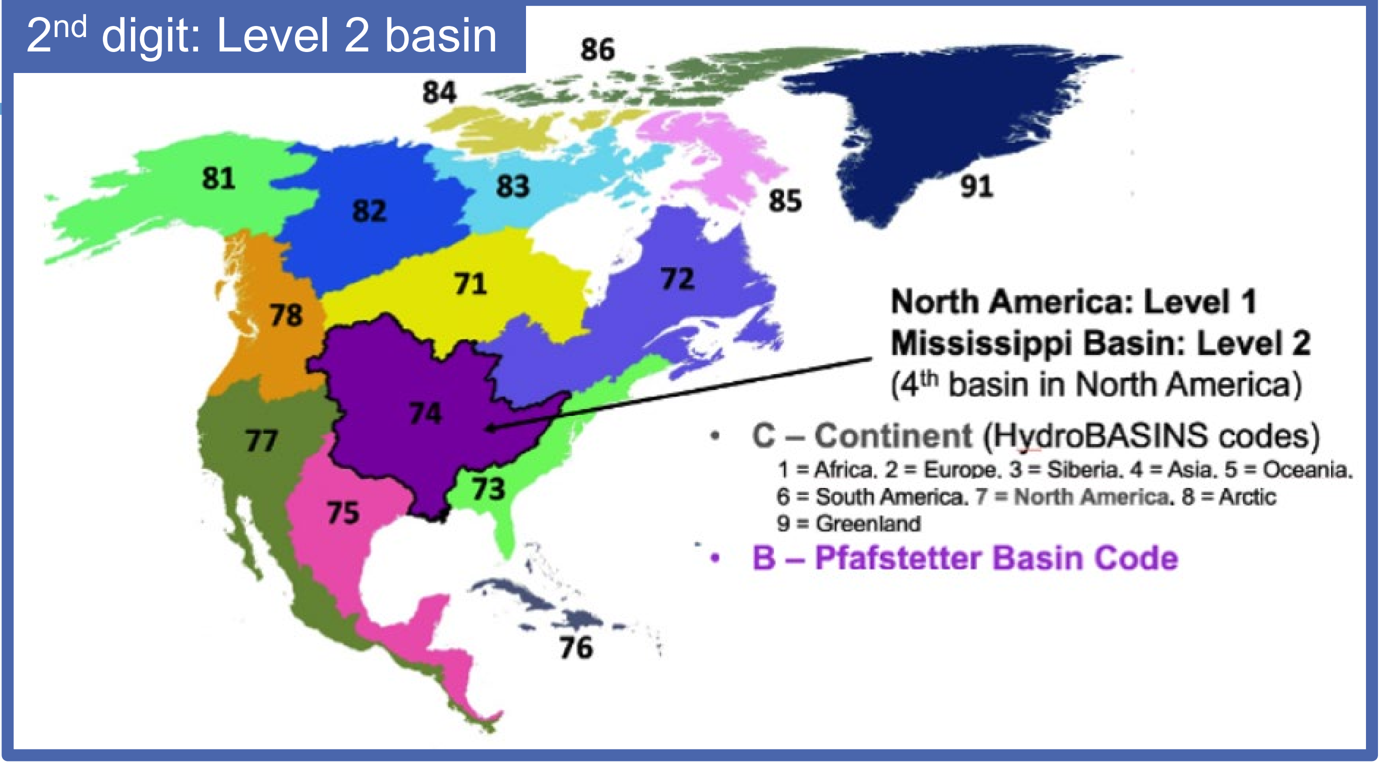


Lake identifier: Pfafstetter basin coding system

public.lake		
lake_id	integer	« pk »
basin_id	integer	
names	text	
res_id	integer	
reach_id_list	text	
lon	real	
lat	real	
ref_area	real	
ref_area_u	real	
ref_wse	real	
ref_wse_u	real	
date_t0	text	
ds_t0	real	
storage	real	
ice_clim_flag	text	
ice_clim_flag2	text	
nb_ice_clim	integer	
ice_dyn_flag	integer	
pass_full_cal	text	
nb_pass_full_cal	integer	
pass_part_cal	text	
nb_pass_part_cal	integer	
pass_full_nom	text	
nb_pass_full_nom	integer	
pass_part_nom	text	
nb_pass_part_nom	text	
cycle_flag_cal	integer	
cycle_flag_nom	integer	
min_dist	integer	
min_dist_lake	integer	
min_dist_lake_id	integer	
min_dist_river	integer	
min_dist_river_id	integer	
sources	text	
geometry	polygon	
basin_id_basin	integer	« fk nn »
lake_id_lake_catchment	integer	« fk uq »
lake_id_lake_influence	integer	« fk uq »

1st digit: Level 1 continent

2nd digit: Level 2 basin



Lake identifier: Pfafstetter basin coding system

public.lake		
lake_id	*integer*	« pk »
basin_id	*integer*	
names	*text*	
res_id	*integer*	
reach_id_list	*text*	
lon	*real*	
lat	*real*	
ref_area	*real*	
ref_area_u	*real*	
ref_wse	*real*	
ref_wse_u	*real*	
date_t0	*text*	
ds_t0	*real*	
storage	*real*	
ice_clim_flag	*text*	
ice_clim_flag2	*text*	
nb_ice_clim	*integer*	
ice_dyn_flag	*integer*	
pass_full_cal	*text*	
nb_pass_full_cal	*integer*	
pass_part_cal	*text*	
nb_pass_part_cal	*integer*	
pass_full_nom	*text*	
nb_pass_full_nom	*integer*	
pass_part_nom	*text*	
nb_pass_part_nom	*text*	
cycle_flag_cal	*integer*	
cycle_flag_nom	*integer*	
min_dist	*integer*	
min_dist_lake	*integer*	
min_dist_lake_id	*integer*	
min_dist_river	*integer*	
min_dist_river_id	*integer*	
sources	*text*	
geometry	*polygon*	
basin_id_basin	*integer*	« fk nn »
lake_id_lake_catchment	*integer*	« fk uq »
lake_id_lake_influence	*integer*	« fk uq »

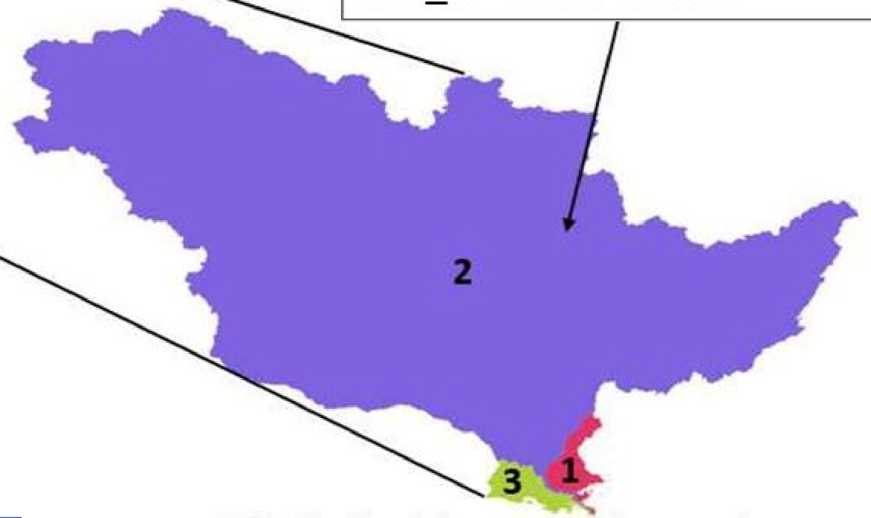
North America (continent ID = 7)
Sub-basins (level 2)

2nd digit: Level 2 basin

CBBNNNNNNT (*lake_id*)

- **C – Continent**
- **B – Basin code**
- **N – Lake index in the basin**
- **T – Type**

Example for a lake in this basin:
lake_id = 7420469602

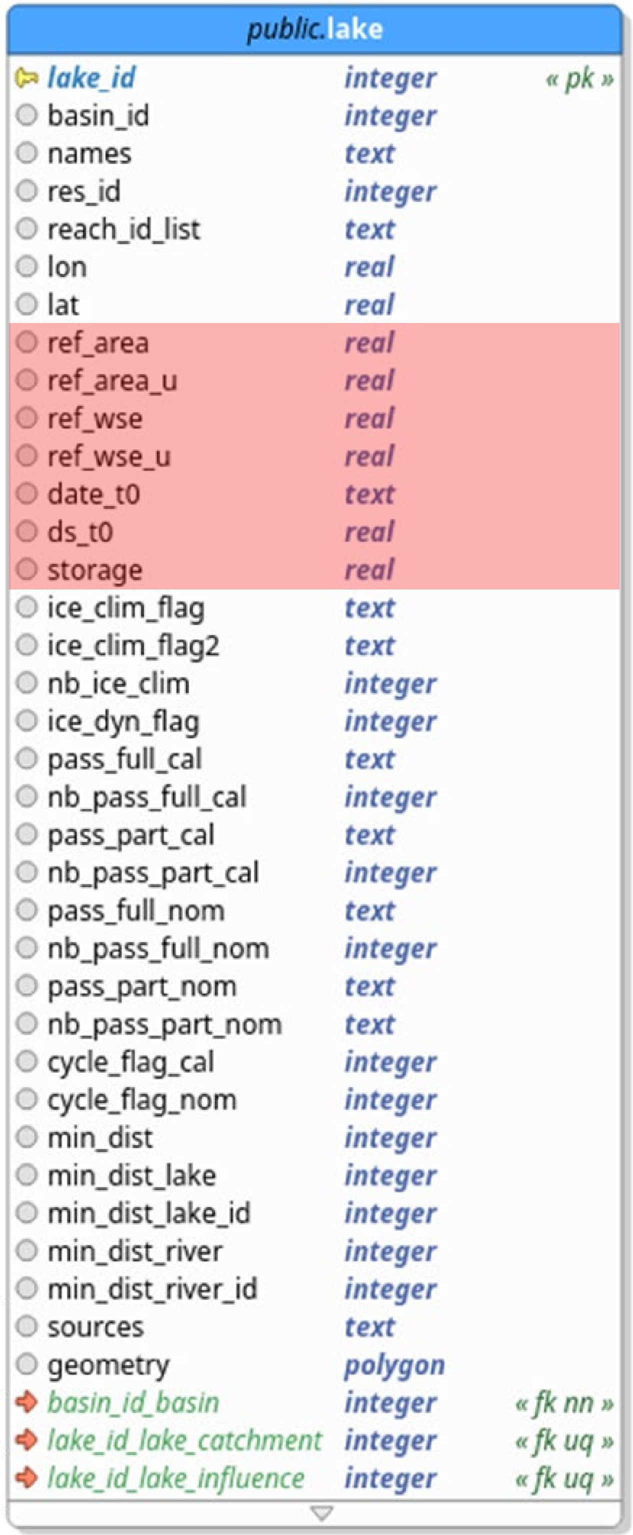


Mississippi (basin 1st digit = 4)
Sub-basins (level 3)

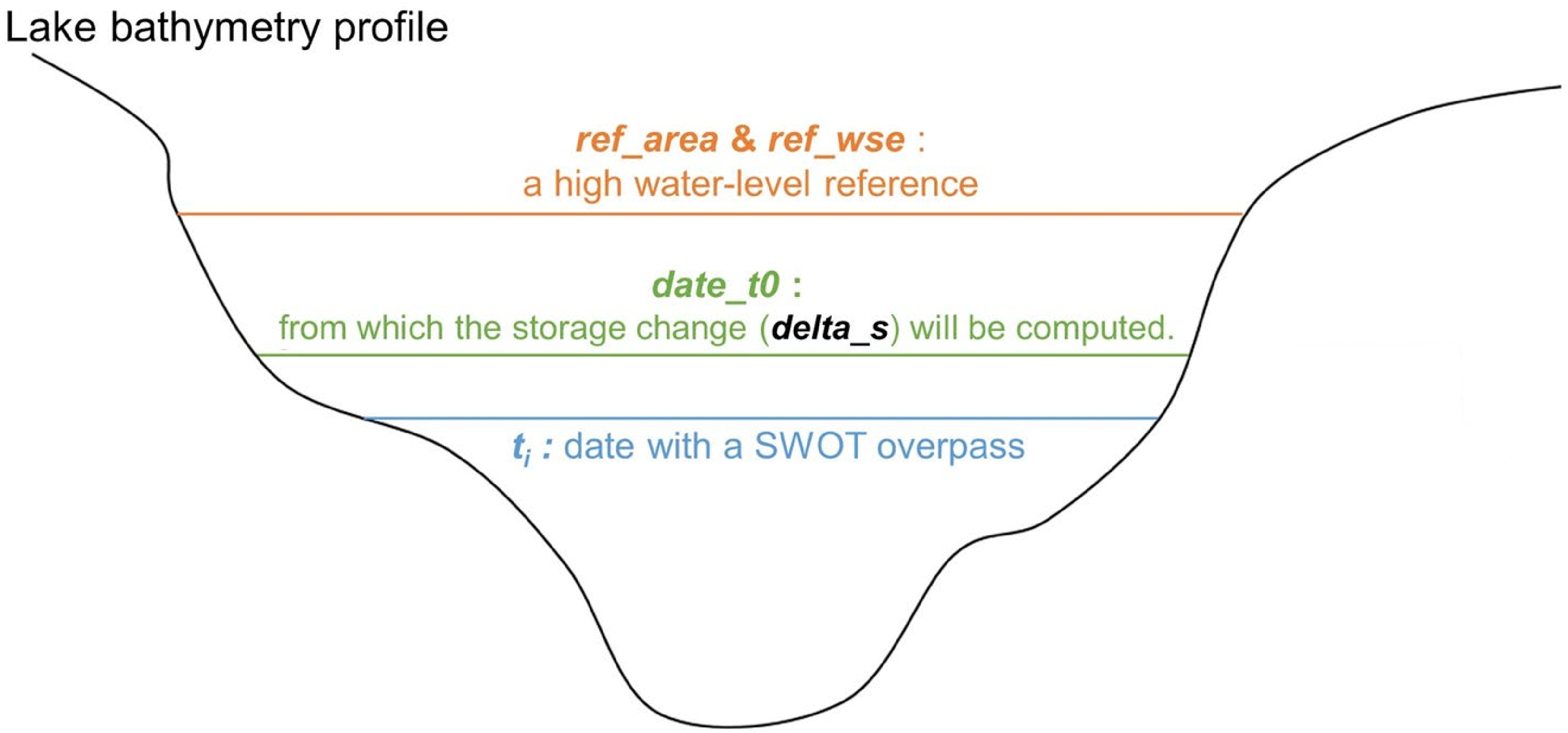
3rd digit: Level 3 basin

Last digit in *lake id*	Water body type
1	River (x)
2	Disconnected lake
3	River-connected lake
4	Dam (x)
5	No topology (x)

Prior parameters for computing storage changes

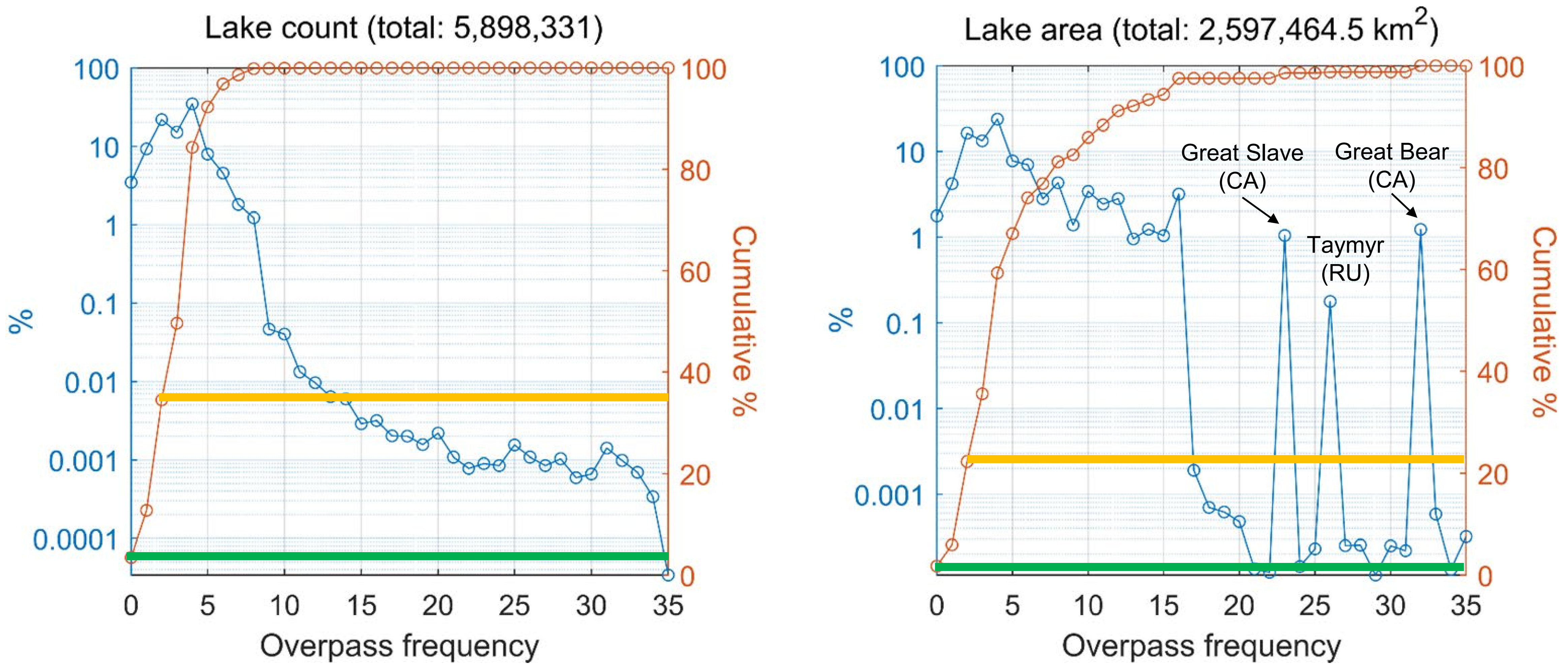


public.lake		
lake_id	integer	« pk »
basin_id	integer	
names	text	
res_id	integer	
reach_id_list	text	
lon	real	
lat	real	
ref_area	real	
ref_area_u	real	
ref_wse	real	
ref_wse_u	real	
date_t0	text	
ds_t0	real	
storage	real	
ice_clim_flag	text	
ice_clim_flag2	text	
nb_ice_clim	integer	
ice_dyn_flag	integer	
pass_full_cal	text	
nb_pass_full_cal	integer	
pass_part_cal	text	
nb_pass_part_cal	integer	
pass_full_nom	text	
nb_pass_full_nom	integer	
pass_part_nom	text	
nb_pass_part_nom	text	
cycle_flag_cal	integer	
cycle_flag_nom	integer	
min_dist	integer	
min_dist_lake	integer	
min_dist_lake_id	integer	
min_dist_river	integer	
min_dist_river_id	integer	
sources	text	
geometry	polygon	
basin_id_basin	integer	« fk nn »
lake_id_lake_catchment	integer	« fk uq »
lake_id_lake_influence	integer	« fk uq »



The calculated storage change (***delta_s***) in the L2 lake product is relative to the condition on the date of **the first valid SWOT measurement (*date_t0*).**

Lake overpass frequency within each SWOT nominal (21-day) orbit cycle



"SWOT shall collect data over a minimum of 90% of all ocean and land area covered by the orbit inclination for 90% of the operation time."
(SWOT Science Requirements Document, 2018).

- 96.5% of the global lakes, covering 98.2% of the total lake area, are observed at least **once per orbit**.
- More than 65% of the global lakes, covering nearly 80% of the total lake area, are observed at least **weekly** on average (3 times per cycle).
- 3.5% of the global lakes, or 1.8% of the total lake area, may never be observed due to nadir gaps and orbit intertrack gaps.

Plan for next deliveries

<u>End of August (to be used in forward processing in October 2024)</u>

- Update with missing or inaccurate lakes, identified by ST and us
  - Do not hesitate to contact us to ask for modifications!

<u>End of November (to be used in forward processing in January 2025)</u>

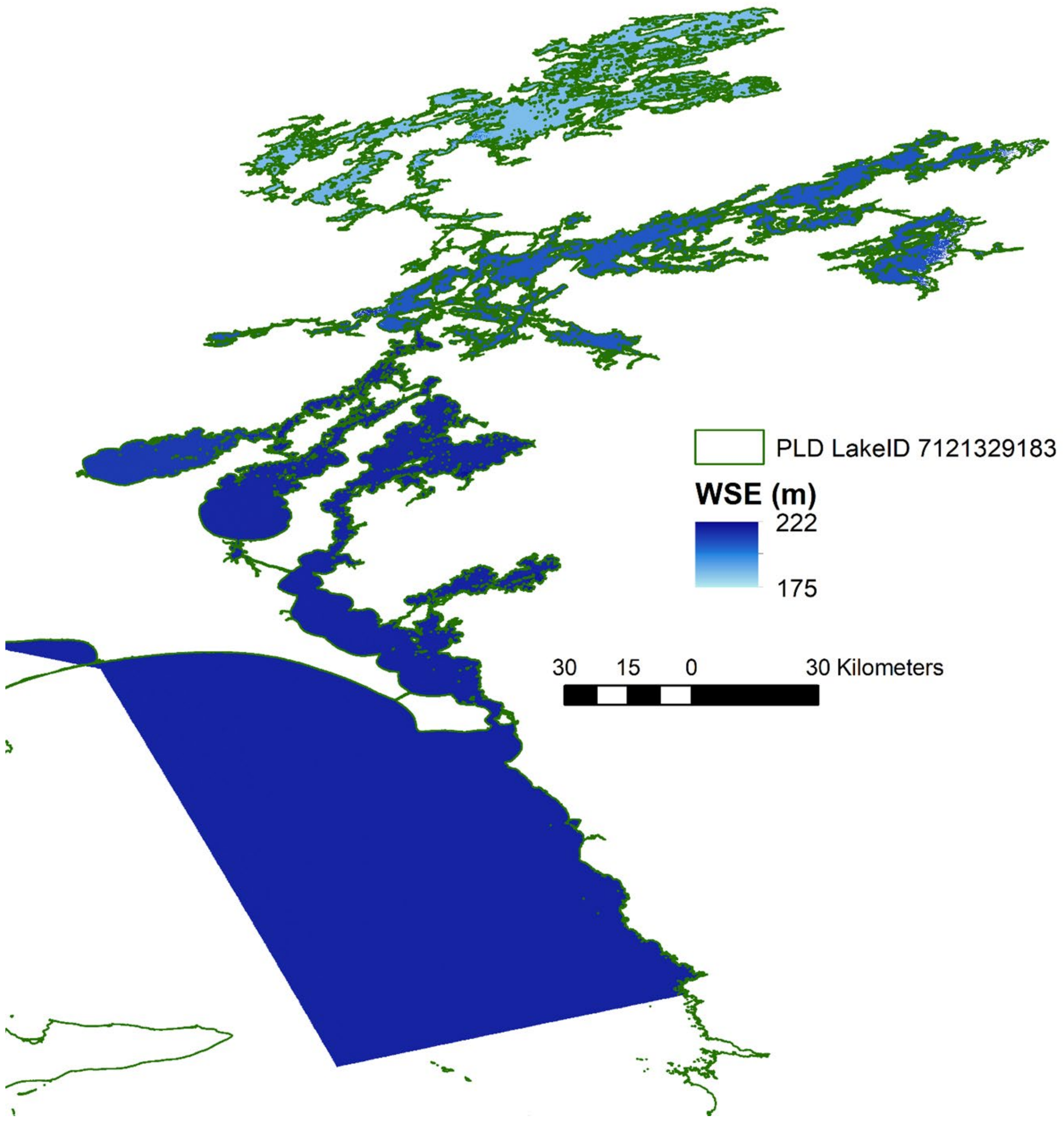
- Analysis of storage change values computed for the 26 lakes that currently have bathymetry
- Addition of other external bathymetries
- Extension to some other lakes with *p_ref_height* computed from SWOT

<u>Beyond (what is in the pipe):</u>

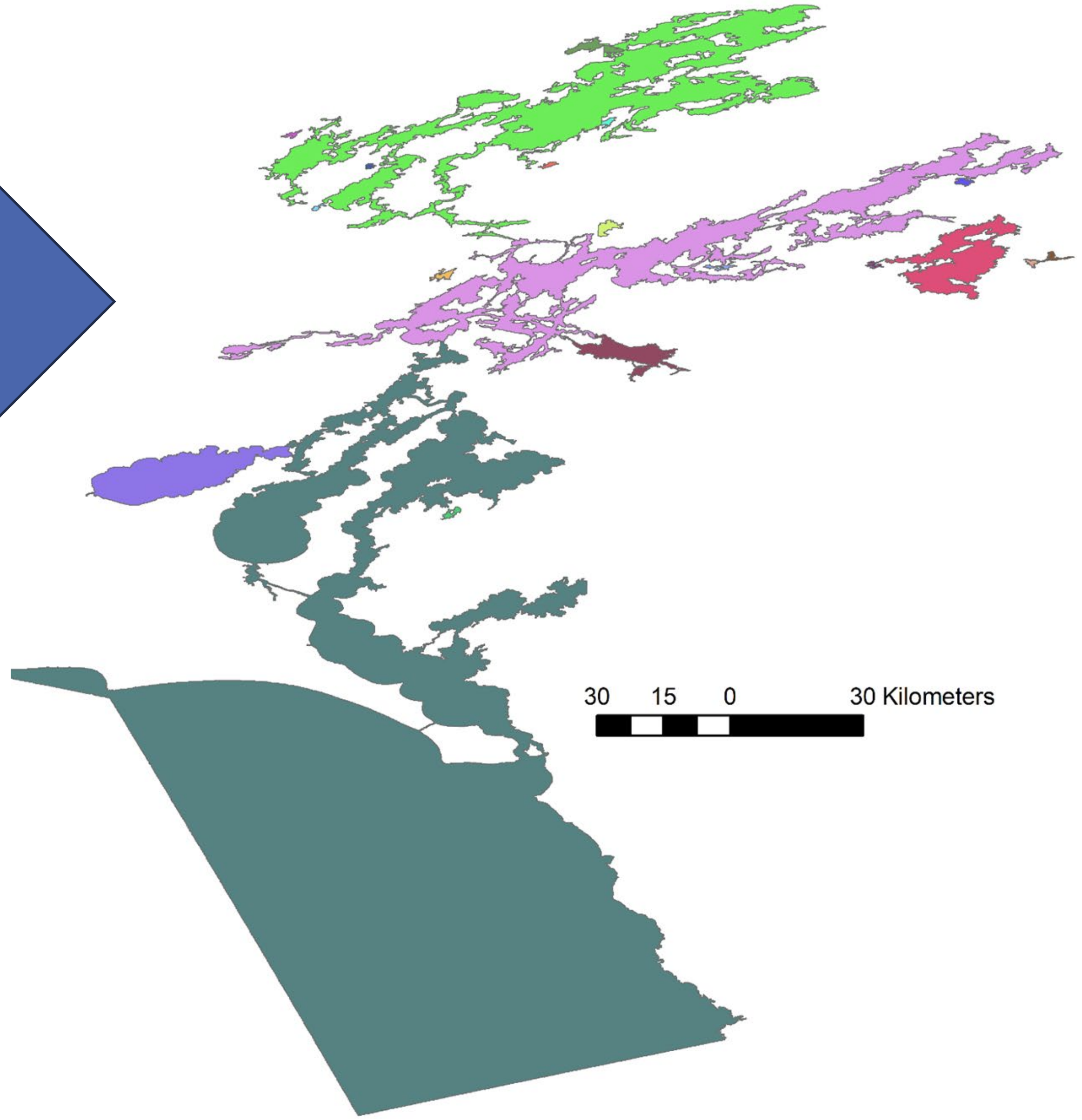
- Initial version of the *hypso_curve* table for some lakes (table will be periodically updated)
- Update geometries
- To add new prior lakes absent from the previous PLD versions, using the _Unassigned product.

  => need for the LakeSP product to be improved before

Update on geometries



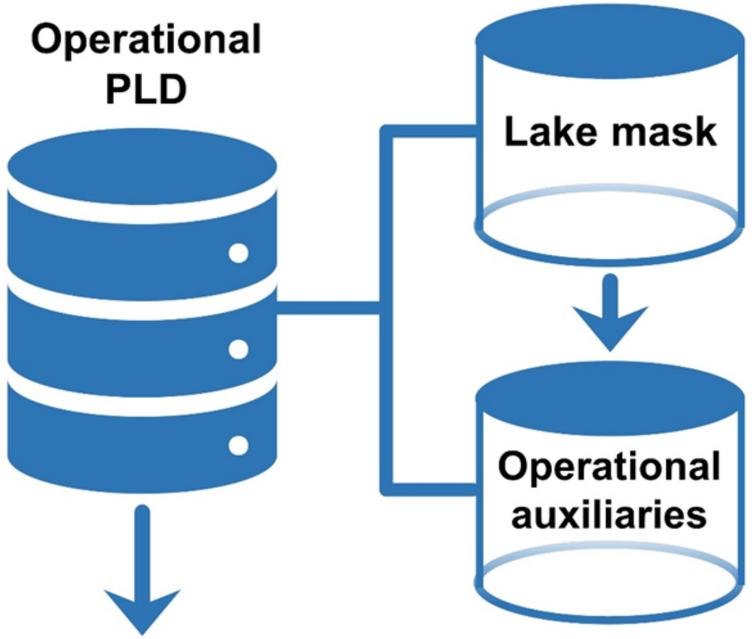
Segmentation based on SWOT WSE



Wnnipeg Lake, Canada

Melanie Trudel, U Sherbrooke

In progress: from **operational** to **scientific**

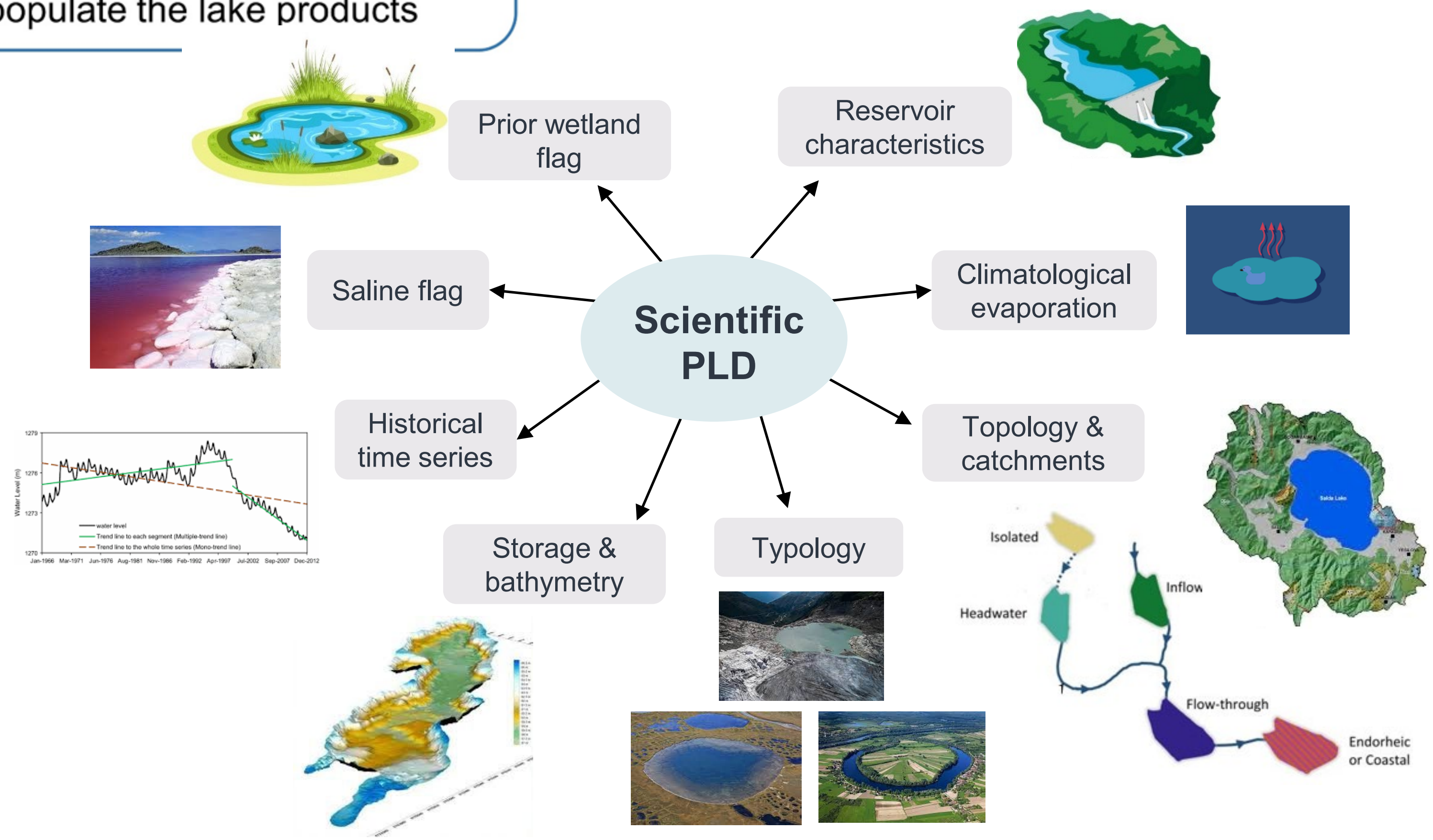


1. High-resolution global lake and reservoir polygons
- 6 million prior lakes >1 ha (SWOT science goal)
- Representative inundation extents based on the UCLA Circa-2015 Global Lake Dataset
- Appended by recently constructed reservoirs

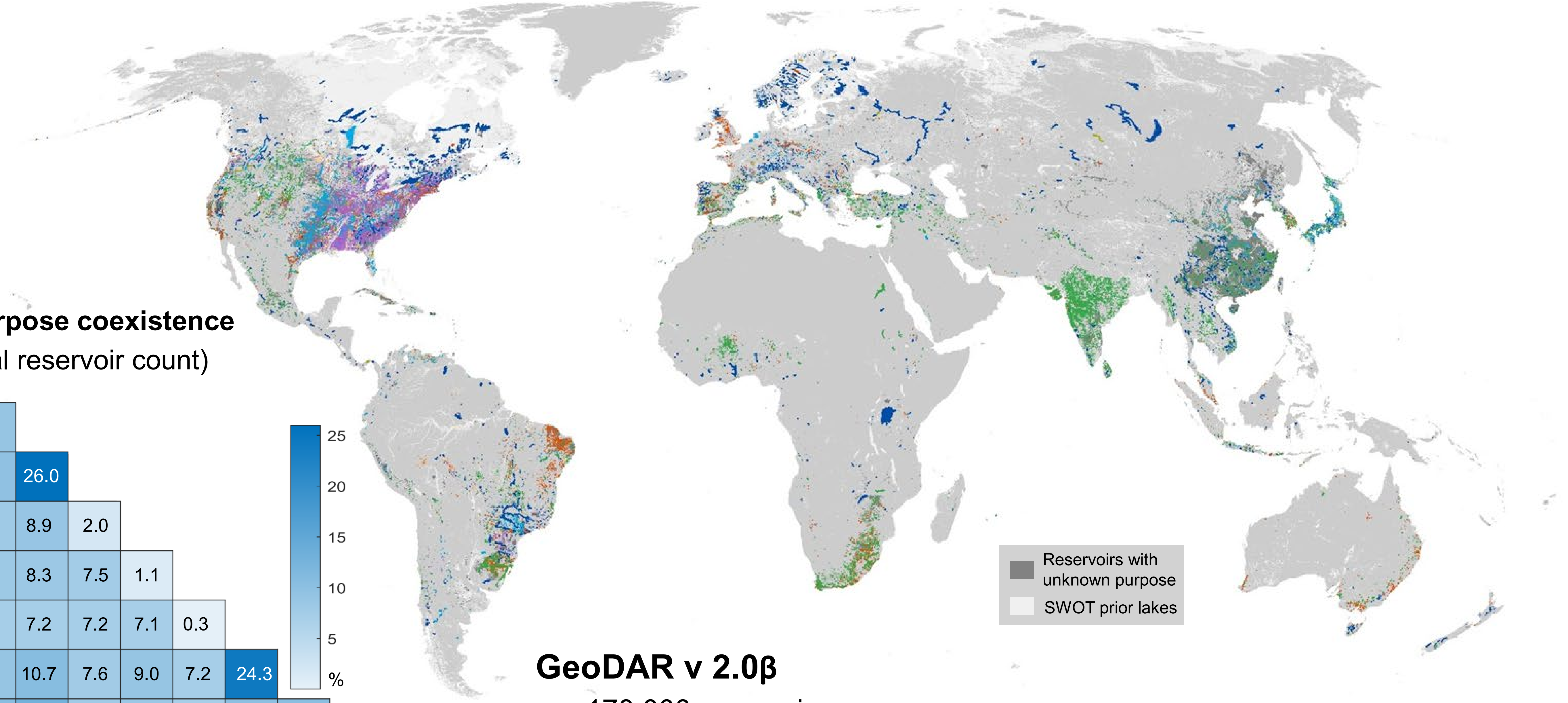
2. Prior information to ease SWOT lake data production
- To link SWOT observations to prior lakes
- To enable computation of lake storage changes
- To ease lake processing and populate the lake products

3. Scientific attributes and geometries to facilitate SWOT product applications
- Reservoir flag and characteristics
- Saline lake flag and salinity
- Prior wetland flag
- Lake storage and bathymetry
- Climatological evaporation
- Typology
- Lake topology and catchments
- …

In progress, expected around late 2024 …



Reservoir flags and characteristics



Matrix of purpose coexistence
(% of global reservoir count)

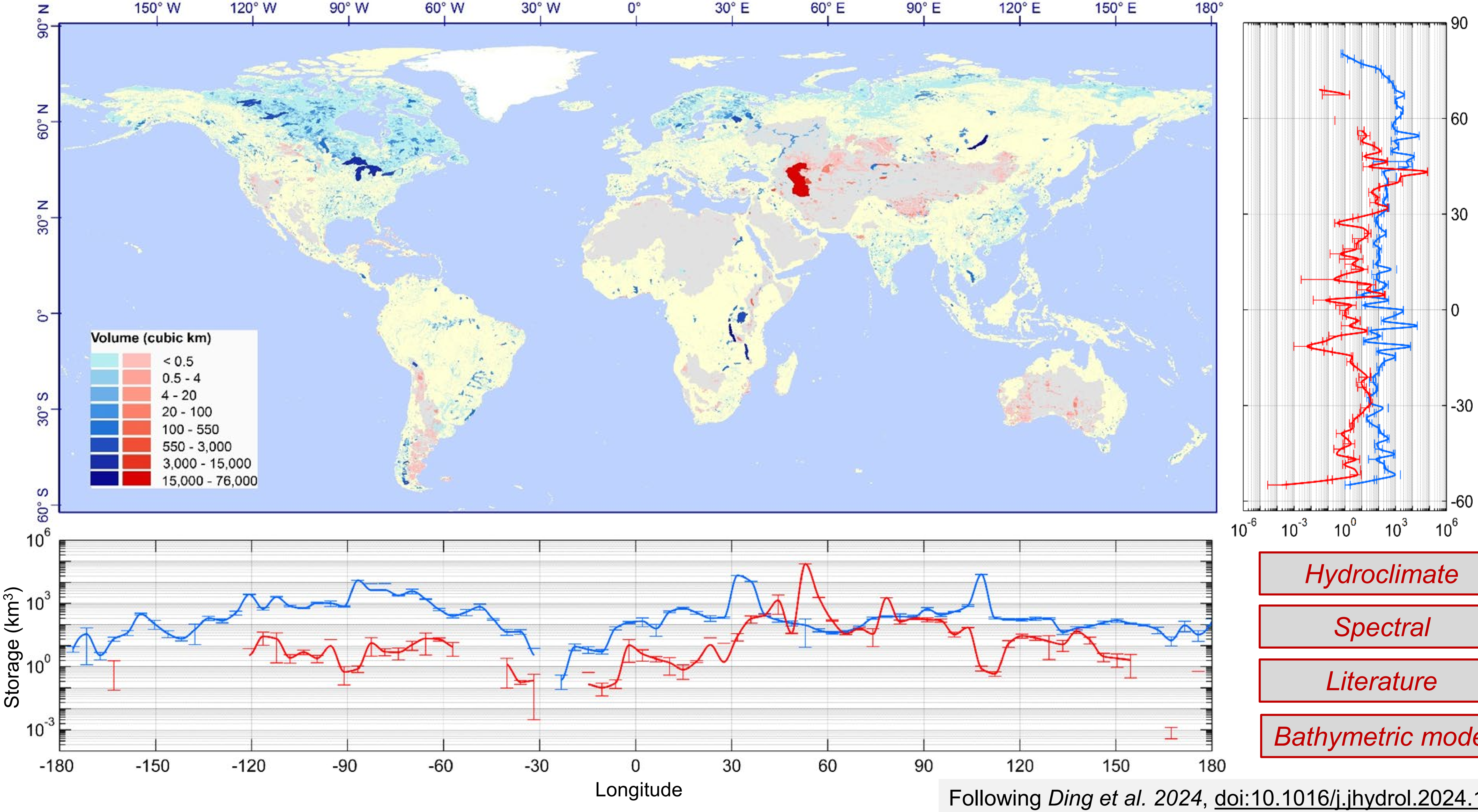
	Flood control	Irrigation/water supply	Hydropower	Fisheries	Navigation	Recreation	Others
Flood control	9.0						
Irrigation/water supply	9.7	26.0					
Hydropower	8.0	8.9	2.0				
Fisheries	8.2	8.3	7.5	1.1			
Navigation	7.2	7.2	7.2	7.1	0.3		
Recreation	9.0	10.7	7.6	9.0	7.2	24.3	
Others	10.5	13.0	7.9	8.5	7.2	10.1	10.3

GeoDAR v 2.0β

- ~170,000 reservoirs
- Harmonized from multiple dam registers and inventories
- Open access to a wide suite of attributes
- Integrated to the PLD

Upgraded from GeoDAR v1.1 (*Wang et al., 2022*, doi:10.5194/essd-14-1869-2022)

Freshwater and saline lakes (with estimated mean depth and volume)



Hydroclimate

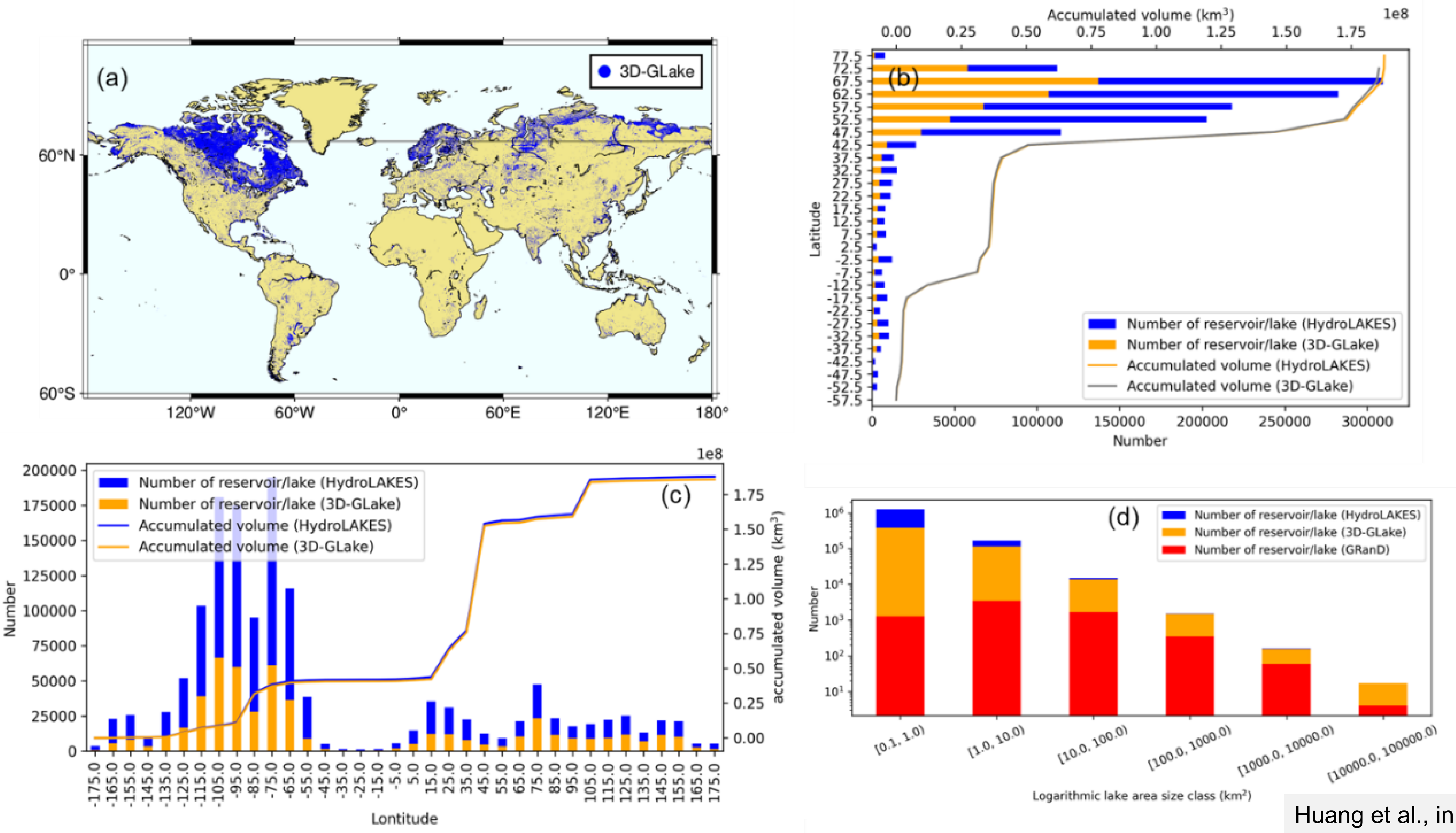
Spectral

Literature

Bathymetric model

Following *Ding et al. 2024*, doi:10.1016/j.jhydrol.2024.130704

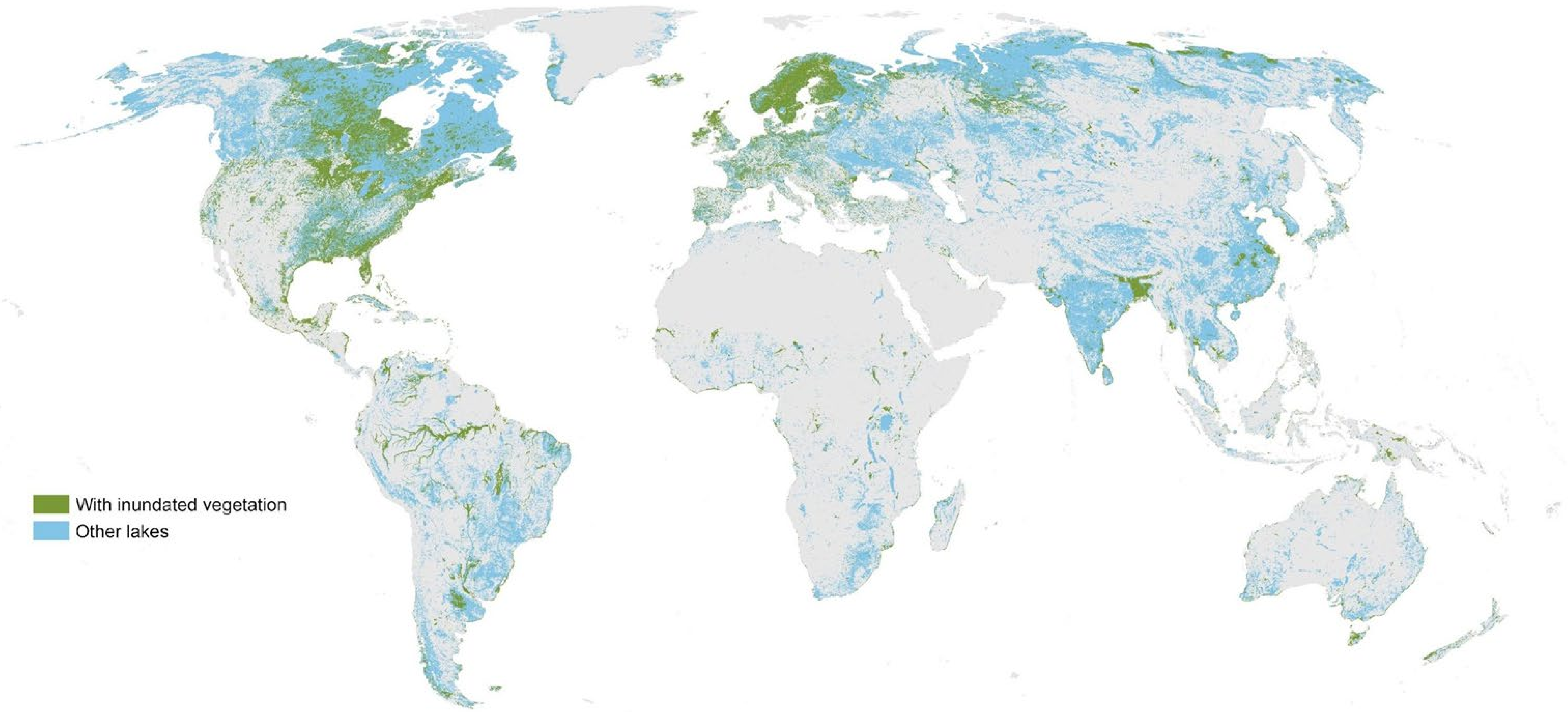
A global dataset (3D-Glakes) contains A-E relationships and 3D bathymetry for 0.5 million locations (derived from Landsat and ICESat-2)



Huang et al., in preparation

Huilin Gao, Texas A&M

Wetland flag



Fernando Jaramillo, Stockholm University

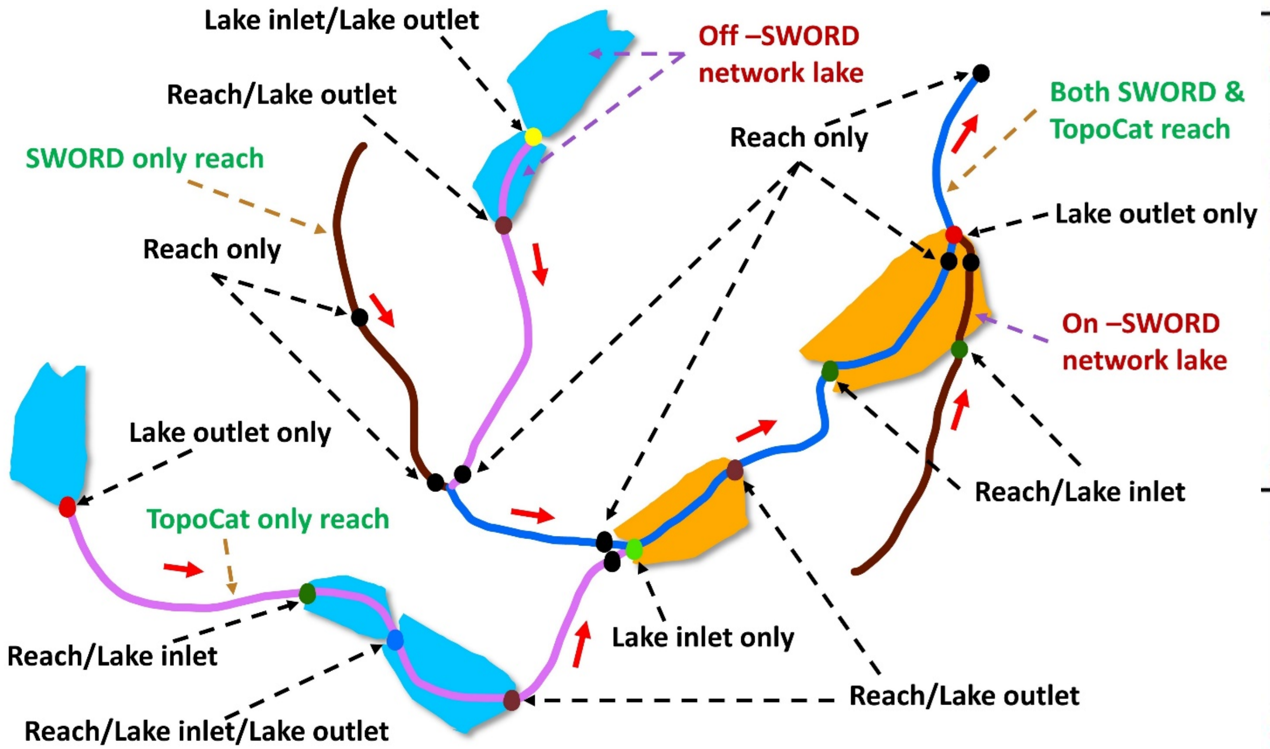
Collecting ST feedback: **Harmonized SWORD-PLD**

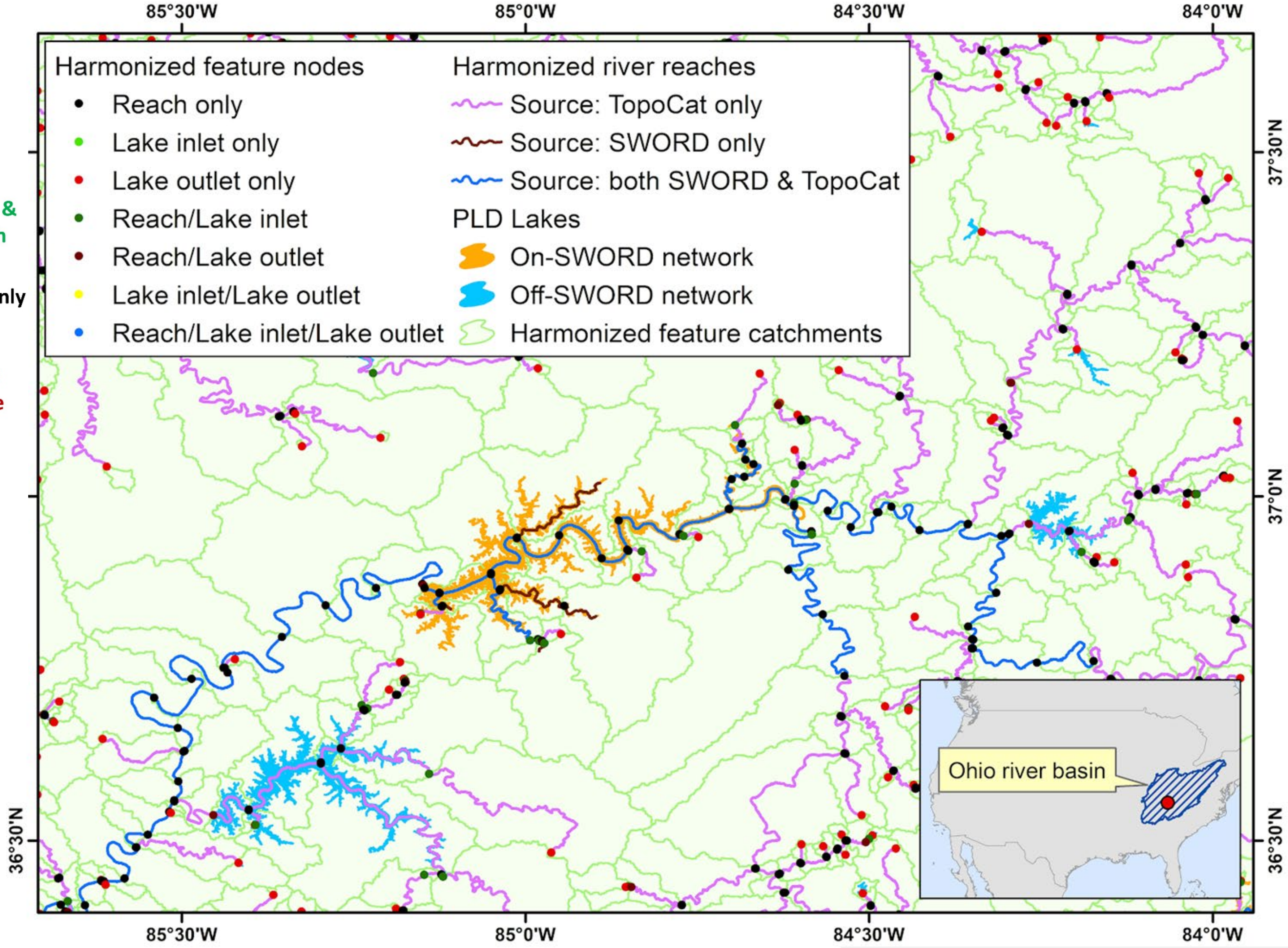
Scan to access the Harmonized SWORD-PLD database *version 1.0 beta*. Suggestions are welcome.

Purposes:

- Allow SWOT's lake and river data products to benefit each other
- Facilitate coupled lake-river modeling

Feature definition





Outcome:

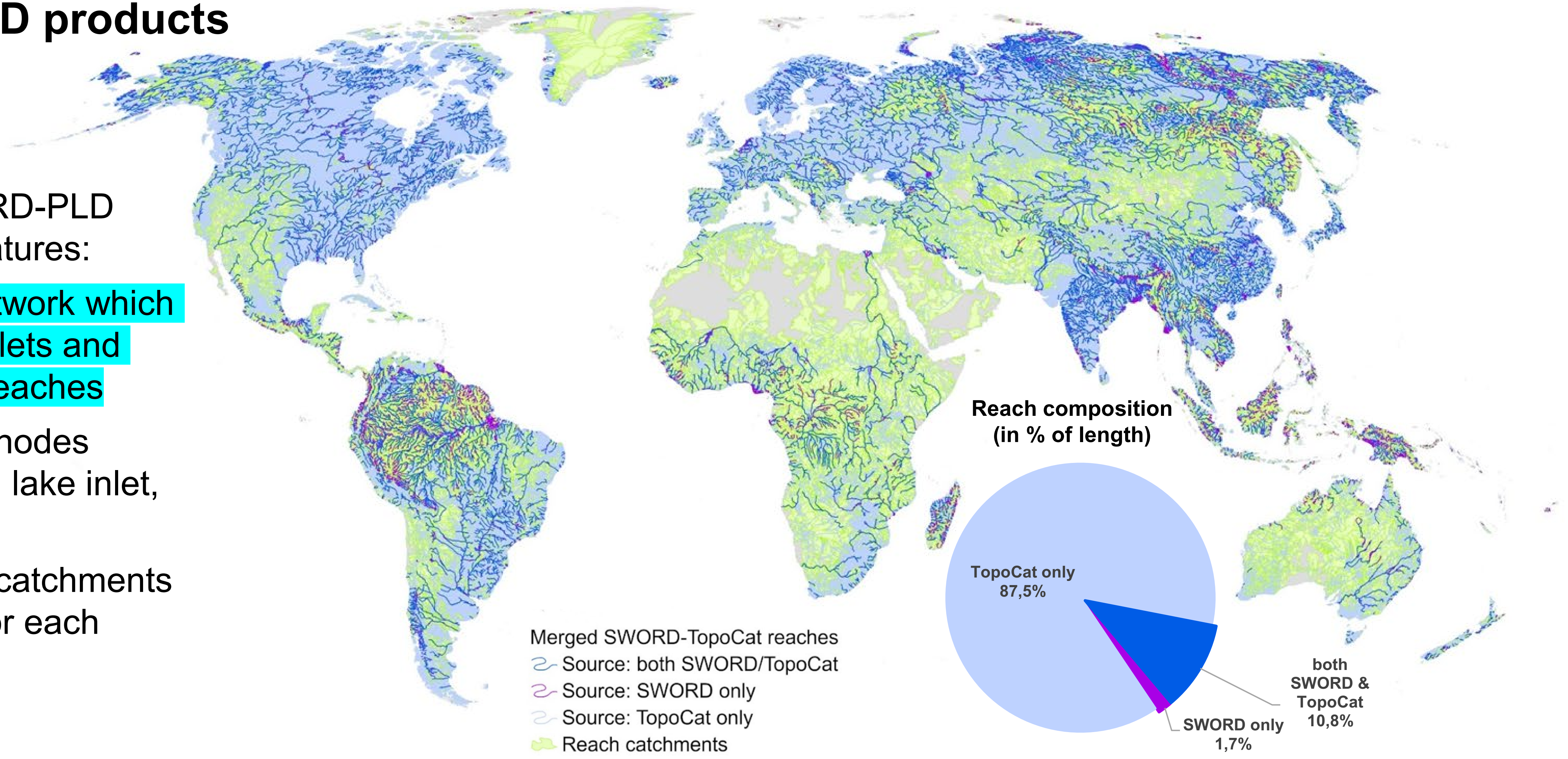
- **Nodes**: lake inlets, outlets, and reach ends
- **Reaches**: harmonized river network (19,487,626 km, ~9 times of the length of SWORD)
- **Catchments**: 15,945,655 sub-basins covering ~90% of the land area excluding Antarctica

Safat Sikder et al., in preparation

Harmonized SWORD-PLD products

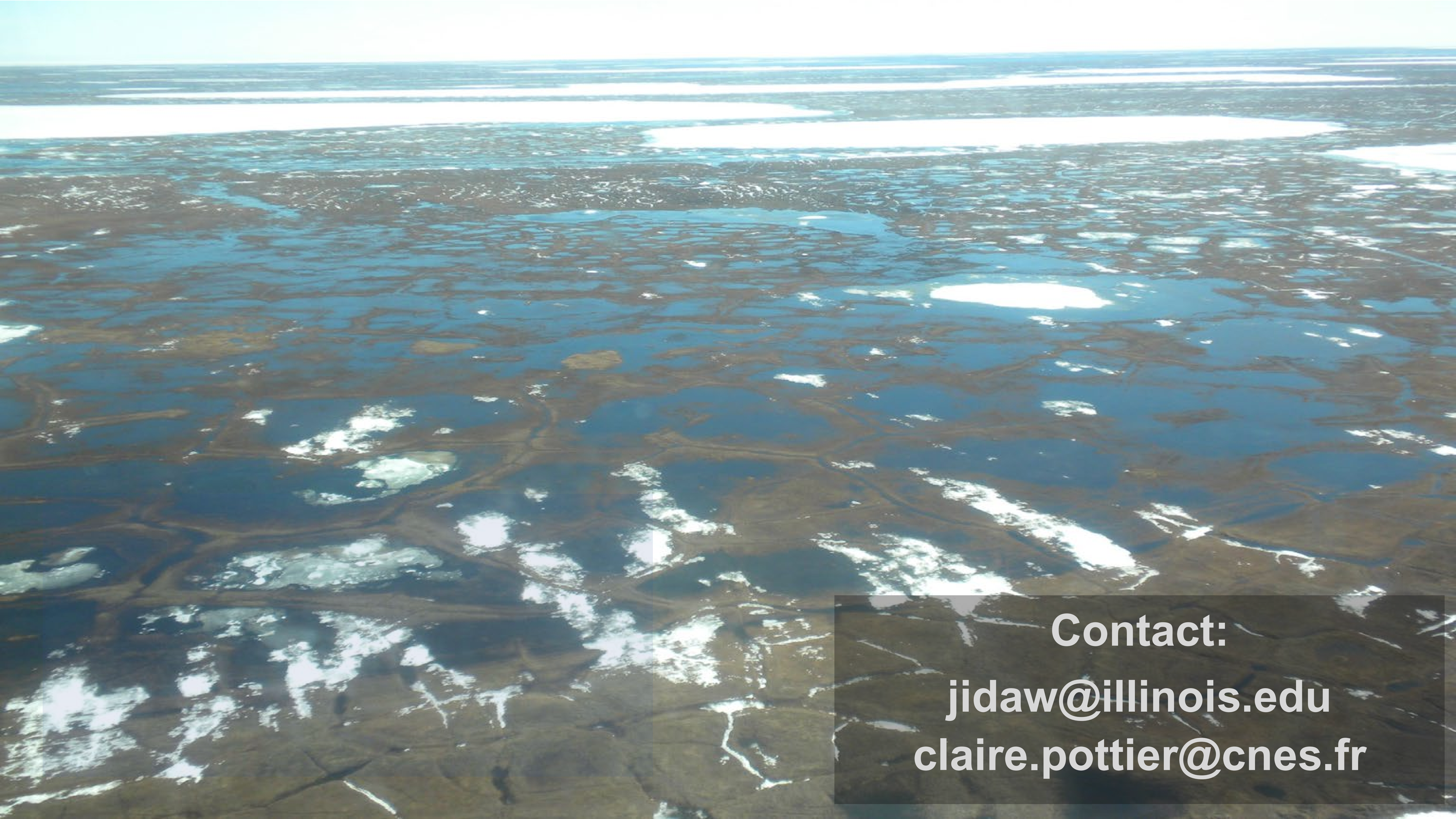
The fully harmonized SWORD-PLD dataset consists of three features:

1. the harmonized river network which connects all PLD lake inlets and outlets to the SWORD reaches
2. the harmonized feature nodes (reach downstream end, lake inlet, and lake outlet), and
3. the harmonized feature catchments (unit/local catchments for each feature)



The Harmonized network:

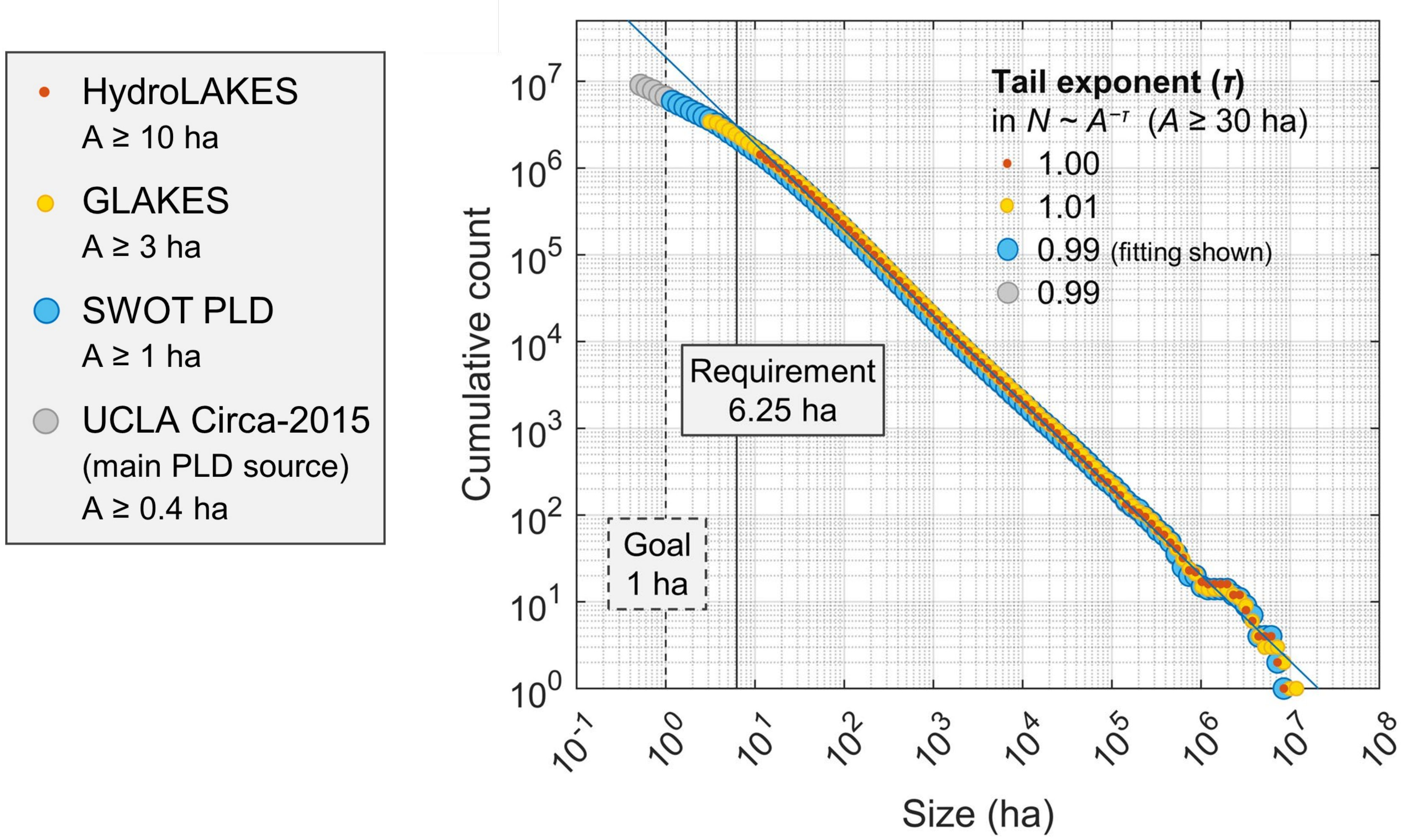
- The harmonized river network (with typology error flags on SWORD) is the merged SWORD reaches and inter-lake reaches of the PLD lakes (from PLD-TopoCat; *Sikder et al., 2023*)
- Total length of the harmonized network is about 9 times larger than the SWORD network.
- A total of 12,544,835 reaches connects the inlets and/or outlets of 5,892,853 PLD lakes into the SWORD network



Contact:

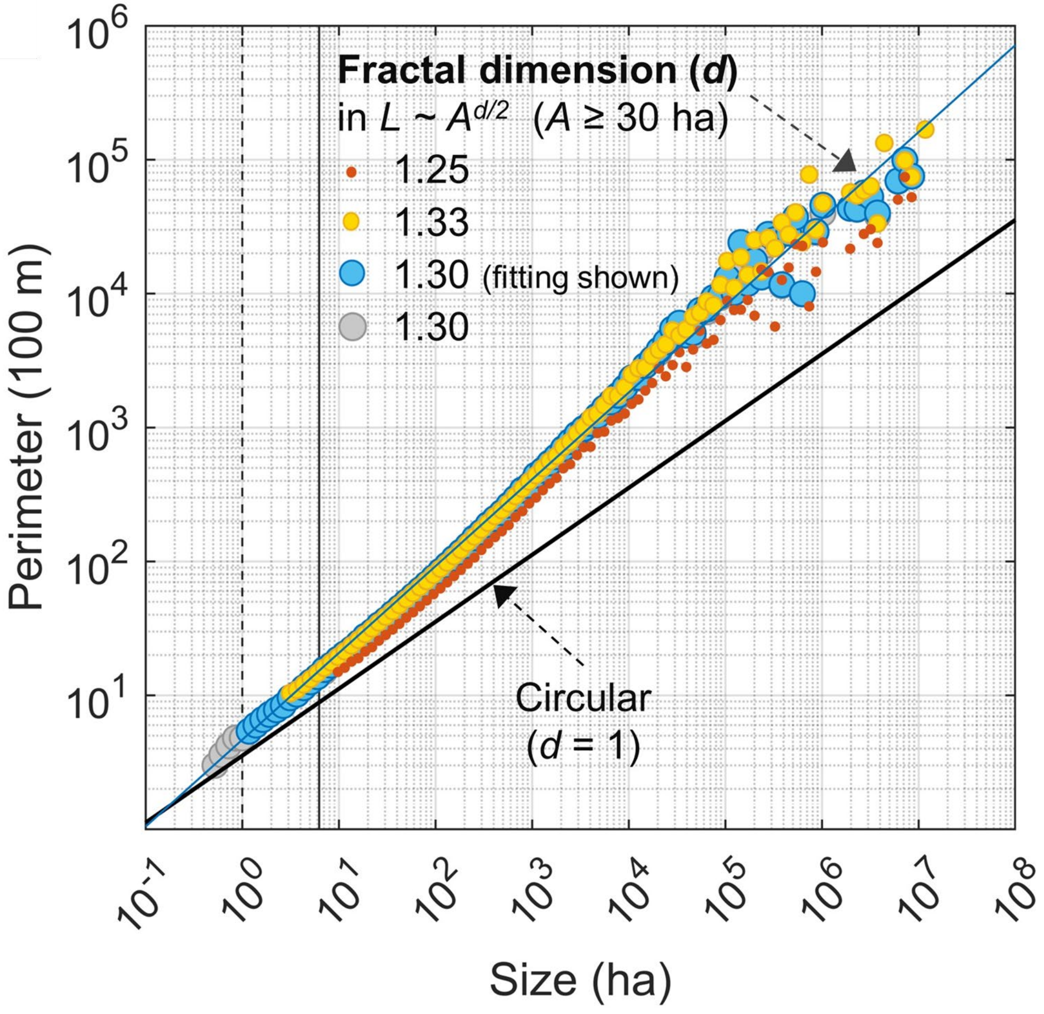
jidaw@illinois.edu

claire.pottier@cnes.fr



Size-abundance relationship:

- Pareto distribution (scale-invariant) until lake size is smaller than ~30 ha:
  - Tail exponents (τ) for all data are consistent (~1.0), conforming to the percolation theory.
- Lakes < 30 ha are less self-similar, highlighting the value of PLD for monitoring small lake dynamics.

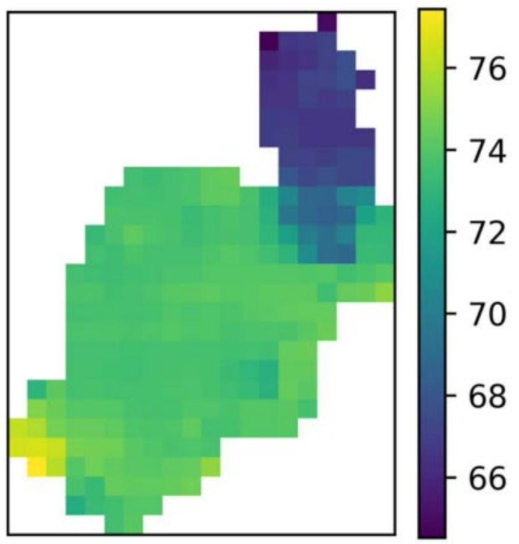


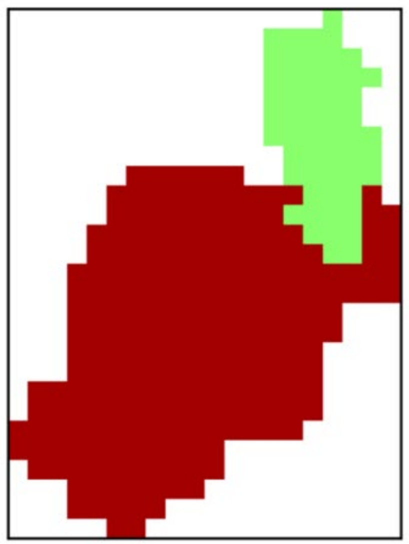
Perimeter-area scaling:

- Scale-invariant for lakes > 30 ha:
  - Fractal dimensions (d) for all data are overall consistent (~1.3).
  - Lower d for HydroLAKES indicates simplified lake boundaries.
- For lakes < 30 ha, d tends to approach 1.0 (more circularly shaped)
- Highlighting the advantage of PLD in representing reliable shoreline morphology for both sizable and small lakes.

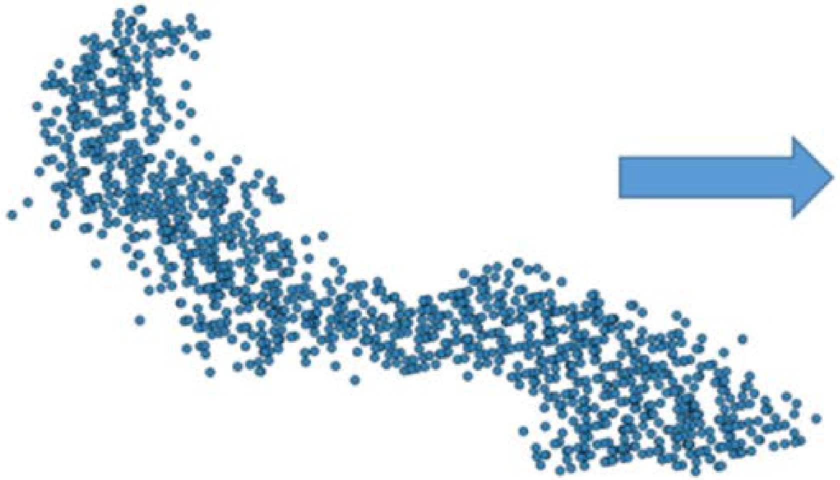
How the PLD helps link SWOT observations

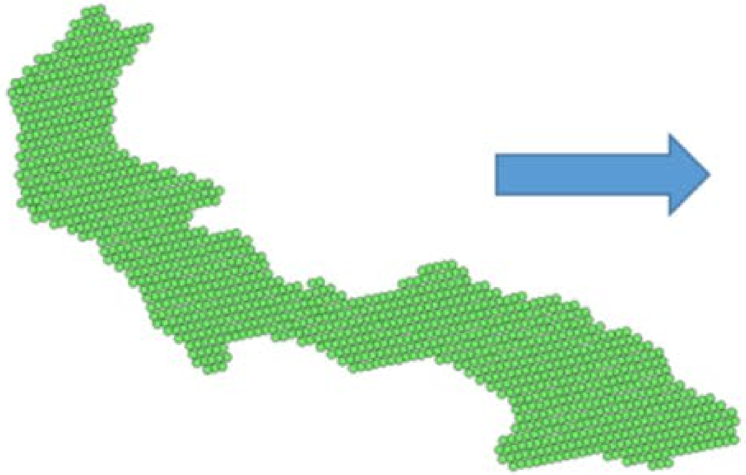
Water surface elevation (WSE) in the SLC image





Segmented water regions based on WSE statistics





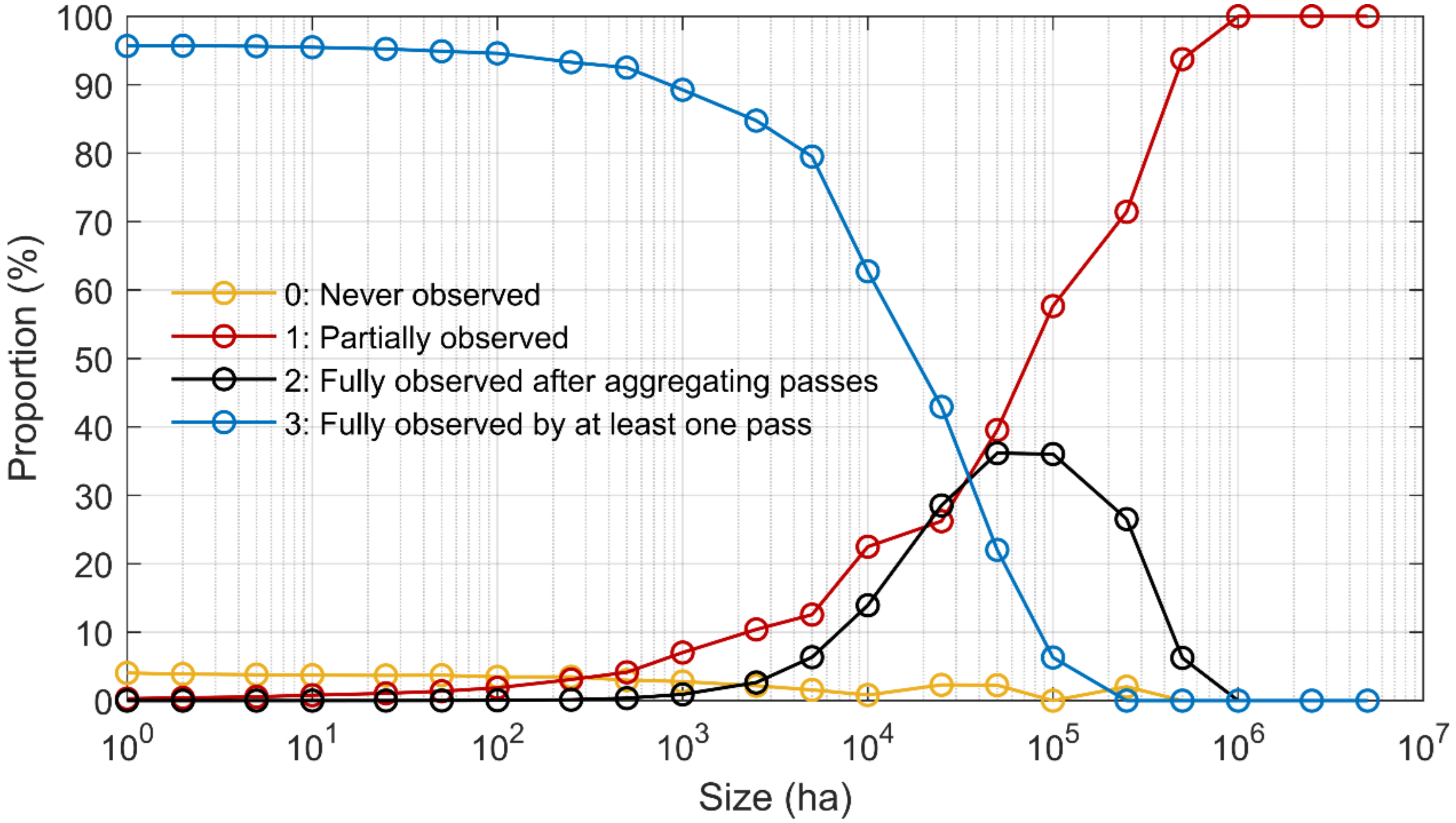
Original PIXC (pixel cloud) for any water region

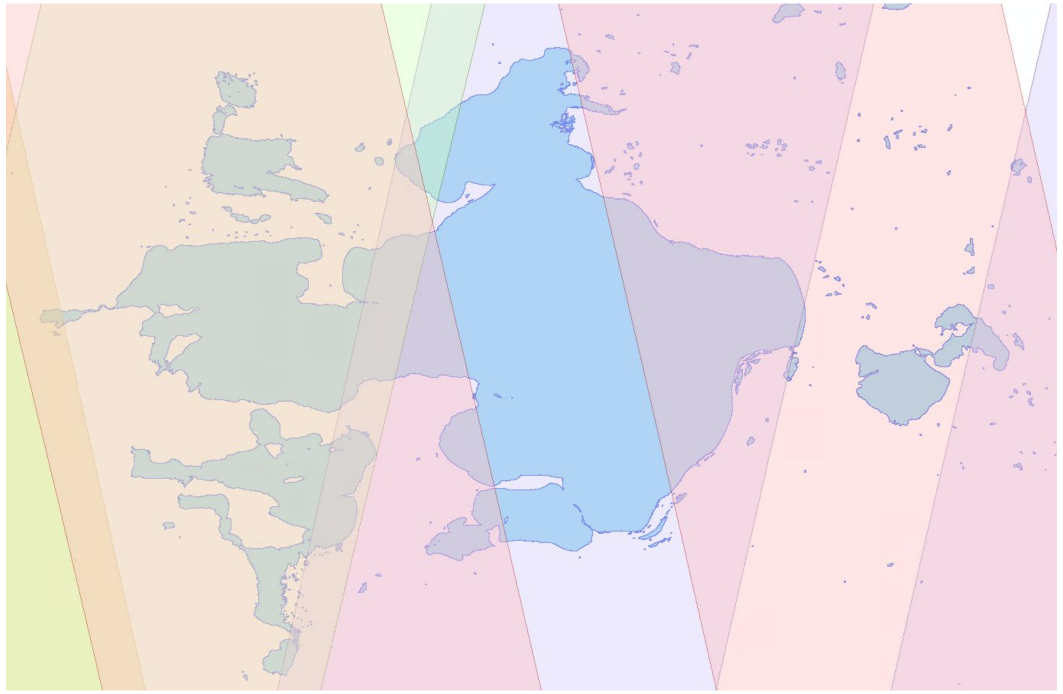
PIXC with height-constrained geolocation (PIXCVec)

Vectorized water feature (SWOT-observed water feature)

L2_HR_LakeSP ATBD and product description
CNES Internal Document, 2022, 2023

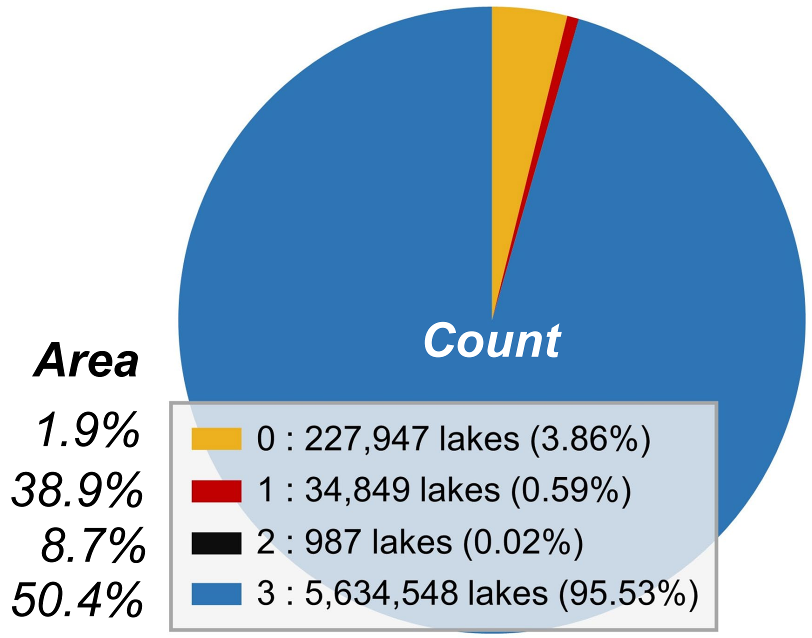
Lake spatial coverage during each SWOT nominal (21-day) orbit cycle





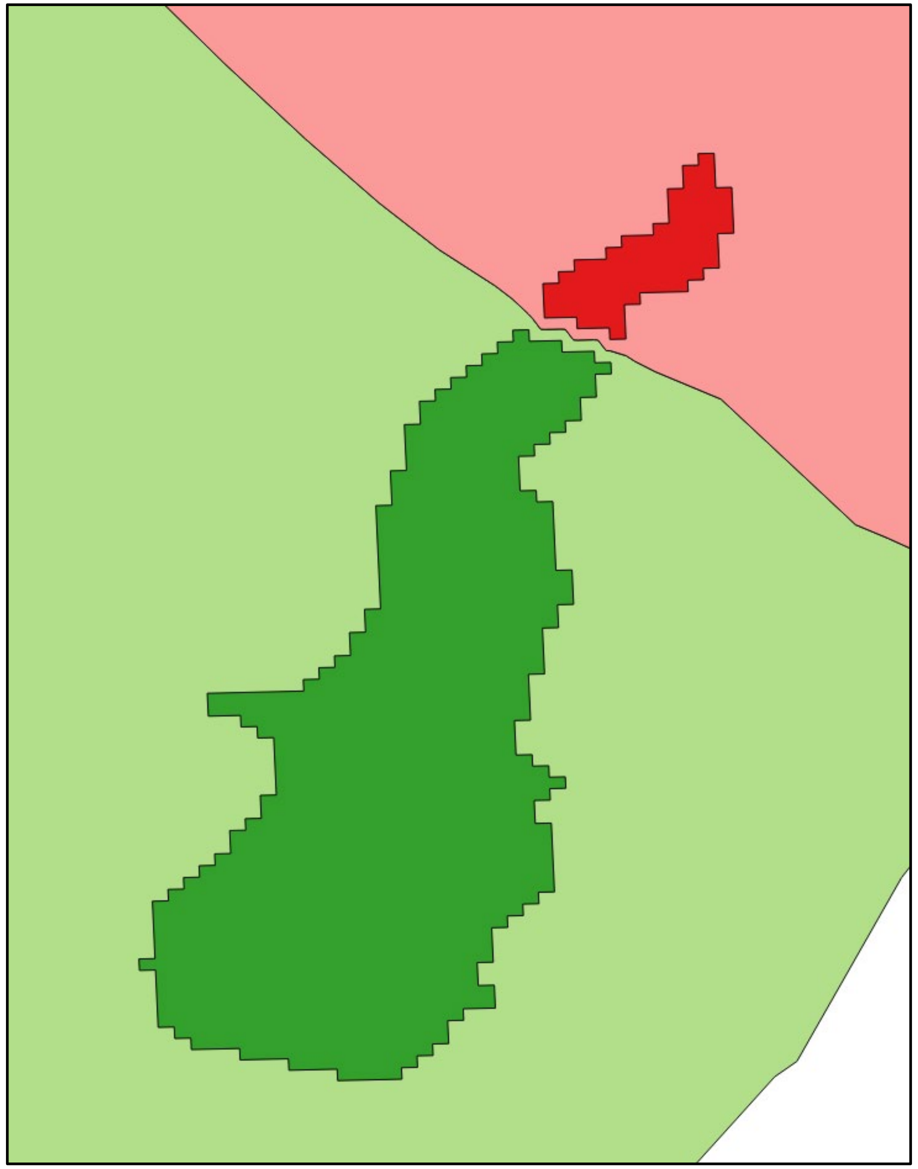
LakeSP
(partially observed)

LakeAvg
(fully observed)

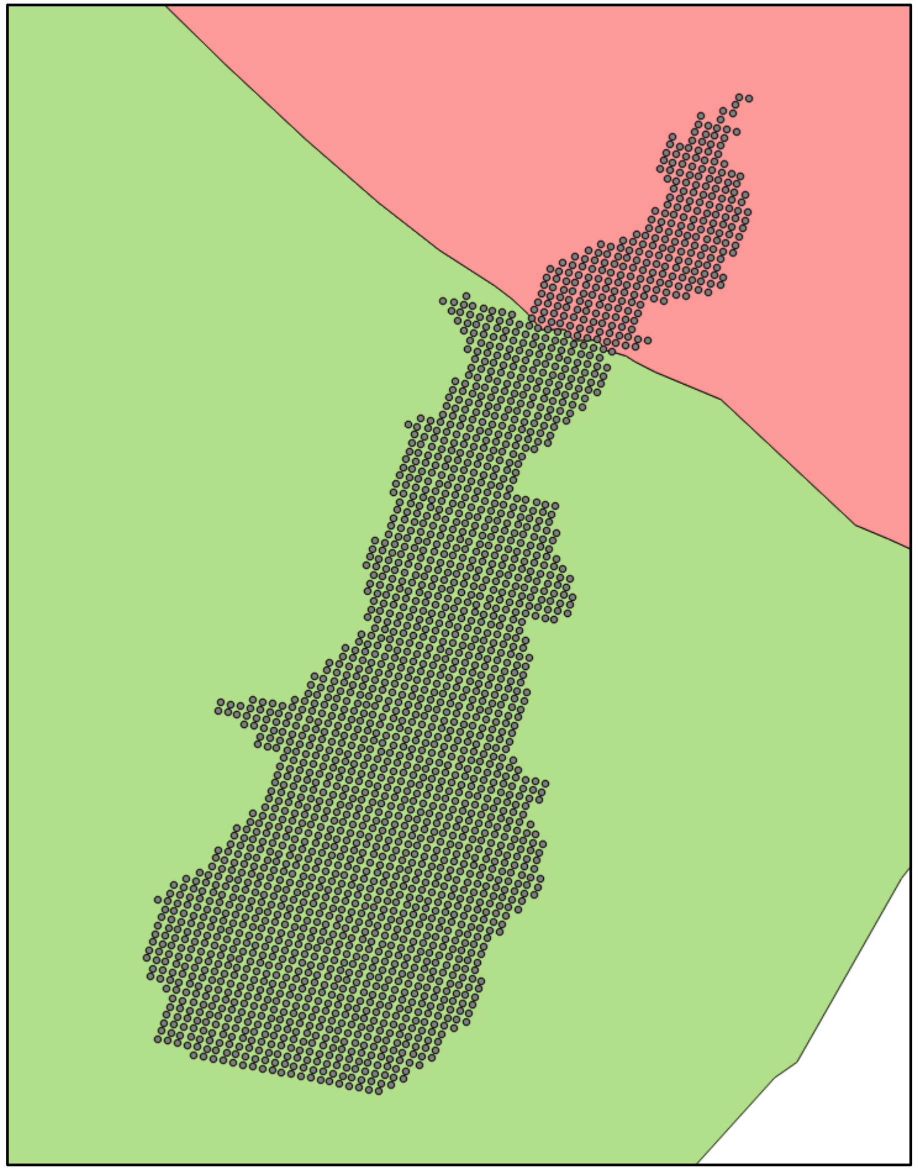


- ~90% of the lakes < 10 km^2 (1000 ha) are fully observed at least once per cycle.
- All three observation scenarios (1, 2, 3) cross at ~300–500 km^2, about 33% for each.
- Lakes > 500 km^2 are gradually dominated by scenario 1, i.e., partially observed, despite a very high overpass frequency.

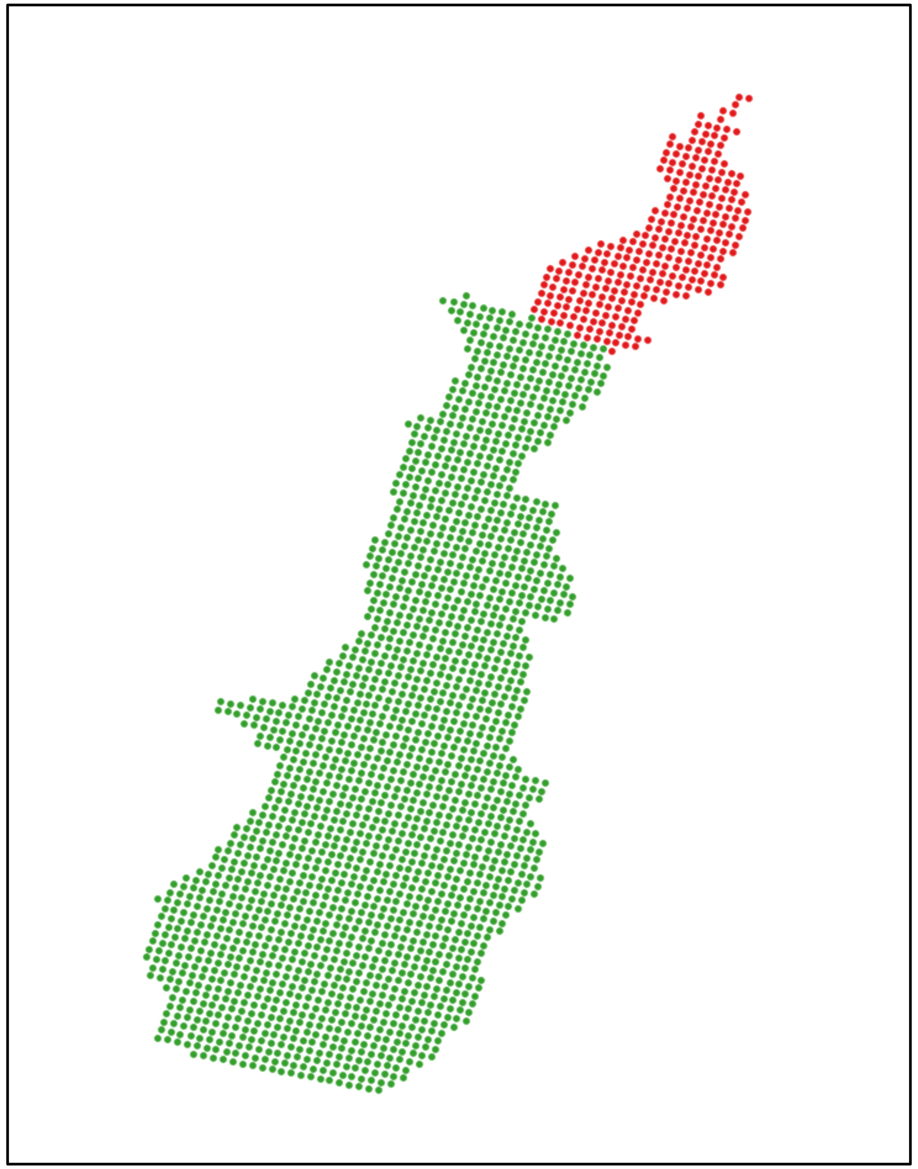
Link PIXC pixels to prior lakes if an observed feature matches more than one prior lake



PLD lakes and their influence area



PIXCVec and lakes influence area



PIXCVec linked to appropriate PLD lake

Courtesy: Claire Pottier, CNES