Amazon water levels with deep learning

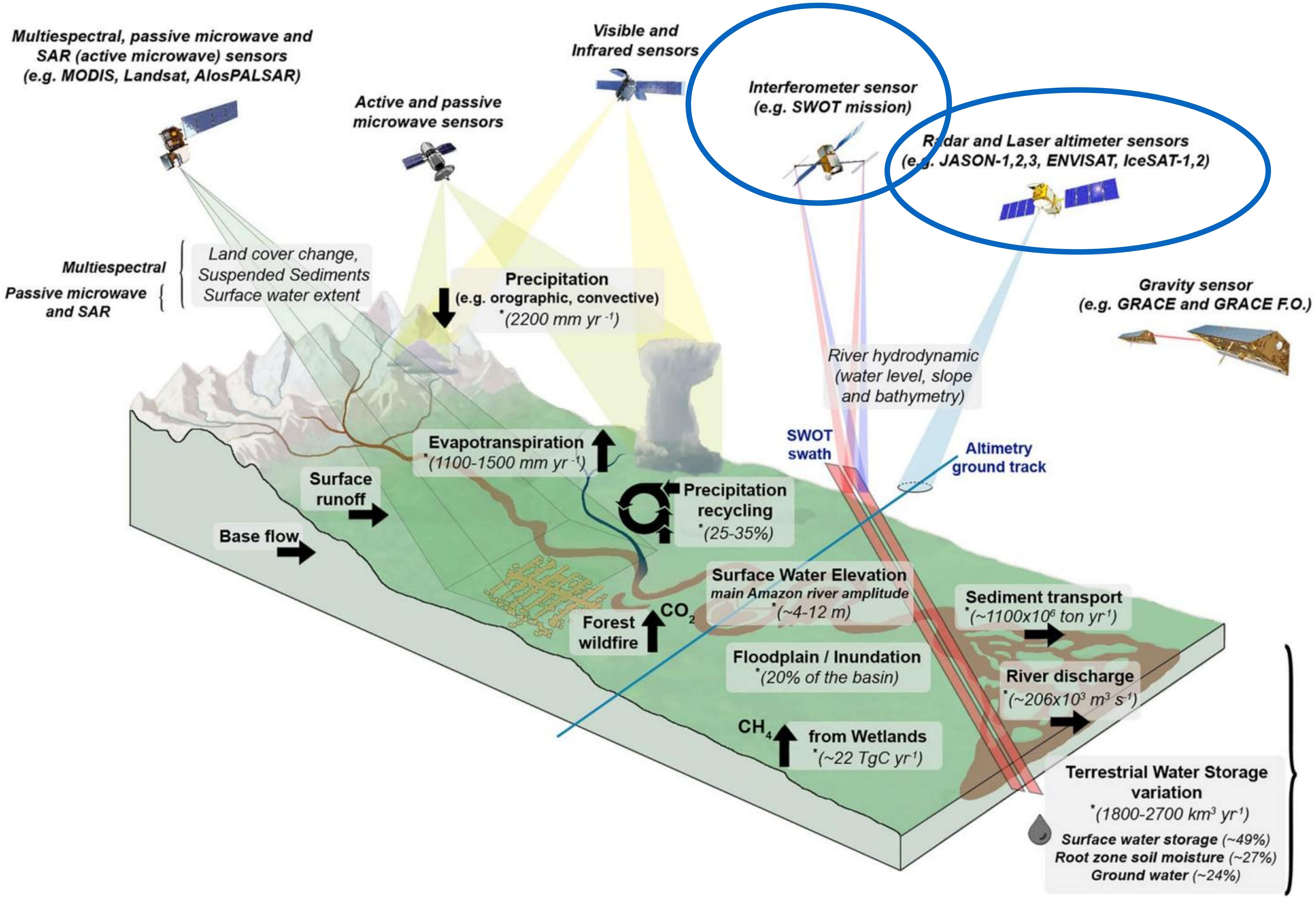
Ruben Cartuyvels,
Karim Douch, Gabriele Bertoli, Mounia El Baz, Sébastien Lefèvre, Diego Fernández-Prieto

Φ-lab & Science Hub, ESA

Collaborators

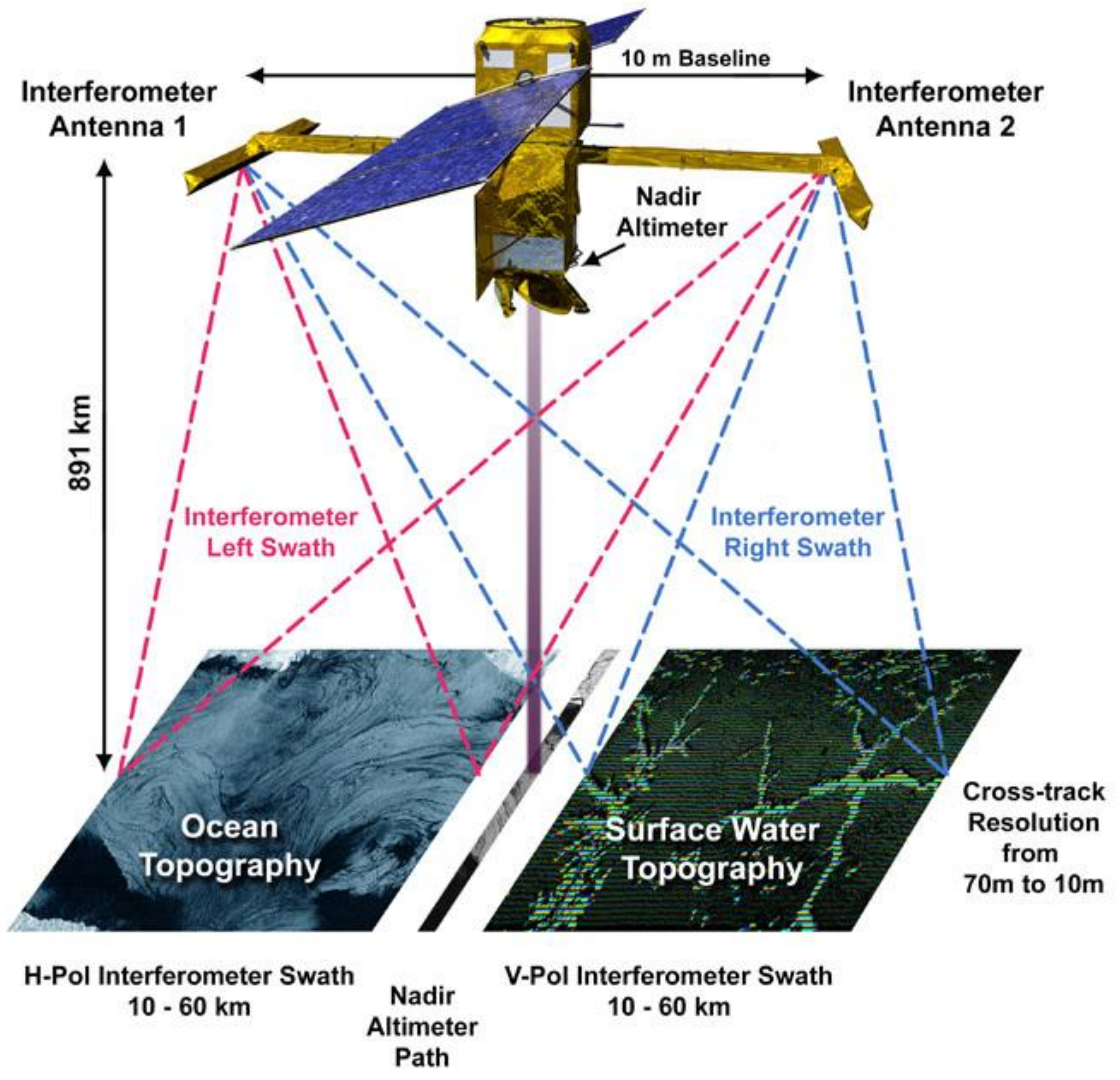
Ruben Cartuyvels	Mounia El Baz	Karim Douch	Diego Fernández-Prieto	Gabriele Bertoli	Sébastien Lefèvre
					

Intro



Fassoni-Andrade, A. C., Fleischmann, A. S., Papa, F., Paiva, R. C. D. D., Wongchuig, S., Melack, J. M., ... & Pellet, V. (2021). Amazon hydrology from space: scientific advances and future challenges. *Reviews of Geophysics*, *59*(4), e2020RG000728.

SWOT mission



Article | Open access | Published: 20 March 2024

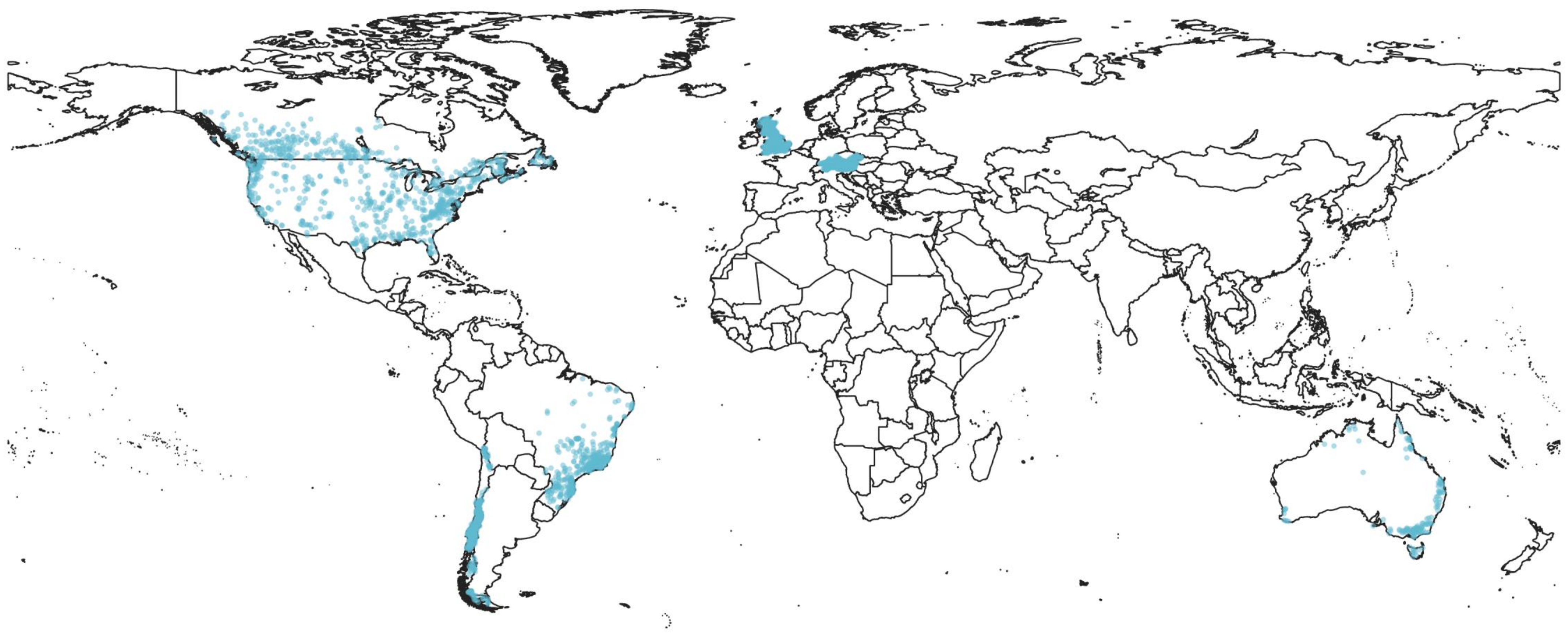
Global prediction of extreme floods in ungauged watersheds

Grey Nearing, Deborah Cohen, Vusumuzi Dube, Martin Gauch, Oren Gilon, Shaun Harrigan, Avinatan Hassidim, Daniel Klotz, Frederik Kratzert, Asher Metzger, Sella Nevo, Florian Pappenberger, Christel Prudhomme, Guy Shalev, Shlomo Shenzis, Tadele Yednkachw Tekalign, Dana Weitzner & Yossi Matias

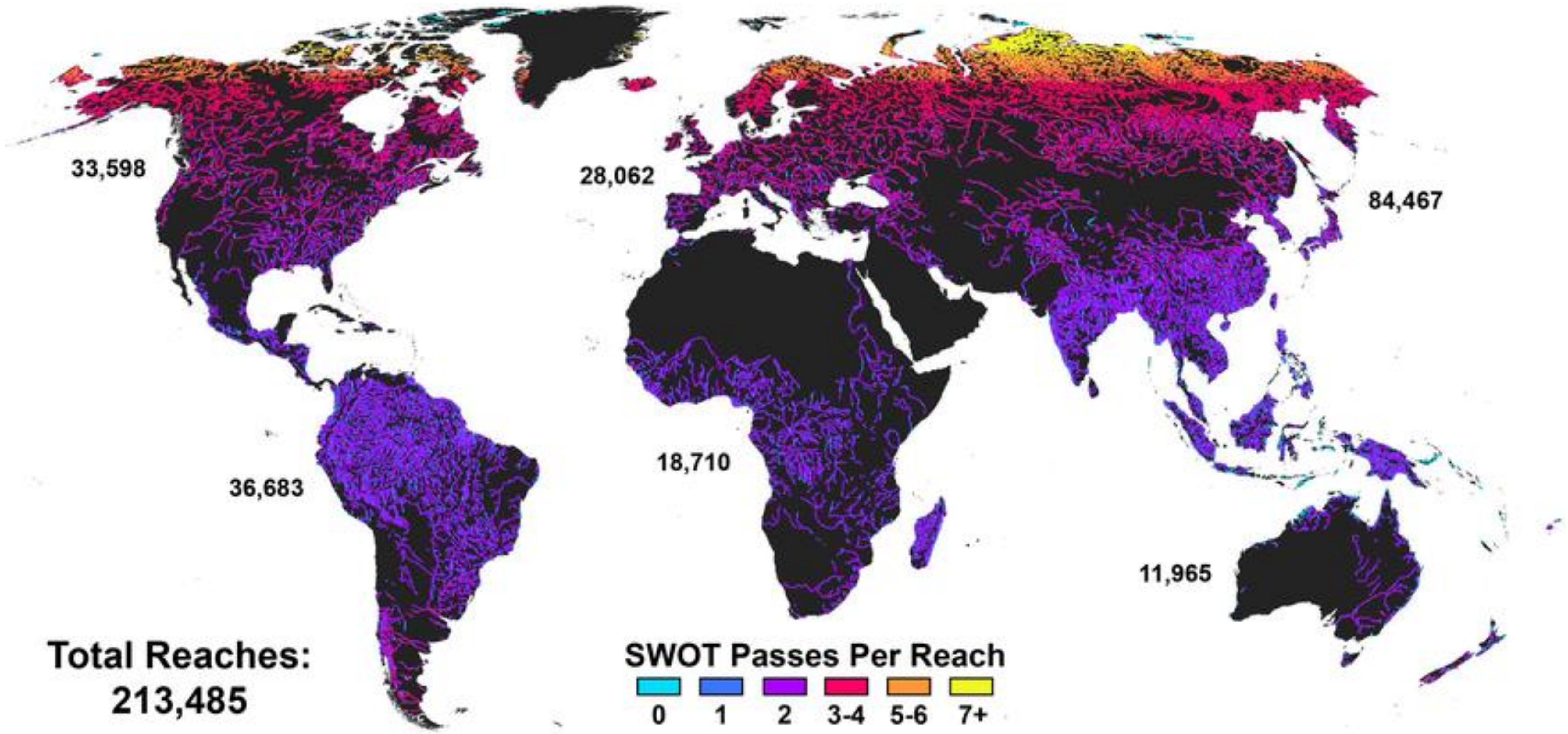
Nature **627**, 559–563 (2024) | Cite this article

Abstract

Floods are one of the most common natural disasters, with a disproportionate impact in developing countries that often lack dense streamflow gauge networks[1]. Accurate and timely warnings are critical for mitigating flood risks[2], but hydrological simulation models typically must be calibrated to long data records in each watershed. Here we show that artificial intelligence-based forecasting achieves reliability in predicting extreme riverine events in



"5,680 streamflow gauges"



Altenau, E. H., Pavelsky, T. M., Durand, M. T., Yang, X., Frasson, R. P. D. M., & Bendezu, L. (2021). The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A global river network for satellite data products. Water Resources Research, 57(7), e2021WR030054.

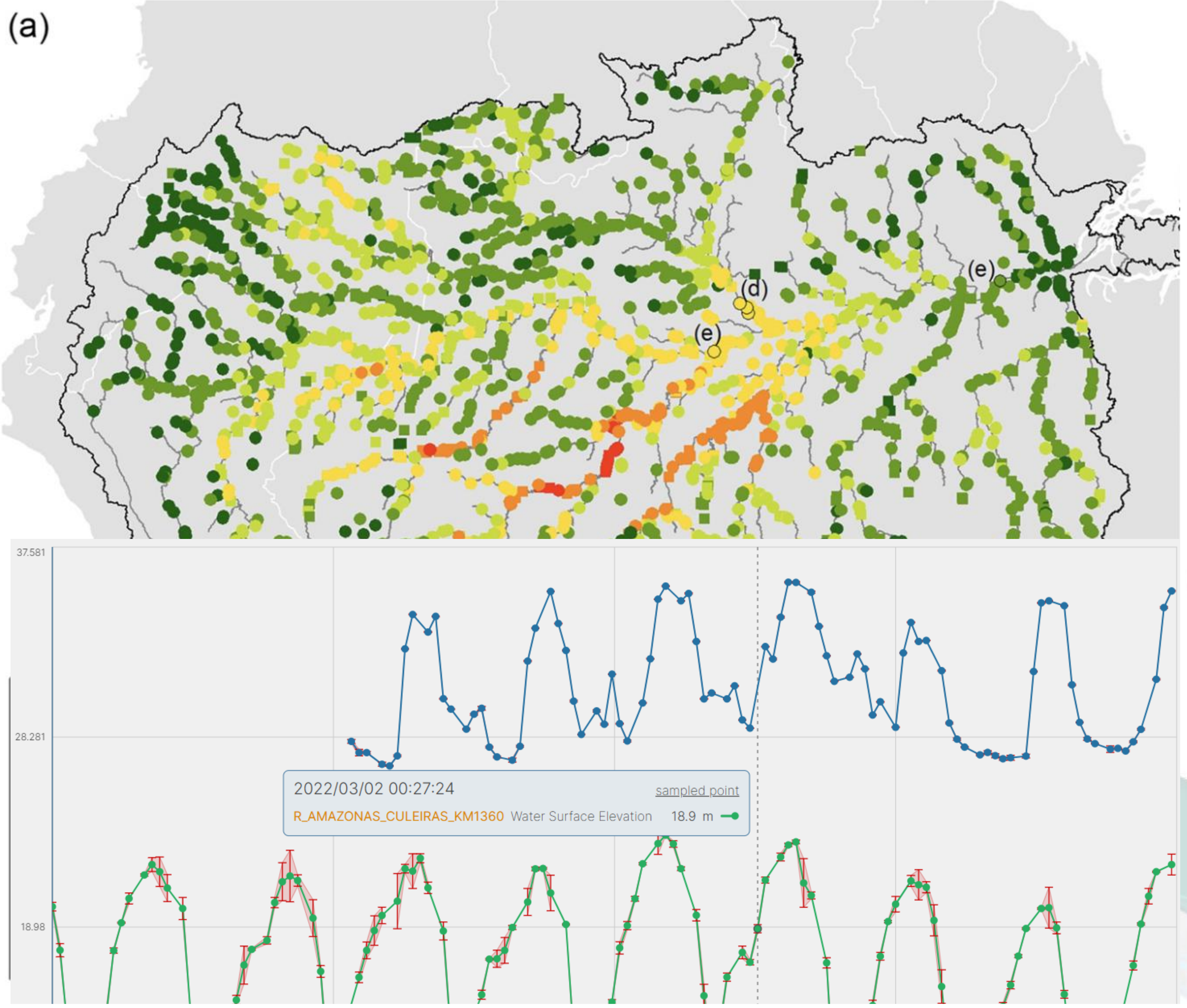
Problem overview

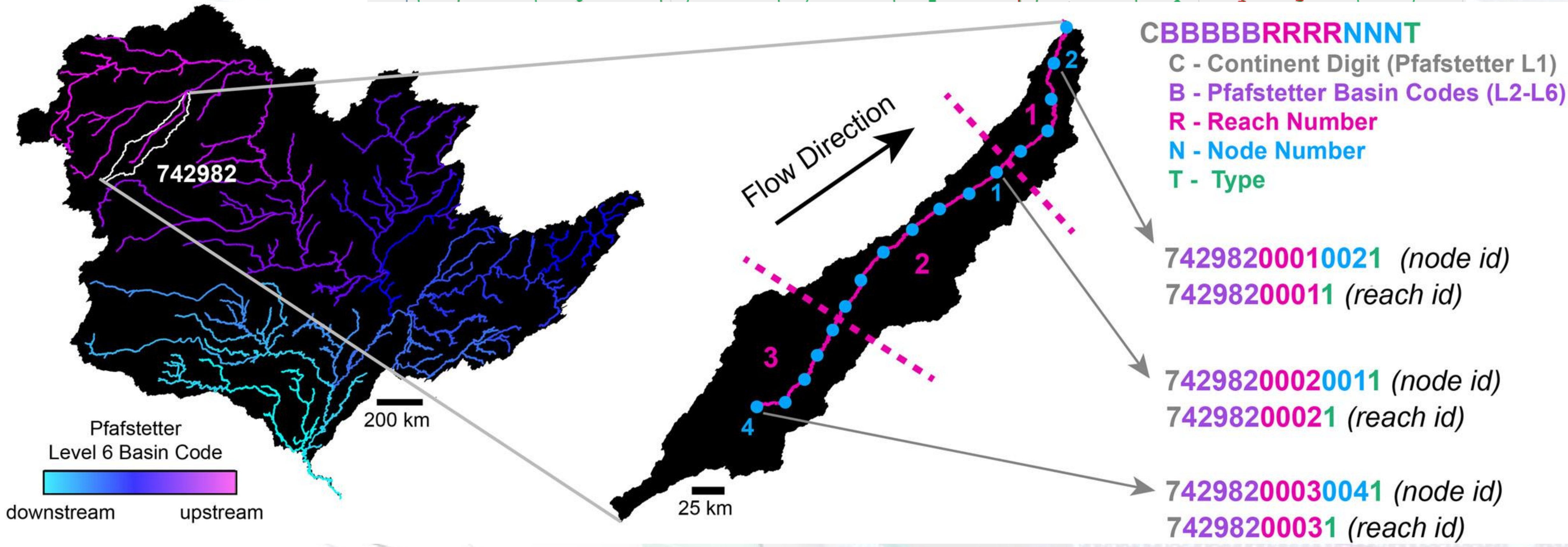
Water surface elevation

- Timeseries for non-gridded (lat, lon) locations

Sources

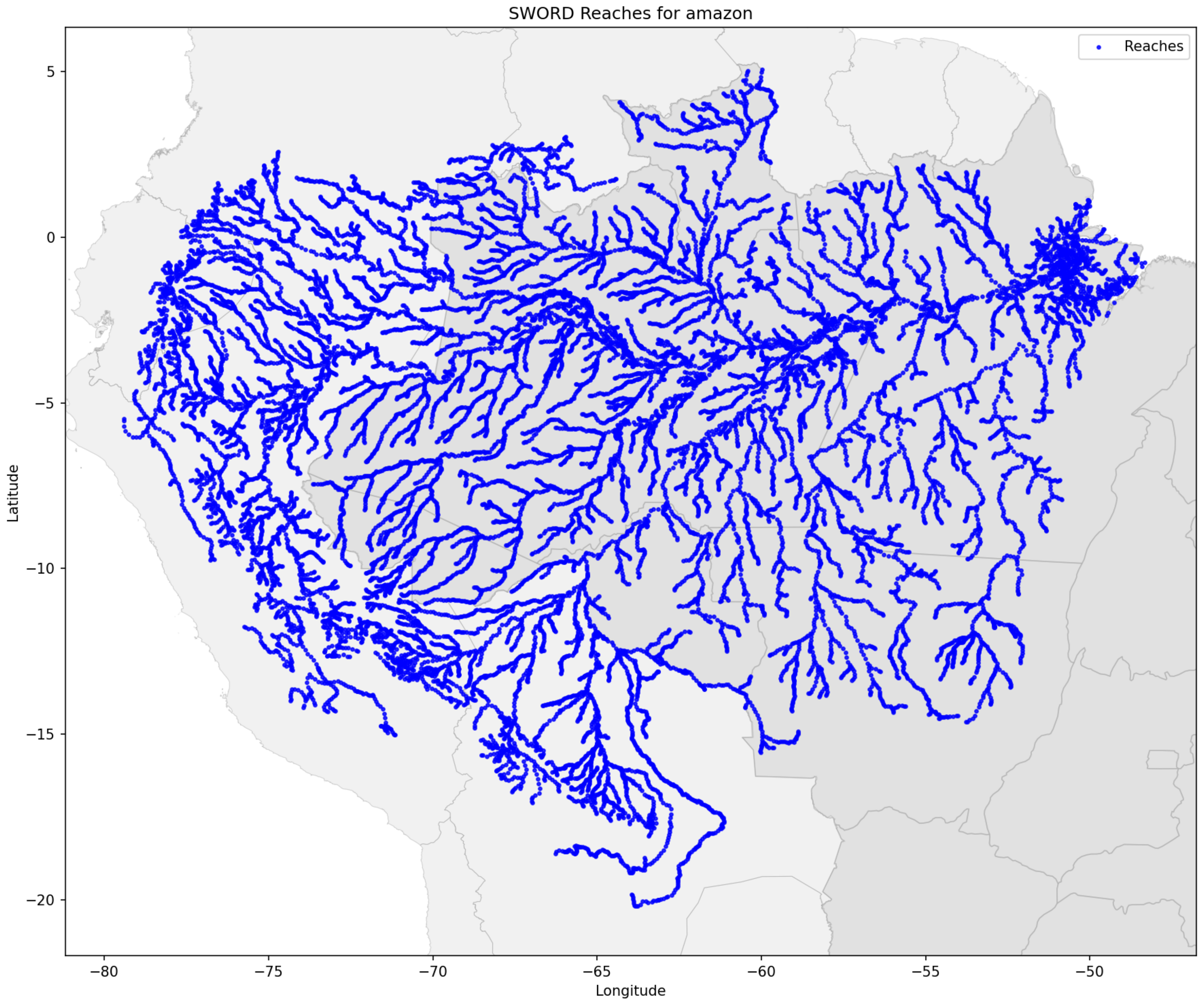
- **SWOT**: 2023-2026, dense in space / sparse in time
- **Classical altimetry**: 1993-2026, sparse in space & time
  - HydroWeb: ERS-1, Topex/Poseidon, ERS-2, GFO, Jason-1, Envisat, Jason-2, Saral/Altika, Jason-3, Sentinel-3A, Sentinel-3B, Sentinel-6A and SWOT
- **In-situ gauges**: 19??-2026, very sparse in space / dense in time
  - From Brazilian ANA API
- **SWORD**: static river database with topology

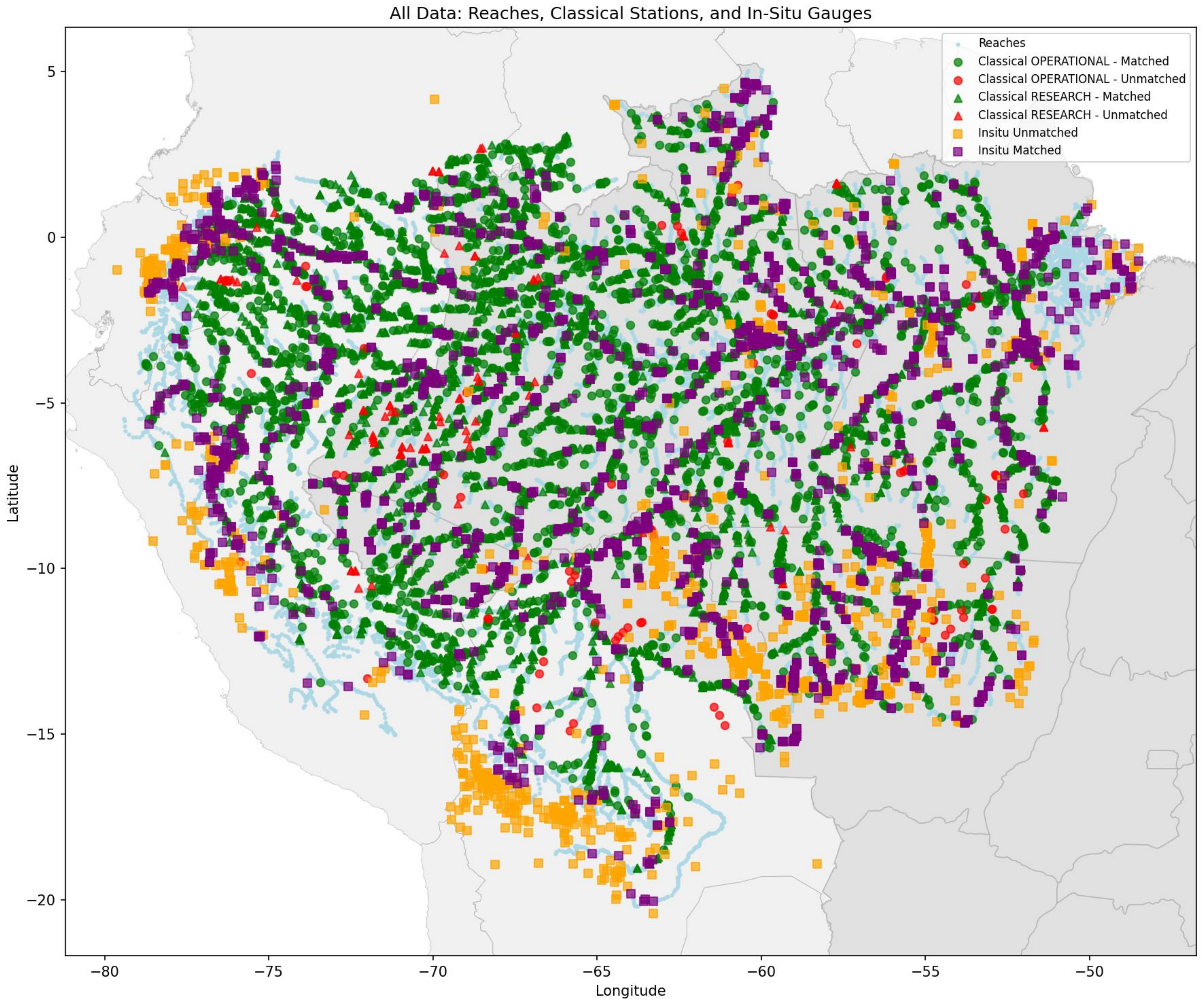




https://www.theia-land.fr/en/blog/product/water-levels-of-rivers-and-lakes-hydroweb/
Normandin, C., Frappart, F., Diepkilé, A. T., Marieu, V., Mougin, E., Blarel, F., ... & Ba, A. (2018). Evolution of the performances of radar altimetry missions from ERS-2 to Sentinel-3A over the Inner Niger Delta. Remote Sensing, 10(6), 833.
https://www.ana.gov.br/hidrowebservice/swagger-ui/index.html#/

Problem overview



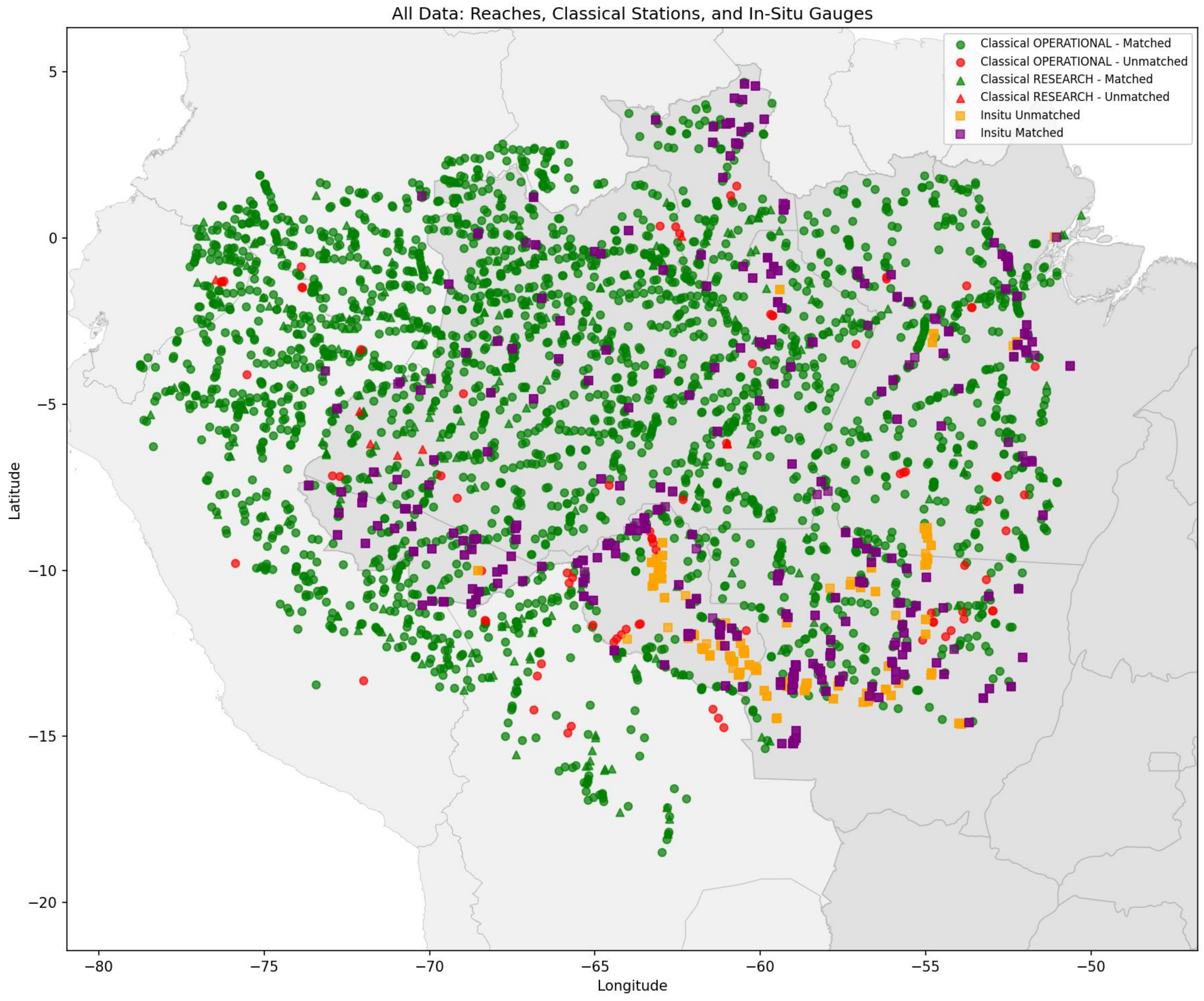


Statistics:

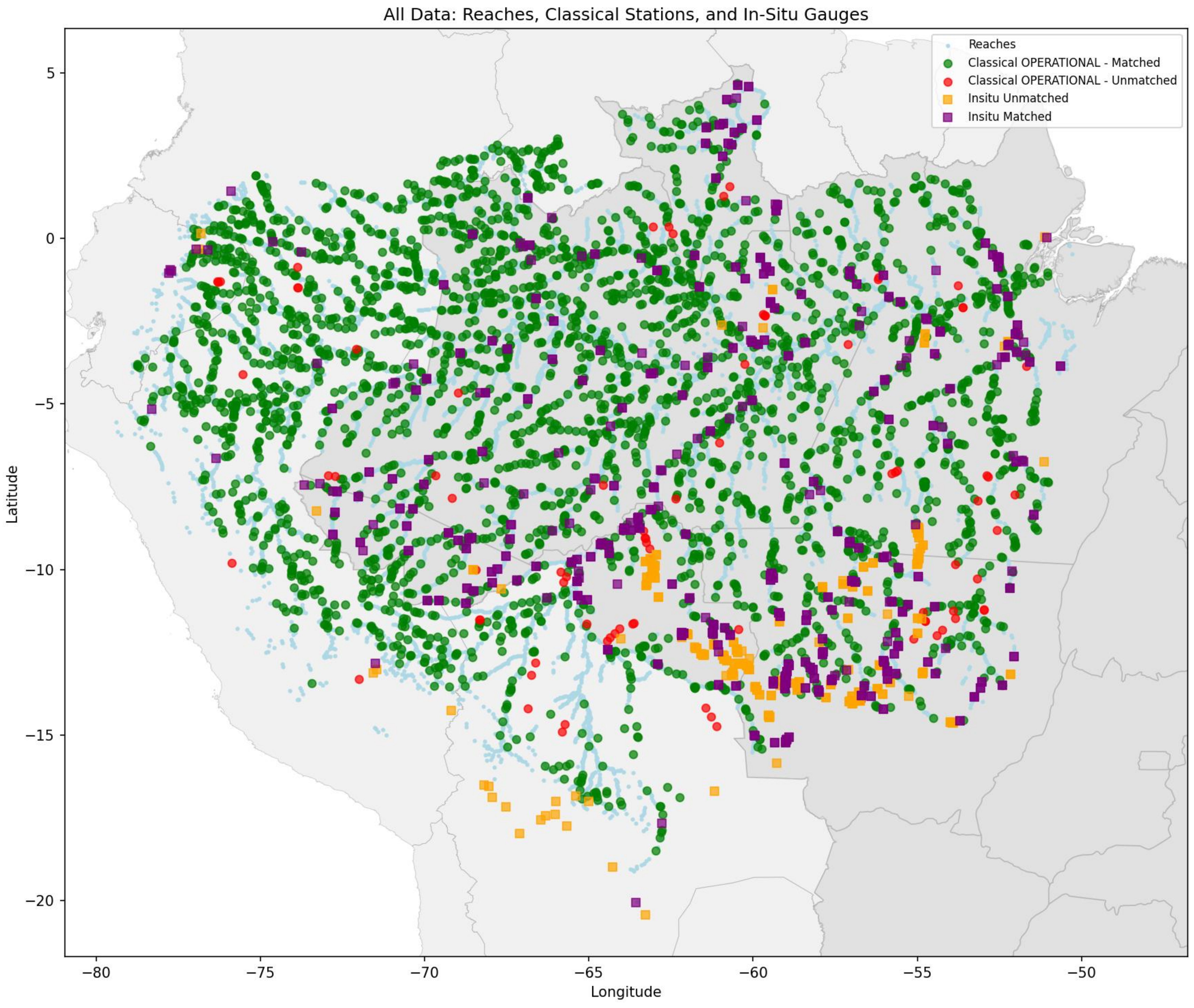
- 17K /19K SWORD reaches are observed by SWOT
- 10K pass our quality filters
- 3K classical altimetry virtual stations
- 100-300 insitu gauges

Problem overview

2016-2023

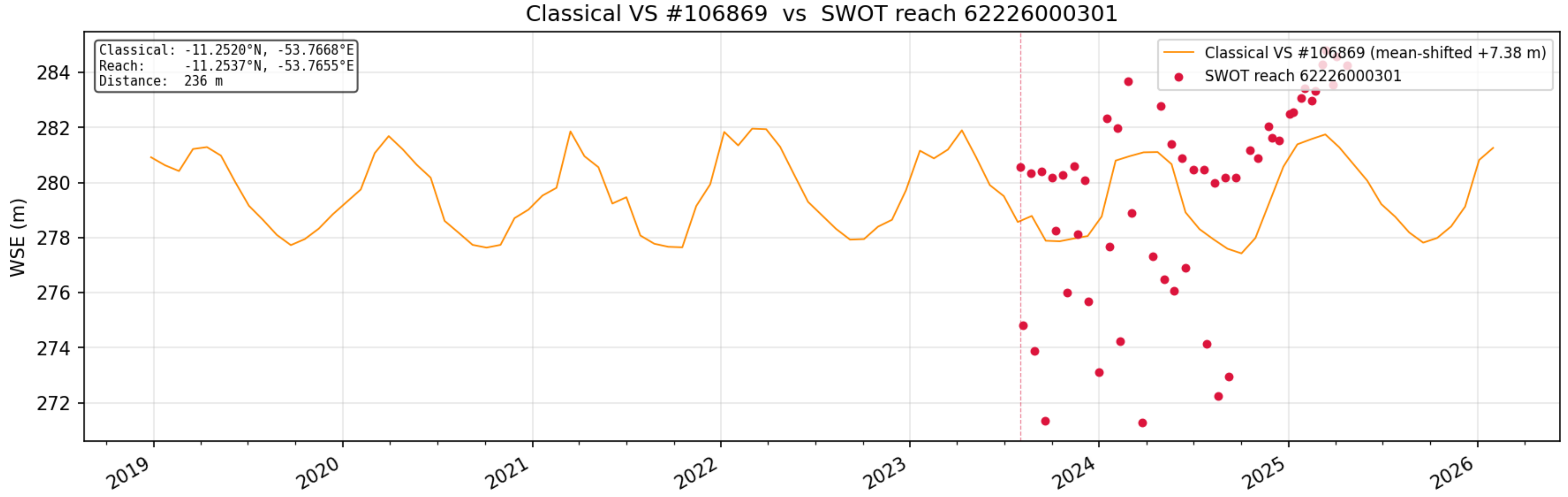


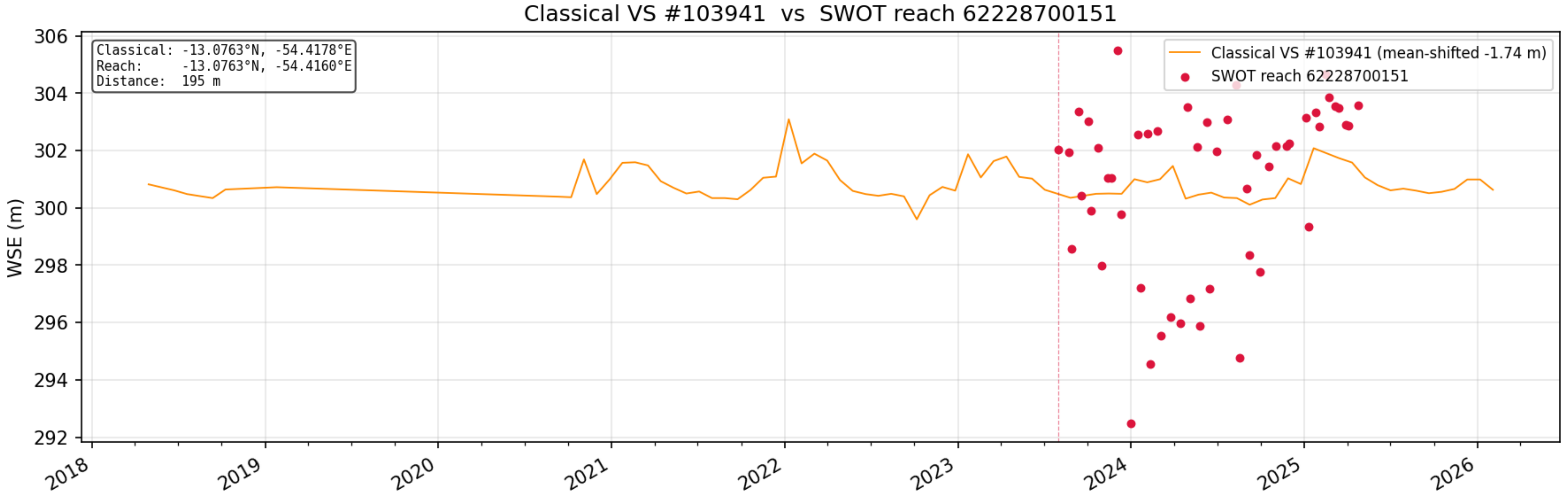
2023-…



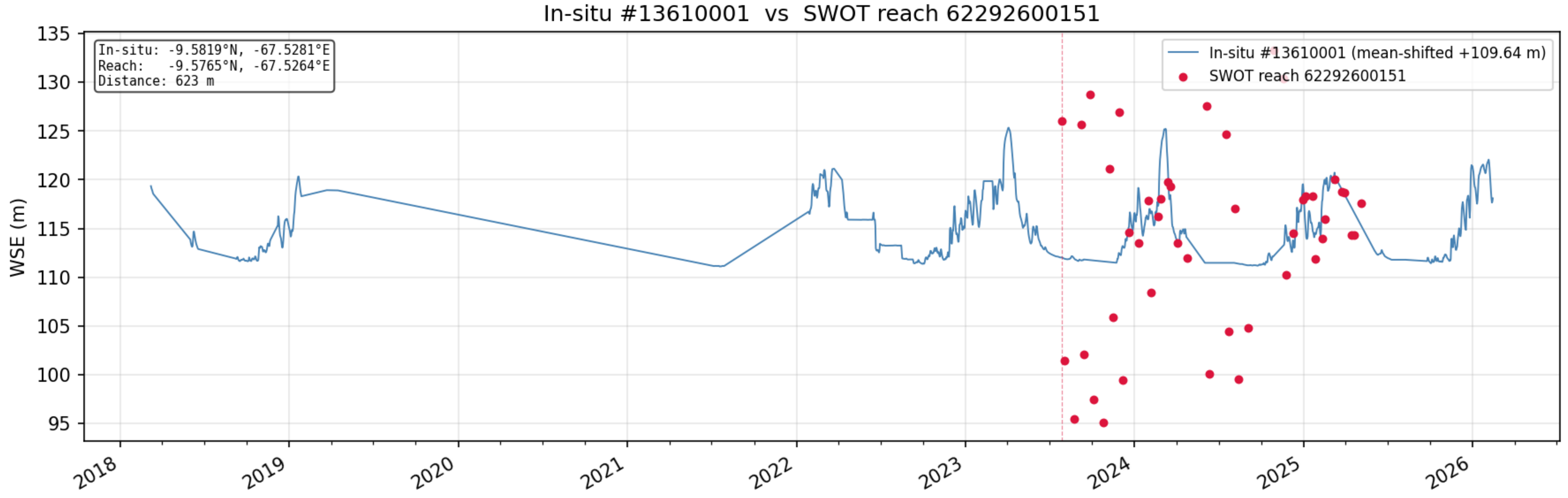
L2 HR RiverSP – Reach version C

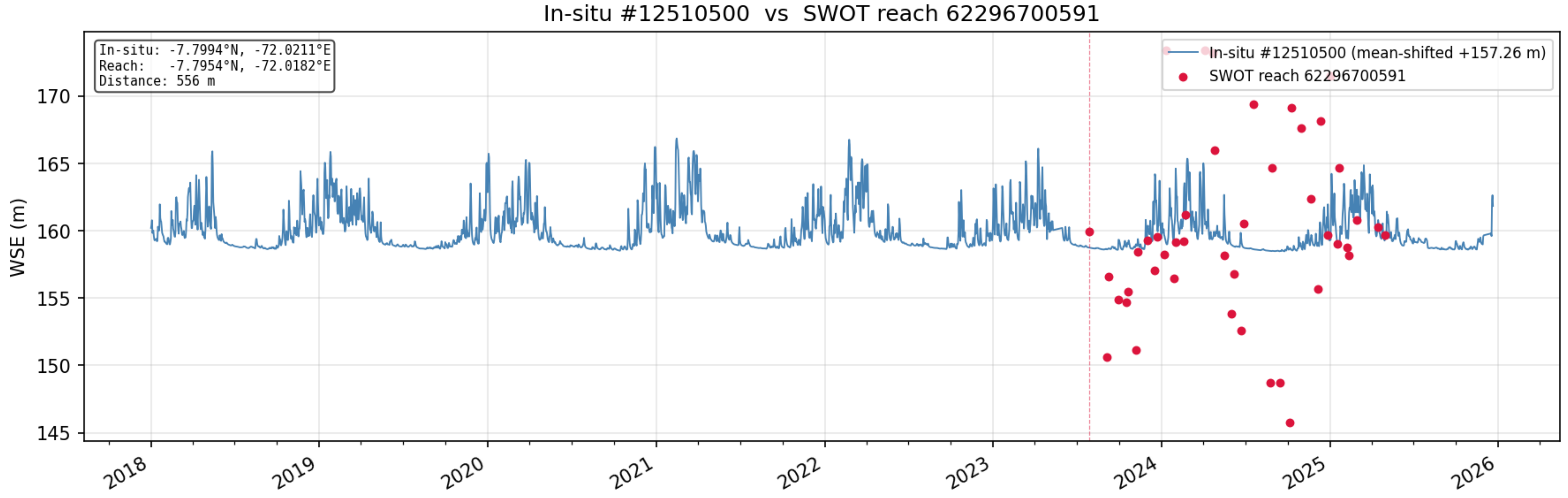
SWOT vs Classical altimetry





SWOT vs In-situ gauges

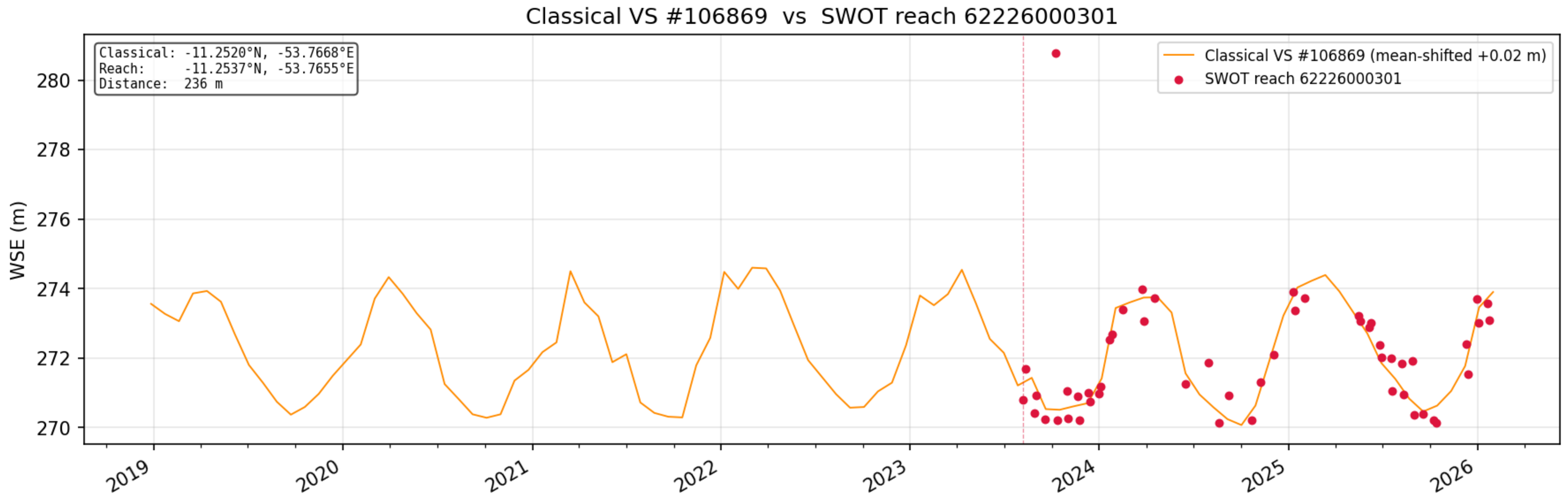


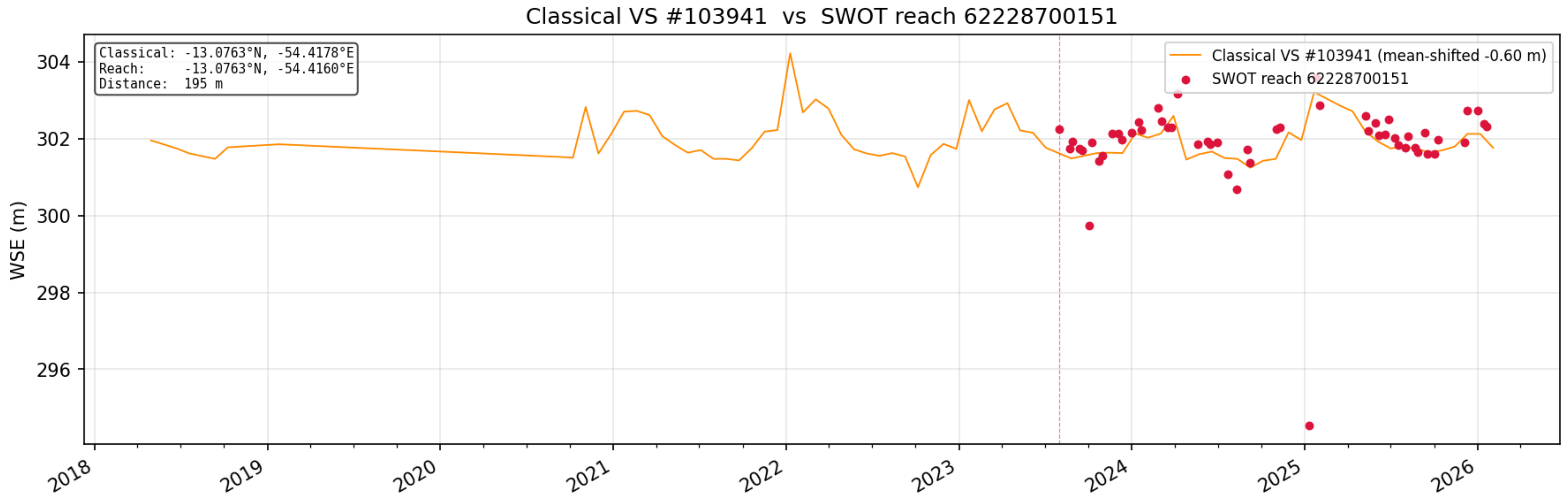


https://www.earthdata.nasa.gov/data/catalog/pocloud-swot-l2-hr-riversp-2.0-2.0

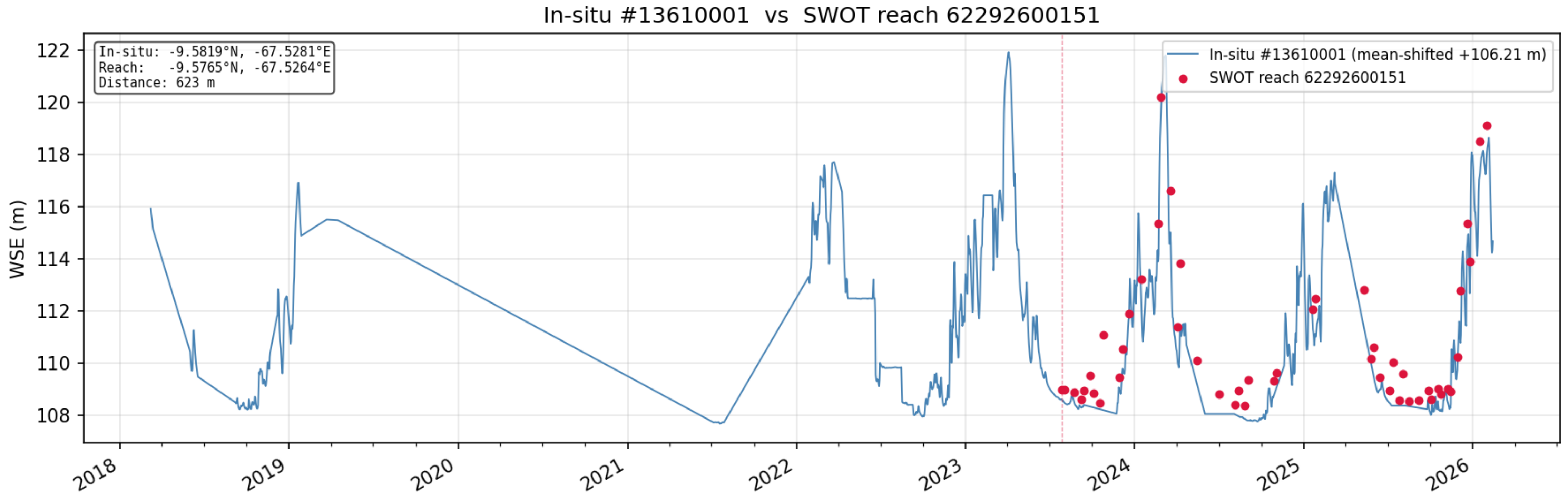
L2 HR RiverSP – Reach version D

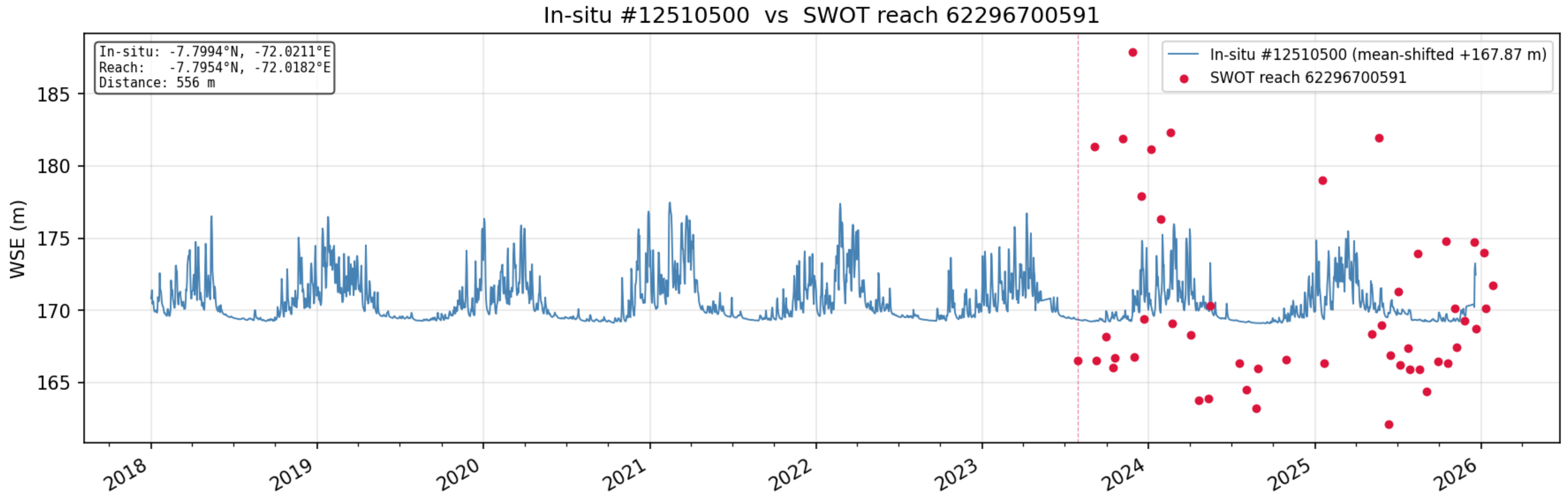
SWOT vs Classical altimetry





SWOT vs In-situ gauges



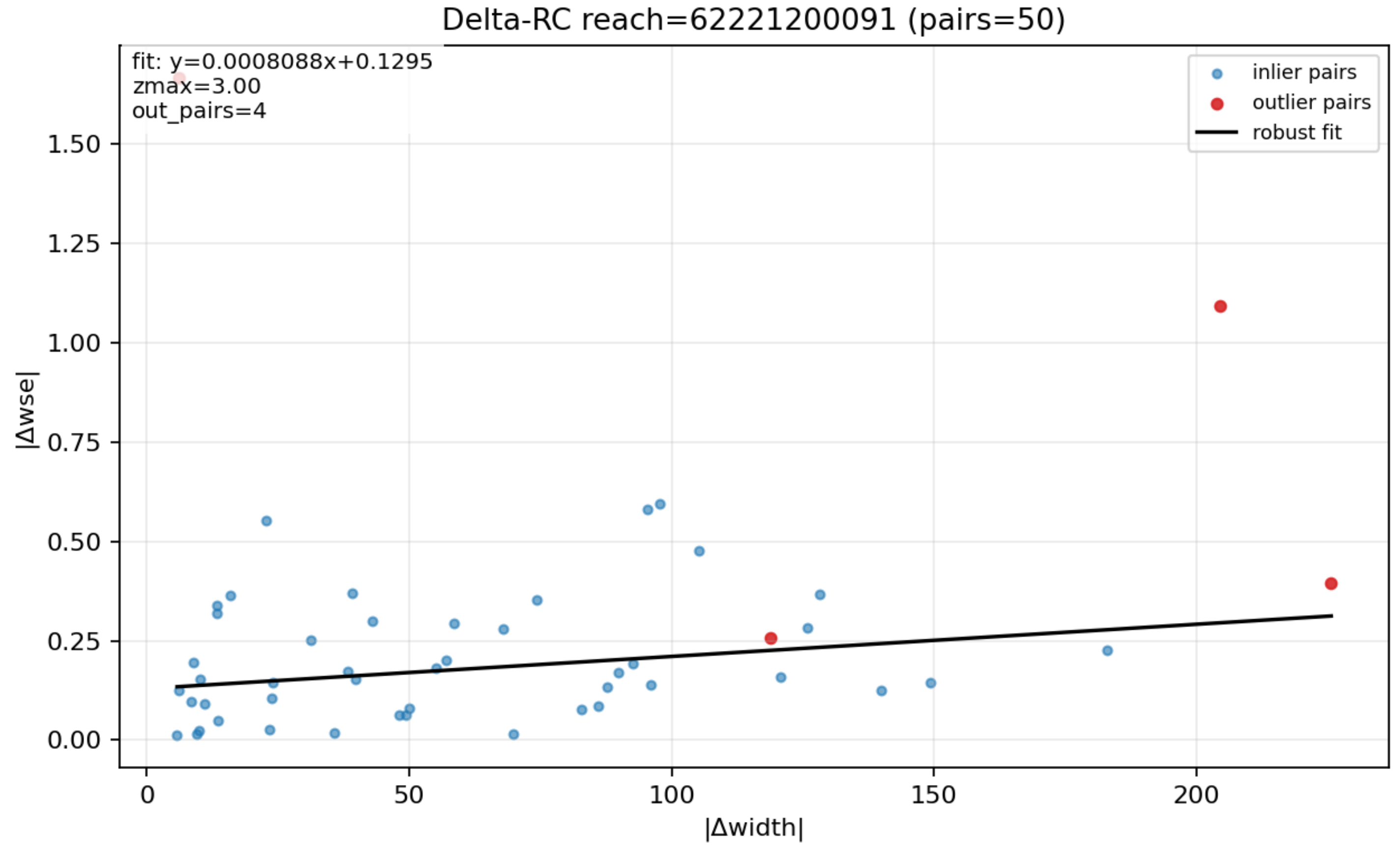


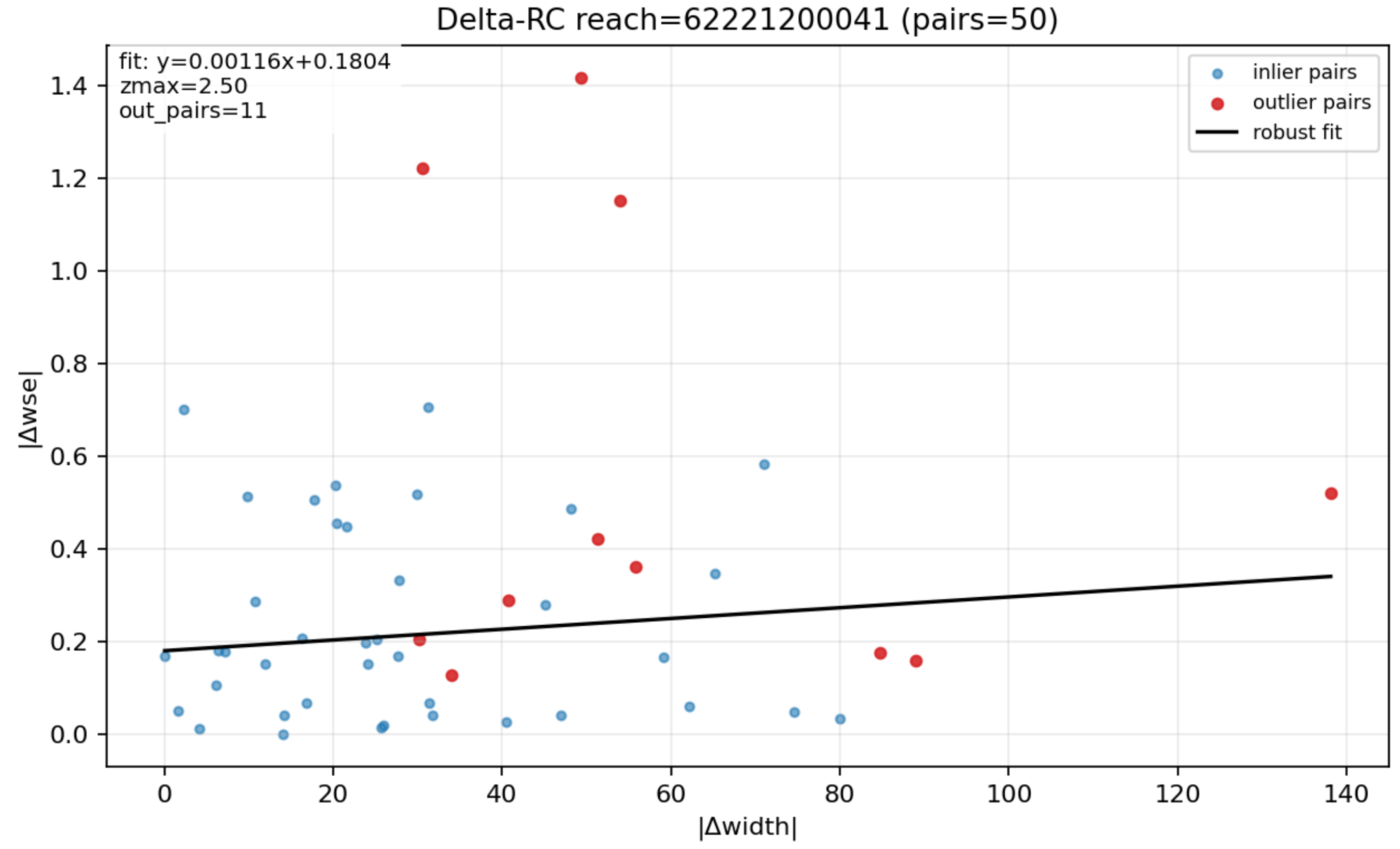
https://www.earthdata.nasa.gov/data/catalog/pocloud-swot-l2-hr-riversp-d-d

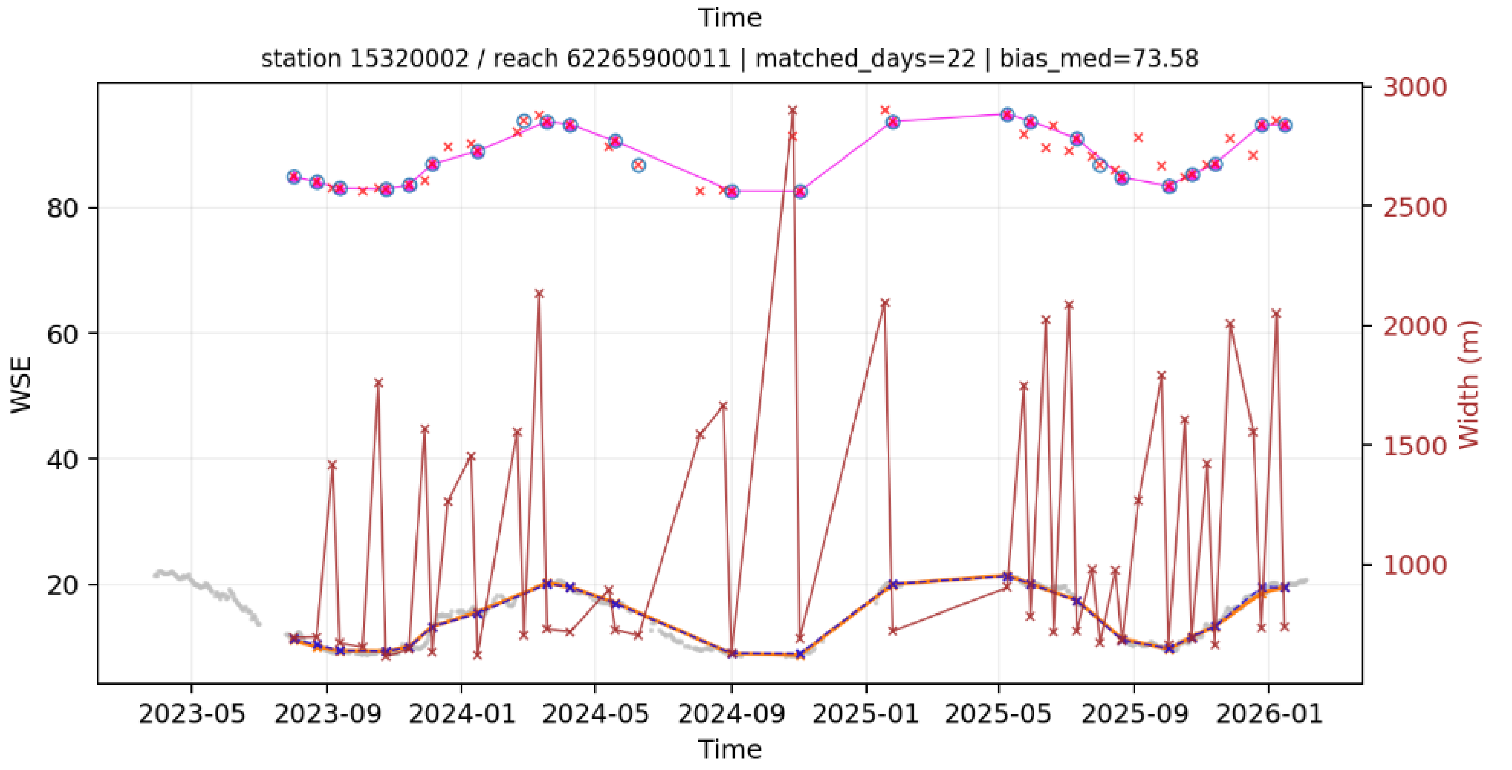
SWOT quality filtering

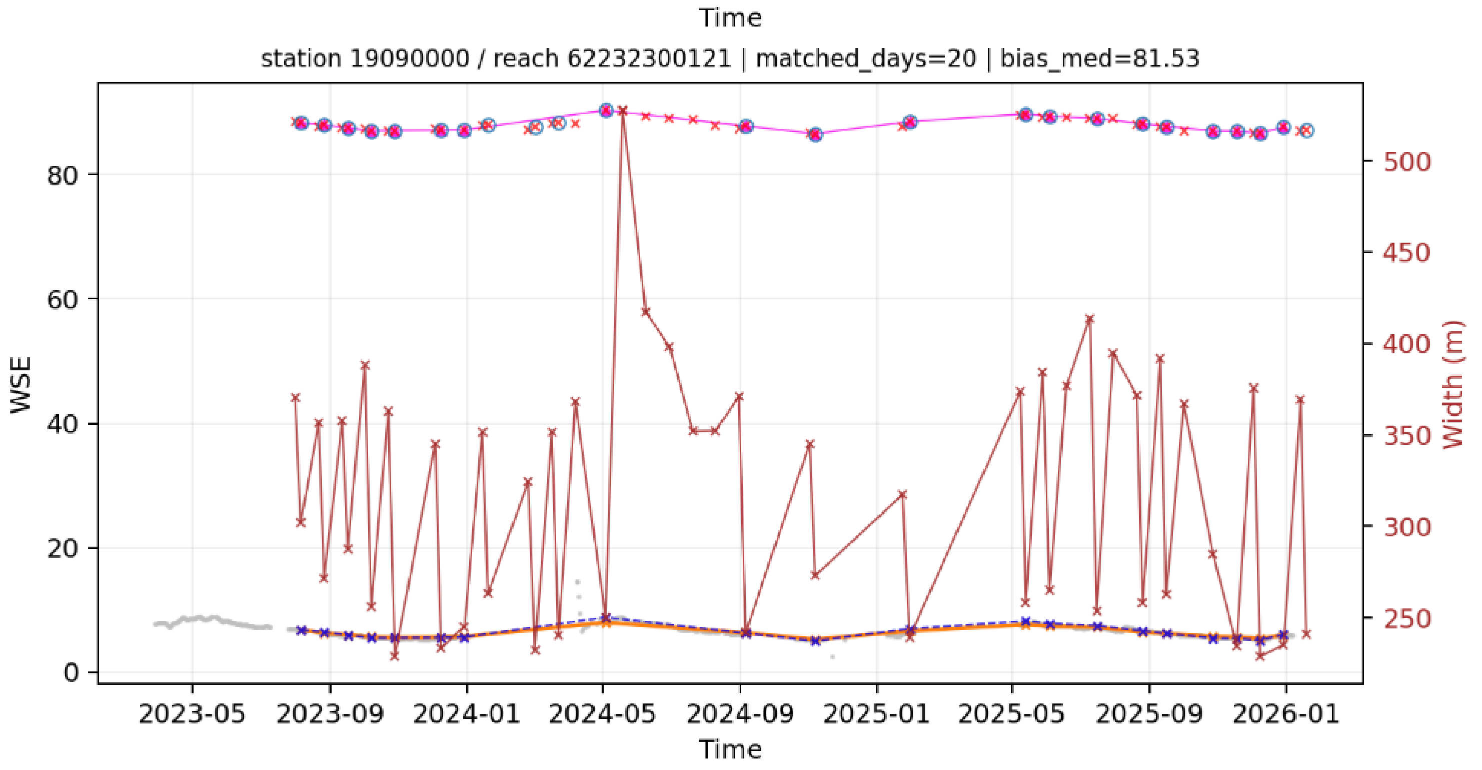
- **Filters inspired by (Andreadis et al., 2025)**
  - Estimated observed precision derived from dark_fraction, minimum width, reach/measurement quality flag, random component of uncertainty, cross-over calibration flag, distance to nadir, reach topology

- **Rating curve filter: width vs. WSE**





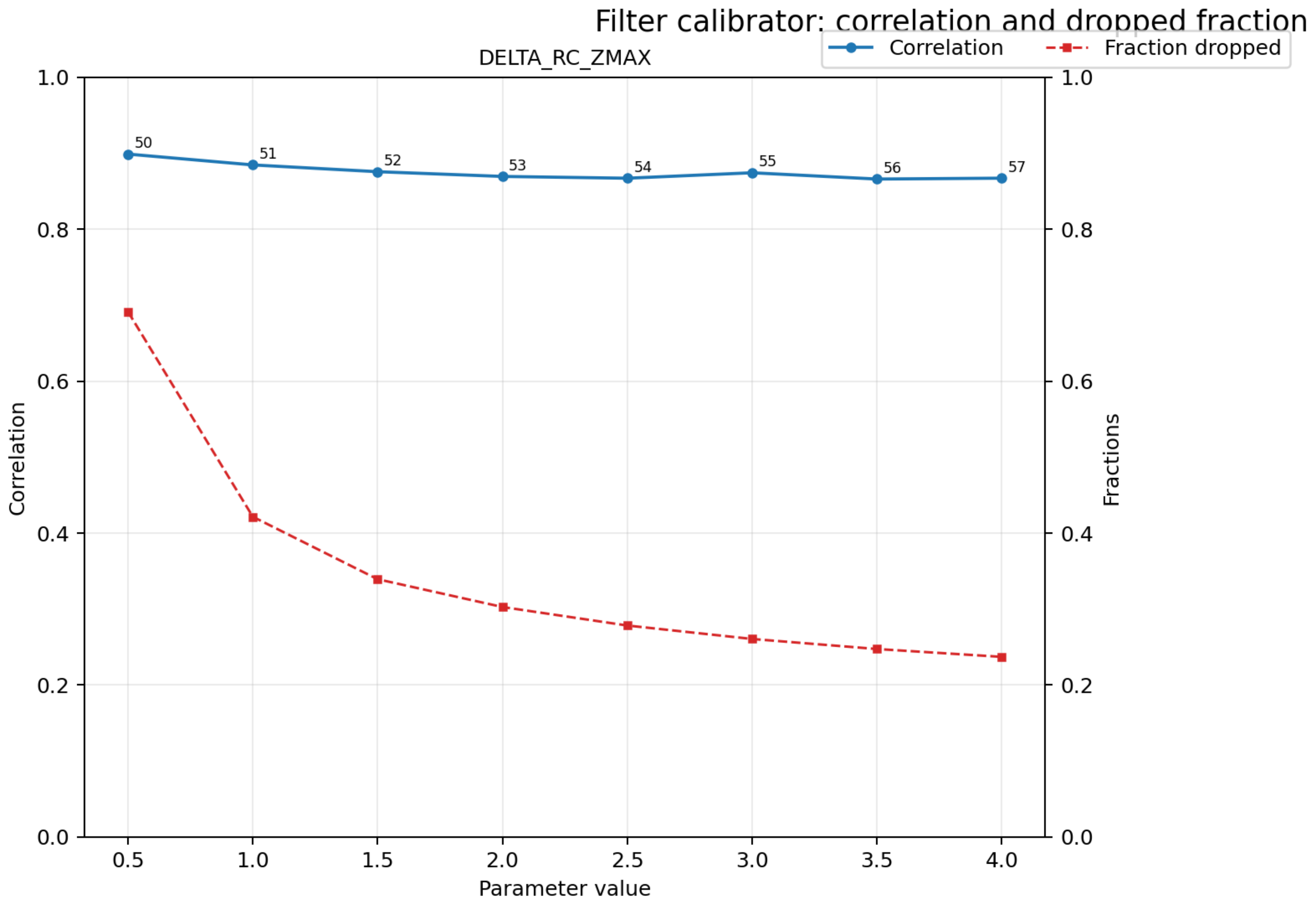


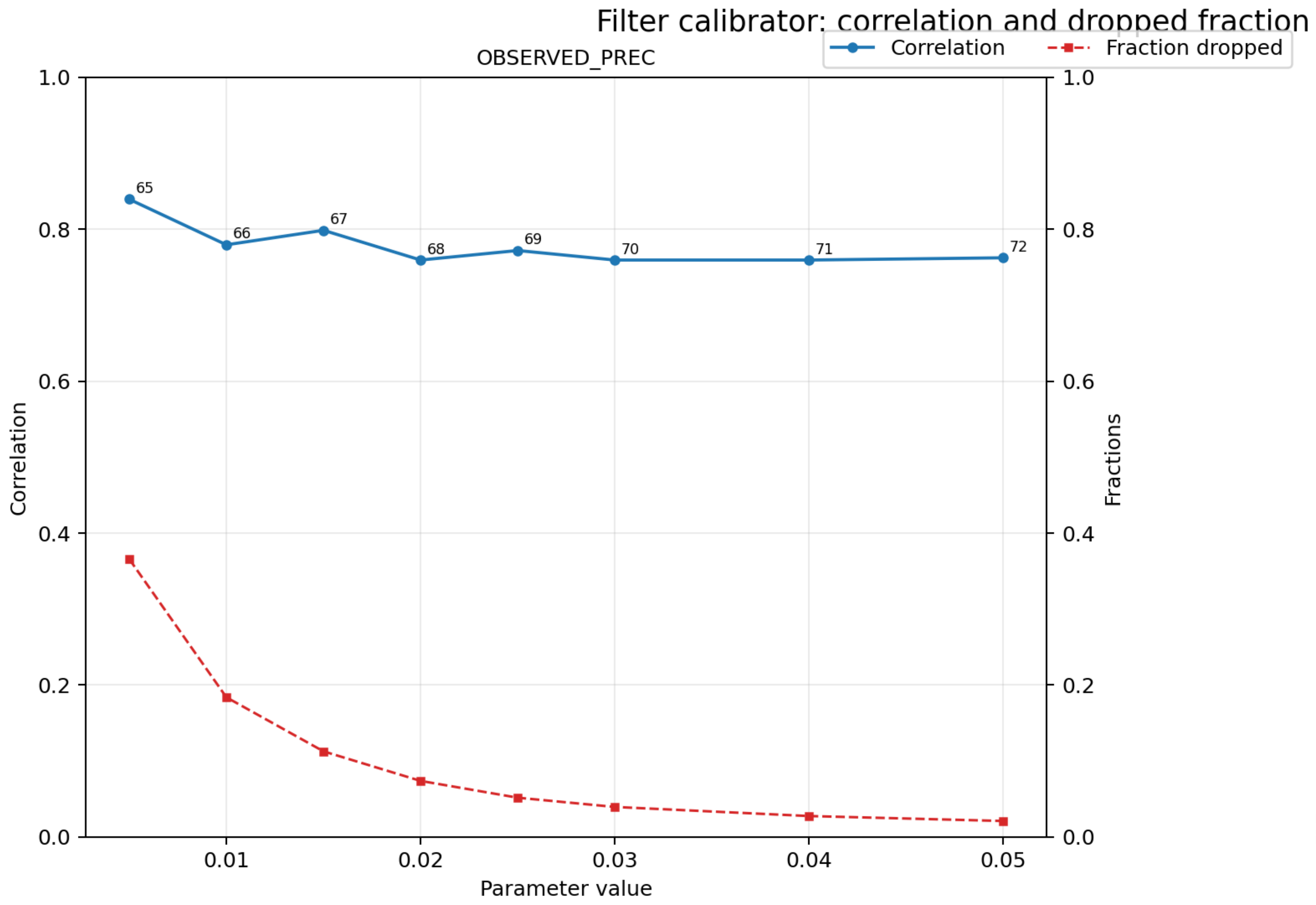


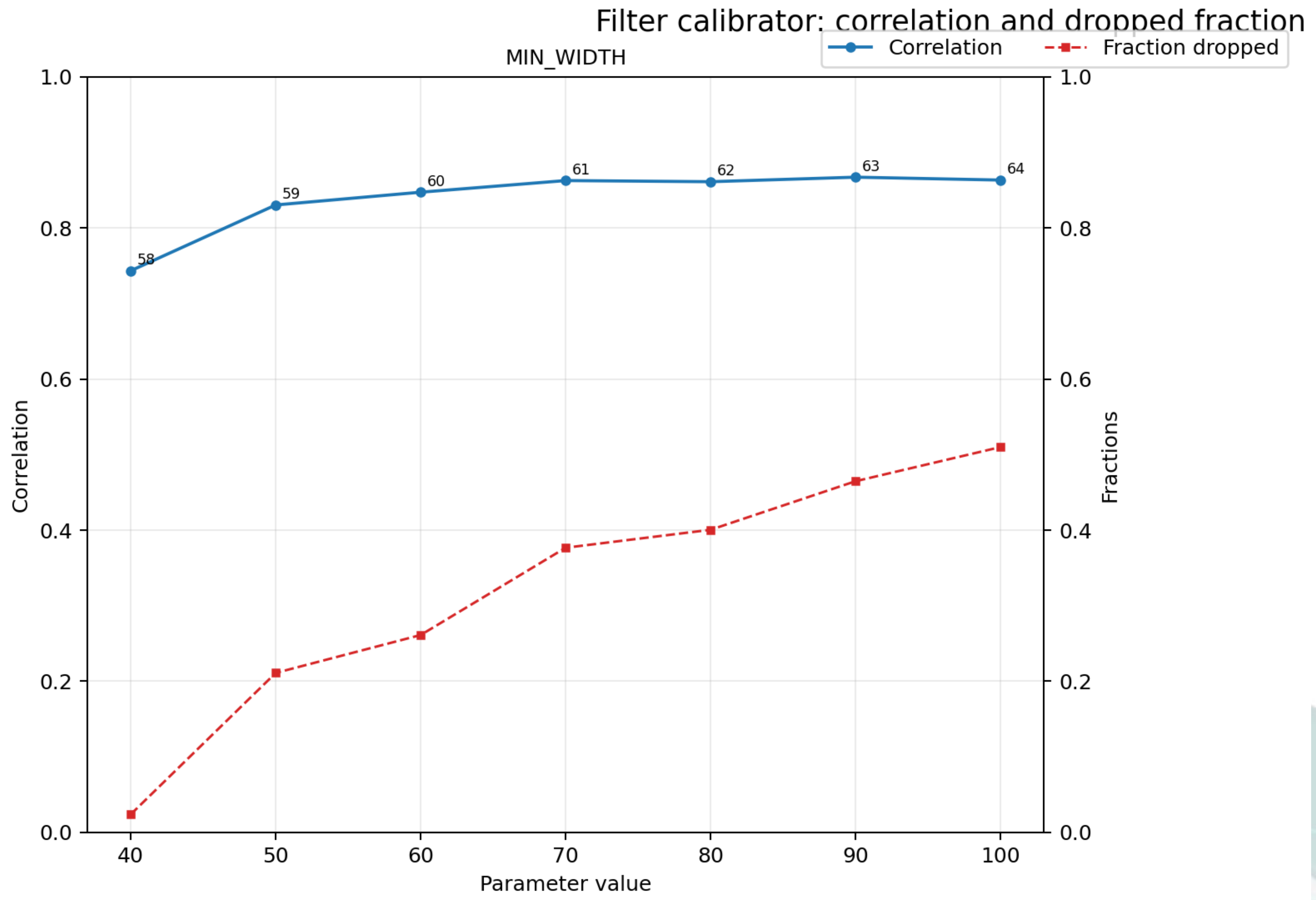
Andreadis, K. M., Coss, S. P., Durand, M., Gleason, C. J., Simmons, T. T., Tebaldi, N., ... & Yadav, B. (2025). A first look at river discharge estimation from SWOT satellite observations. Geophysical Research Letters, 52(9), e2024GL114185.

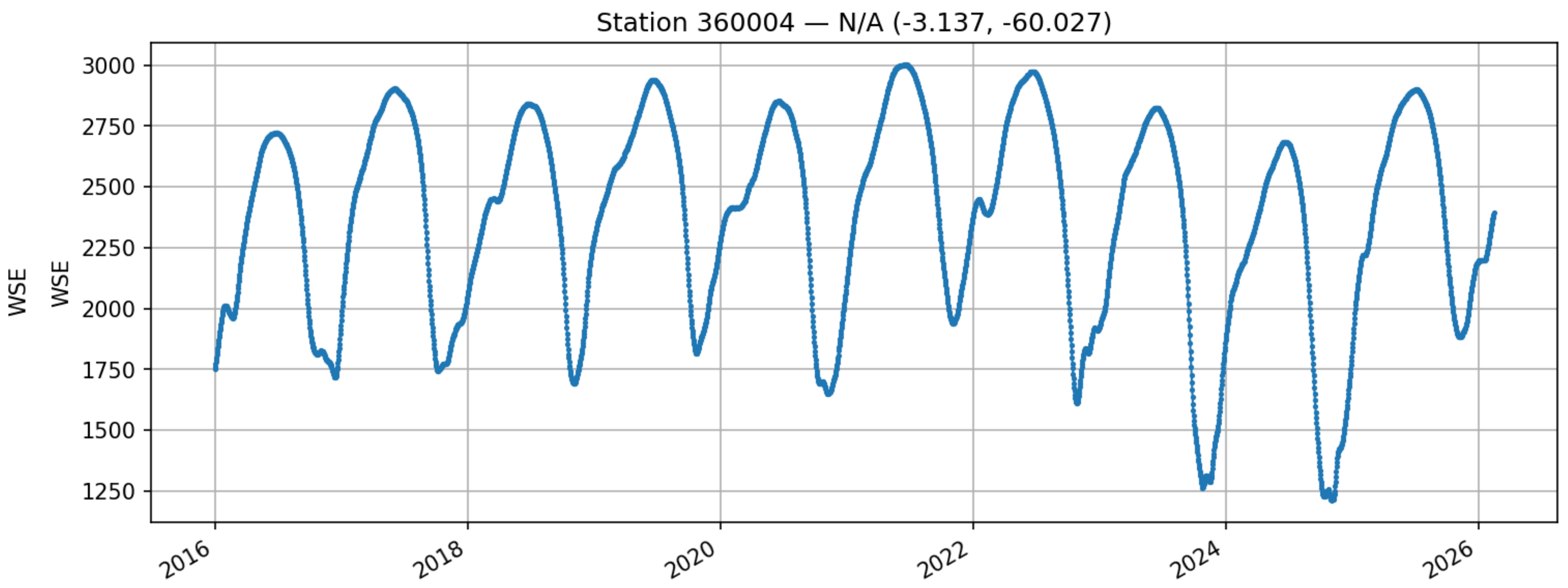
SWOT quality filtering

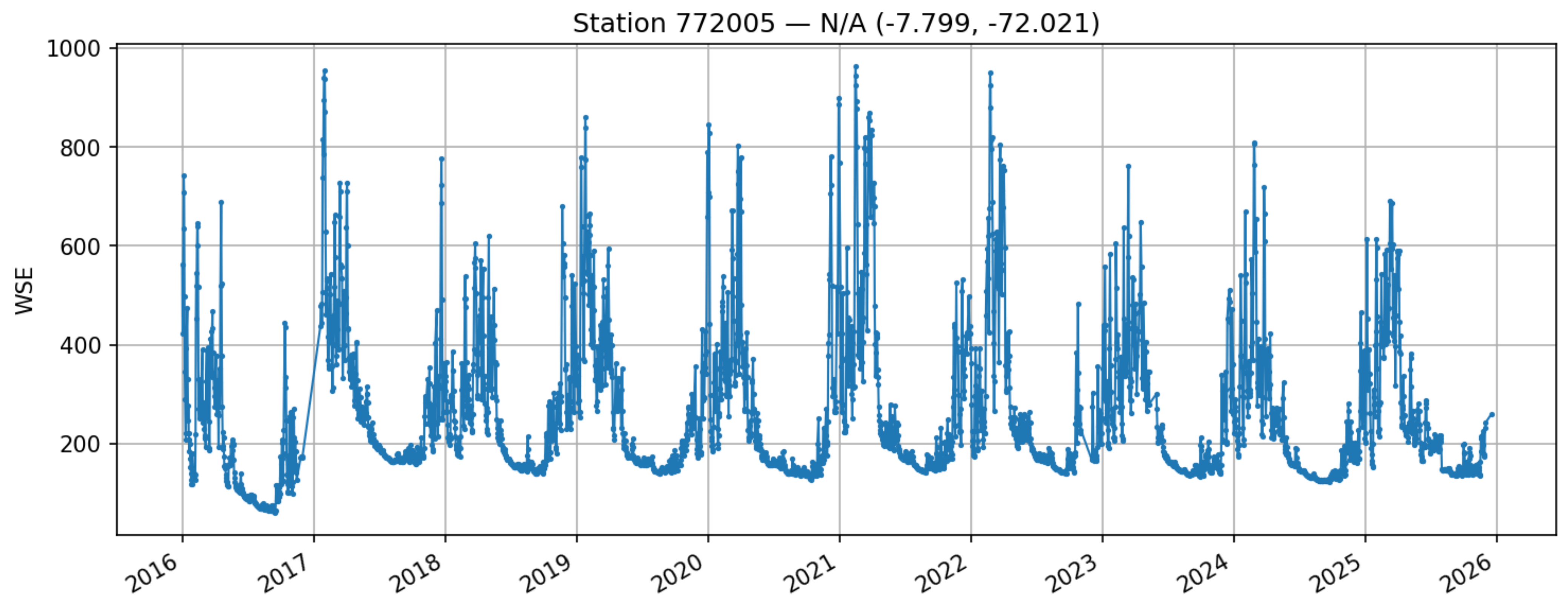
- **Setting filter threshold values by maximizing correlation with insitu gauge timeseries**



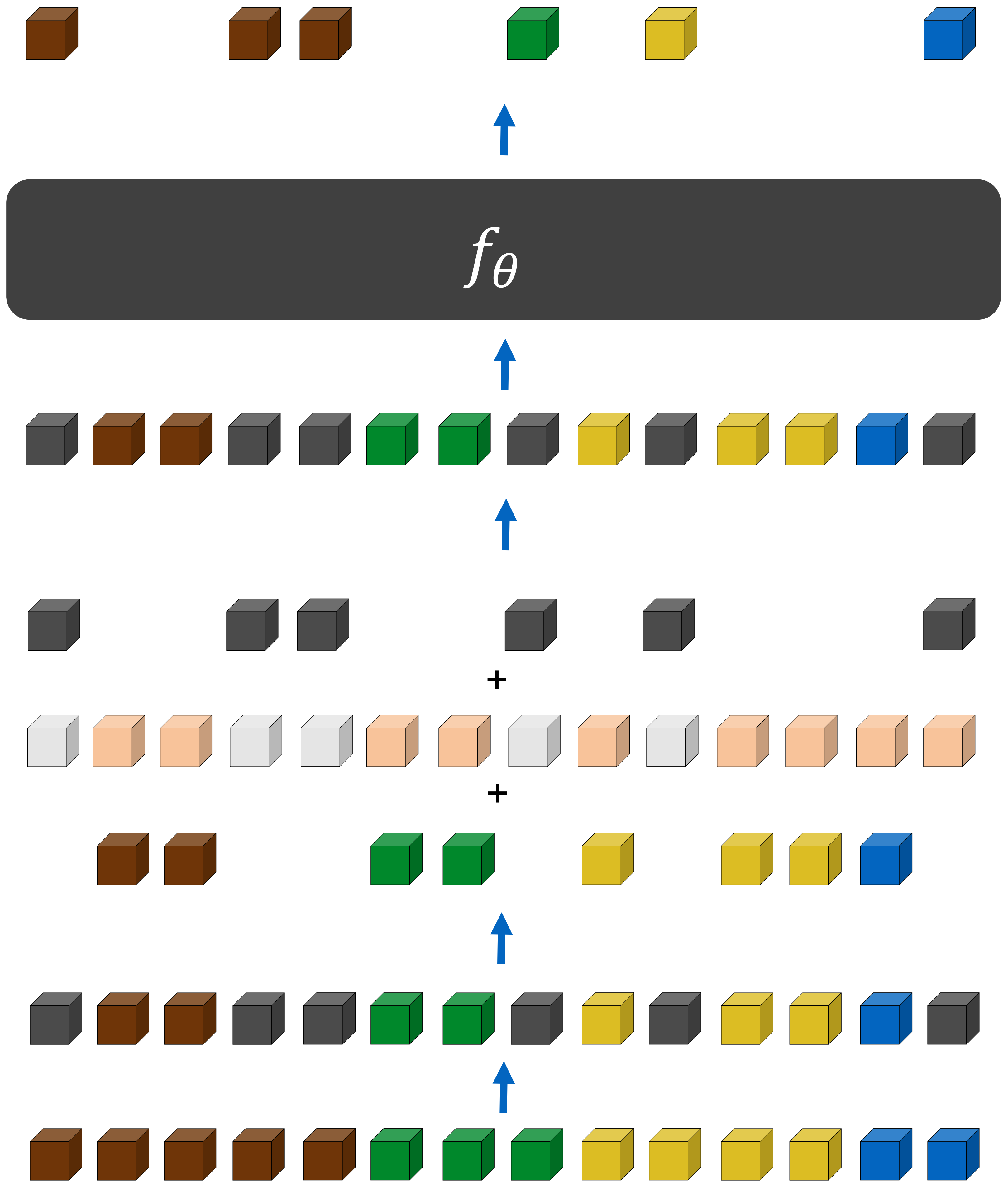








Model & training



1-dim

Decoder
- Sees non-masked and masked inputs
- Projects masked inputs to predicted WSE

192-dim

Add query/placeholder vectors

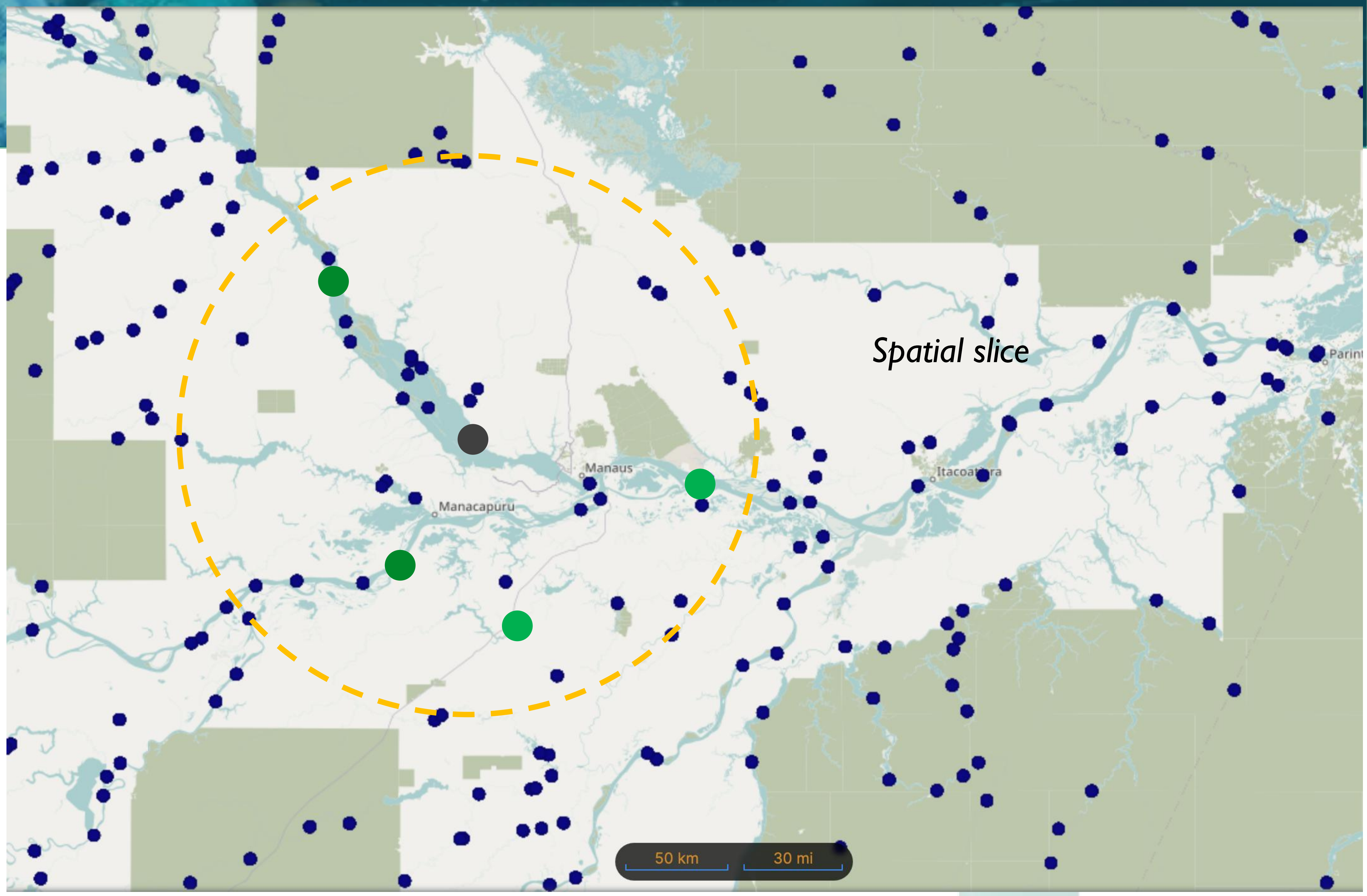
Add metadata: lat/lon, timestamp, position on river graph, (sub)basin

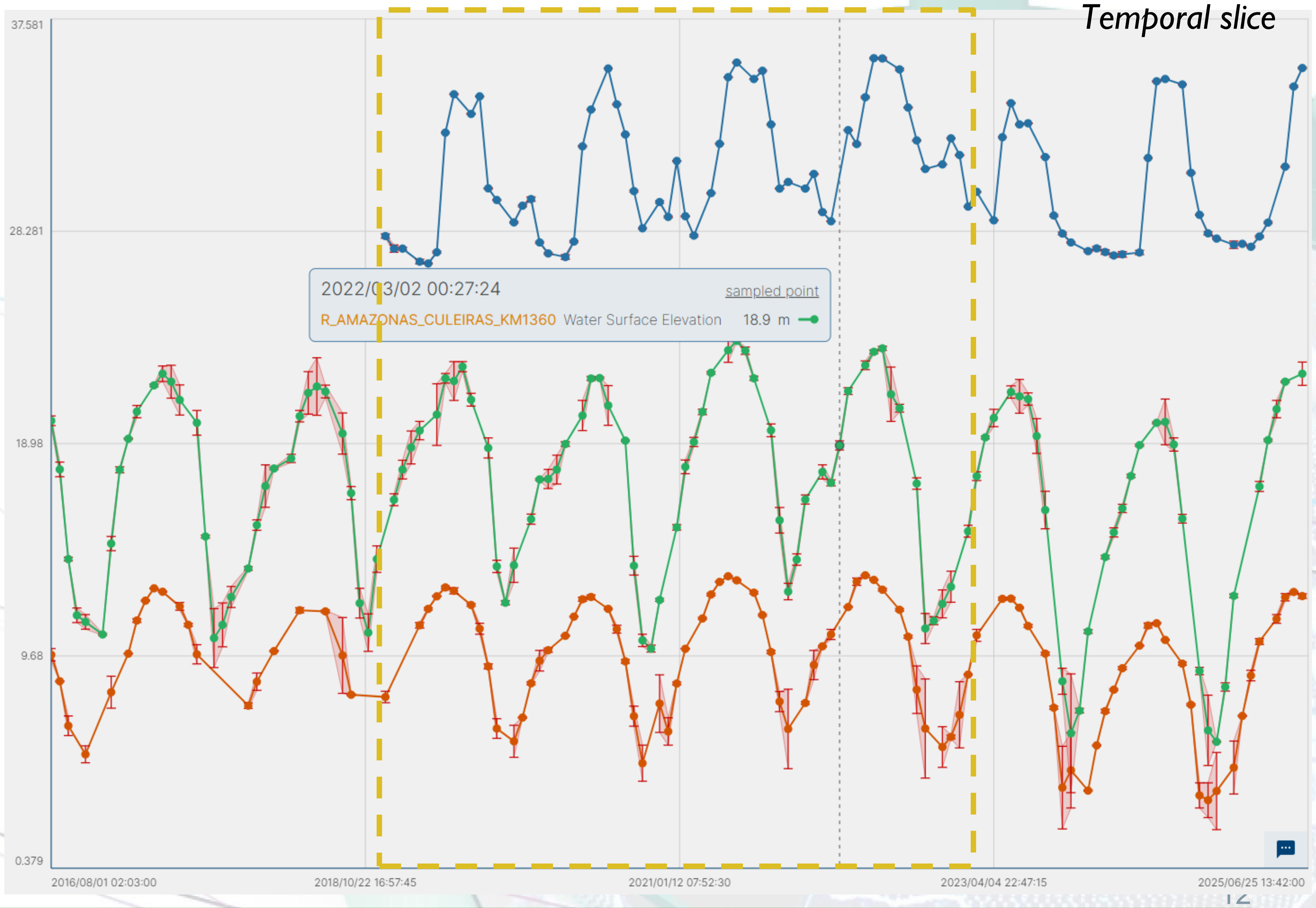
192-dim

Apply mask

1-dim

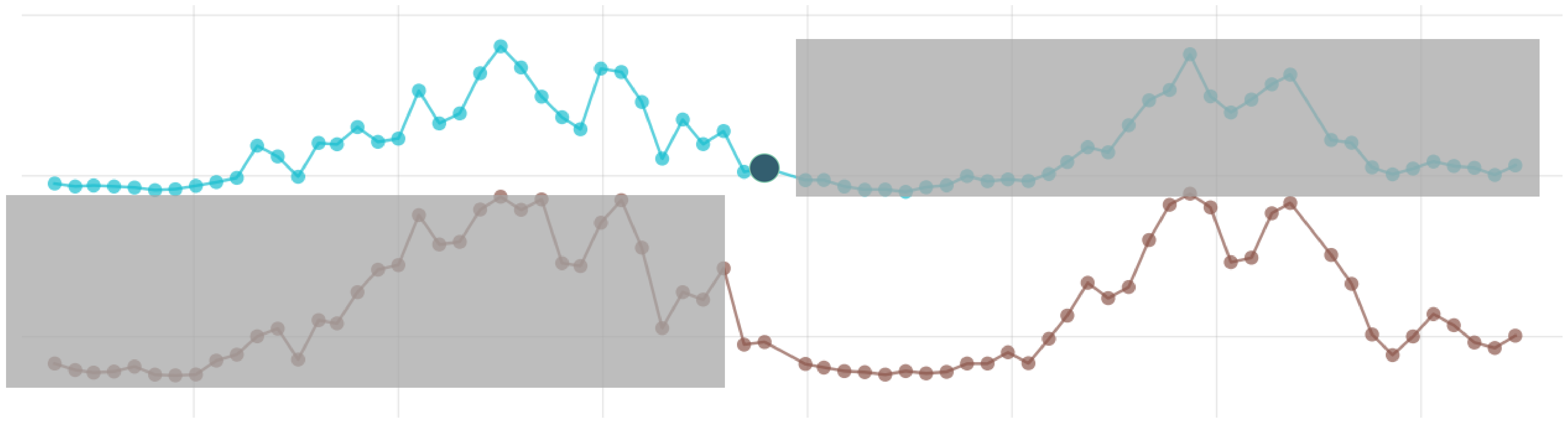
Sparse water level timeseries





Model & samples

- Model can see [upstream and prior in time] and optionally [downstream and posterior in time]



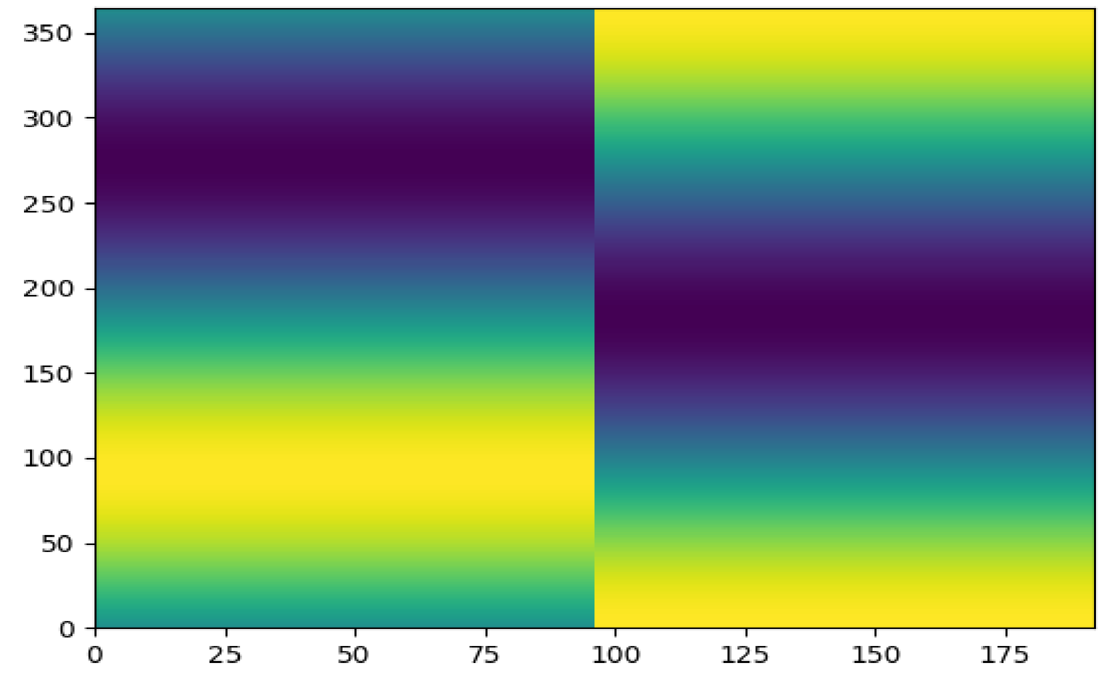
- Parallel transformer / autoregressive transformer
  - XLNet inspired architecture (Yang et al., 2019)
  - Following [prior → posterior in time]
  - and [upstream → downstream] order

- (Bidirectional) LSTM

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. Advances in neural information processing systems, 32.

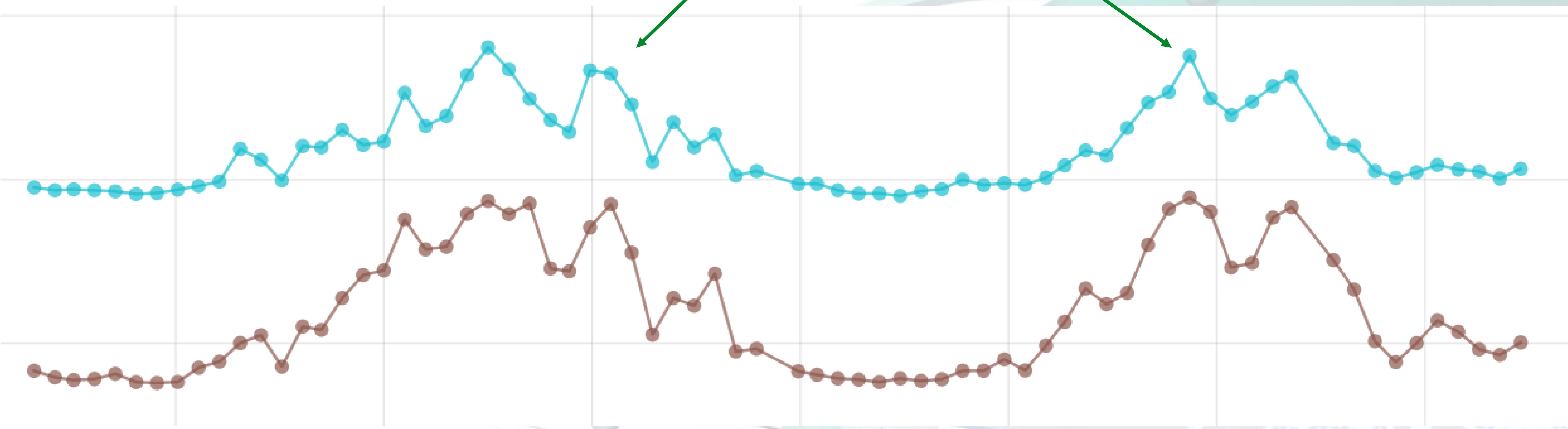
Metadata encoding: temporal

- **Month** of the year: 12-cyclical sine/cosine



- **Relative day differences**: sine/cosine (Vaswani et al., 2017)
  - Measurement with lowest date: 0
  - Measurement with N-th lowest date: N

D(t_1, t_2) = +80 days



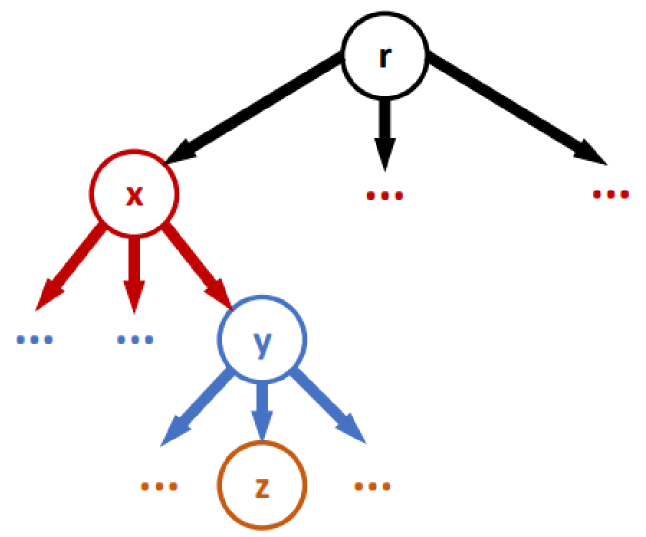
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

Metadata encoding: spatial

- **Latitude / longitude**: linear project
- W.r.t. to river outlet + w.r.t. sampled anchor

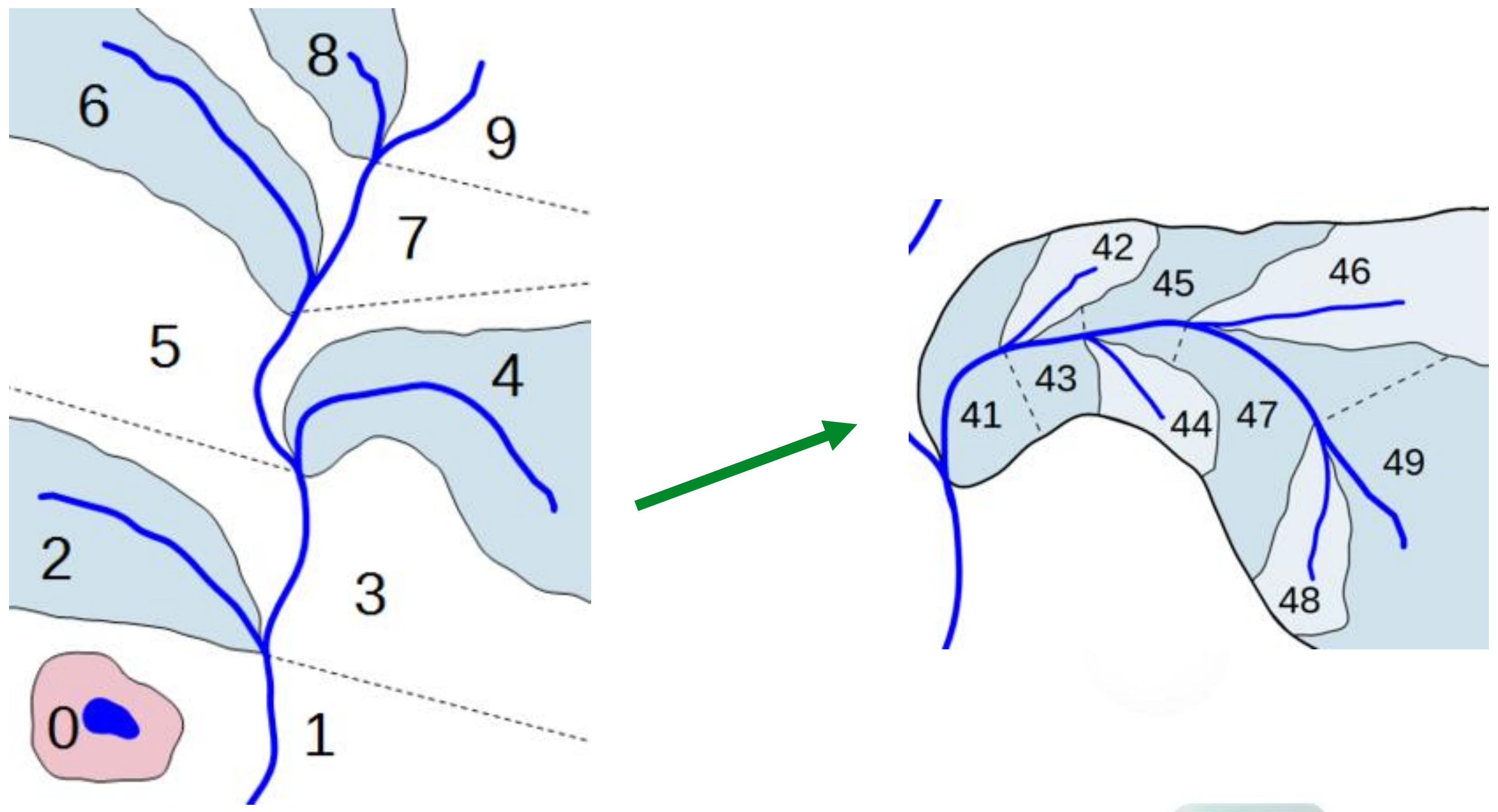
- **Sub-basin from SWORD:** C **BBBBB** RRRR T
- Using (Shiv & Quirk, 2019): encode path from root to node

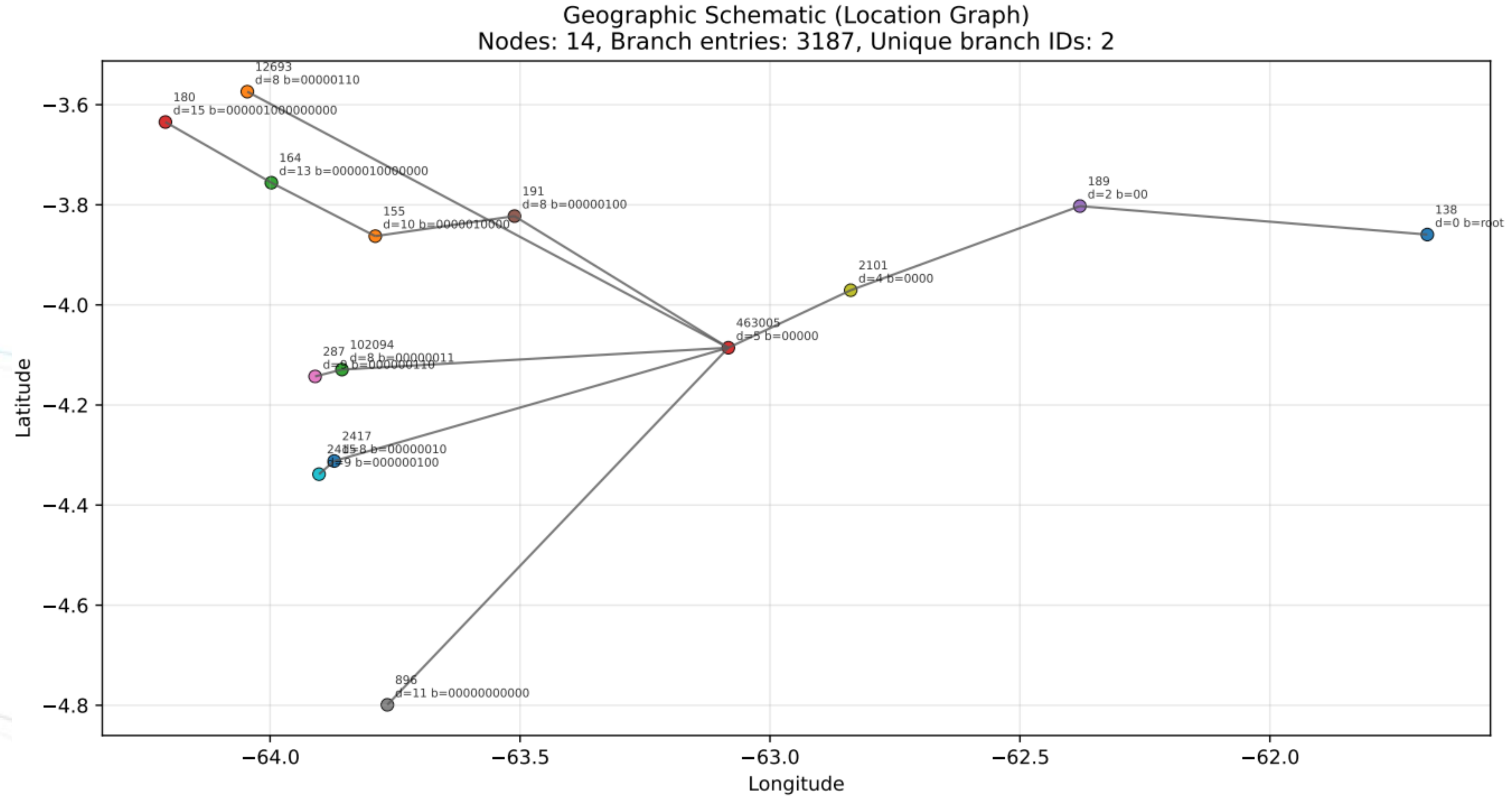
$$
\begin{aligned}
r &= & \mathbf{0} &= [0,0,0,0,0,0,0,0,0]\\
x &= & D_0\mathbf{0} &= [1,0,0,0,0,0,0,0,0]\\
y &= & D_2D_0\mathbf{0} &= [0,0,1,1,0,0,0,0,0]\\
z &= & D_1D_2D_0\mathbf{0} &= [0,1,0,0,0,1,1,0,0]
\end{aligned}
$$



- **Path from anchor to sampled locations:** C BBBBB **RRRR** T
- Through river topology taken from SWORD

- **Mean WSE** → proxy for elevation / topography

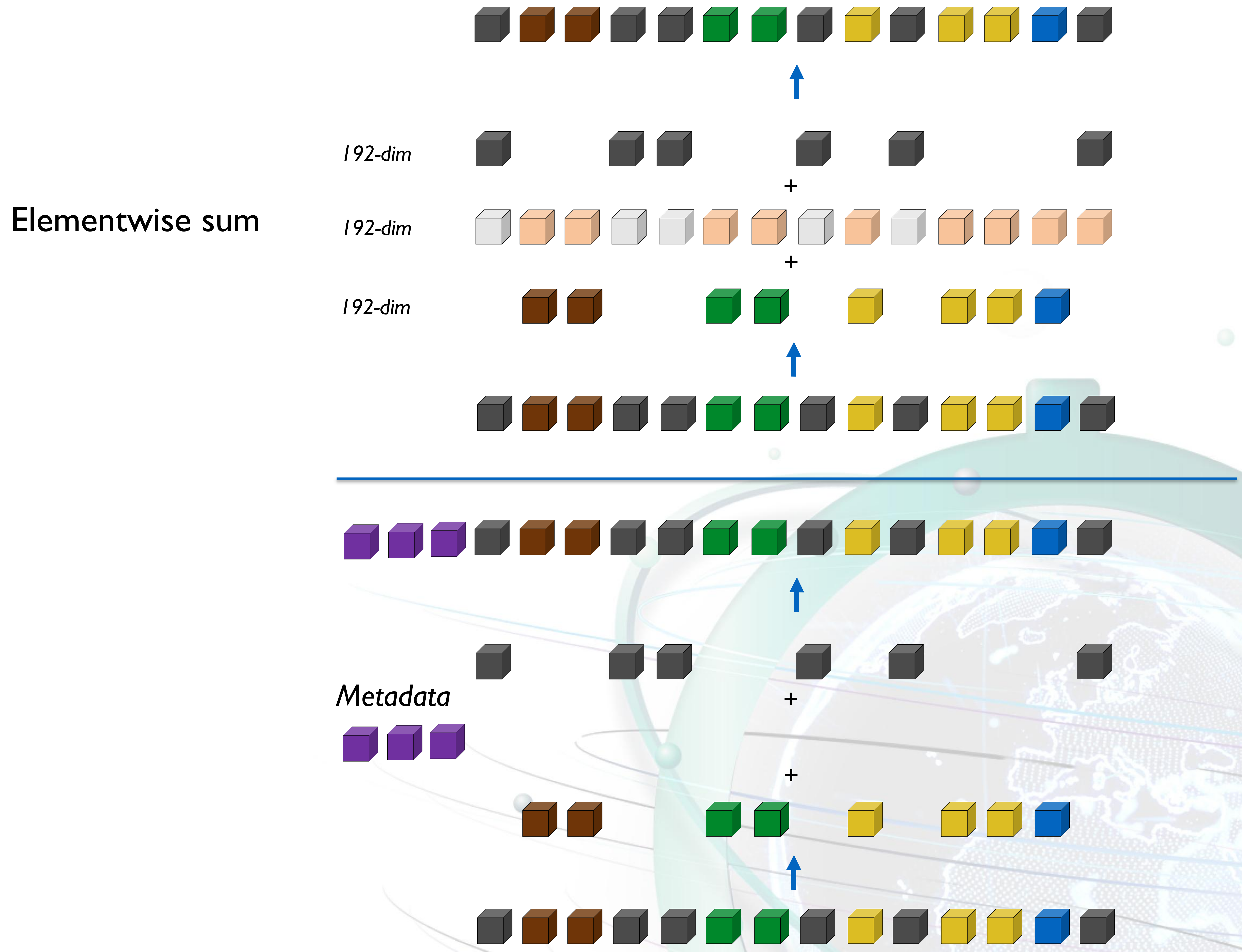




Shiv, Vighnesh, and Chris Quirk. "Novel positional encodings to enable tree-based transformers." *Advances in neural information processing systems* 32 (2019).

Metadata encoding

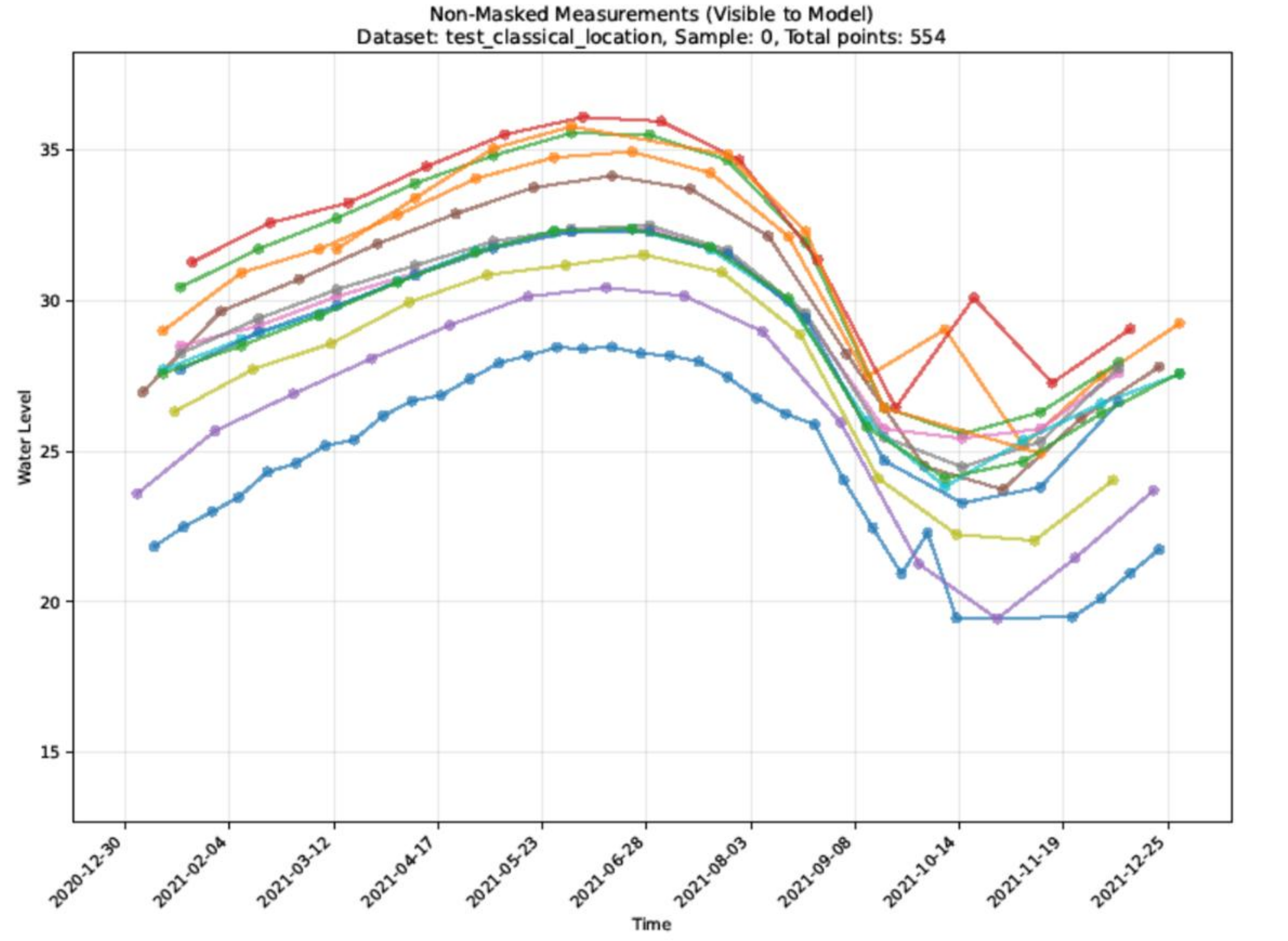
- As positional encoding embeddings
- As metadata tokens
- Ablation experiments: combination works best

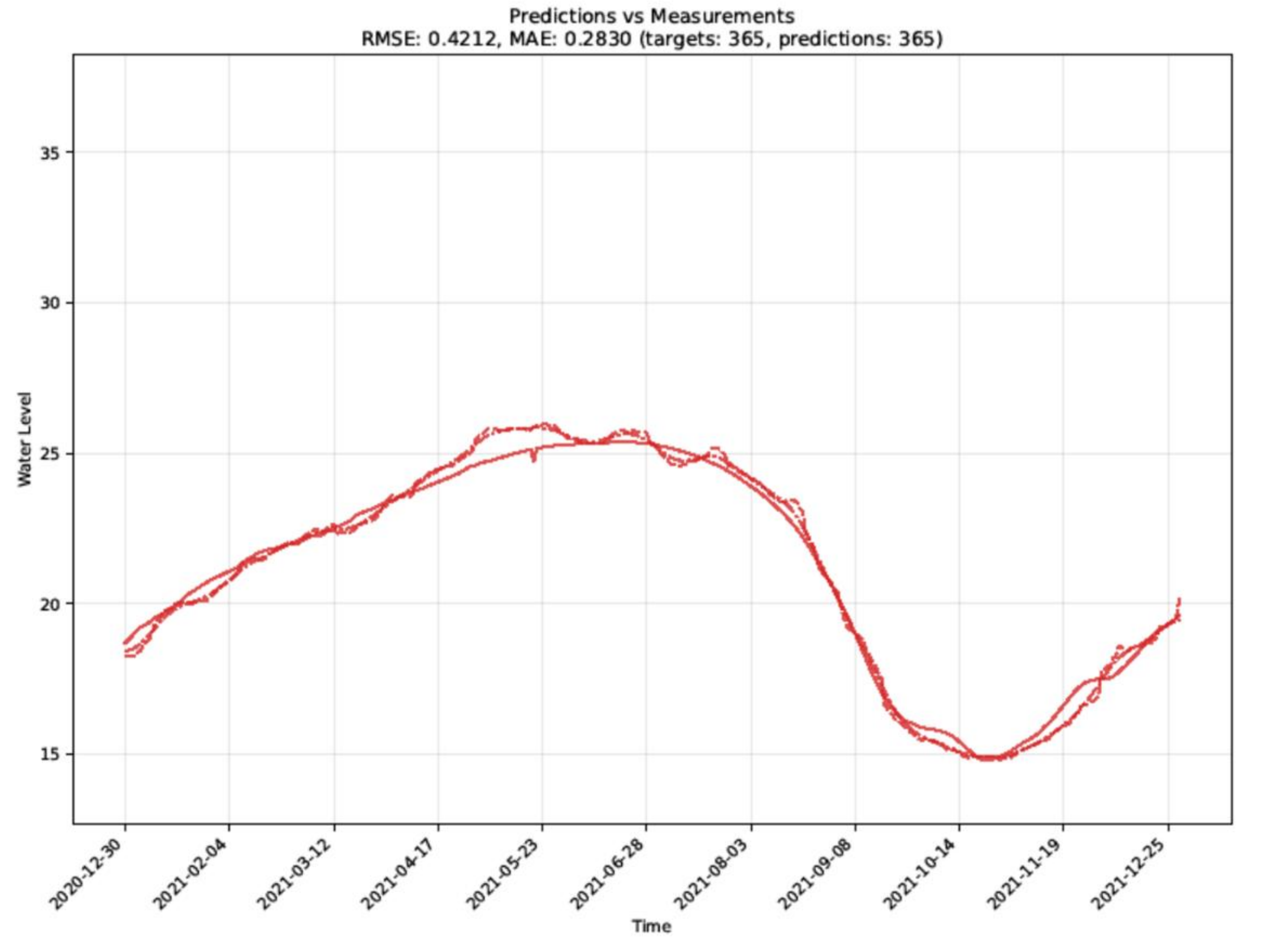


Preliminary results

- Bi-LSTM:
  - Nash-Sutcliffe model Efficiency (NSE) [-inf, 1.0]: 0.92
  - RMSE [0.0, inf]: 0.91m

- XLNet transformer:
  - NSE: 0.83
  - RMSE: 1.34m

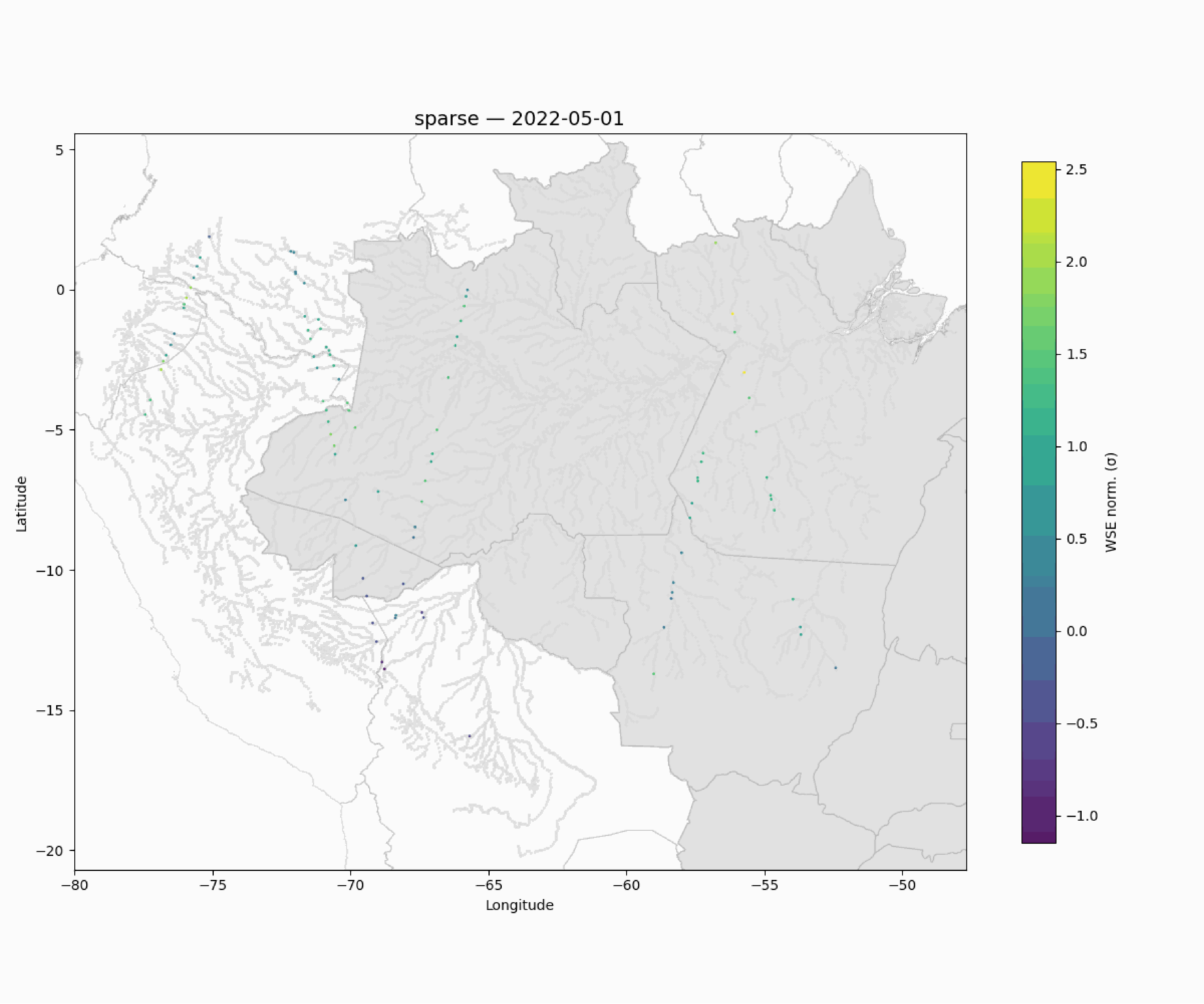




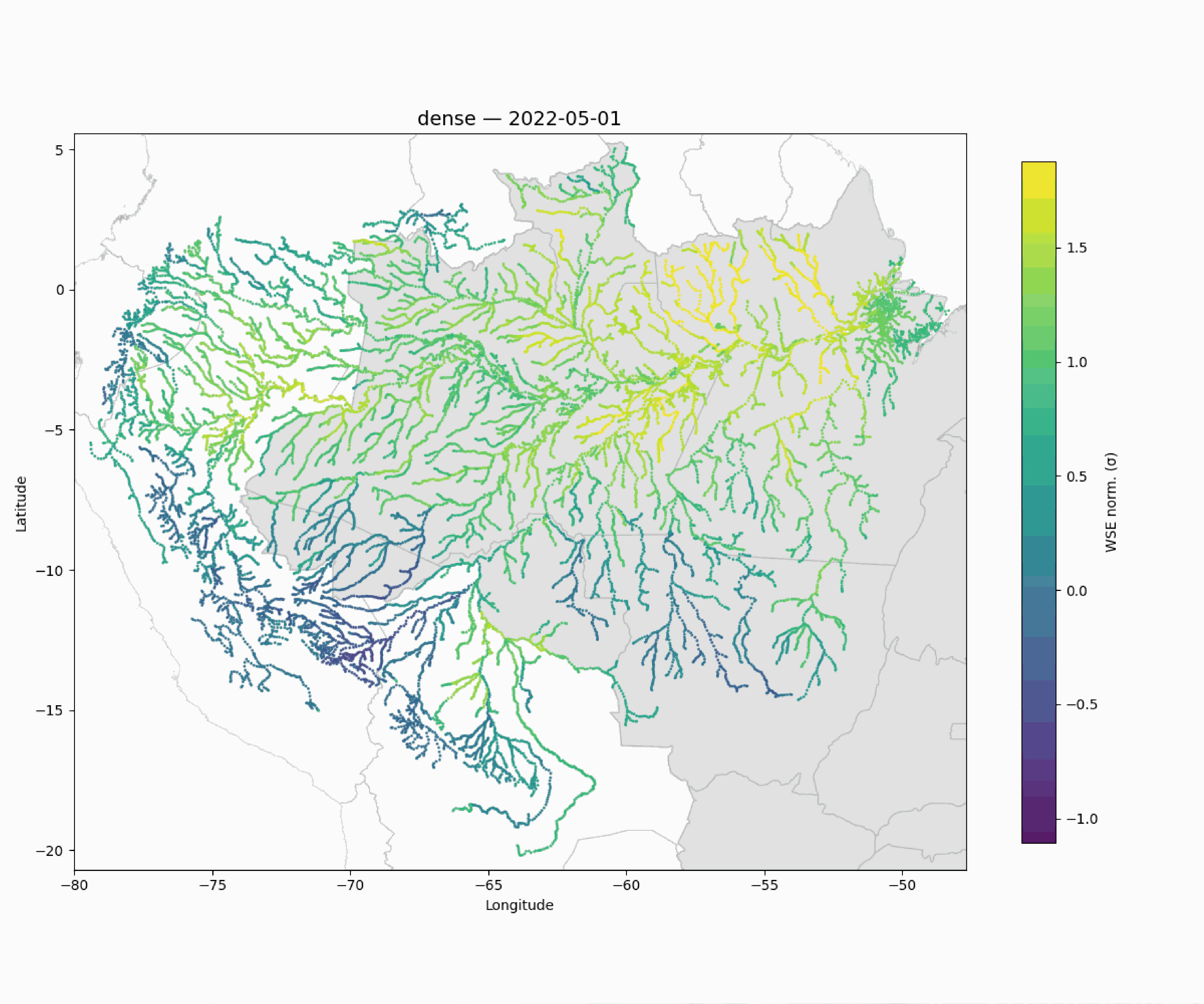
Preliminary results

- 2022-05-01 – 2022-07-31

Classical altimetry



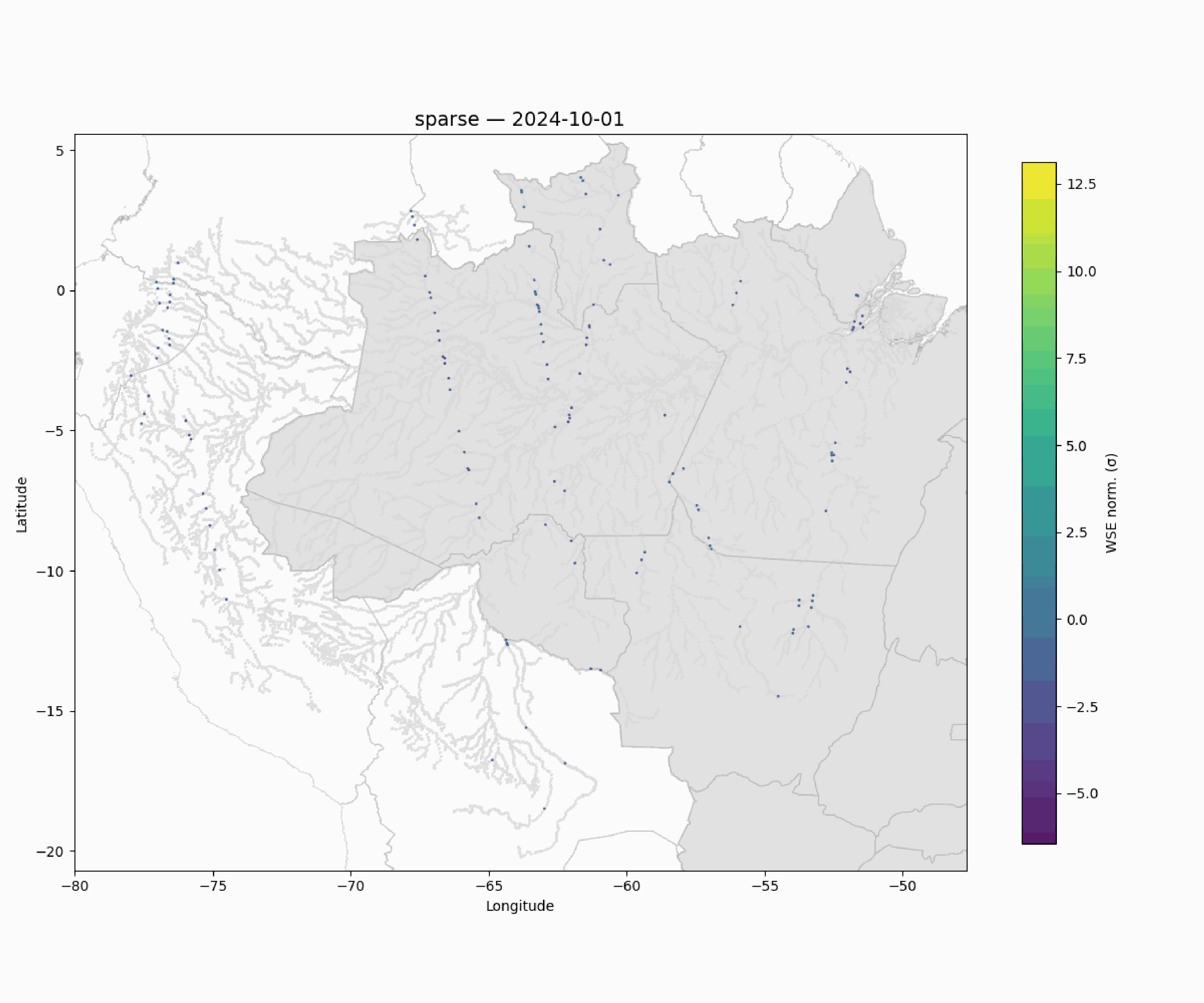
Densified version



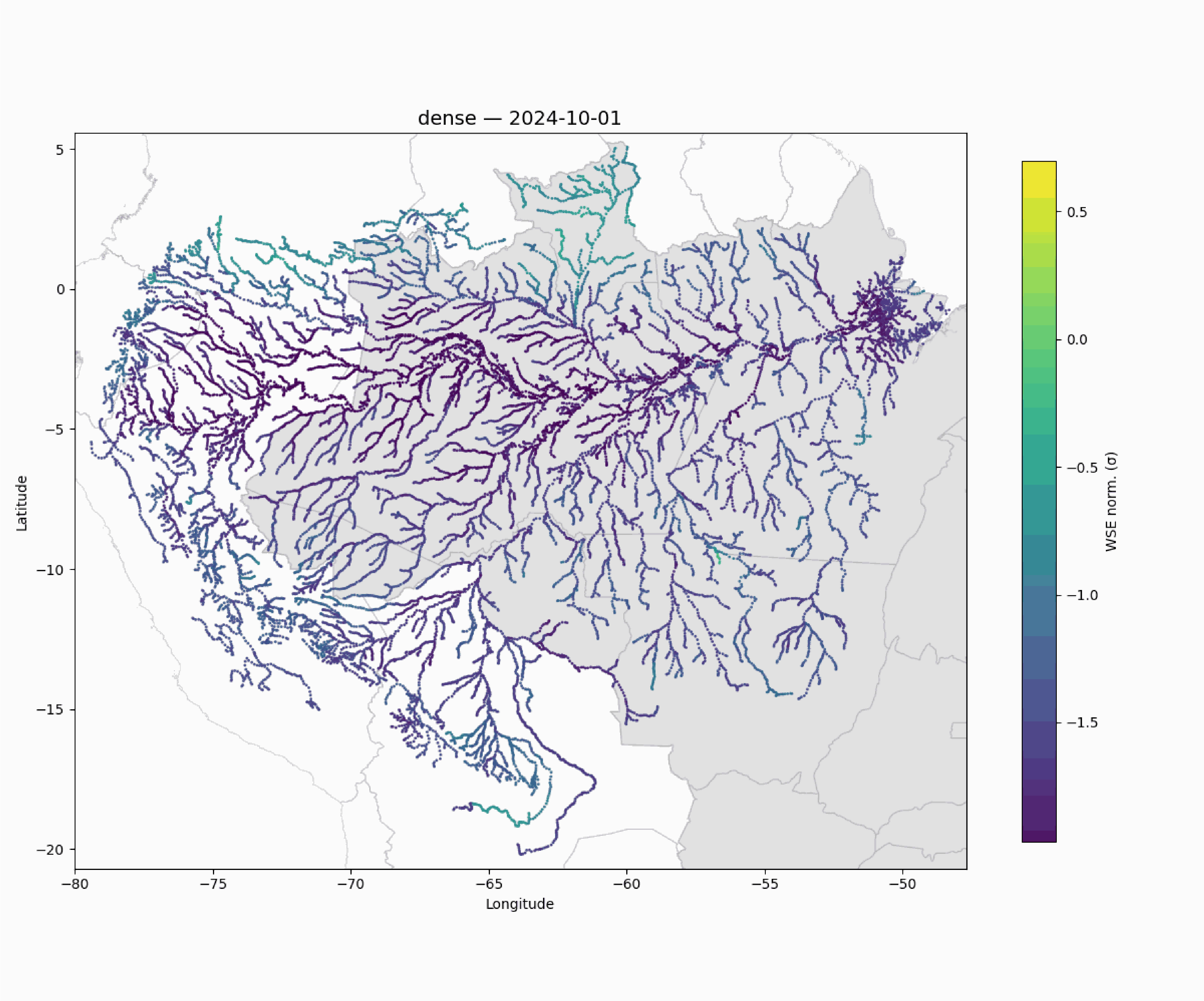
Preliminary results

- 2024-10-01 – 2024-12-31

Classical altimetry + SWOT

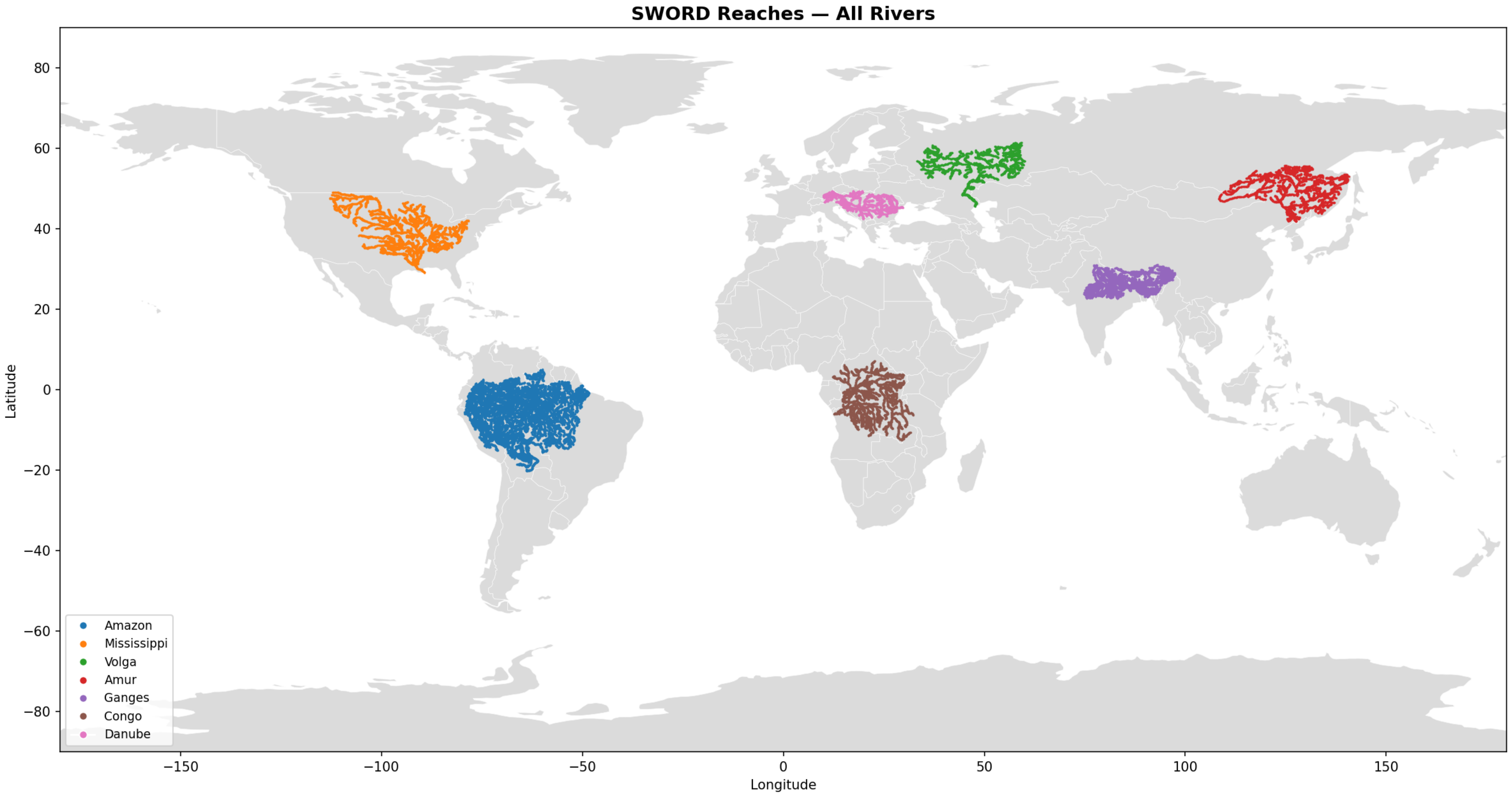


Densified version



What's next

- Other major basins
  - Amazon, Mississippi, Volga, Amur, Ganges, Congo, Danube
  - 45K SWOT reaches, 14K classical altimetry virtual stations



What's next

- Incorporate other variables:
  - River geometry: width
  - Water cycle: precipitation

- Compare to statistical / physical model baselines
  - Don't take into account river branches
    - Halicki, M., Niedzielski, T., Schwatke, C., Scherer, D., & Dettmering, D. (2026). Daily river water levels from multi-mission altimetry: A reach-based regression method using the unique SWOT data geometry. Available at SSRN 6037735.

Find me ☺

For suggestions, questions, to collaborate, …

- ruben.cartuyvels@esa.int