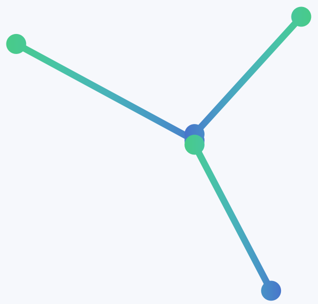
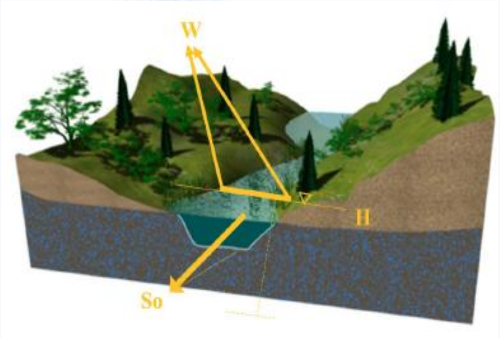
Vector-Based River Hydrography and Enhanced Attributes

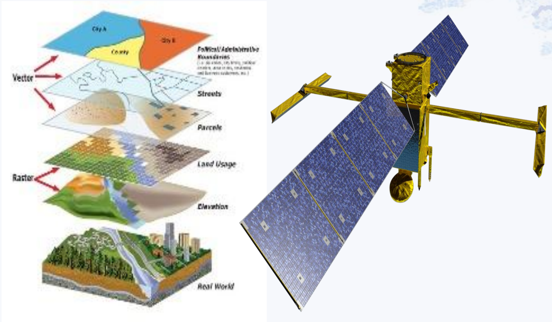
北京大學 PEKING UNIVERSITY

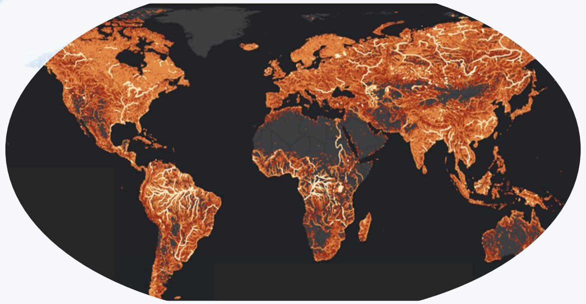
Peirong Lin, Ph.D.
Assistant Professor, Peking University

Collaborators: Dai Yamazaki, Ming Pan, Yuan Yang, Cedric David, etc.
GeoWater Lab Members: Zimin Yuan, Haomei Lin, Ziyun Yin, Jie Xu, Fenghe Zhang, etc.







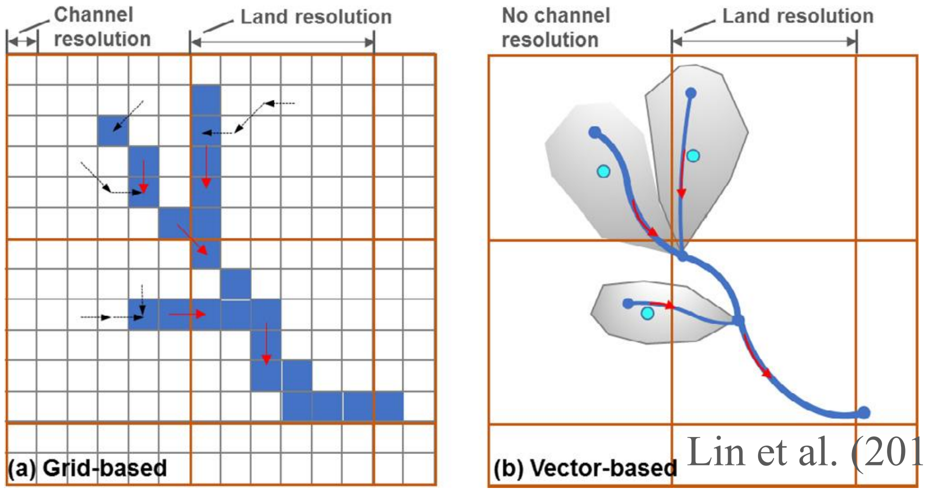


Outline

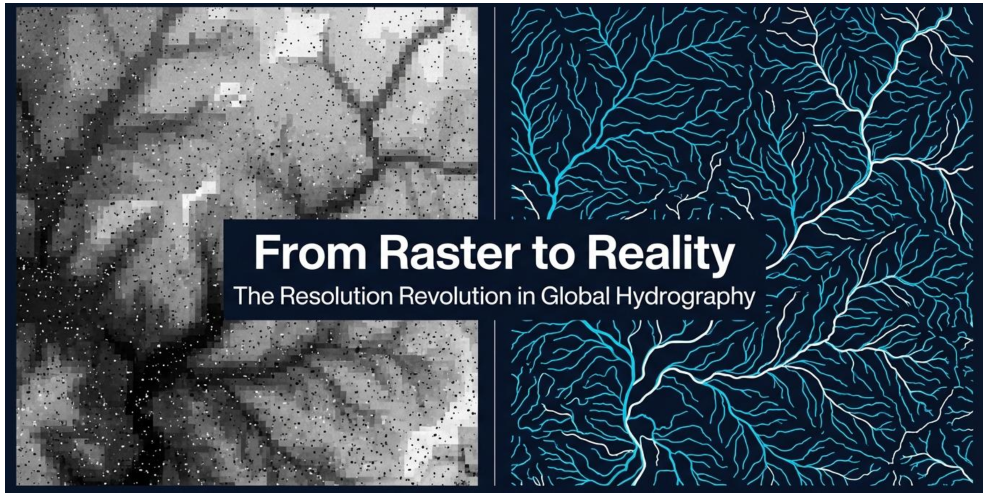
- **Vector hydrography**
  - ✓MERIT-Basins, MERIT-Hydro-Vector
- **Mapping hydrography of different sources**
  - ✓MERIT-Basins to GRWL, HydroSHEDS/ATLAS, SWORD
- **Enhanced attributes**
  - ✓Channel shape parameters, time-evolving attributes

Why vector-based hydrography?

- **Vector hydrography better depicts the real world, vital inputs to large-scale river models**
  - **① realistic centerlines and planform geometry, ② flexibility to integrate with other features (e.g., dams), ③ better alignments with remote sensing for DA, and ④ computational efficiency**



Lin et al. (2018)



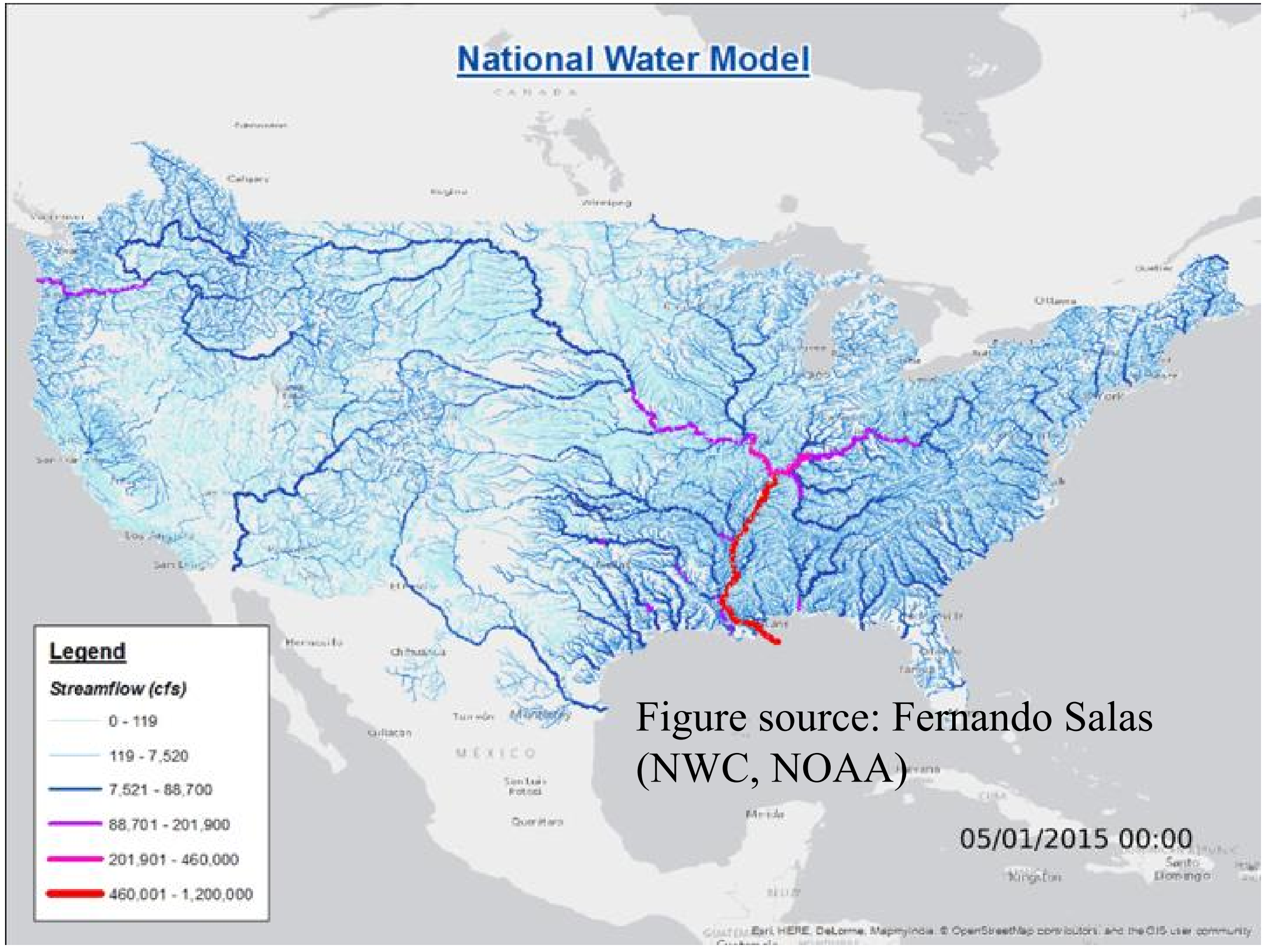
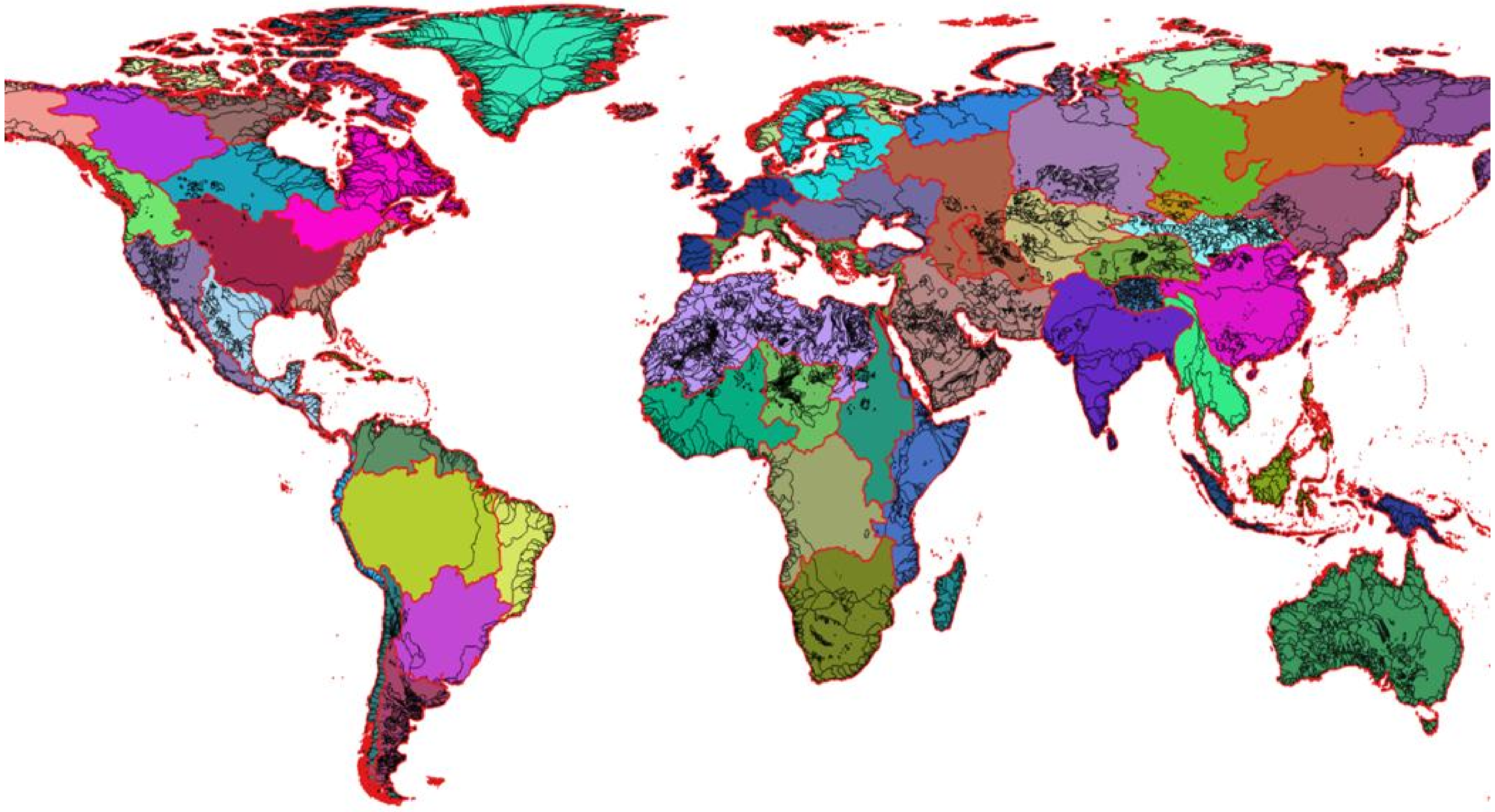


Figure source: Fernando Salas (NWC, NOAA)

MERIT-Basins: a vectorized hydrography dataset built on MERIT Hydro

- **Channelization threshold: 25 km^2, a total of 2.94 million river reaches**
  - Data organized following HydroSHEDS at level 1 & 2
    - ✓ 9 continents (1: Africa, 2: Europe, 3: North Asia, 4: East Asia, 5: Oceania, 6: South America, 7: North America, 8: Arctic, 9: Greenland); HydroSHEDS level 2: 61 basins



River flowlines (polylines), median length = 6.8 km

✓ riv_pfaf_*_MERIT_Hydro_v07_Basins_v01.shp

Unit catchments (polygons), local contributing area, median area = 36.8 km^2

✓ cat_pfaf_*_MERIT_Hydro_v07_Basins_v01.shp

MERIT-Basins: attributes

- **Basic attributes that come with MERIT-Basins**
  - ✓ **COMID:** identification number of the river reach; → inherited from NHDPlusV2
  - ✓ **lengthkm:** river length in km;
  - ✓ **lengthdir:** length directly connecting start/end node;
  - ✓ **sinuosity:** sinuosity of the river segment (unitless);
  - ✓ **slope:** mean slope of the river segment (unitless);
  - ✓ **uparea:** upstream drainage basin area in km^2;
  - ✓ **order:** Strahler-Horton stream order;

  → attributes useful to calculate routing parameters

  - ✓ **NextDownID:** COMID of the downstream reach;
  - ✓ **maxup:** number of upstream river reaches (max=4)
  - ✓ **up1:** upstream reach #1 COMID
  - ✓ **up2:** upstream reach #2 COMID
  - ✓ **up3:** upstream reach #3 COMID
  - ✓ **up4:** upstream reach #4 COMID

  → river network topology (inherited from the US NWM)

Global models built on top of MERIT-Basins

2019 — Lin et al. (2019, WRR) GRADES: naturalized discharge
- VIC + RAPID
- Daily
- 1979-2014

2021 — Yang et al. (2021, BAMS): GRFR river floods
- VIC + RAPID
- 3-hourly
- 1980-2019

2021 — Mizukami et al. (2021, JAMES)
- CLM/VIC + MizuRoute
- Daily
- 1980-2013

2021 — Feng et al. (2021, NC): RADR arctic Q
- VIC/GloFAS + HRR
- daily
- 1984-2018

2024 — Collins et al. (2024, Nat. Geo.): meanDRS
- GLDAS + RAPID
- monthly
- 1980-2009

2024 — Feng et al. (2024, Science): GRDR
- VIC/GloFAS + HRR
- Daily
- 1984-2018

2025 — Yang et al. (2025, WRR) GRADES-hydroDL
- LSTM + RAPID
- daily
- 1980-near present

How do we map MERIT-Basins with other hydrography datasets?

- **Different hydrography datasets have their own advantages, how to leverage the strengths of different datasets?**

✓ **MERIT-Basins:** global models

✓ **GRWL:** Landsat river width observations (Allen & Pavelsky, 2018, Science)

✓ **HydroSHEDS/ATLAS:** rich river attributes (Linke et al., 2019, Sci Data)

✓ **SWORD:** SWOT observations (Altenau et al., 2024, WRR)

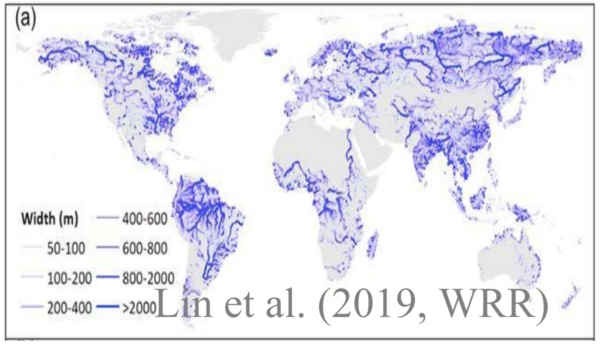
✓ **GRIT:** bifurcating channels (Wortmann et al., 2025, WRR)

✓ ……

Mapping MERIT-Basins with other hydrography datasets

GRWL to MERIT-Basins

- ✓ > 58 million GRWL cross sections
- ✓ Find the middle point as the cross section location
- ✓ Find the nearest river reach in MERIT-Basins
- ✓ Remove any mapping if the closest distance is >100 m
- ✓ >10 GRWL data suggests meaningful mapping



Lin et al. (2019, WRR)

MERIT-Basins to hydroATLAS

one-on-one mapping

- ✓ Find the middle point of each one of the MERIT-Basins river reach
- ✓ Find the nearest river reach in hydroATLAS
- ✓ Remove any mapping if the closest distance is >5 km

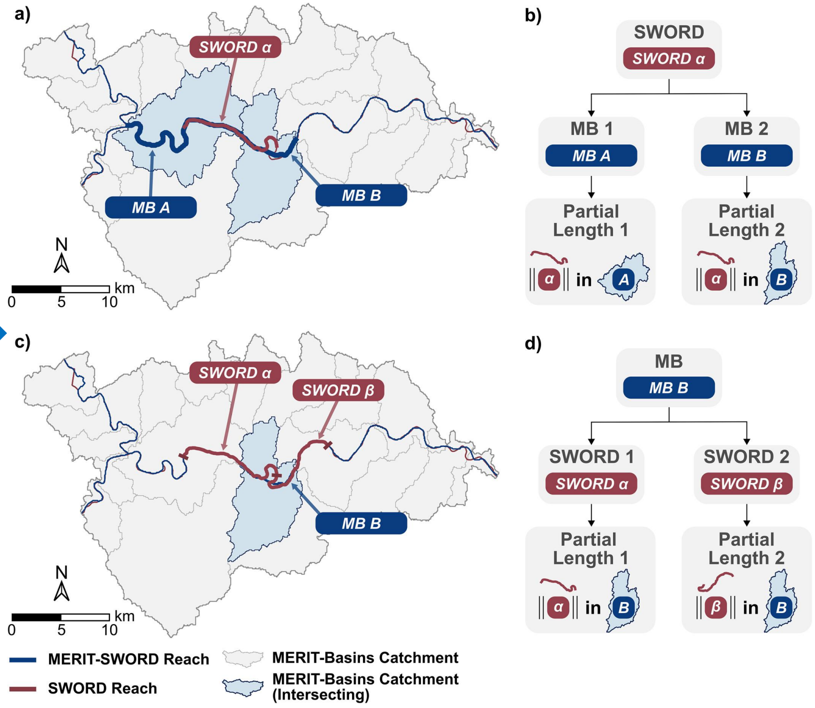
shared code for mapping: https://github.com/dry-rivers-rcn/G4/blob/master/src/join_GRADES_HydroATLAS.py

Krabbenhoft et al. (2022, NS)

more sophisticated spatial relationships

MERIT-Basins to SWORD

one-to-many translations



Wade et al. (2026, WRR)

MERIT-Basins and SWORD bilateral mapping

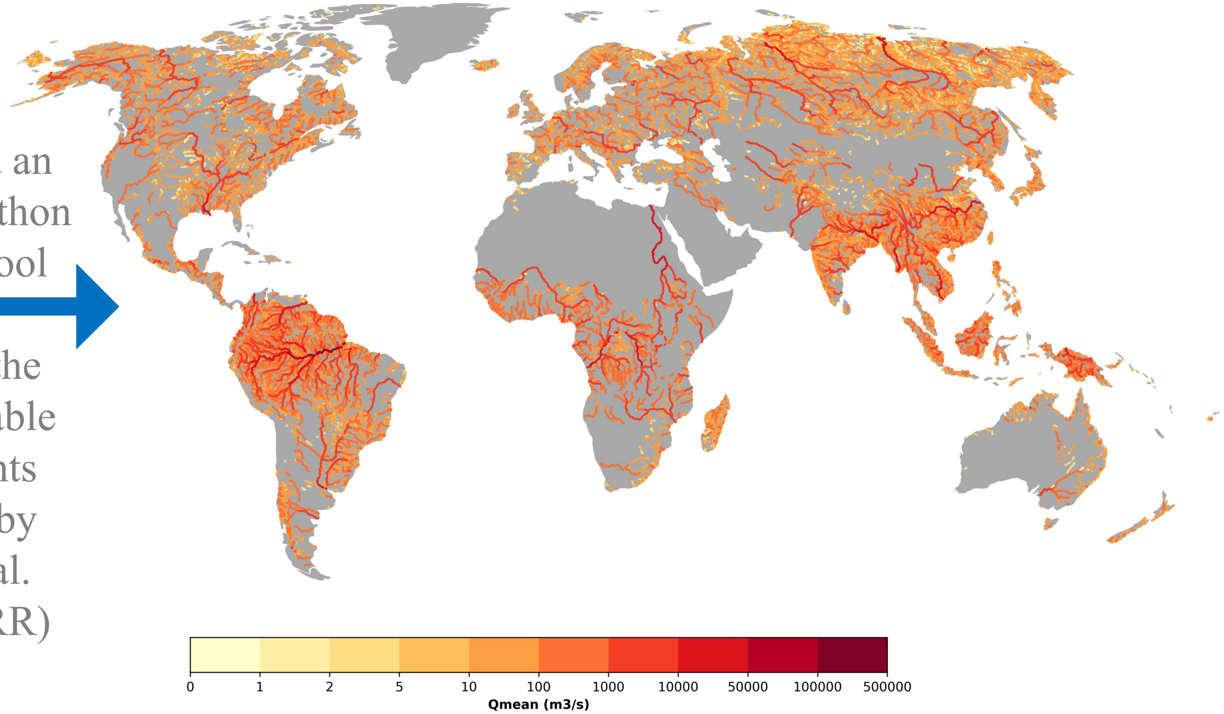
mean annual discharge on MERIT-Basins

mean annual discharge on SWORD

Developed an efficient Python mapping tool

based on the mapping table and weights provided by Wade et al. (2026, WRR)

Global river discharge mapped on SWORD



prior knowledge

modelled Q

remote sensing data

constraints

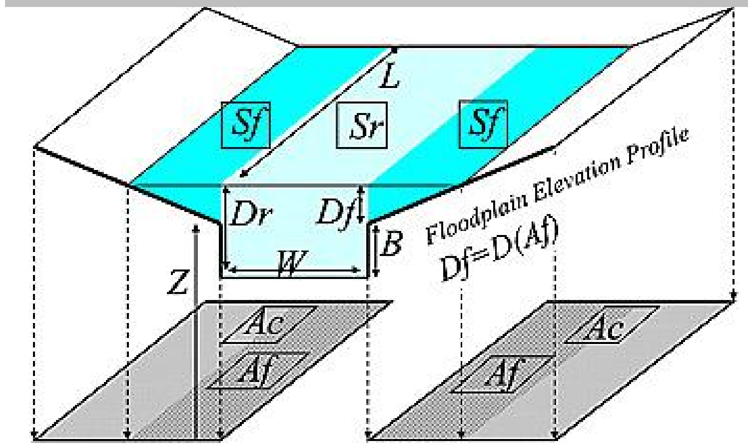
Outline

- **Vector hydrography**
  - ✓MERIT-Basins, MERIT-Hydro-Vector
- **Mapping hydrography of different sources**
  - ✓MERIT-Basins to GRWL, HydroSHEDS/ATLAS, SWORD
- **Enhanced attributes**
  - ✓Channel shape parameters, time-evolving attributes

Enhanced properties: Cross-sectional channel shapes

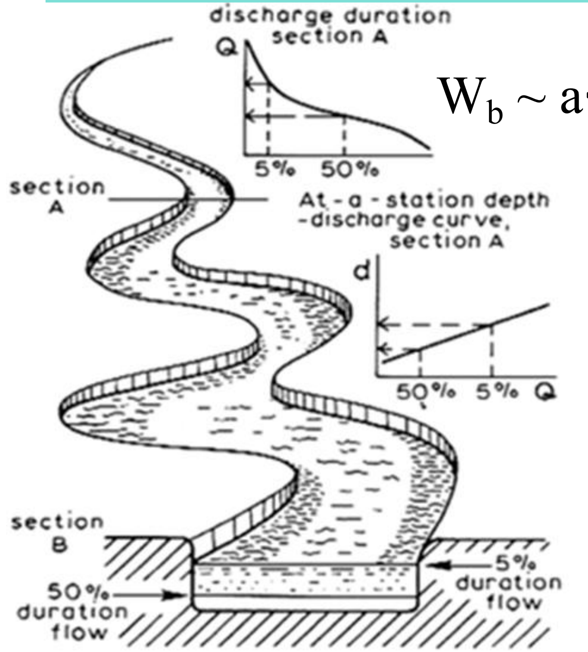
- **A revisit to the shape assumptions and spatial scaling of global river models**

Rectangle, bankfull width/depth: W/B



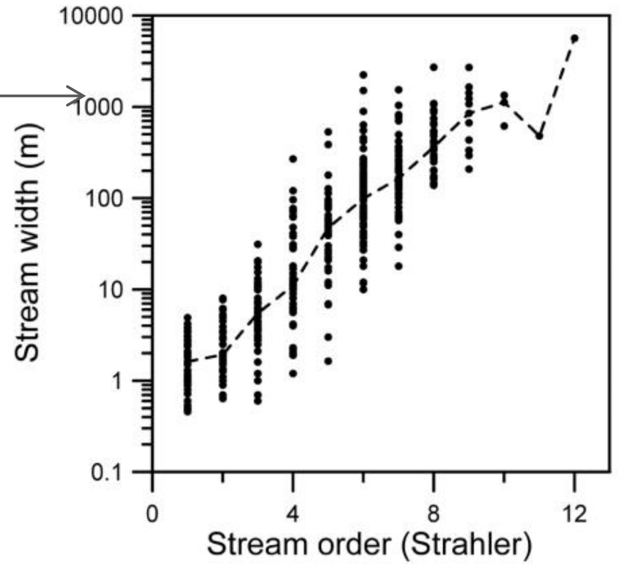
Yamazaki et al. (2011, WRR)

Scaling by discharge (DHG)

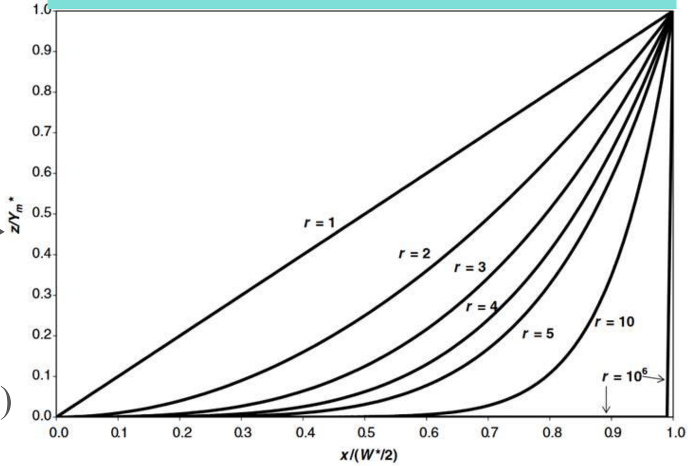


$W_b \sim a \cdot Q^b$

Scaling by stream order



Continuous shape parameters

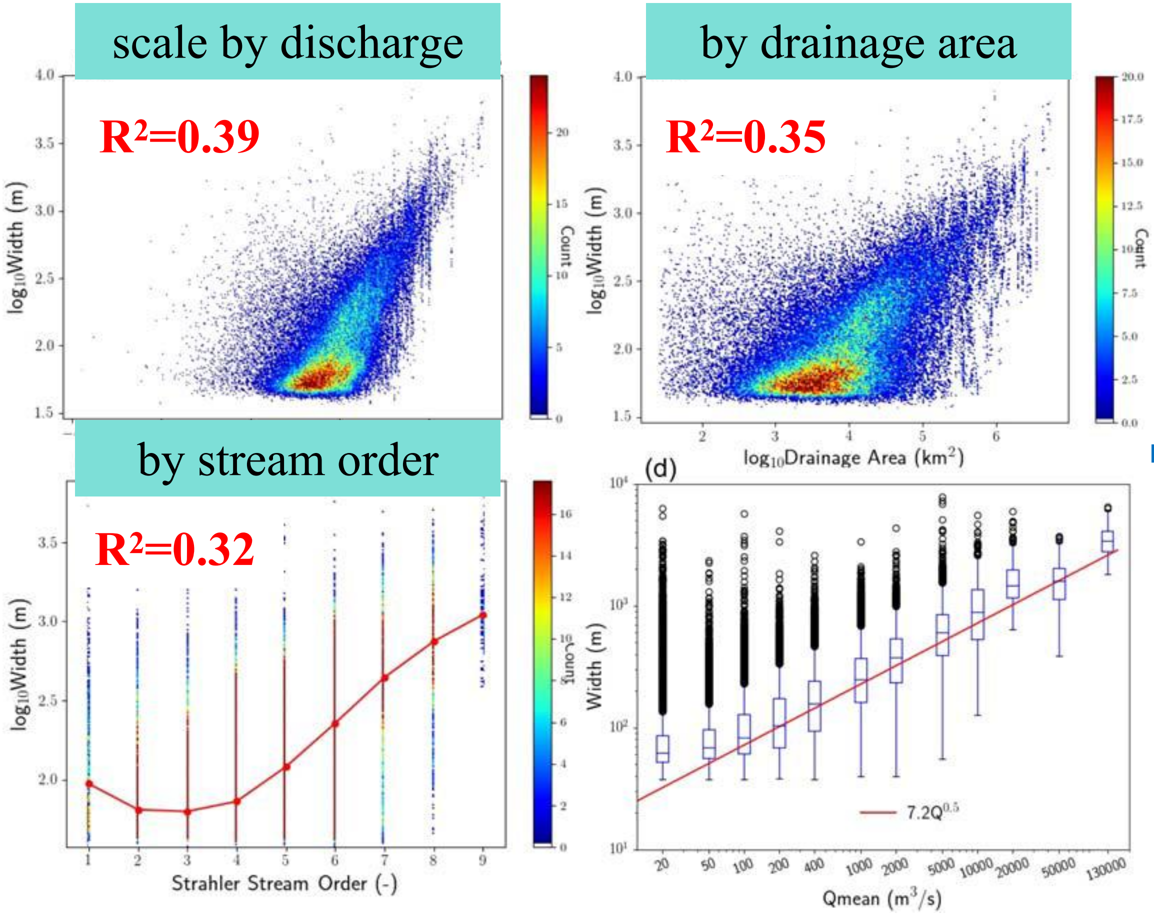


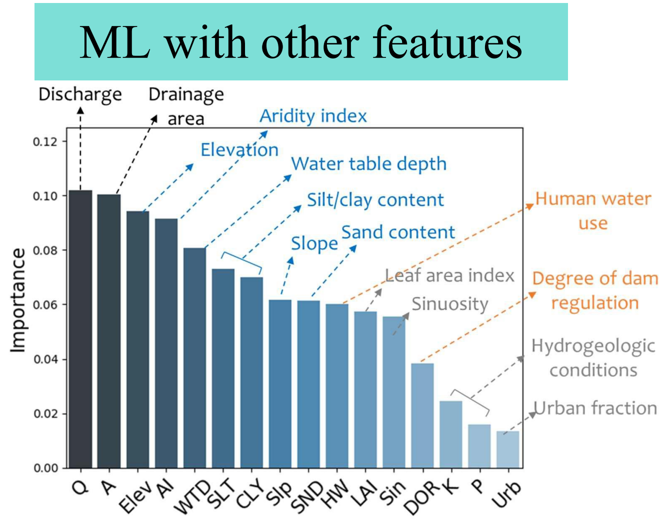
Dingman (2007s)

What does satellite data tell about global channel shapes?

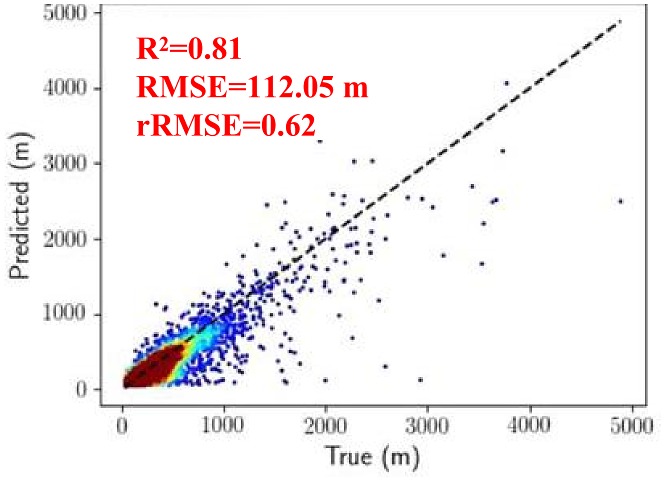
Revisiting the scaling relation across global river networks

- **Traditional scaling can only capture ~30-40% variance, but ML captures ~80% of the river width variance**





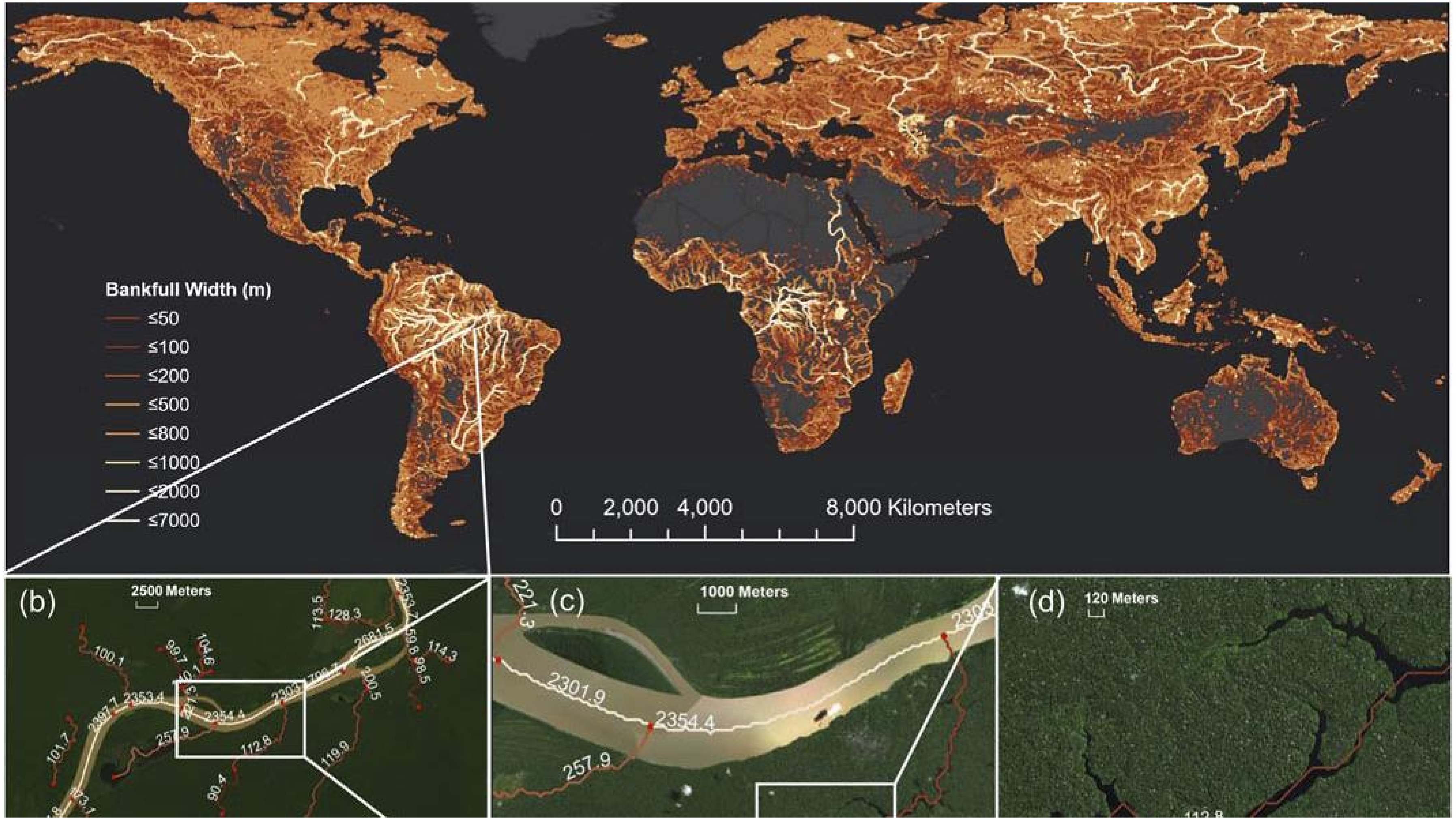
Elev, AI, WTD, soil composition, human water use, matter for width variance!



Lin et al. (*2020, GRL*)

Enhanced properties - channel shape parameters

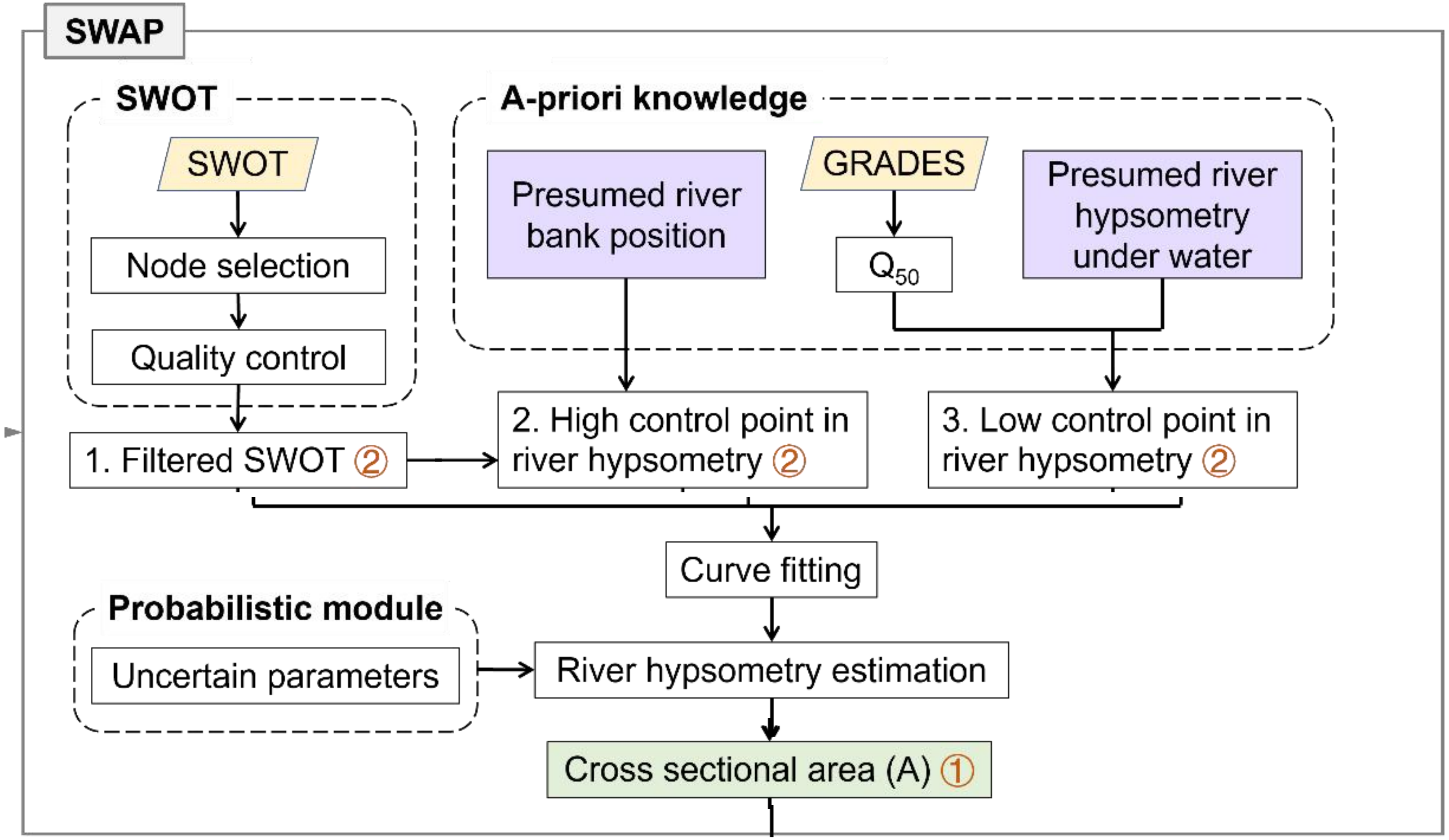
- Data-driven prediction of the global bankfull river width
  - ✓ ML as a feasible tool to estimate spatial variability of channel geometry parameters

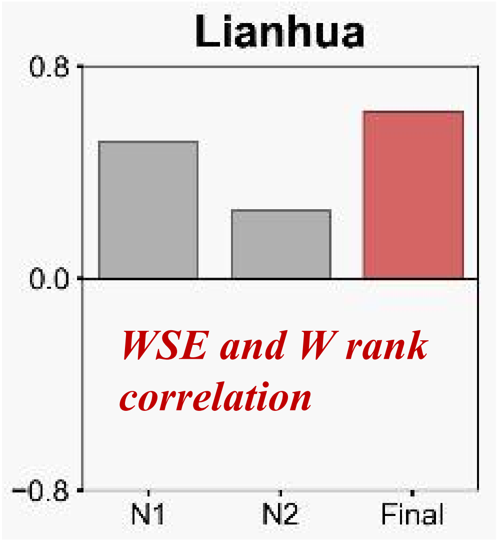


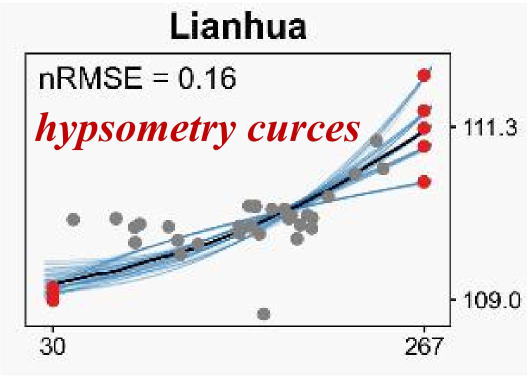
Lin et al. (*2020, GRL*)   https://zenodo.org/records/3552776

Using SWOT observations to estimate channel shapes

- **“SWOT and A-Priori information (SWAP)” Constrained Hypsometry Curves**



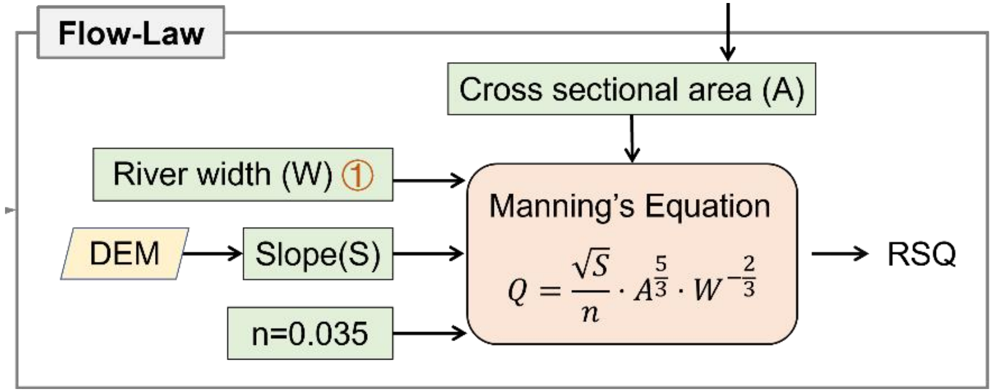




Xu, Yuan, Xu, Lin* (*2026, ISPRS, in revision*)

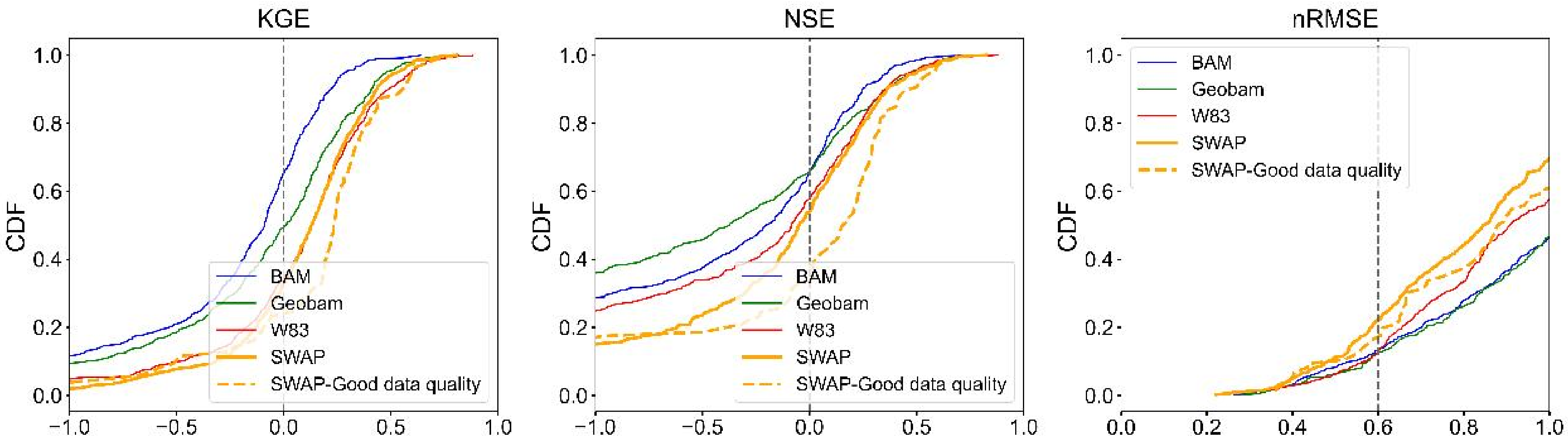
Using SWAP-constrained hypsometry curves to improve Q estimates

- **Developed the method at 15 northern China, and validated it at 557 global sites**



✓ **SWAP (orange lines) outperforms BAM/geoBAM/W83 estimates**

✓ **Even better performance when SWAP comes with good data quality (dashed orange lines)**



Xu, Yuan, Xu, Lin* (*2026, ISPRS, in revision*)

Using SWOT observations to estimate channel side slope

Config 1 — **SWOT Quality Tag Filter**

Quality Tag	Filter
Node quality	good / suspect
Few WSE observations	False
Cross track distance	10 - 60 km
Dark water fraction	≤ 40%
Climatological ice cover flag	≤ 1
WSE uncertainty	≤ 0.5 m
Relative width uncertainty	≤ 10%

Config 2 — **Additional Filters**

Inversion filter inconsistent width-WSE relation → **Outlier filter** median ± 3×std

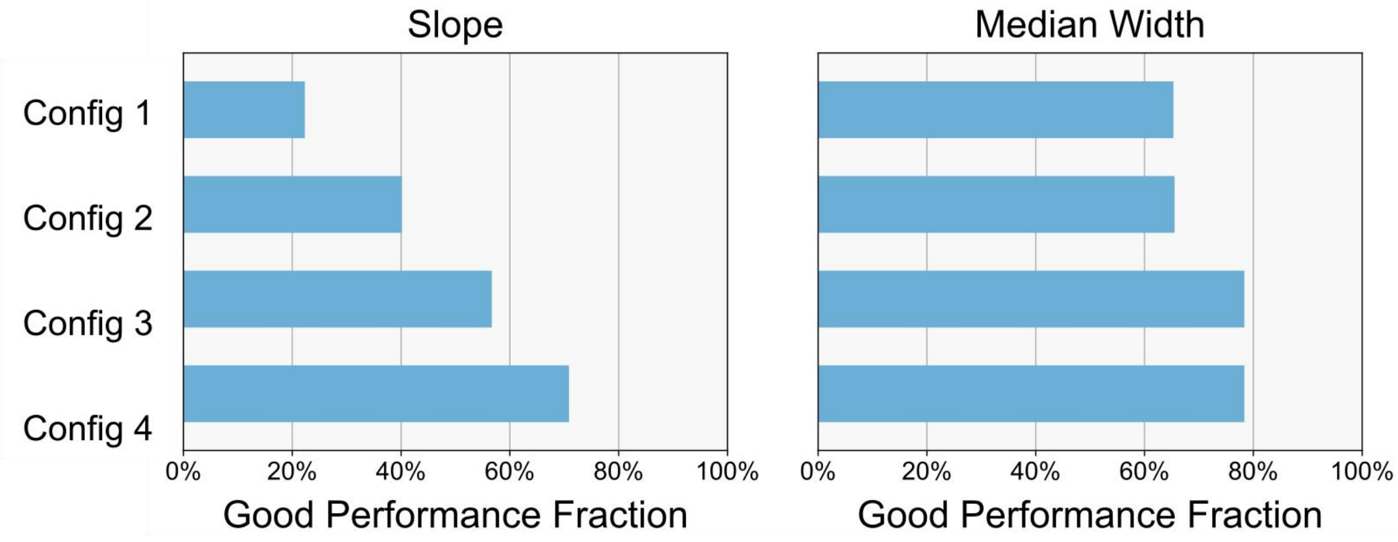
Config 3 — **Width Calibration**

GLOW → Resample to SWORD reaches → **CDF matching** weighted by inverse monthly frequency

Config 4 — **Fitting**

Fit with line or trapezoid + Total Least Squares (TLS)

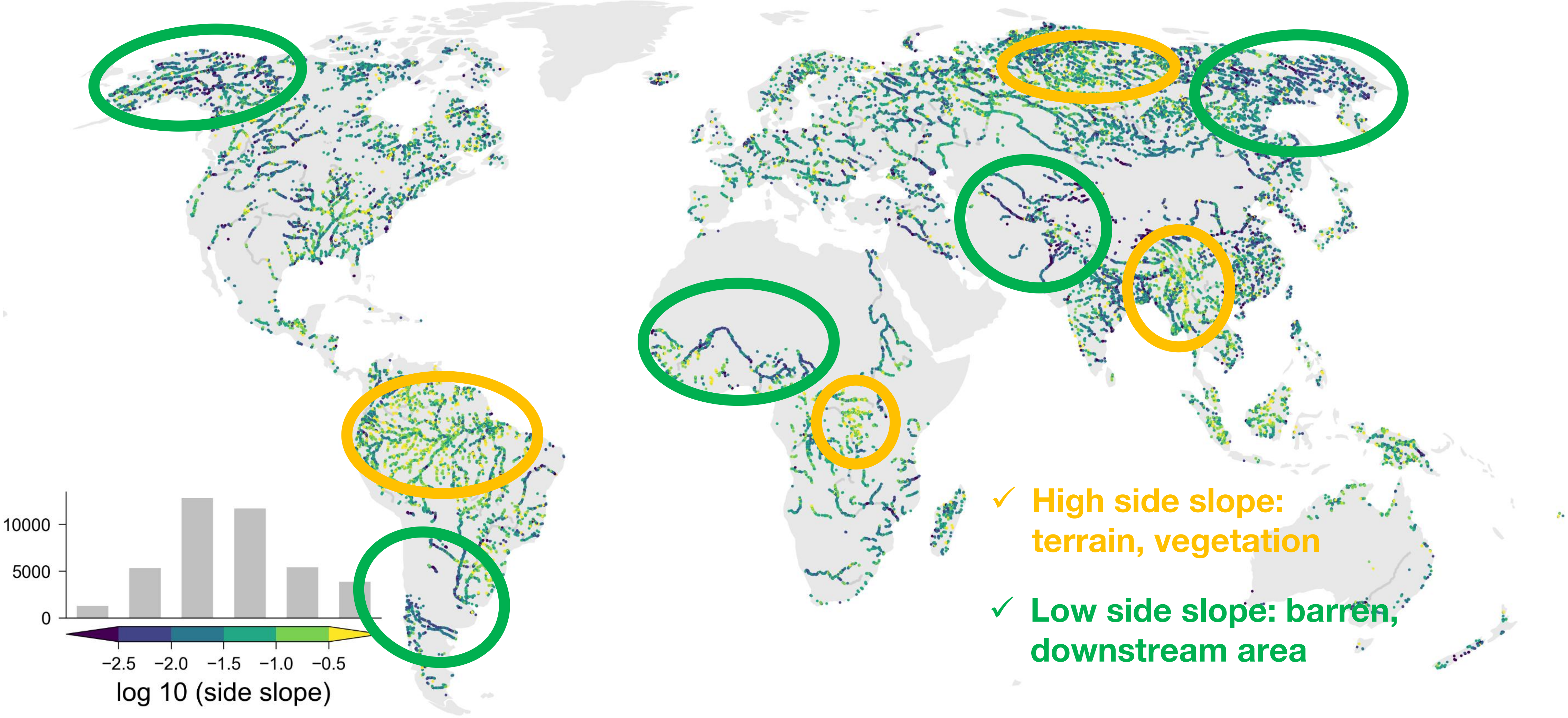
	SWOT Filter	Additional Filter	Width Calibration	OLS	TLS
Config 1	✔			✔	
Config 2	✔	✔		✔	
Config 3	✔	✔	✔	✔	
Config 4	✔	✔	✔		✔



Yuan, Lin* (*2025 AGU; manuscript in prep.*)

SWOT-observed channel side slope

- 40,475 reaches, 87% of which ranges between 0.003 ~ 0.32



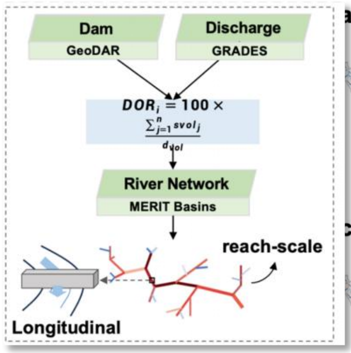
✓ **High side slope: terrain, vegetation**

✓ **Low side slope: barren, downstream area**

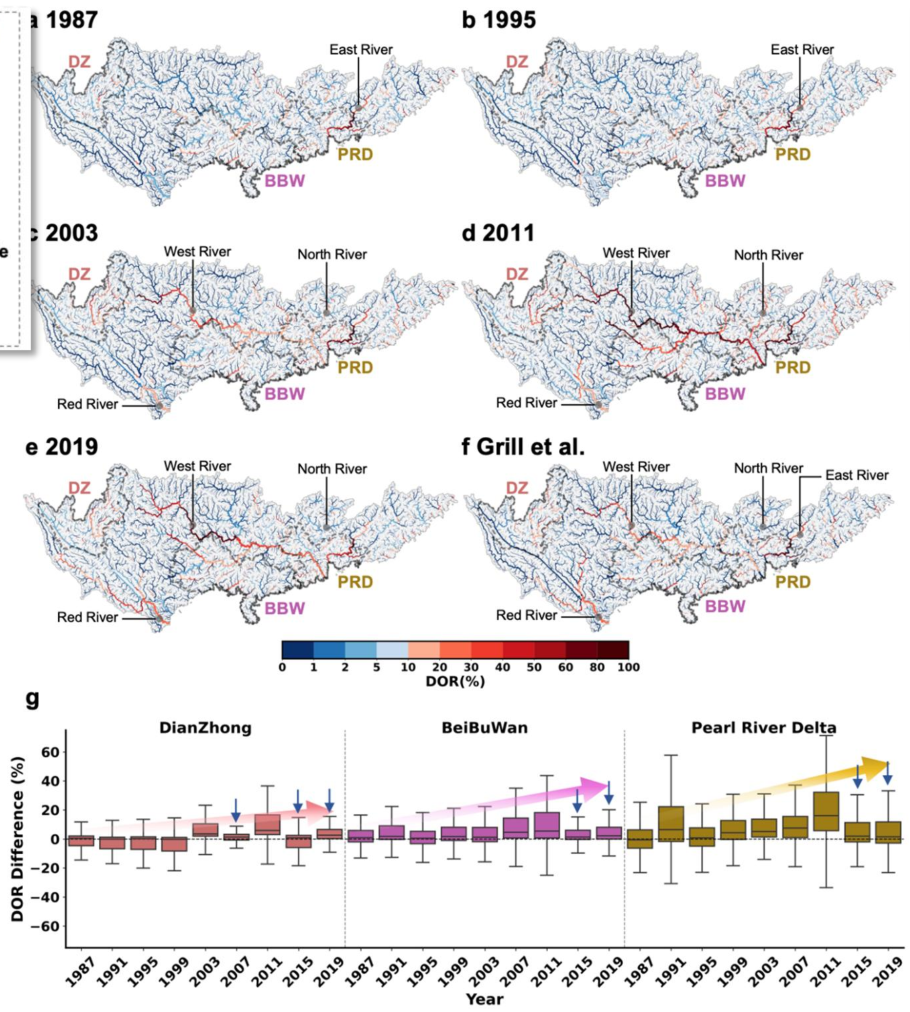
Yuan, Lin* (*2025 AGU; manuscript in prep.*)

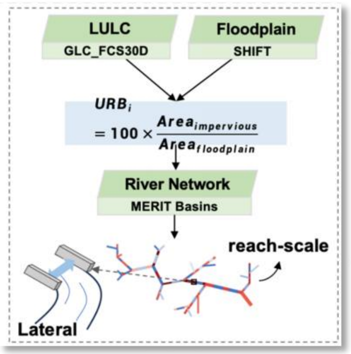
Enhanced properties - River Confinements

- **Time-varying longitudinal and lateral confinements to rivers**

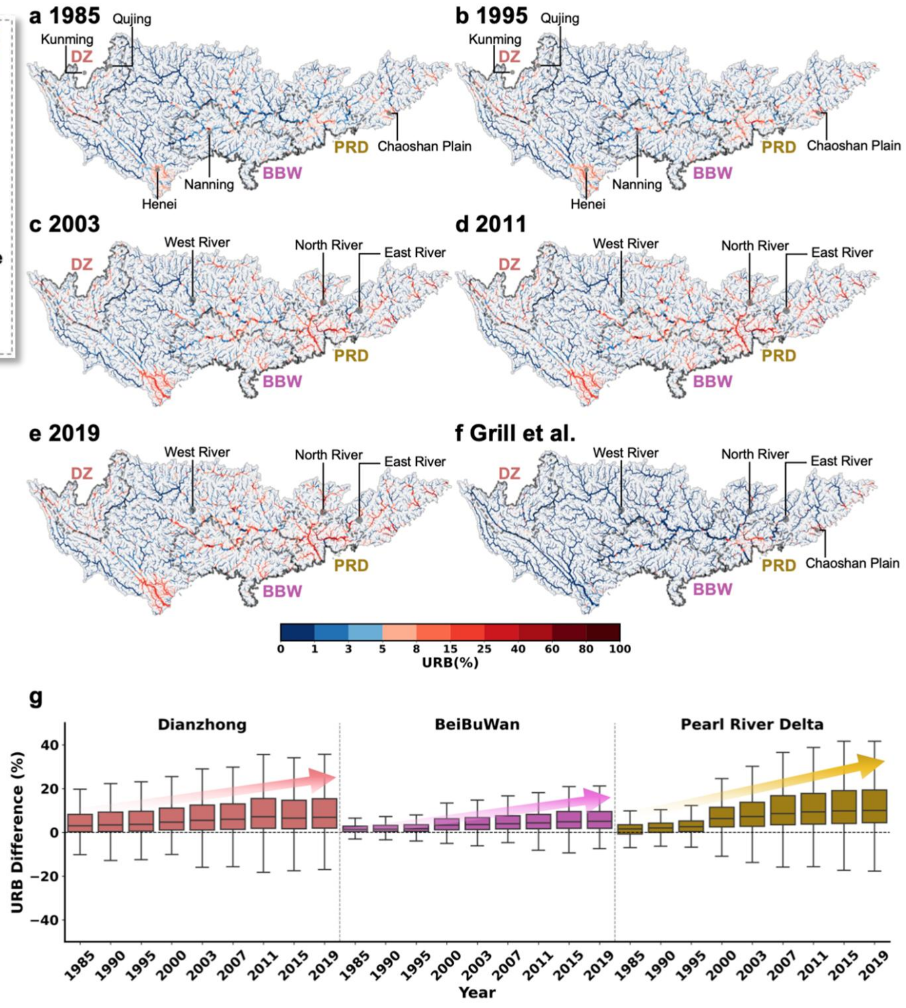


dams attenute flood peaks





levees elevate WSE and flood wave celerity

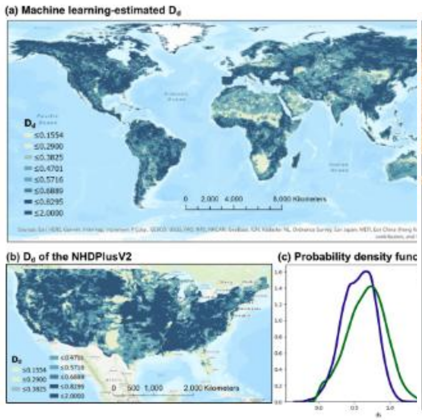


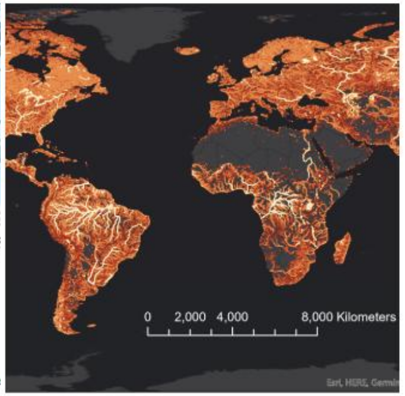
Lin, Lin*, Zheng (*2026, Sustainable Horizon*)

https://zenodo.org/records/15752866

Data sources

reachhydro.org
https://www.reachhydro.org/home/params





MERIT-BASINS

Vector river reaches and catchments derived from MERIT-Hydro

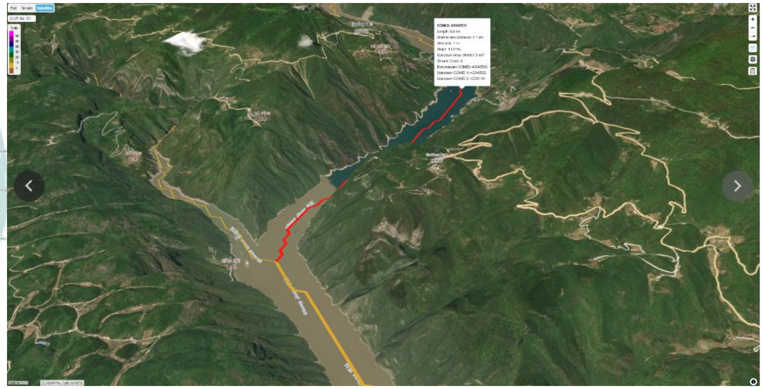
DRAINAGE DENSITY

Global Drainage Density

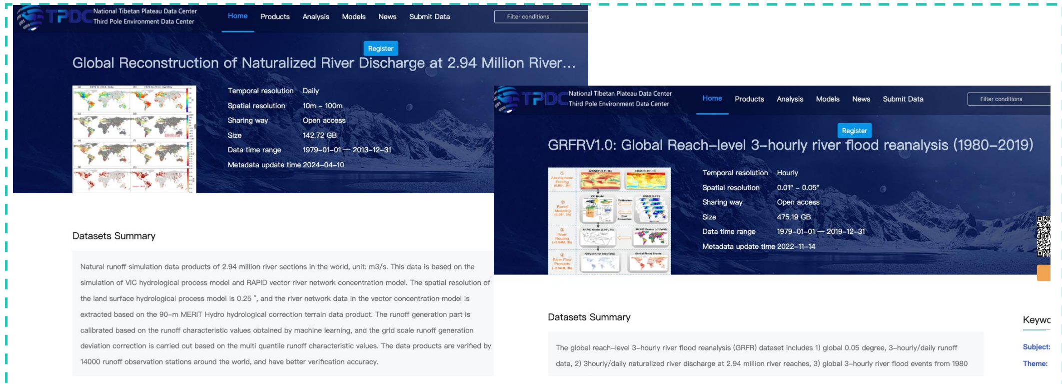
BANKFULL RIVER WIDTH

Global Reach-level Bankfull River Widths derived from Machine Learning

Credit to Ming Pan, Yuan Yang for interactive tools



National Tibetan Plateau Data Center (TPDC)
https://data.tpdc.ac.cn/home



zenodo

- More enhanced attributes (e.g., channel shape parameters) will be made available.
- Warmly welcome community contributions and discussions.

Summary and Outlooks

- **MERIT/MERIT Hydro/MERIT-Basins:**
  - ✓ backbones for a number of global routing models
- **Increasing need for interoperability of different hydrography datasets**
  - key to mapping satellite observations with models
  - key to creating rich attributes for use in improving model development
- **Community discussions: essential for standardizing global river modeling experiments**

Thanks for your attention!

Peirong Lin, Ph.D.

@ GeoWater lab @ Peking University

For questions, please contact: peironglinlin@pku.edu.cn