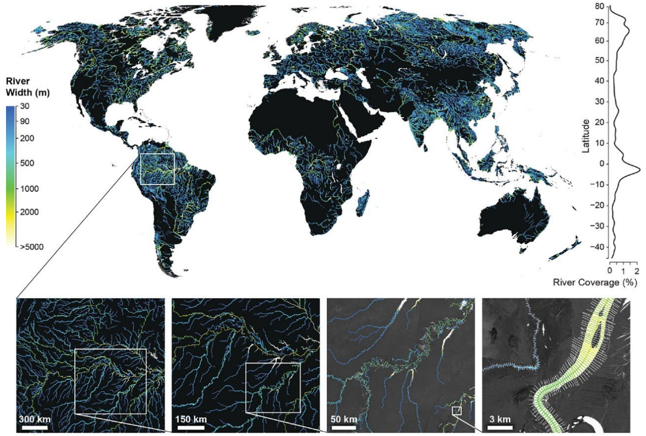
SWOT River Database (SWORD) Update

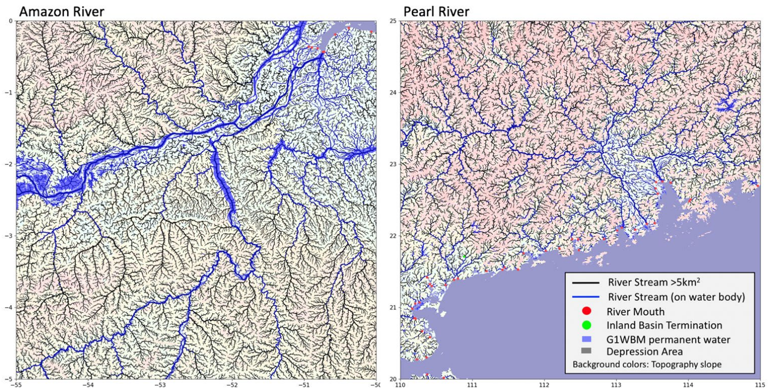
Elizabeth Altenau,
Tamlin Pavelsky,
Elyssa Collins
University of North Carolina at Chapel HIll

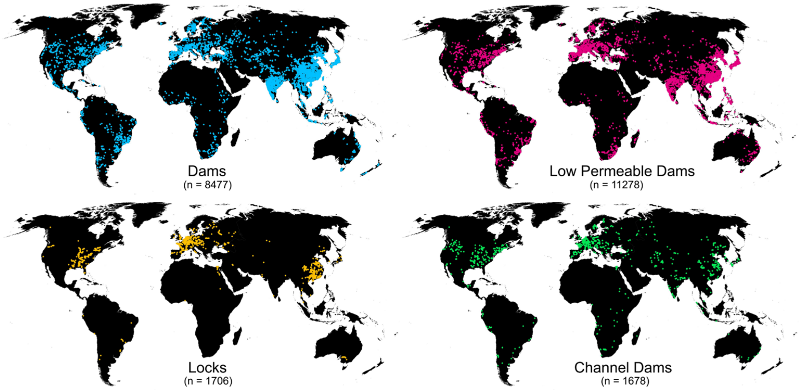
June 17–21, 2024
SWOT Science Team Meeting
Chapel Hill, NC

SWORD Background - Datasets

DATASET	ATTRIBUTE CONTRIBUTION
Global River Widths from Landsat (GRWL) (*Allen & Pavelsky, 2018*)	Provides river centerline locations at 30 m resolution and associated width, water body type, and number of channels attributes.
MERIT Hydro (*Yamazaki et al., 2019*)	Provides elevation and flow accumulation at 3 arc-second resolution (~90 m at the equator).
HydroBASINS (*Lehner & Grill, 2013*)	Provides Pfafstetter nested basin codes up to level 6.
Global River Obstruction Database (GROD) (*Whittemore et al., 2020, Yang et al., in review*)	Provides global locations of anthropogenic river obstructions along the GRWL river network.
Global Delta Maps (*Tessler et al., 2015*)	Provides the spatial extent of 48 of the world's largest river deltas.
SWOT Orbits (*https://www.aviso.altimetry.fr/en/missions/future-missions/swot/orbit.html*)	Provides polygons containing SWOT track coverage for each pass throughout the 21-day cycle orbit.
HydroFALLS (*http://wp.geog.mcgill.ca/hydrolab/hydrofalls/*)	Provides global locations of waterfalls and natural river obstructions.

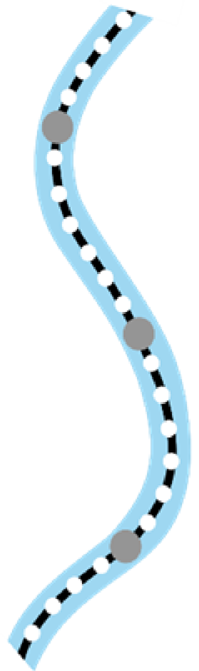






SWORD Background - Structure

SWORD Structure



Dimensions

- ○ Centerline: ~30 m
- ● Node: ~200 m
- ▬ Reach: ~10 km

SWORD Variables (44+)

Centerline-Node-Reach Shared Variables

- cl_id
- lon
- lat
- reach_id
- node_id

Node-Reach Shared Variables

- wse
- wse_var
- width
- width_var
- max_width
- n_chan_max
- n_chan_mod
- obstr_type
- grod_id
- hfalls_id
- lakeflag
- dist_out
- facc
- river_name
- edit_flag
- trib_flag
- path_freq
- path_order
- path_segs
- main_side
- stream_order
- end_reach
- network

Node Variables

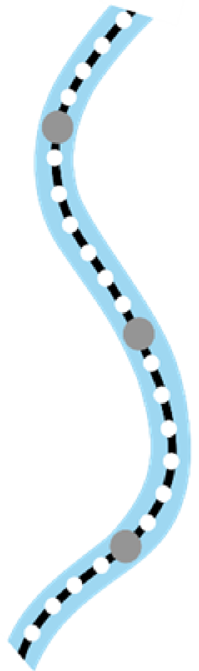
- node_length
- width_coef
- ext_dist_coef
- meander_length
- sinuosity

Reach Variables

- reach_length
- n_nodes
- slope
- n_rch_up
- n_rch_down
- rch_id_up
- rch_id_down
- ice_flag

SWORD Background - Structure

SWORD Structure



Dimensions

- ○ Centerline: ~30 m
- ● Node: ~200 m
- — Reach: ~10 km

SWORD Variables (44+)

Centerline-Node-Reach Shared Variables

- cl_id
- lon
- lat
- reach_id
- node_id

Reach Location Change

Node-Reach Shared Variables

- wse
- wse_var
- width
- width_var
- max_width
- n_chan_max
- n_chan_mod
- obstr_type
- grod_id
- hfalls_id
- lakeflag
- dist_out
- facc
- river_name
- edit_flag
- trib_flag
- path_freq
- path_order
- path_segs
- main_side
- stream_order
- end_reach
- network

Node Variables

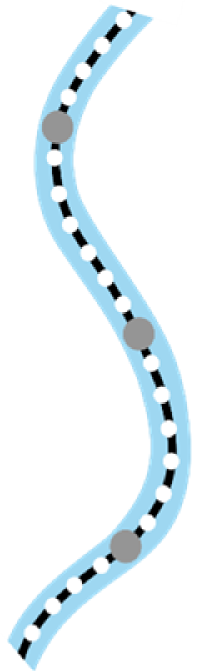
- node_length
- width_coef
- ext_dist_coef
- meander_length
- sinuosity

Reach Variables

- reach_length
- n_nodes
- slope
- n_rch_up
- n_rch_down
- rch_id_up
- rch_id_down
- ice_flag

SWORD Background - Structure

SWORD Structure



Dimensions

- ○ Centerline: ~30 m
- ● Node: ~200 m
- — Reach: ~10 km

SWORD Variables (44+)

Centerline-Node-Reach Shared Variables

- cl_id
- lon
- lat
- reach_id
- node_id

Reach Location Change

Node-Reach Shared Variables

- wse
- wse_var
- width
- width_var
- max_width
- n_chan_max
- n_chan_mod
- obstr_type
- grod_id
- hfalls_id
- lakeflag
- dist_out
- facc
- river_name
- edit_flag
- trib_flag
- path_freq
- path_order
- path_segs
- main_side
- stream_order
- end_reach
- network

Node Variables

- node_length
- width_coef
- ext_dist_coef
- meander_length
- sinuosity

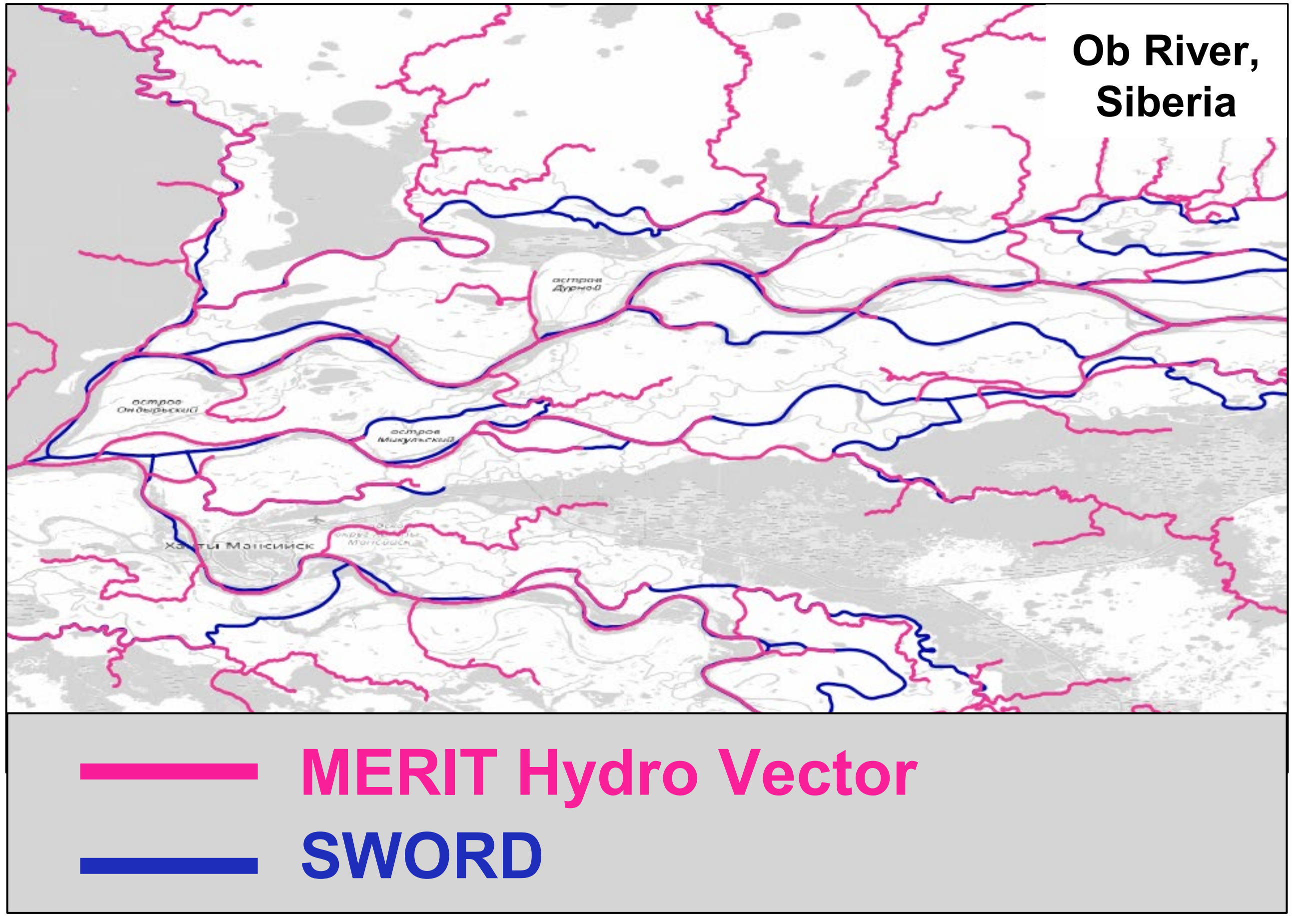
Reach Variables

- reach_length
- n_nodes
- slope
- n_rch_up
- n_rch_down
- rch_id_up
- rch_id_down
- ice_flag

SWORD Topology Challenges

When first developing SWORD:

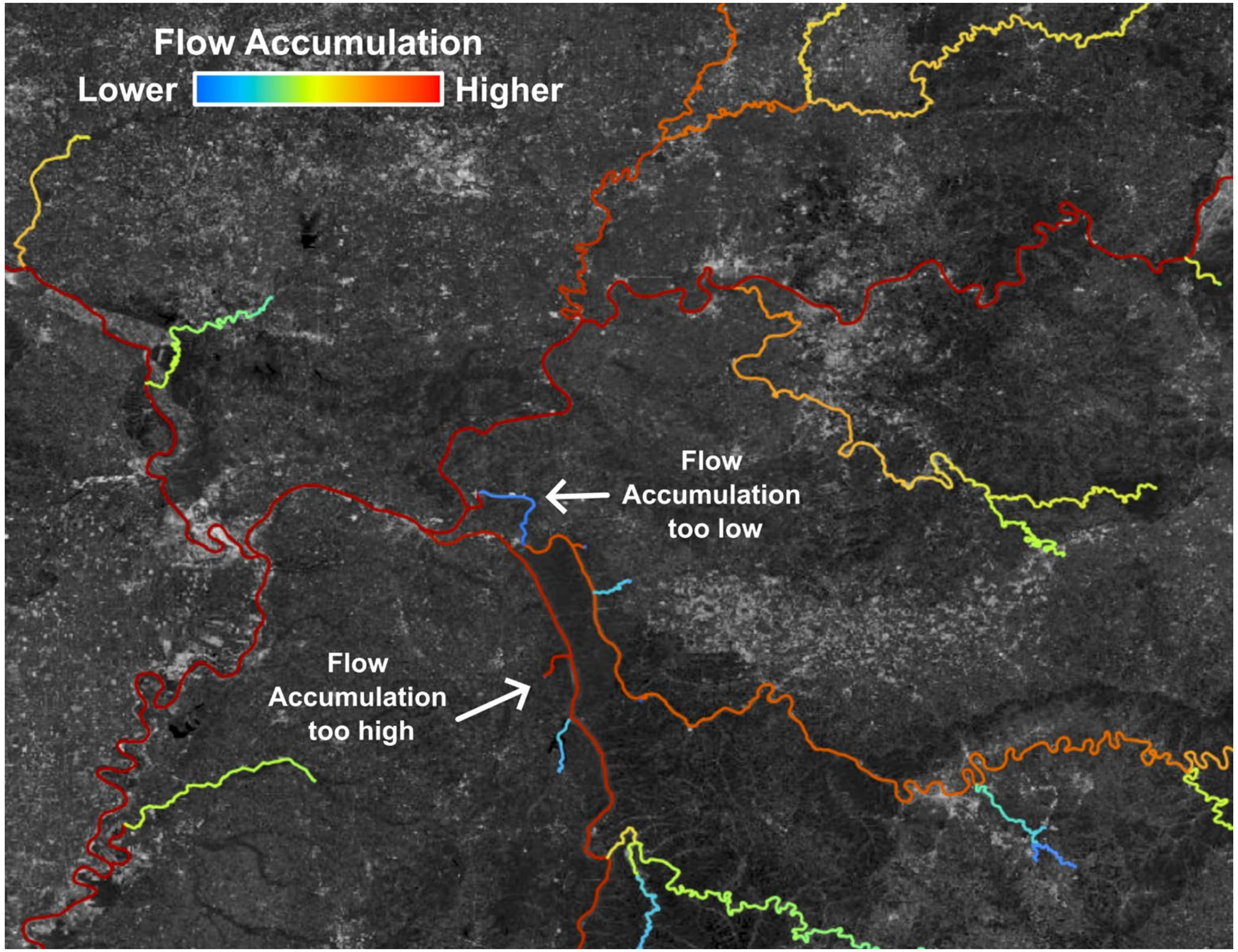
- **Chose optically derived centerline to try and represent rivers as the are rather than how a DEM represents them.**
- **Global hydrography datasets were limited - MERIT-Hydro was not published yet.**
- **Topology was not considered a critical variable**



SWORD Topology Challenges

Consequences of optically-derived centerlines:

- **Typical variables used for topology have errors / inconsistencies due to merging problems in areas where centerlines don't match well.**
- **Discontinuities are more common in the river network.**
- **Small localized errors can lead to large propagations upstream.**

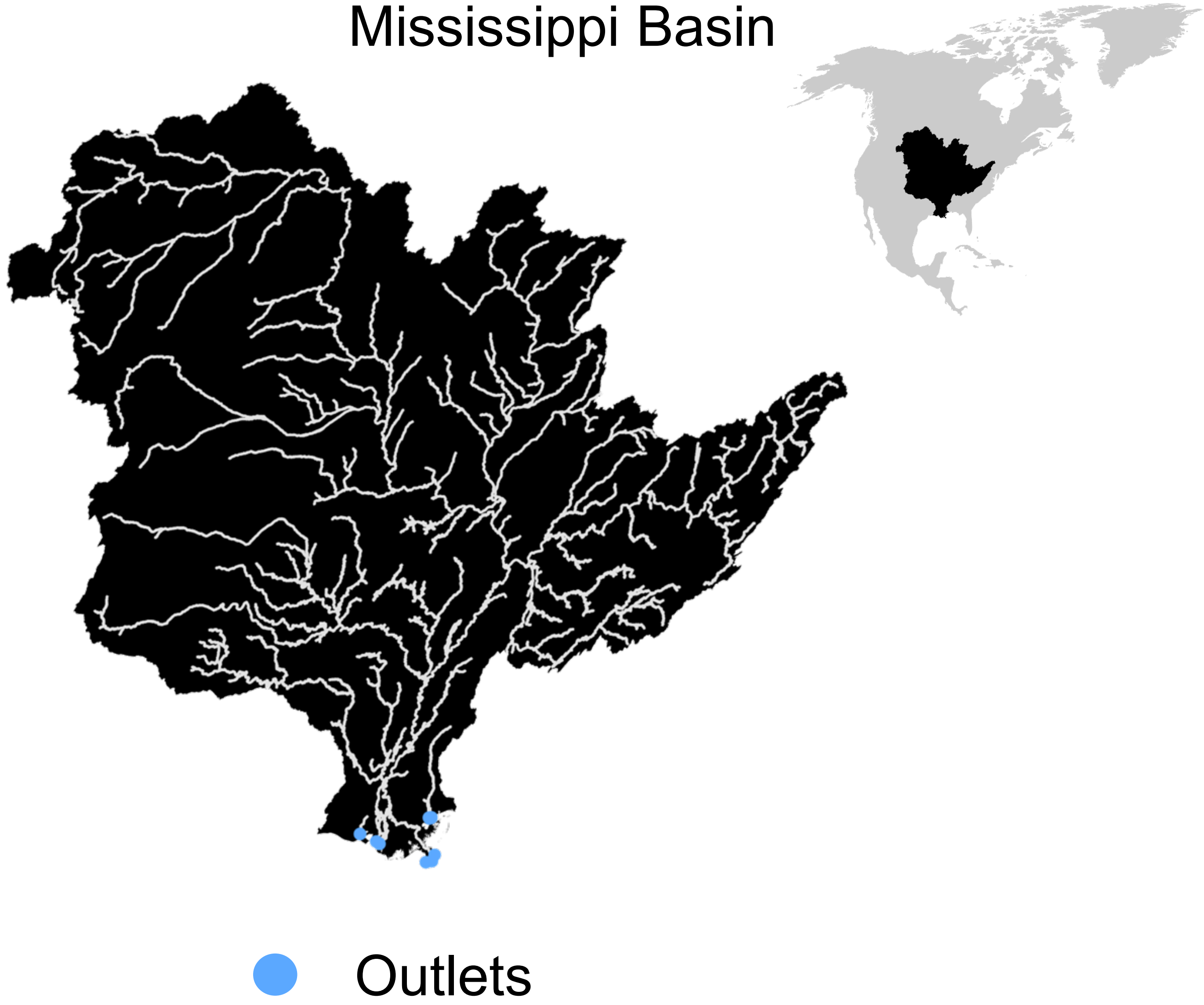


Needed a SWORD based variable for calculating topology.

Solution:
Build new distance from outlet based on pathways from outlets to headwaters.

Used a shortest path algorithm to map all paths from every outlet to all associated headwaters

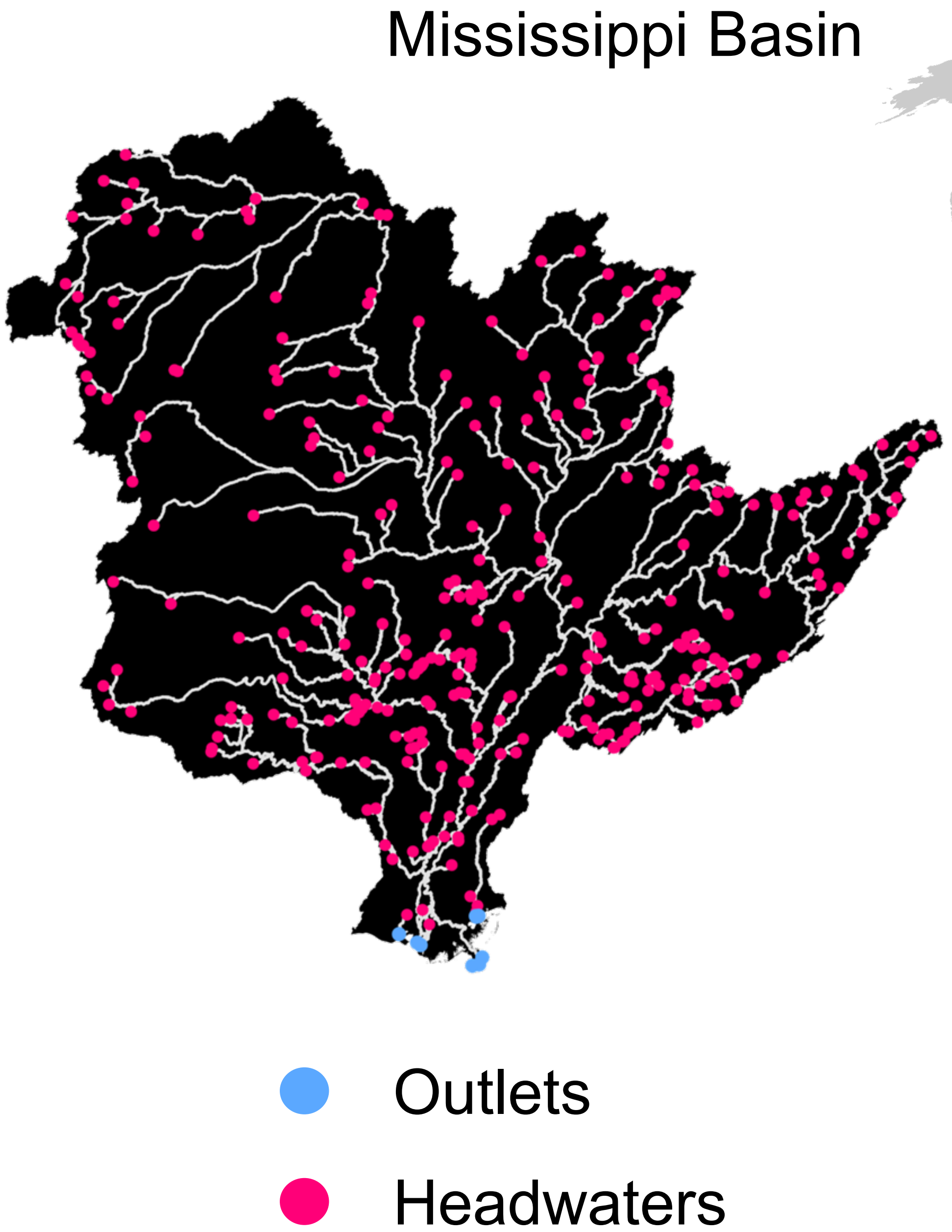
Mississippi Basin



Needed a SWORD based variable for calculating topology.

Solution:
Build new distance from outlet based on pathways from outlets to headwaters.

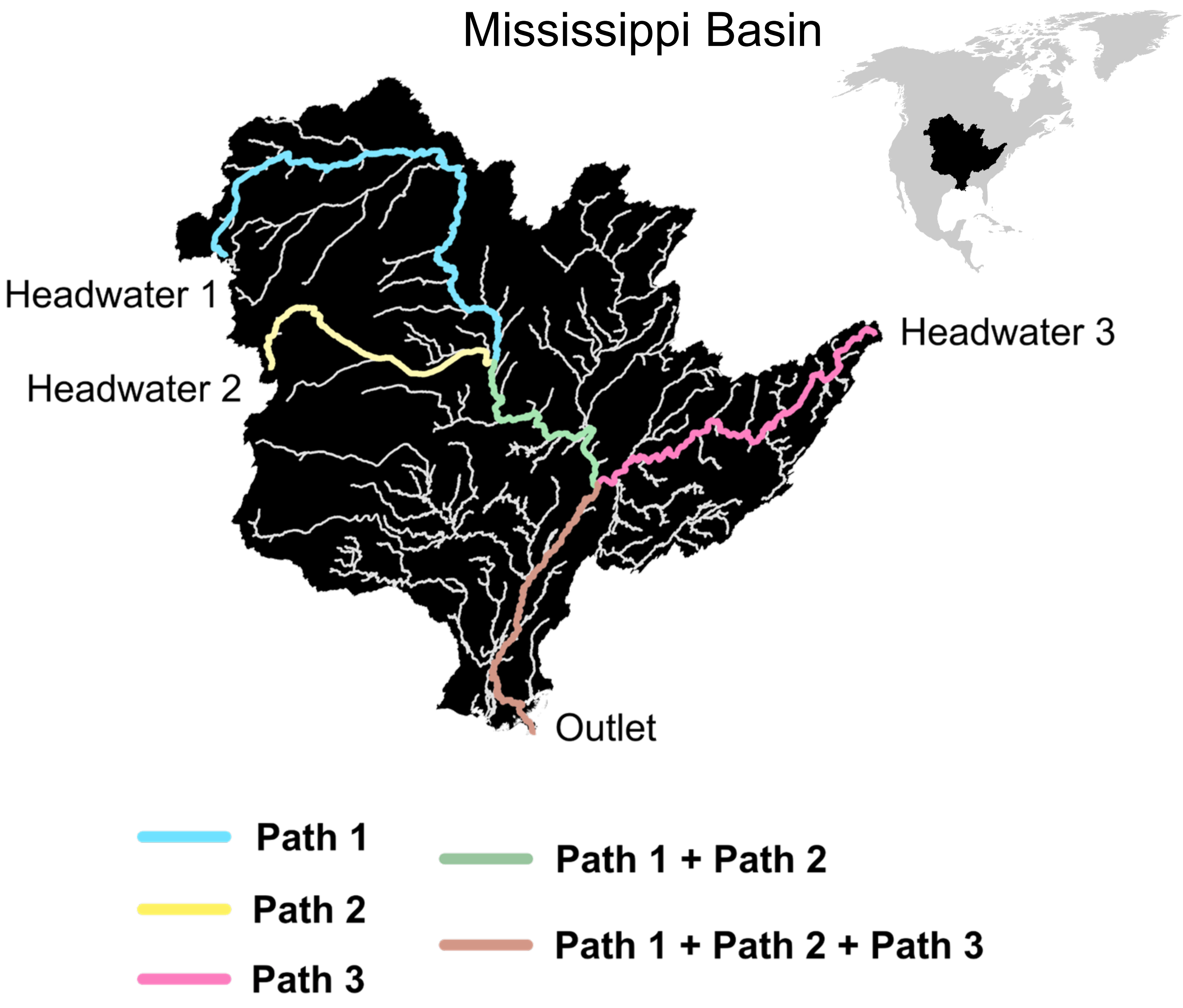
Used a shortest path algorithm to map all paths from every outlet to all associated headwaters

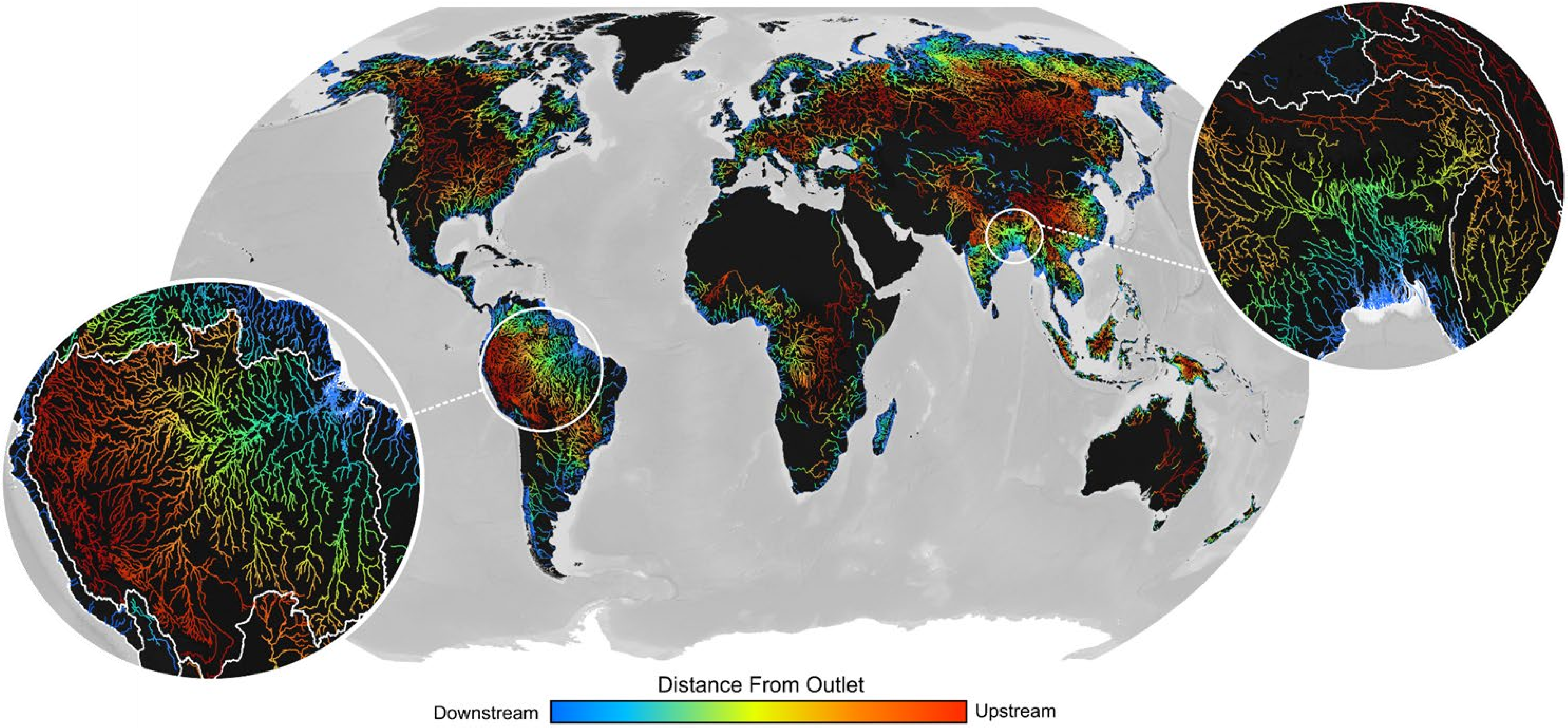


Needed a SWORD based variable for calculating topology.

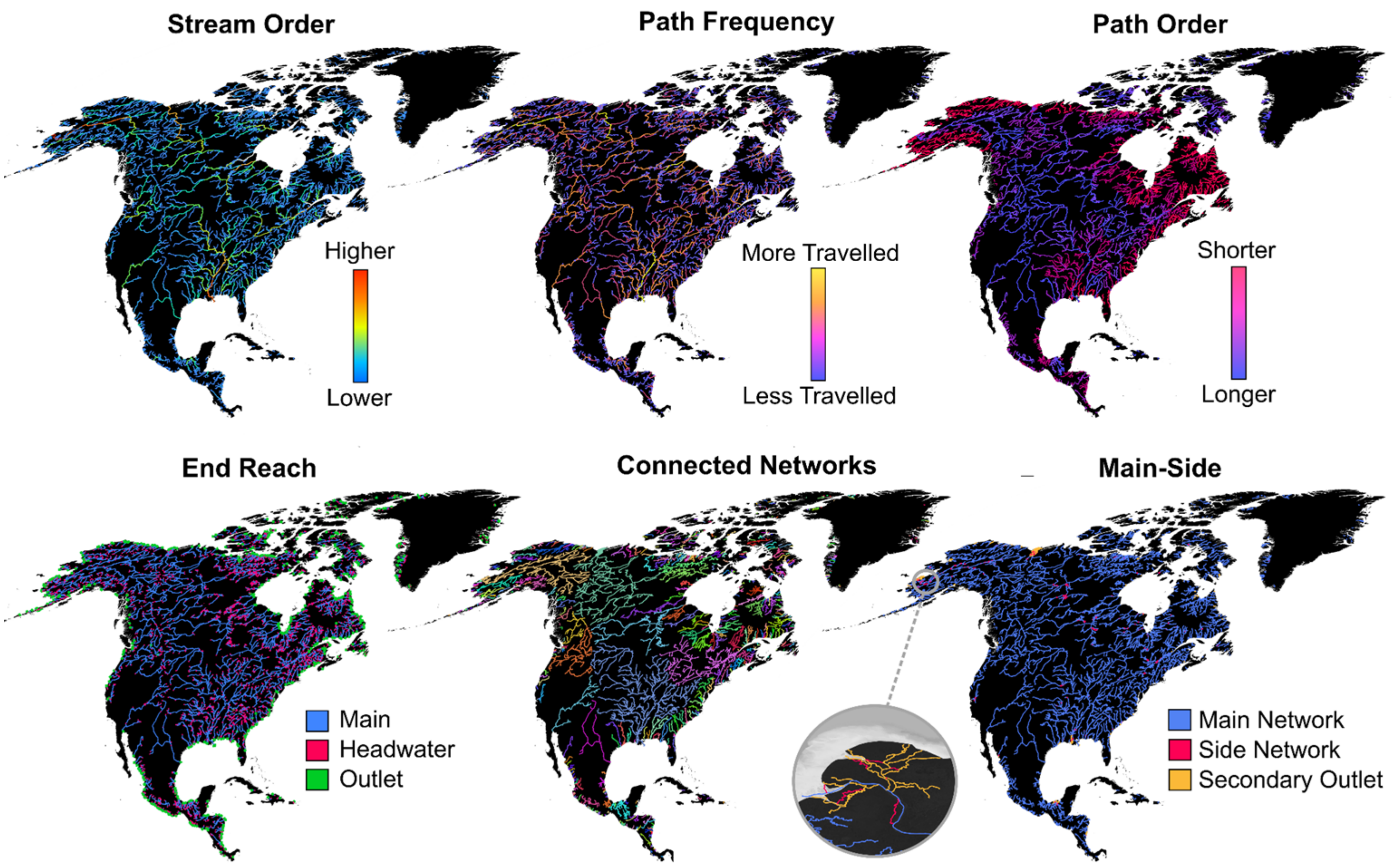
Solution:
Build new distance from outlet based on pathways from outlets to headwaters.

Used a shortest path algorithm to map all paths from every outlet to all associated headwaters



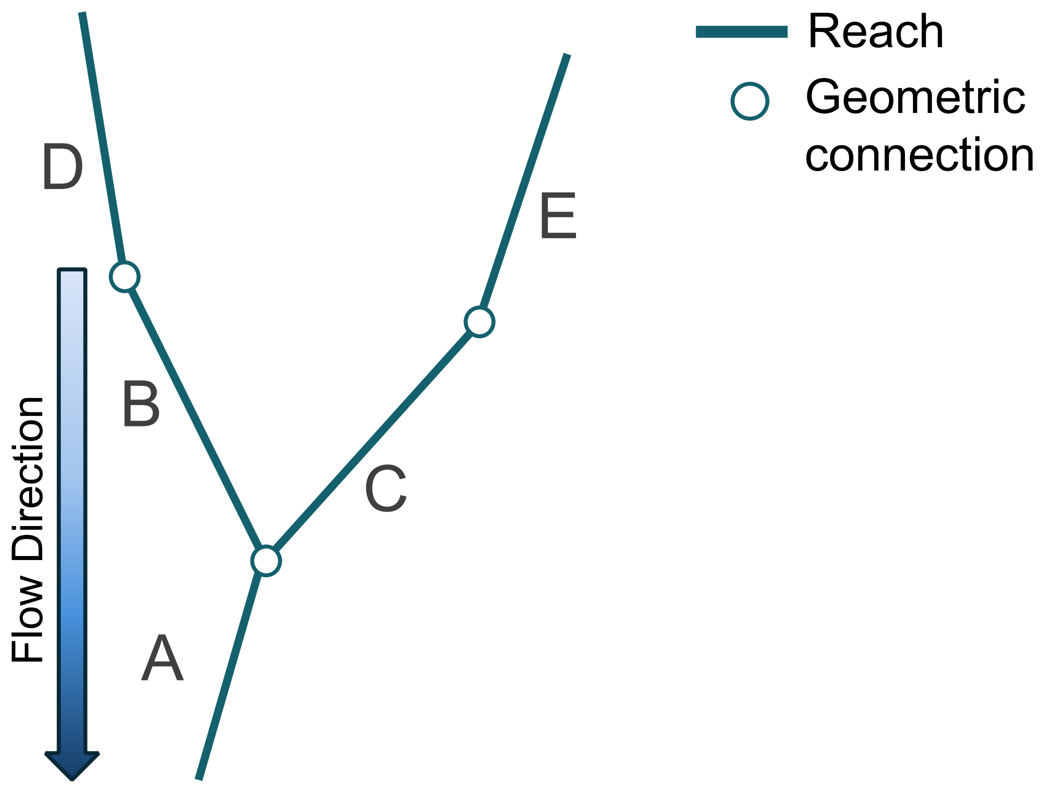


Distance From Outlet
Downstream
Upstream



Stream Order
Higher
Lower
Path Frequency
More Travelled
Less Travelled
Path Order
Shorter
Longer
End Reach
Main
Headwater
Outlet
Connected Networks
Main-Side
Main Network
Side Network
Secondary Outlet

Find geometric connections

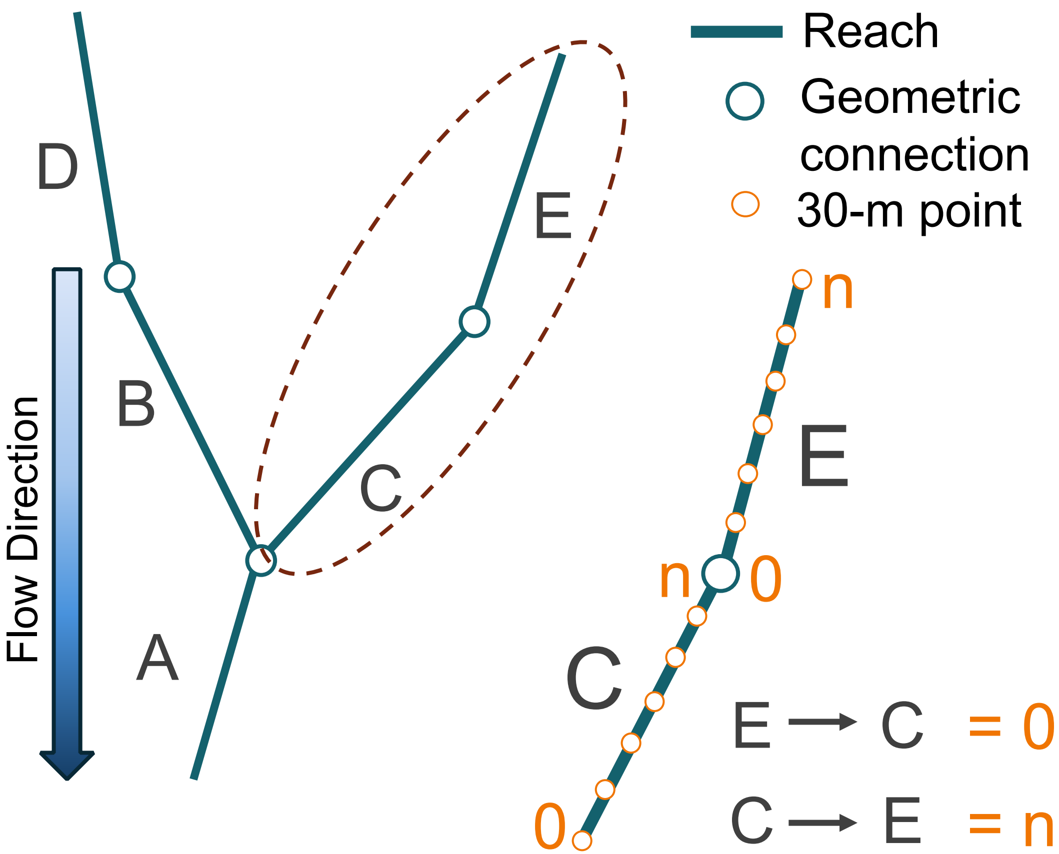


Geometric connections

A ⟷ B B ⟷ D
A ⟷ C C ⟷ E
B ⟷ C

Slides Courtesy of Dr. Elyssa Collins

Find geometric connections



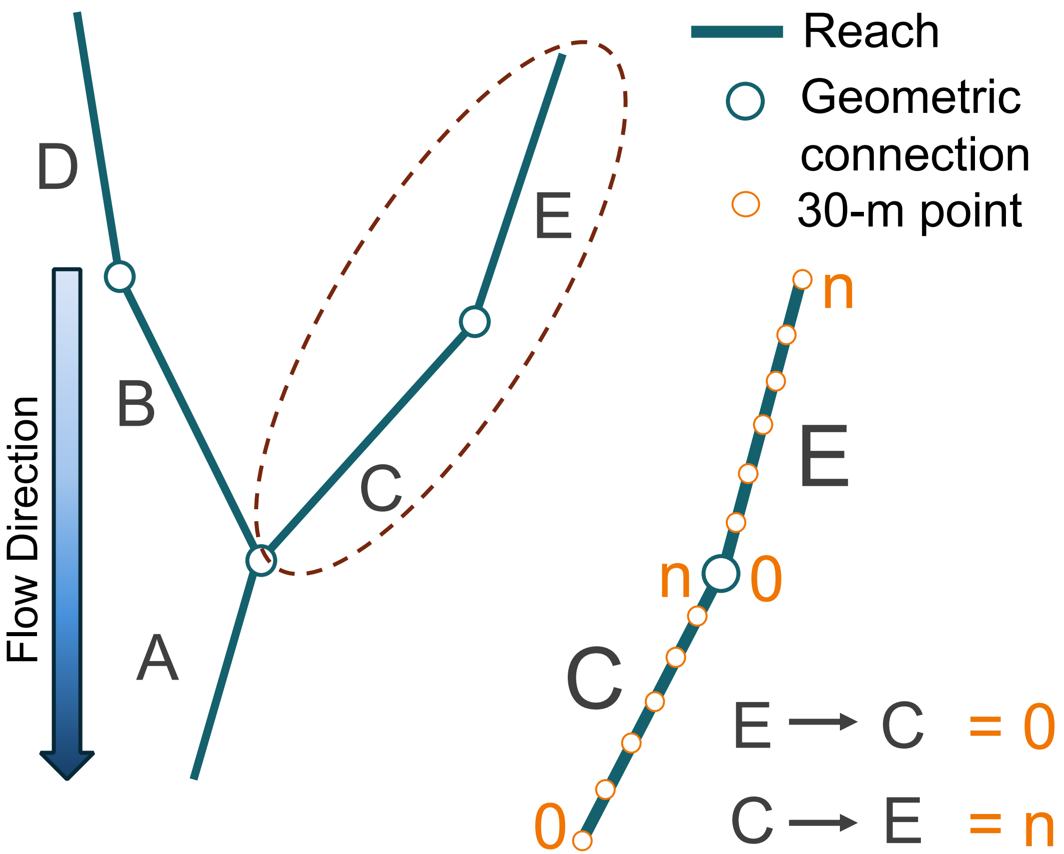
Geometric connections

A ↔ B B ↔ D
A ↔ C C ↔ E
B ↔ C

Here, reach E intersects reach C at the 30-m point in the 0 index. Reach C intersects reach E at the 30-m point in the last (n) index.

Slides Courtesy of Dr. Elyssa Collins

Find geometric connections



Geometric connections

A ↔ B   B ↔ D

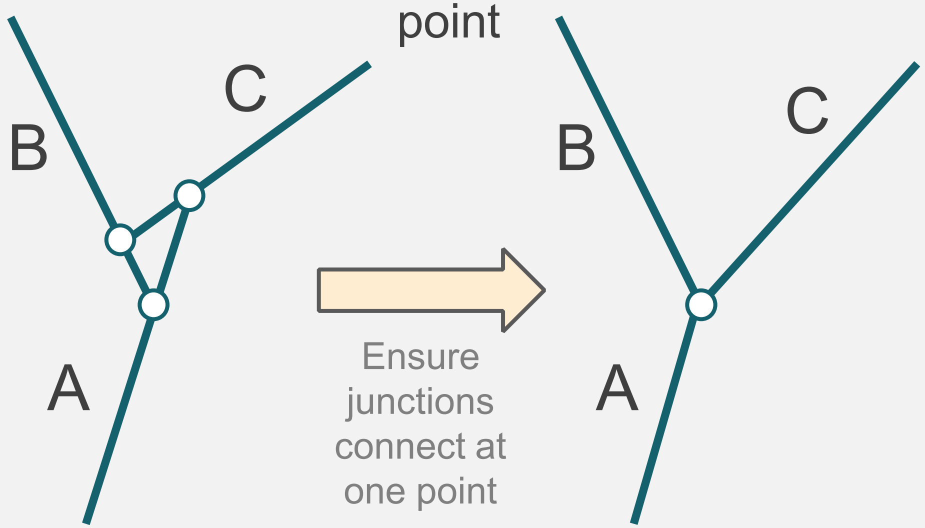
A ↔ C   C ↔ E

B ↔ C

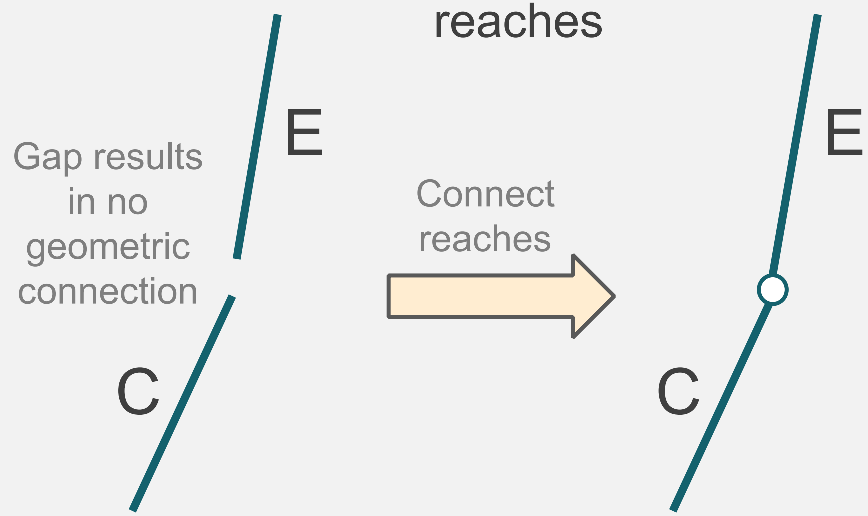
Here, reach E intersects reach C at the 30-m point in the 0 index. Reach C intersects reach E at the 30-m point in the last (n) index.

Find and fix geometry problems

Junctions that don't meet at one point



Incorrect gaps between reaches

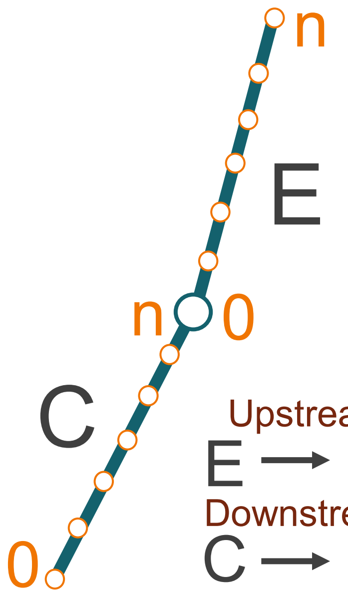


Slides Courtesy of Dr. Elyssa Collins

Geometry to topology – Local

In an ideal case…

Reach
Geometric connection
30-m point



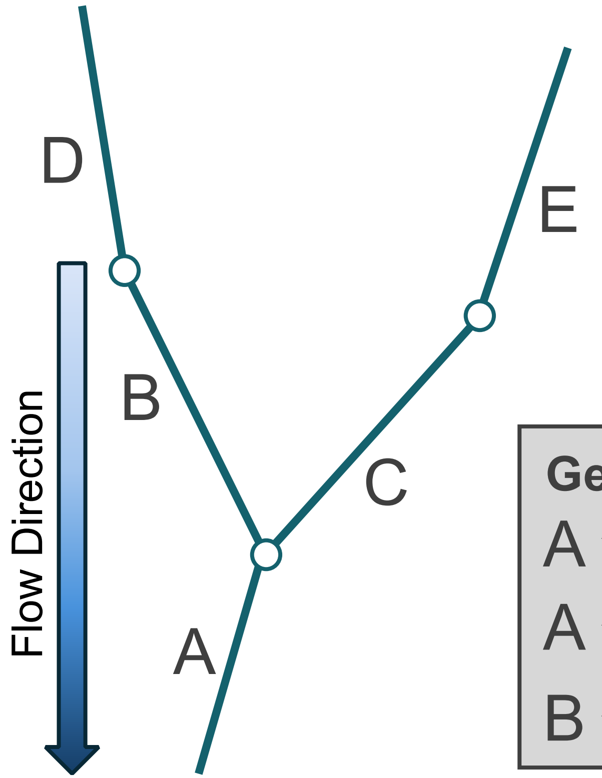
Upstream
E → C = 0

Downstream
C → E = n

Intersection at the 0 index:
- The reach is upstream
  - E is upstream of C

Intersection at the last (n) index:
- The reach is downstream
  - C is downstream of E

Geometry to topology – Global



Using this method, we can go through all geometric connections and define topology

Geometric connections

A ⟷ B
B ⟷ D
A ⟷ C
C ⟷ E
B ⟷ C

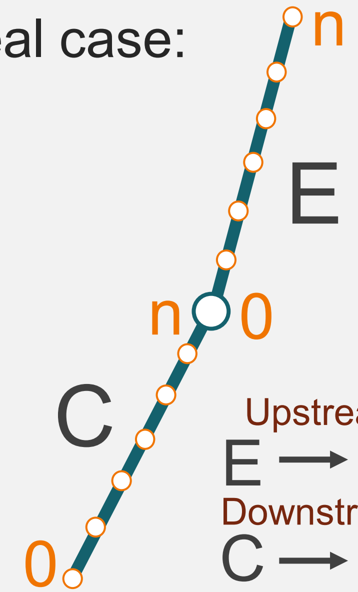
Topology

Downstream	Upstream	Downstream	Upstream
A → B	B → A	B → D	D → B
Downstream	Upstream	Downstream	Upstream
A → C	C → A	C → E	E → C

****Slides Courtesy of Dr. Elyssa Collins****

Geometry to topology – Problem areas

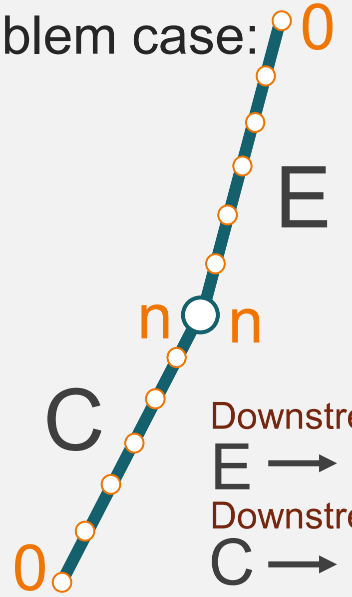
Ideal case:



Reach
Geometric connection
30-m point

Upstream
E ⟶ C = 0
Downstream
C ⟶ E = n

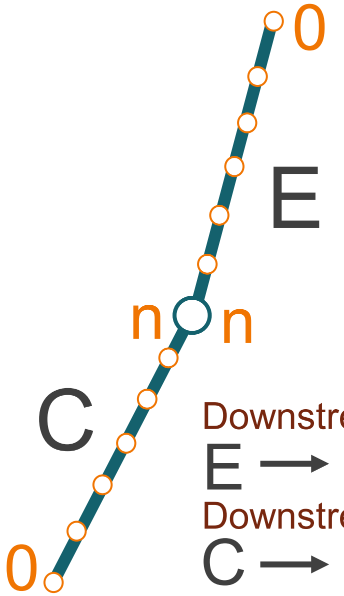
Problem case:



Downstream
E ⟶ C = n
Downstream
C ⟶ E = n

Sometimes the reach LineStrings are not oriented in the correct direction, which causes incorrect and/or impossible topology

Fixing problem areas



Downstream
E ⟶ C = n
Downstream
C ⟶ E = n

Outlet Distance
C = 1000 m
E = 2000 m

$E_{distance} > C_{distance}$

Visit all problem areas and label reach topology based on both geometry and outlet distance

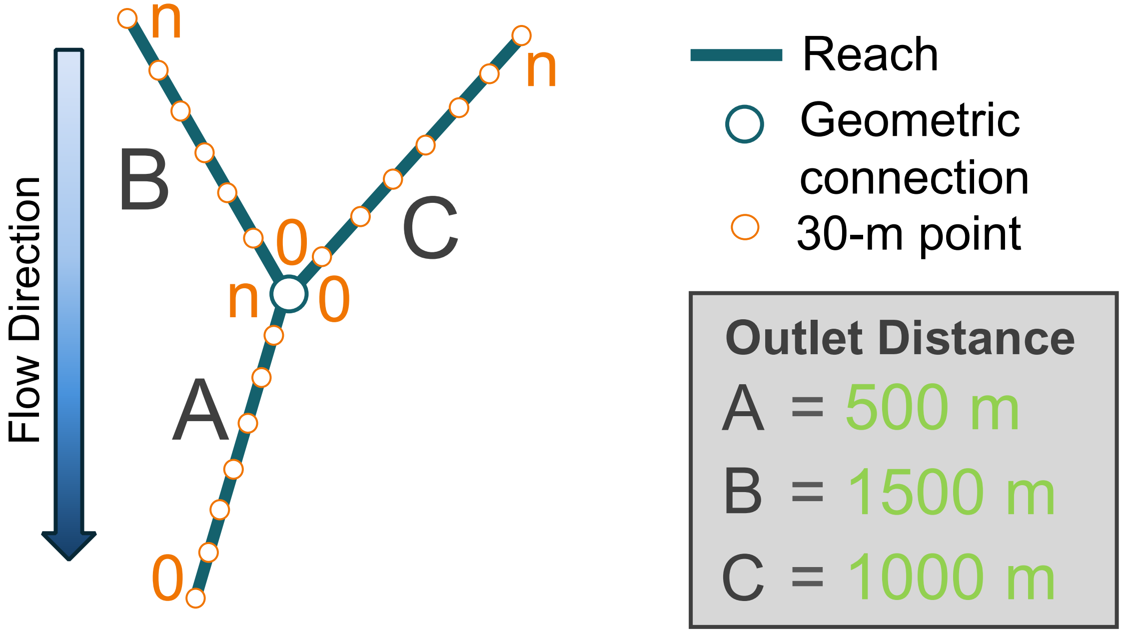
Geometry		Outlet Distance	
Downstream E ⟶ C	= n	Upstream E ⟶ C	2000 m
Downstream C ⟶ E	= n	Downstream C ⟶ E	1000 m

Incongruencies are found for connection E ⟶ C

Therefore, reach E needs to be reversed

Slides Courtesy of Dr. Elyssa Collins

Complicated areas – Junctions



Upstream
C ⟶ A = 0

Upstream
B ⟶ A = 0

Upstream
B ⟶ C = 0

Downstream
A ⟶ C = n

Downstream
A ⟶ B = n

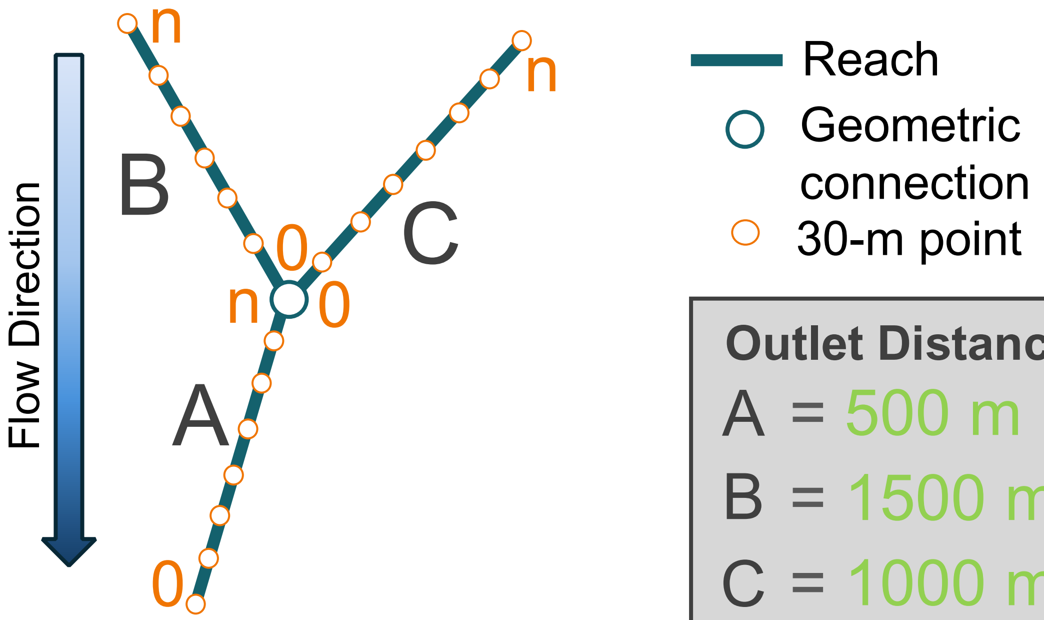
Upstream
C ⟶ B = 0

In this case, reaches B and C would be flagged as a potentially problematic connection.

***We know from visually looking at the network that the geometries for reaches B and C are not problematic. However, we have to determine this via code logic because sometimes LineStrings can be reversed at junctions.

Slides Courtesy of Dr. Elyssa Collins

Complicated areas – Junctions



Outlet Distance

A = 500 m

B = 1500 m

C = 1000 m

Upstream C → A = 0	Upstream B → A = 0	Upstream B → C = 0
Downstream A → C = n	Downstream A → B = n	Upstream C → B = 0

In this case, reaches B and C would be flagged as a potentially problematic connection.

***We know from visually looking at the network that the geometries for reaches B and C are not problematic. However, we have to determine this via code logic because sometimes LineStrings can be reversed at junctions.

Dealing with junctions

Outlet Distance

A $\overset{Min}{=}$ 500 m

B $\overset{Max}{=}$ 1500 m

C = 1000 m

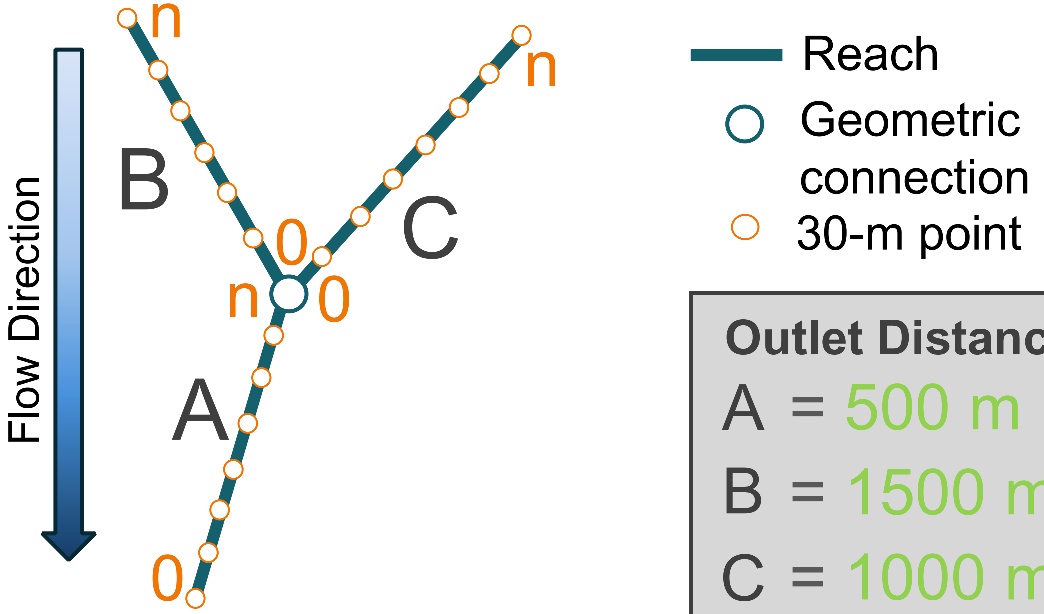
Use outlet distance to label each reach in the junction: the reach with the minimum outlet distance is downstream (A) and the reach with the maximum distance is upstream (B).

? C → A	Upstream B → A	? B → C
? A → C	Downstream A → B	? C → B

However, via this method alone we are unable to label reach C, and therefore its relationship to A and B, as it could be either an additional upstream or downstream reach in the junction.

Slides Courtesy of Dr. Elyssa Collins

Complicated areas – Junctions



Outlet Distance

A	= 500 m
B	= 1500 m
C	= 1000 m

Upstream
C → A = 0
Downstream
A → C = n

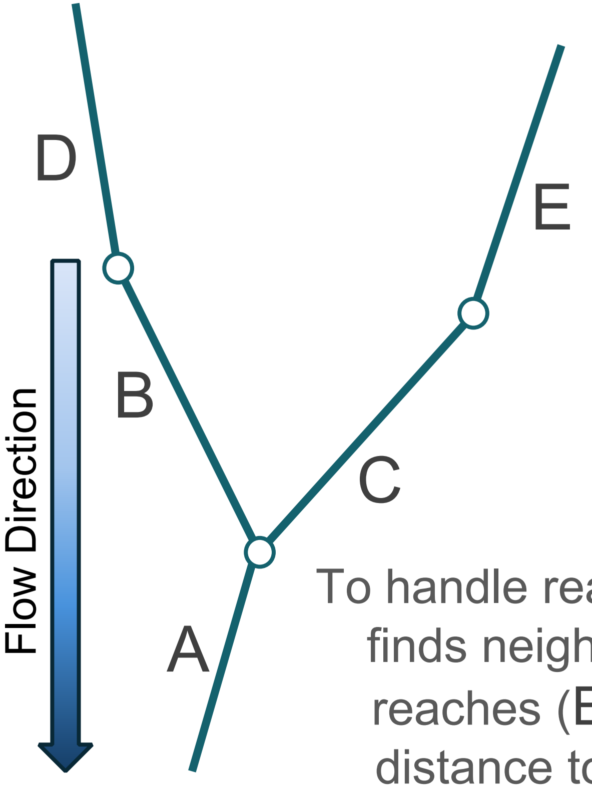
Upstream
B → A = 0
Downstream
A → B = n

Upstream
B → C = 0
Upstream
C → B = 0

In this case, reaches B and C would be flagged as a potentially problematic connection.

***We know from visually looking at the network that the geometries for reaches B and C are not problematic. However, we have to determine this via code logic because sometimes LineStrings can be reversed at junctions.

Dealing with junctions



Outlet Distance

A	= 500 m
B	= 1500 m
C	= 1000 m
D	= 1800 m
E	= 2000 m

To handle reach C, the algorithm finds neighboring connected reaches (E) and uses outlet distance to define topology.

$E_{distance} > C_{distance}$ tells us that reach C is upstream of reach A and therefore reaches B and C make a multi-upstream junction with reach A downstream.

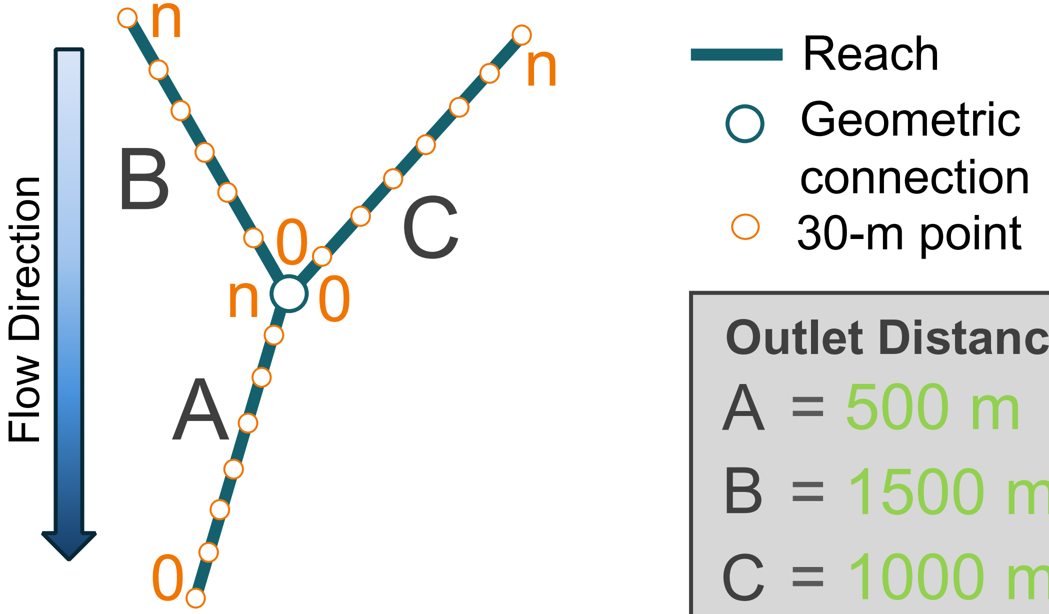
Upstream
C → A = 0
Downstream
A → C = n

Upstream
B → A = 0
Downstream
A → B = n

Junction – upstream
B → C = 0
Junction – upstream
C → B = 0

Slides Courtesy of Dr. Elyssa Collins

Complicated areas – Junctions



Outlet Distance

A = 500 m

B = 1500 m

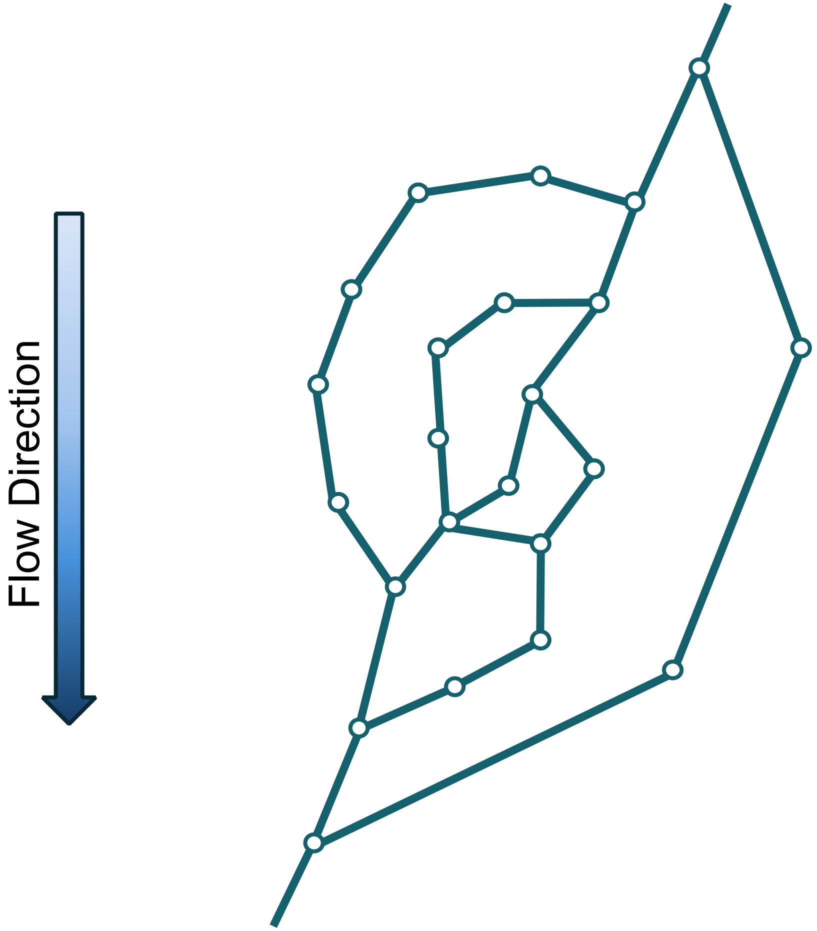
C = 1000 m

Upstream C → A = 0	Upstream B → A = 0	Upstream B → C = 0
Downstream A → C = n	Downstream A → B = n	Upstream C → B = 0

In this case, reaches B and C would be flagged as a potentially problematic connection.

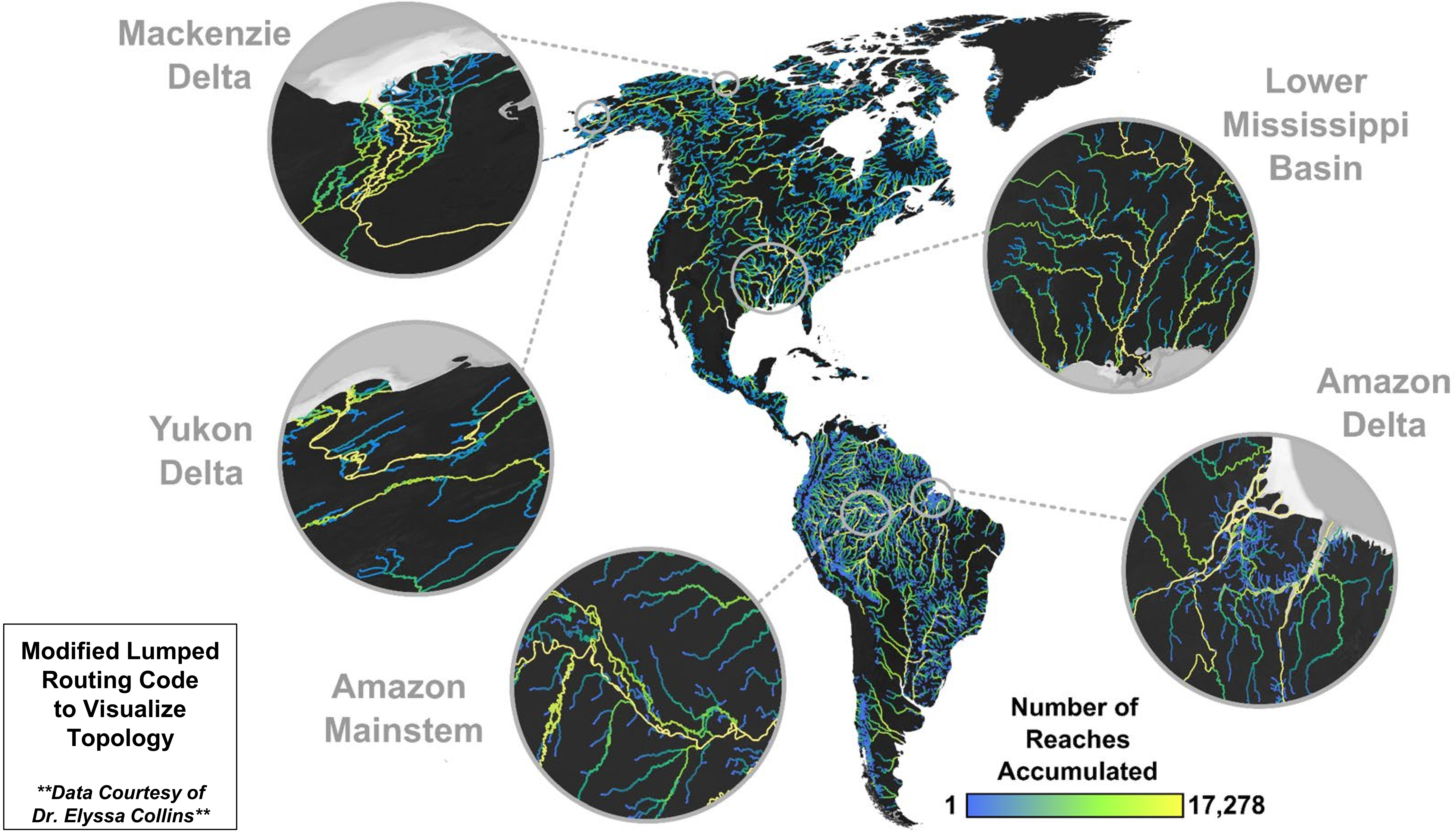
***We know from visually looking at the network that the geometries for reaches B and C are not problematic. However, we have to determine this via code logic because sometimes LineStrings can be reversed at junctions.

Super complex junctions



- Expand out to larger areas than junctions.
- LOTS of logical conditions.
- Rare cases - must manually define topology.
  - 42 reaches in South America needed manual definitions - 0.1%.

Slides Courtesy of Dr. Elyssa Collins


Mackenzie Delta
Lower Mississippi Basin
Yukon Delta
Amazon Delta
Amazon Mainstem
Number of Reaches Accumulated
1
17,278
Modified Lumped Routing Code to Visualize Topology
Data Courtesy of Dr. Elyssa Collins

North & South America SWORD v17 Beta Download:

Updates:

- Final global update estimated Fall 2024
- Topological updates to ensure consistency.
- Distance from outlet re-calculation based on shortest paths between outlets and headwaters.
- New variables for nodes and reaches: “path_freq”, “path_order”, “path_segs”, “main_side”, “stream_order”, “end_reach”,” network”.
- Improved geometry for reach shapefiles.
- Additional channels added for improved network connectivity.
- Centerline shifts in some areas where geolocation errors were present.
- New reach and node ids that reflect improved topology.
- Corrected node lengths to match reach lengths when summed.

Linking MERIT Hydro Basins to the SWOT River Database

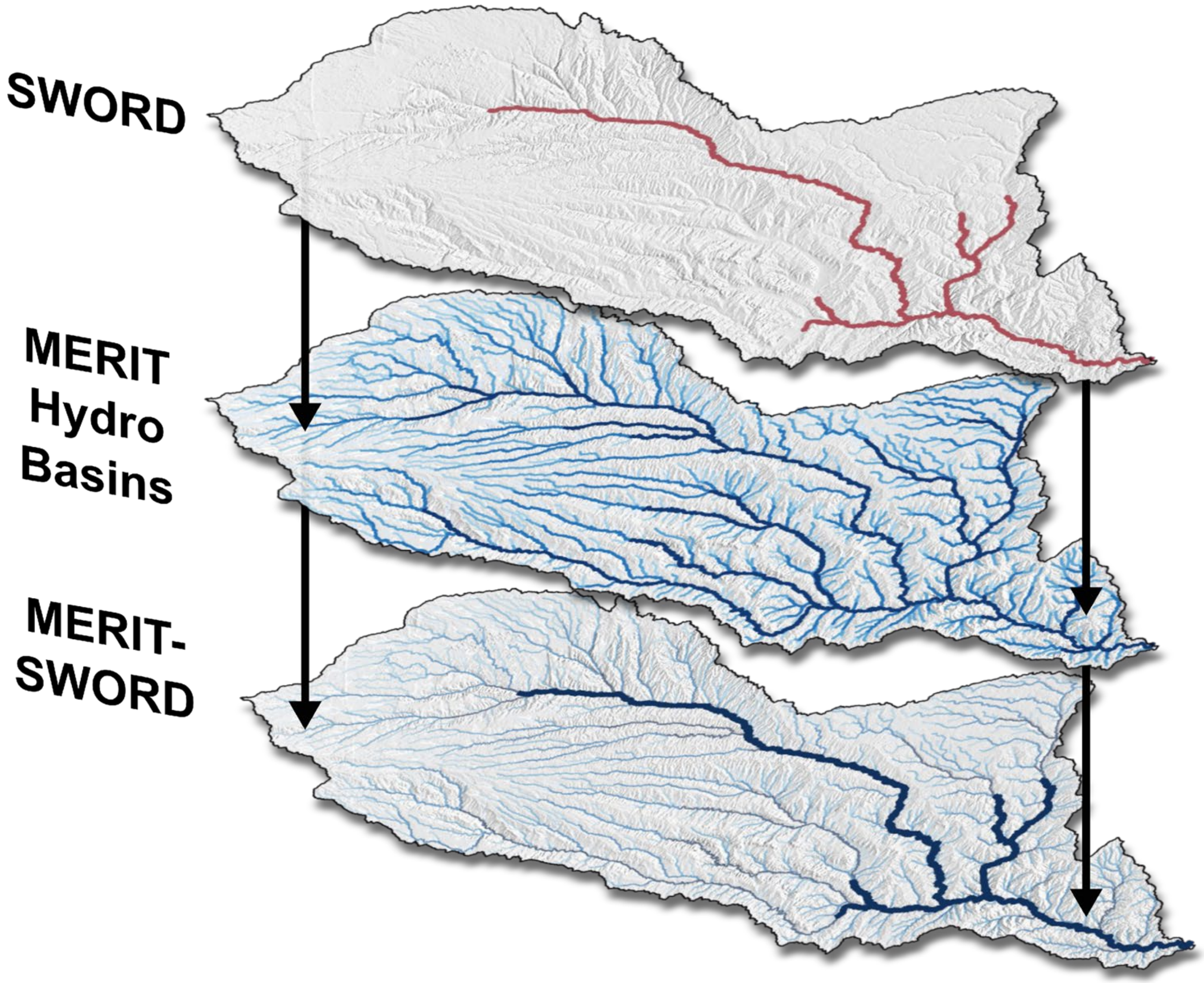
Jeffrey Wade, Cédric David

329F | Water and Ecosystems

Jet Propulsion Laboratory
California Institute of Technology

Jeffrey Wade, Cédric David
Jet Propulsion Laboratory
California Institute of Technology

The MERIT-SWORD Dataset

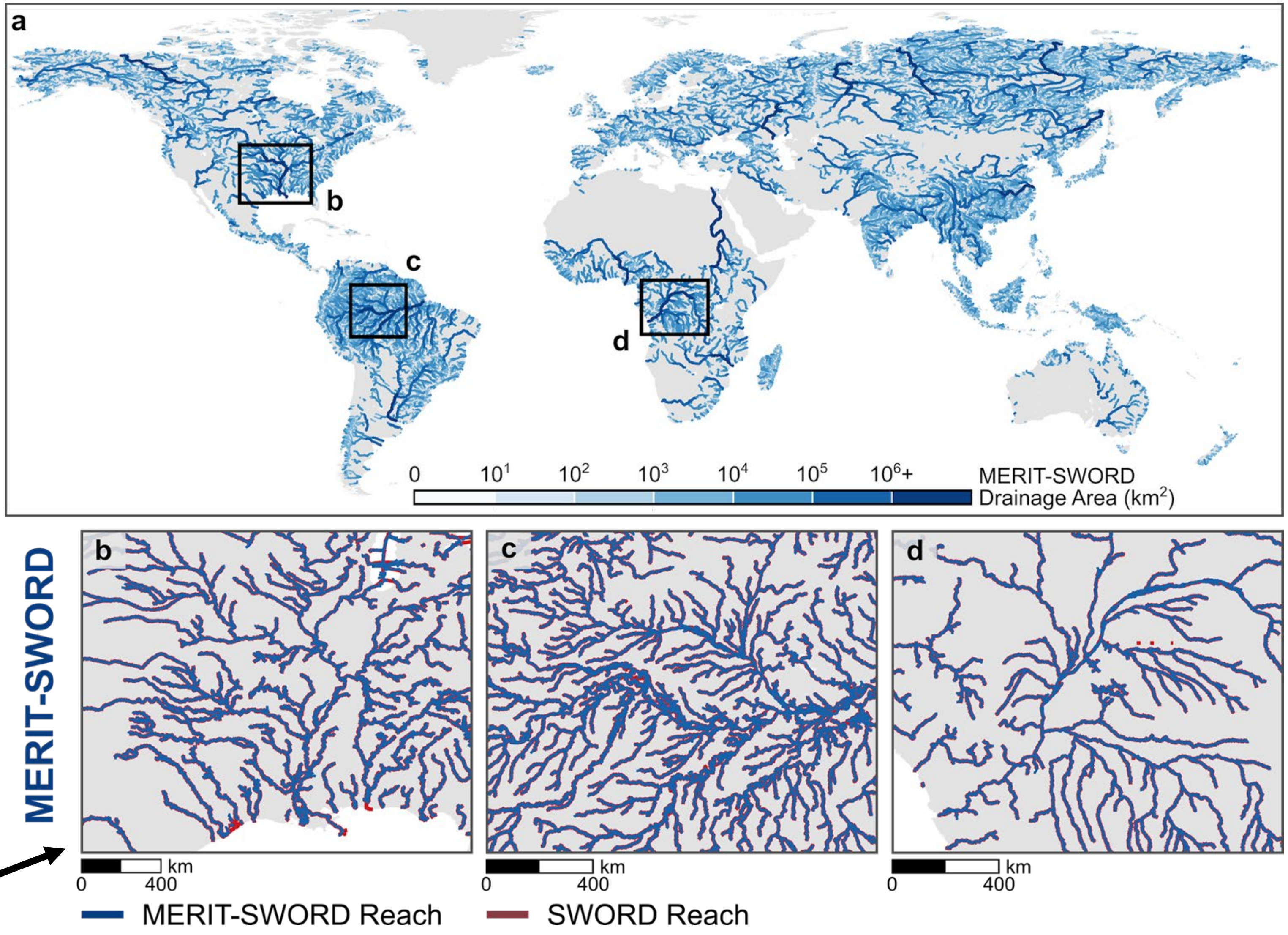


MERIT-SWORD is a data product that links the **MERIT-Hydro Basins** vector dataset to the SWOT River Database (**SWORD**) to enable information transfer between networks, facilitating opportunities for **data assimilation** and **improved a priori estimates**.

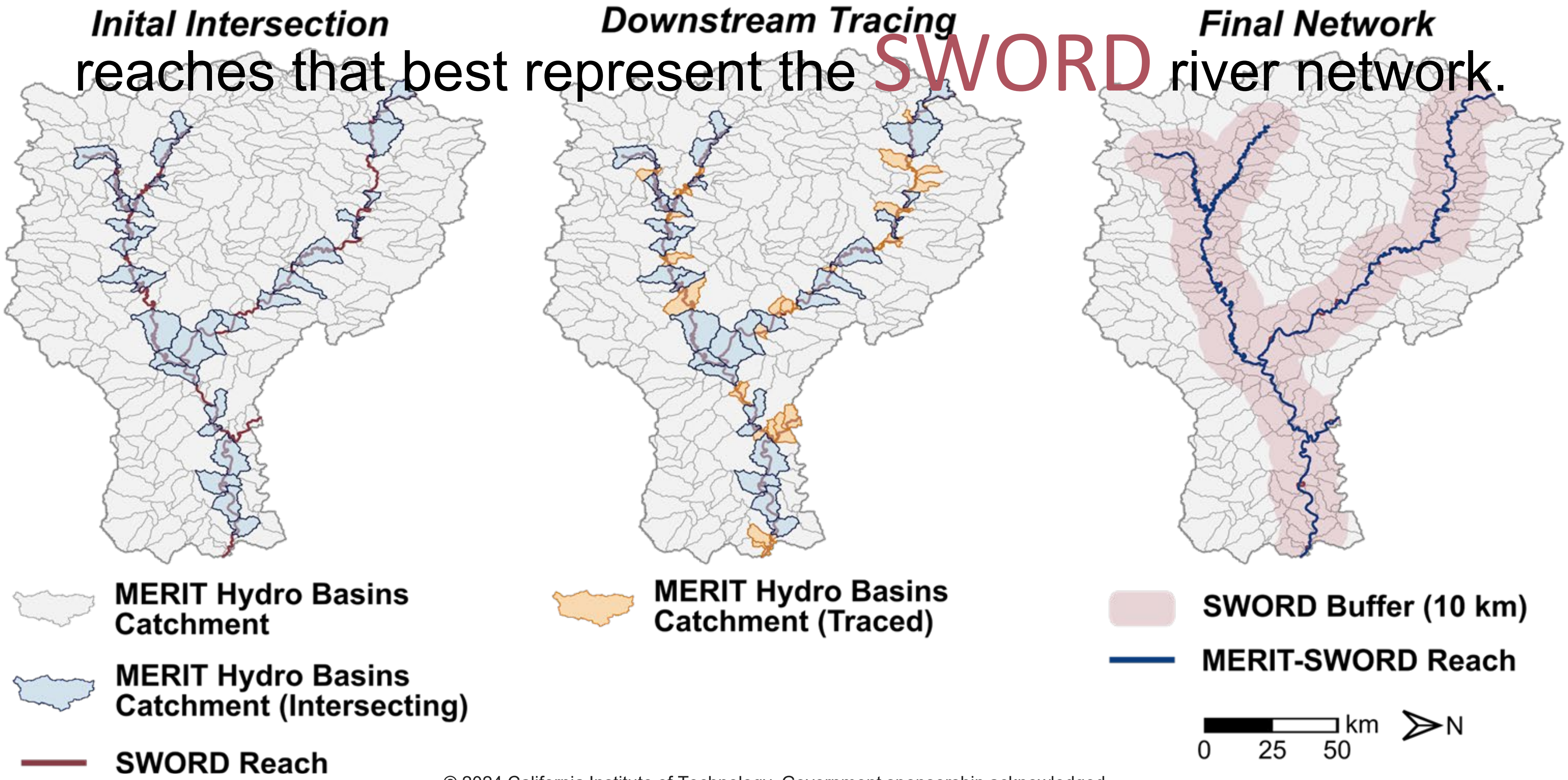
MERIT-SWORD

consists of MERIT Hydro Basins reaches that directly correspond to SWORD reaches and one-to-many

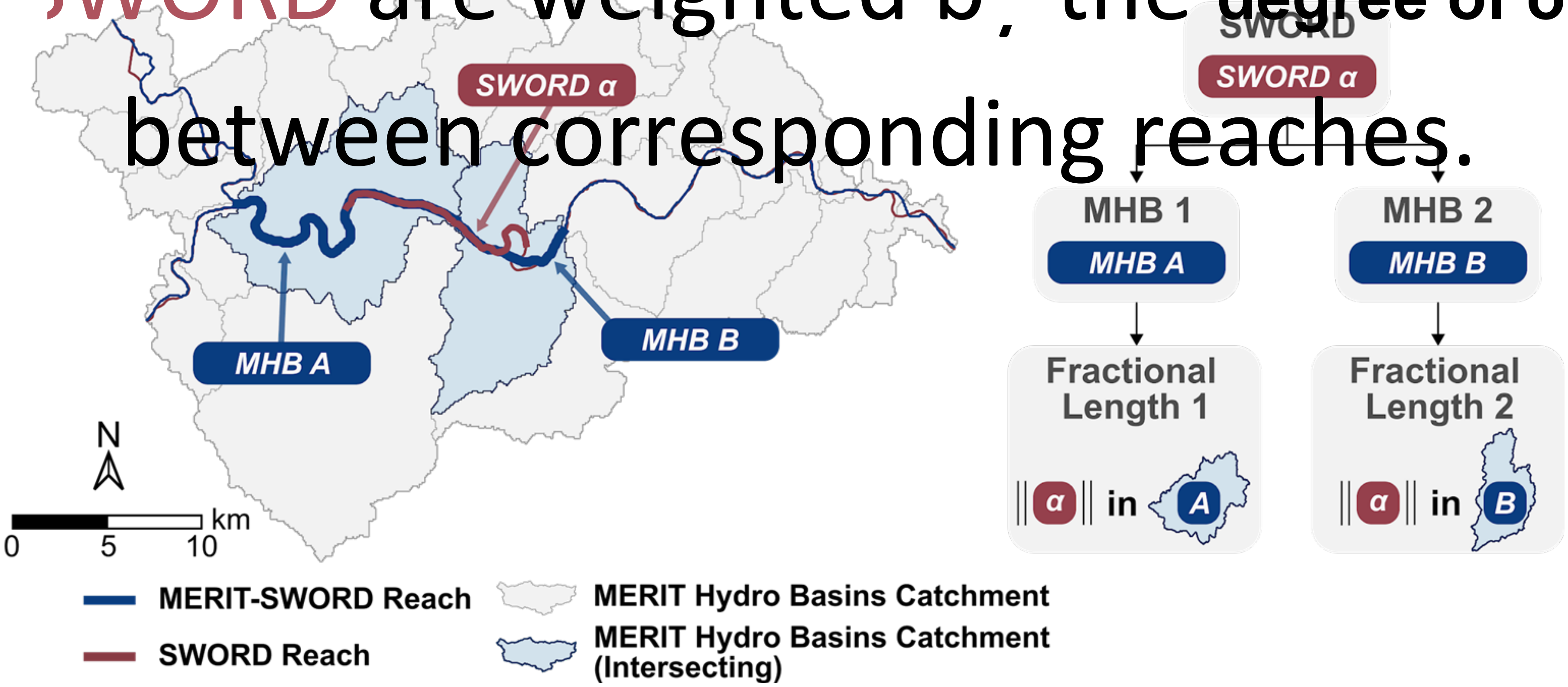
MERIT-SWORD provides excellent coverage of **SWORD** reaches!



We use a combination of **intersection**, **downstream tracing**, and **buffering** to select **MERIT Hydro Basins** reaches that best represent the SWORD river network.



One-to-many translations between MERIT Hydro Basins and SWORD are weighted by the **degree of overlap** between corresponding reaches.



MHB □ SWORD and **SWORD □ MHB translations** and **diagnostic flags** are stored as NetCDF files.