https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Non-dendritic hydrologic networks: foundational patterns, algorithms, and a global empirical demonstration

David Blodgett[1], Taher Chegini[2], Ryan Gillespie[3], and Scott Kehoe[4]

[1]U.S. Geological Survey, Integrated Modeling and Prediction Division, Reston, VA, USA
[2]Resilience Solutions Department, Dewberry, Fairfax, VA, USA
[3]U.S. Geological Survey, National Geospatial Technical Operations Center, Reston, VA, USA
[4]Contractor to the U.S. Geological Survey, National Geospatial Technical Operations Center, Reston, VA, USA

Correspondence: David Blodgett (dblodgett@usgs.gov)

Abstract. Diverted, non-dendritic flow is widespread in hydrologic networks and influences flood forecasting, water-quality assessments, and aquatic habitat analysis, yet this type of flow is inadequately supported by traditional dendritic data structures and algorithms. The hydrofabrics (defined as holistic systems of data that represent the cascade of catchments, waterbodies, flow connectivity, and linked points of interest) that do represent non-dendritic flow have developed independent encodings without shared formalization. This study addresses this gap by making three contributions toward a common foundation. First, it presents a small collection of graph and database patterns — self-referencing, junction, and catchments-as-edges tables, plus primary-path designation — that span the structural choices observed across hydrofabrics internationally. Second, it develops reusable algorithms for non-dendritic networks: network navigation with primary-path filtering, divergence-aware stream order, and a bridge-based total-accumulation procedure. The accumulation procedure uses Tarjan's bridge-finding algorithm to partition the network into a dendritic backbone and diversion groups, then runs bounded local searches within each group. This gives effectively linear scaling on networks dominated by dendritic structure, where the brute-force alternative is quadratic. Formal correctness and complexity arguments are provided, and the procedure complements prior spectral graph-theoretic methods developed for delta channel networks by targeting total drainage accumulation at hydrofabric scale. Third, it applies the patterns and algorithms across six hydrofabrics — NHDPlusV2, NHDPlusHR, the Australian Geofabric, Canada's National Hydro Network, the pan-European EU-Hydro, and the global GRIT (Global RIver Topology) — to give the first, to the authors' knowledge, cross-fabric measurement of non-dendritic prevalence. The measurements show that diverted flow is a pervasive feature of digital hydrography globally. These patterns and algorithms are realized in implementations such as a hydrography data model and a software library. Rather than proposing a comprehensive framework, the paper offers a minimum yet sufficient toolkit: a shared vocabulary for comparing, integrating, and extending non-dendritic hydrologic software across organizational and national boundaries.

1 Introduction

A hydrologic geospatial fabric (hydrofabric) is a holistic system of data that represents the cascade of catchments, waterbodies, flow connectivity, and linked points of interest (e.g., Bondelid et al., 2010; Moore et al., 2025). Diverted, non-dendritic

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

flow, as illustrated in Figure 1, is common in hydrologic networks, yet hydrologic analysis has largely developed around the assumption of strictly dendritic, tree-like connectivity. This assumption breaks down across a range of operational and scientific applications. Continental-scale flood and streamflow forecasting in coastal-plain, deltaic, and reservoir-managed settings depends on routing flow through divergences and managed splits that single-path models cannot reproduce (e.g., Cosgrove et al., 2024; Wang et al., 2021). Pollutant source attribution and basin-scale water-quality accounting require accurate upstream contributing areas. Large diversions such as the Old River Control Structure, which routes roughly a third of Mississippi River streamflow and the nutrient load it carries to the Atchafalaya River, redistribute upstream contributions in ways that dendritic-only mass balances cannot represent (e.g., Schwarz et al., 2006; Robertson and Saad, 2021). In aquatic ecology, side channels, anabranches, and distributary reaches act as connectivity refugia and habitat that disappear from purely dendritic representations, biasing fish-passage and habitat-connectivity assessments (e.g., Fullerton et al., 2010; Peterson et al., 2013). Hydrofabric implementations that represent non-dendritic flow have developed their own graph and database structures without shared formalization (e.g., McKay et al., 2012; Bureau of Meteorology, 2015; INSPIRE Thematic Working Group Hydrography, 2024; Wortmann et al., 2025).

Drainage basins tend toward dendritic, tree-like structure (Leopold et al., 1964), but non-erodible geology, sediment build-up, fluvial geomorphic processes, and engineered diversions cause flow to diverge from this pattern. The locations at which splitting occurs are *divergences*, and the flowlines downstream of a divergence that have not recombined with the flow from which they diverged are *diversions*. At coarse resolutions non-dendritic flows are often negligible because most diversions return within a local basin and inter-basin diversions are small relative to dendritic flow (Brakebill et al., 2020; Leopold et al., 1964); however at the resolution of national and global hydrofabrics they are not negligible (see Section 5.2). *We use "dendritic" throughout in the graph-theoretic sense (tree-structured: every node has at most one downstream neighbor), distinct from the drainage-pattern classification sense (Howard, 1967), in which "dendritic" describes planform shape and is contrasted with trellis, radial, rectangular, or parallel patterns regardless of the presence of diversions.*

Purely dendritic hydrography products openly acknowledge their omission of non-dendritic structure: MERIT Hydro states that "the bifurcating channels in the delta region cannot be fully represented because only one downstream direction was assumed at each pixel" (Yamazaki et al., 2019). Brazil's national hydrography (Base Hidrográfica Ottocodificada, BHO) is published as "uma rede unifilar… em formato de grafo (árvore) e sem confluências duplas" — a single-line, tree-shaped graph without double confluences (Agência Nacional de Águas, 2017). Recent work directly counters this gap: GRIT (Global RIver Topology) delivers a bifurcating global hydrography with tens of thousands of explicit divergences (Wortmann et al., 2025), and the Enhanced NHDPlusV2 (e2NHDPlus) of Brakebill et al. (2020) documents that mishandling its more than 64,000 divergences "can result in large errors in watershed parameters derived from the network."

This paper makes three contributions:

1. A small set of graph and database patterns — self-referencing, junction, and catchments-as-edges tables, plus primary-path designation — that span the structural choices observed across hydrofabrics internationally.
2. Reusable algorithms for non-dendritic networks at hydrofabric scale, including a bridge-based total-accumulation procedure that runs in effectively linear time on networks dominated by dendritic structure.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

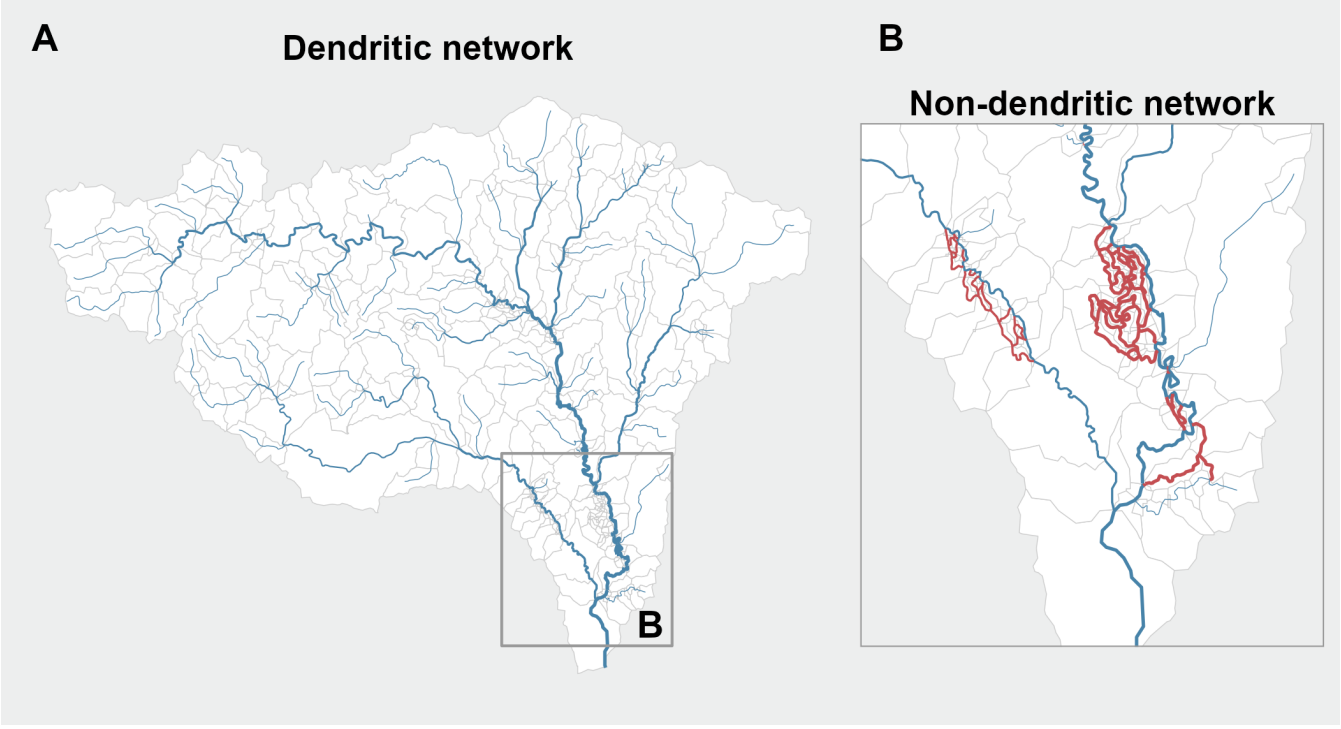

Figure 1. A) A dendritic hydrologic network from the NHDPlusV2 dataset (McKay et al., 2012) with flowlines scaled by stream order. B) The same network showing non-dendritic flowlines (red) that are removed when the network is simplified to a purely dendritic representation. Non-dendritic flowlines are identified as those where the divergence-aware stream order (see Section 4.3) differs from the classical Strahler order (Strahler, 1957), indicating secondary flow paths through diversions.

3. A cross-fabric measurement of non-dendritic prevalence and diversion-group size distribution applied across six hydrofabrics spanning three continents and a global compilation, demonstrating that diversions are a pervasive feature of digital hydrography.

The three contributions build on each other as a cohesive set of tools. The patterns of Section 3 supply a shared vocabulary that lets one set of algorithm implementations work with any of the underlying representations. The algorithms of Section 4 are reusable and scale to continental and global hydrography. Cross-fabric measurements of Section 5 apply a single methodology to six independently developed hydrofabrics, demonstrating how the patterns and algorithms transfer across data models. As a whole, these three contributions are offered as a foundation for further study of diverted hydrology. Interpretation of the empirical results is intentionally limited — Section 7 leaves their implications and cross-fabric questions open.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

2 Background

Digital hydrology has historically treated drainage systems as directed dendritic networks: 1) cell-to-cell digital elevation model (DEM) routing (O'Callaghan and Mark, 1984), 2) edge-node stormwater models (Metcalf & Eddy, Inc. et al., 1971), and 3) graph-theoretic analyses of dendritic topology (Tarboton et al., 1988), that all assumed tree-like structure. Horn (1986) established the foundational network topology still used in modern hydrofabrics (directed flowlines, called "reaches" in that lineage, connected by from-node/to-node identifiers, with a divergence code at divergences classifying flowlines as major or minor flow recipients). This pattern persisted through ERF1 (Nolan et al., 2002), the National Hydrography Dataset (U.S. Geological Survey in cooperation with the U.S. Environmental Protection Agency, 1999), the ArcHydro node-edge schema (Maidment et al., 2001), and Pfafstetter hierarchical coding (Pfafstetter, 1989). Nolan et al. (2002) introduced a nearest-neighbor catchment data model capable of explicit non-dendritic semantics. NHDPlusV1 (Bondelid et al., 2010) was among the first broadly applied frameworks to pair elevation-derived catchments with both dendritic and non-dendritic attributes. NHDPlusV2 (McKay et al., 2012) and NHDPlusHR (Moore et al., 2025) retain the same network-representation data model and provide the foundational scope for this paper. NHDPlusHR adds the NHDPlus attribute model and catchments to the high-resolution NHD geometry; the two are related but not synonymous.

2.1 NHDPlus non-dendritic representation

The NHDPlus data model (Moore et al., 2025) centers on two feature types: flowlines (linear geometry) and catchments (polygons). Most flowlines pair with a catchment and vice versa, with exceptions for sinks (catchments without a flowline) and small or pipeline-like flowlines that do not fall within a distinct catchment.

NHDPlus encodes connectivity through from-node and to-node identifiers (`FromNode` and `ToNode` in NHDPlusV2) on each flowline: two flowlines connect when the upstream feature's to-node equals the downstream feature's from-node. Although node identifiers are unique, NHDPlus maintains no separate node table; nodes exist only as implicit endpoints of flowline geometries. The from-node/to-node construct permits many-to-many relationships in both directions. To distinguish main from non-main paths, NHDPlus adds two attributes: a stream-level value (`StreamLevel` in NHDPlusV2), derived from stream name and relative basin size, which identifies the primary upstream connection at confluences; and a divergence code (`Divergence` in NHDPlusV2), assigned 1 (primary) or 2 (diverted), which identifies the primary downstream connection at divergences. Together, these support upstream and downstream navigation with optional tributary or diversion filtering. Brakebill et al. (2020) corrected primary-path assignments at divergences, producing the Enhanced NHDPlusV2 (e2NHDPlus) and introducing the concept of "diversion groups" (sets of diverted paths whose flow recombines downstream), but did not formalize the underlying graph or database patterns.

2.2 High-resolution NHD and 3DHP non-dendritic representation

The high-resolution NHD (U.S. Geological Survey, National Geospatial Program, 2024) (hereafter "HR NHD"; distinct from the medium-resolution NHD, which uses node identifiers in the NHDPlusV2 style) lacks catchments and represents connec-

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

tivity in a "flow table" with one row per upstream-to-downstream flowline connection. HR NHD uses no node identifiers and does not record primary upstream or downstream connections; headwater and outlet flowlines, which do not connect to flowlines upstream and downstream respectively, are coded as connecting from or to identifier "0". The 3D Hydrography Program (3DHP) data model (Blodgett et al., 2024; U.S. Geological Survey, 2025) extends this flow-table architecture into a "flow network" by adding upmain-and-downmain boolean columns (`upmain` and `downmain` in 3DHP) that flag primary connections, and eliminates the "0" convention by representing headwaters as flowlines that appear only in the from-identifier column and outlets only in the to-identifier column.

2.3 International context

International hydrography datasets exhibit similar patterns. Australia's Hydrological Geospatial Fabric (Bureau of Meteorology, 2015) represents diversions in its fine-grained features but forces dendritic topology at the reporting layer and collapses divergent terminal networks (e.g., deltas) into single "diffuse" nodes. Canada's National Hydro Network (NHN) (Natural Resources Canada, 2024) encodes connectivity via `FROM_JUNCT` / `TO_JUNCT` junction universally unique identifier (UUID) columns with separate `DIRECTION` and `ISOLATED` flags; its HY_Features-aligned successor, the Canadian Hydrospatial Network (Natural Resources Canada, 2025), replicates the NHDPlus divergence-code and primary-flow designation. The pan-European EU-Hydro (Copernicus Land Monitoring Service, 2019) encodes divergences explicitly but designates no primary path, and the underlying European INSPIRE Hydrography specification (INSPIRE Thematic Working Group Hydrography, 2024) provides a node-edge graph model that acknowledges divergences but prescribes no divergence codes. The Global River Topology data product (Wortmann et al., 2025) makes divergences first-class graph elements via multi-edge directed connectivity. The Pfafstetter basin coding system (Pfafstetter, 1989), adopted nationally in Brazil and extended to Australia by Stein (2018), is dendritic by mathematical design and accommodates non-dendritic channels only via ad-hoc workarounds; the purely dendritic products acknowledged in Section 1 (MERIT Hydro, Brazil's BHO) document this restriction explicitly. Across these examples, divergences are 1) omitted due to network simplification, 2) handled using the NHDPlus method, 3) encoded without primary-path designation, 4) made first-class, or 5) omitted entirely — **with no shared graph or database formalization across products**.

2.4 HY_Features conceptual building blocks

The Open Geospatial Consortium (OGC) WaterML2 Part 3: Surface Hydrology Features (HY_Features) standard (Blodgett and Dornblut, 2018) specifies a conceptual model of hydrologic features and the relationships between them. HY_Features introduces an abstract "catchment" as a holistic hydrologic unit that has several conceptual views or "realizations" (a catchment area, a flowpath from inlet to outlet) and nexuses connecting contributing and receiving catchments, but it does not specify implementation logic or physical data model encoding.

The Mainstems logical data model (Blodgett, 2022) interpreted these concepts for cross-scale integration of rivers and drainage basins but deferred discussion of non-dendritic network behavior. Subsequent work on a hydrofabric logical model (Blodgett, 2026b) extended HY_Features concepts to hydrologic modeling and river corridor use cases. These treatments and

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Blodgett et al. (2023) assumed predominantly dendritic connectivity; the present paper provides the patterns and algorithms needed to extend HY_Features to non-dendritic networks.

2.5 Graph-theoretic treatments of non-dendritic networks

A parallel body of research has applied graph theory directly to non-dendritic delta and distributary channel networks. Tejedor et al. (2015a) applied spectral graph theory (eigendecomposition of the weighted Laplacian of a directed acyclic graph) to compute apportioned flux on delta networks and identify, for each outlet, the "exclusive part" of nodes that drain only to that outlet versus the "common part" of nodes that drain to multiple outlets. Tejedor et al. (2015b) extended this with topologic and dynamic complexity metrics for delta comparison and vulnerability assessment. Schwenk and Hariharan (2021) implemented the RivGraph Python package for extracting graph topology from remote-sensing imagery of river and delta channel networks. Aho et al. (2023) used directed acyclic graphs to track connectivity changes in non-perennial networks. These efforts focus on delta morphology and channel-network science; they model flowlines as edges in a directed graph without catchment partitions or hydrofabric integration. The present paper differs in framing (hydrofabric data-model infrastructure rather than delta-science applications), in algorithmic approach (combinatorial bridge-based partitioning rather than spectral decomposition), and in target problem (total drainage-area accumulation, distinct from the apportioned steady-state flux these prior approaches compute).

2.6 Toward a shared formalization

As the review above shows, hydrofabric implementations have repeatedly converged on similar structural choices without naming them as patterns or sharing algorithms across implementations. The graph-theoretic delta-network literature has remained largely separate from the hydrofabric data-model literature. The remainder of the paper closes that gap by developing shared patterns (Section 3), algorithms (Section 4), and a cross-fabric empirical baseline (Section 5).

3 Graph and Database Patterns

When modeled as entities rather than a continuum, hydrologic phenomena are represented using standard database and graph patterns. The patterns presented here are interchangeable building blocks rather than components of a single canonical data model. An implementation may adopt any of the table patterns introduced below (self-referencing, junction, or catchments-as-edges) and still share the algorithms presented in the following section. For algorithm reuse, each pattern must expose the same underlying graph structure.

3.1 Entities

The patterns draw on three vocabularies, listed in Table 1. The hydrologic terms identify the physical features being represented; the graph terms identify the abstract structures used to relate them; and the database terms identify how the relationships are stored.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Table 1: Entities from the three vocabularies. The parenthetical letter denotes the source vocabulary: H = hydrologic, G = graph-theoretic, D = database. Graph-theoretic definitions follow Bondy and Murty (1976) and Cormen et al. (2022); hydrologic terms follow the HY_Features standard (Blodgett and Dornblut, 2018) and NHDPlus usage (McKay et al., 2012); database terms follow standard relational-database usage. V = set of nodes; E = set of edges.

Term	Definition
Catchment (H)	The basic landscape partition; paired with at most one flowpath and at most one inflow and one outflow nexus.
Nexus (H)	A hydrologic connection point; represents confluences, divergences, or more complex hydrologic routing locations.
Flowline (H)	A linear waterbody representation that may or may not have a local catchment.
Flowpath (H)	The single linear path from a catchment's inlet to its outlet; if represented by multiple flowlines, those flowlines form a linear collection along that path (no branches). A flowpath has a known local catchment.
Divergence / diversion (H)	A *divergence* is the splitting node at which one flow path diverges into two or more downstream paths. A *diversion* is a flowline downstream of a divergence that has not recombined with the flow from which it diverged.
Bridge / cut edge (H/G)	A flowline whose removal splits the network into two pieces with no remaining flow path between them; equivalently, an edge that lies on no cycle in the undirected projection. *Bridge* and *cut edge* are synonyms; we use *bridge*.
Diversion group / connected component (H/G)	A *diversion group* is a set of diversion flowlines connected to one another by parallel flow paths (a system of diversions), bounded by the bridges where flow enters and leaves. In graph terms, the non-bridge flowlines of a diversion group form a *connected component* of the non-bridge subgraph; a *connected component* in general is a maximal set of nodes reachable from one another through the graph.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Term	Definition
Bipartite / incidence graph (G)	A *bipartite graph* is one whose nodes can be sorted into two disjoint groups such that every edge connects a node in one group to a node in another. An *incidence graph* is constructed by promoting each edge of a graph to a new node connected to its original endpoints; incidence graphs are always bipartite.
Directed acyclic graph (DAG) (G)	A directed graph that contains no directed cycle. Flow networks are DAGs because water does not flow uphill.
Path / cycle (G)	A *path* is a sequence of edges connecting two nodes through intermediate nodes. A *cycle* is a path that returns to its starting node.
Depth-first search (DFS) (G)	A graph traversal that follows edges as far as possible before backtracking; visits each reachable node and edge once in time proportional to $V + E$.
Topological sort (G)	An ordering of a DAG's nodes such that every edge points forward in the order; produced in time proportional to $V + E$.
Node (G)	A multi-purpose abstract graph entity (could represent anything).
Edge (G)	A relation or connection between nodes.
Junction table / self-referencing table (D)	A *junction table* encodes edges between rows of another table by listing from-and-to pairs; suitable for many-to-many connectivity without row duplication. A *self-referencing table* encodes edges as an attribute on the same table whose rows are the connected entities; requires extra rows wherever a node has multiple outgoing edges.
Attribute (D)	A single value of a specific type.
Geospatial attribute (D)	An attribute containing a geometry with a known coordinate system.
Row (D)	A set of attribute values that together describe one entity (e.g., one catchment or one flowline).
Table (D)	A collection of rows with the same attributes.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

3.2 How hydrologic entities are modeled

With the entities defined, the question is how to represent them computationally. Two complementary modeling decisions determine the answer: which hydrologic entities map to which graph entities (i.e., what the *nodes* are and what the *edges* are), and how those graph relationships are encoded in tabular form. The first decision is a variant of the *primal-vs-dual* choice named in the spatial-network literature (Porta et al., 2006a, b); the second is a database-design choice. The two decisions are orthogonal, and the algorithms of Section 4 are compatible with any combination consistent with them.

3.2.1 Graph patterns

In the spatial-network literature, *primal* refers to the mapping in which the linear feature of the domain is an edge and point-like features are nodes, whereas *dual* refers to the opposite mapping, in which the linear feature is promoted to a node and its intersections become edges — formally, the *line graph* of the primal (Porta et al., 2006a, b). Hydrography has a mixed set of features: flowlines are linear, nexuses are point-like, and catchments are areal. Applied to flowlines, the primal mapping places flowlines as edges and nexuses as nodes; this is the natural mapping for navigation and accumulation where nexus entities are assigned node identifiers. The dual mapping places flowlines as nodes with confluences as edges and has been used in drainage-pattern classification (Yu et al., 2022) and is used by the high-resolution NHD flow table (U.S. Geological Survey, National Geospatial Program, 2024). Catchments are areal, but they have a linear expression — the flowpath from inlet to outlet. By analogy with the flowline case, *catchments-as-edges* places the catchment (via its flowpath) on the edge and nexuses on the nodes, and *catchments-as-nodes* places the catchment on the node and uses edges for the connecting relationships between catchments.

Catchments-as-edges requires nexus identifiers at every connection point, including headwater and outlet nexuses that have no cartographic counterpart; in return, many-to-many connectivity at confluences and divergences is structural — a nexus with three inputs and two outputs simply has five incident edges — and never needs a workaround. *Catchments-as-nodes* leaves nexuses implicit and so avoids the identifier bookkeeping, but when edges are stored on the catchment row itself it requires extra rows wherever a node has multiple downstream connections, because a single row can encode only one downstream edge per catchment. Catchments-as-edges is the natural choice when nexus identifiers are available; catchments-as-nodes is simpler when they are not.

Two constructs from graph theory formalize these structures: *bipartite* graphs and *incidence* graphs. A bipartite graph is one whose nodes can be sorted into two disjoint groups such that every edge connects a node in one group to a node in the other (see Figure 3). An incidence graph is constructed from any graph by promoting each edge to a new node, then connecting that new node to the two original nodes the edge used to join; the result is always bipartite. Applied to the primal graph in which catchments are edges and nexuses are nodes, the incidence-graph construction promotes each catchment to a node joined to its inflow and outflow nexuses, producing a catchment-nexus bipartite graph (see Figure 3). The catchment-to-catchment graph (see Figure 2), in which each directed edge represents flow from one catchment into the next, is the same network with nexuses left implicit: each of its edges corresponds to a pair of catchments that share a nexus.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

The practical consequence is an equivalence: a catchments-as-edges table with columns `id`, `fromnode`, and `tonode` specifies the primal (catchments-as-edges) graph and its bipartite incidence projection simultaneously. The `id` column names an edge in the primal reading and a node in the incidence reading; the two nexus-identifier columns supply both the edge endpoints of the primal graph and the bipartite edges of the incidence graph. The two readings contain the same information at different levels of abstraction; adopting one commits to both.

3.2.2 Database patterns for hydrographic networks

The two decisions of Section 3.2.1 (flowline primal vs dual, catchment edges vs nodes) combine with database-design choices to produce three common table patterns: self-referencing, junction, and catchments-as-edges.

A "self-referencing" table expresses network connectivity by including an attribute that references other rows in the same table. In graph terms, each row is a catchment node and the self-reference defines the connecting edge — a catchments-as-nodes edge list. Self-referencing tables cannot represent a node with multiple downstream connections without adding extra rows that repeat the catchment's other attributes. A "junction table" addresses this by encoding edges in a separate table: one table holds catchments and a second table holds the "from" and "to" connections between them, preserving the catchments-as-nodes mapping while accommodating many-to-many connectivity without row repetition.

The third pattern follows directly from the equivalence of Section 3.2.1. Adding `fromnode` and `tonode` columns to a catchment table turns each row into an edge endpoint specification: an inflow and an outflow nexus identifier per catchment, with the catchment's `id` simultaneously naming the primal edge and the corresponding bipartite node. Many-to-many connectivity at nexuses is carried structurally by shared identifier values. Table 2, the Table 3 / Table 4 pair, and Table 5 illustrate the same seven-catchment non-dendritic network — with a divergence at catchment 3 and recombination at catchment 7 — encoded in each of the three patterns.

Table 2: Self-referencing table. Each row is a catchment; `toid` references another row. Catchment 3 must have two entries to express the divergence to both catchments 4 and 5.

id	toid	color
1	3	blue
2	3	green
3	4	magenta
3	5	magenta
4	6	indigo
5	7	gold
6	7	teal
7		orange

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Table 3: Junction-table pattern — catchment table. Stores catchment attributes; each row is one catchment. Paired with the connections table (Table 4).

id	color
1	blue
2	green
3	magenta
4	indigo
5	gold
6	teal
7	orange

Table 4: Junction-table pattern — connections table. One row per directed edge between catchments in Table 3; carries many-to-many connectivity without duplicating catchment attributes.

from	to
1	3
2	3
3	4
3	5
4	6
5	7
6	7

Table 5: Catchments-as-edges / bipartite table. Each catchment appears once with inflow (`fromnode`) and outflow (`tonode`) nexus identifiers. Shared node identifiers encode the many-to-many topology. The `id` column names an edge in the catchments-as-edges reading and a node in the incidence-graph reading of the same data.

id	fromnode	tonode	color
1	a	c	blue
2	b	c	green
3	c	d	magenta
4	d	e	indigo

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

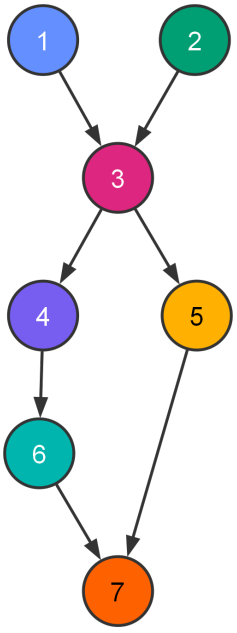

Figure 2. Catchment-to-catchment graph representation of the network in Table 2–Table 5.

id	fromnode	tonode	color
5	d	h	gold
6	e	h	teal
7	h	i	orange

The catchments-as-edges / bipartite form supports non-dendritic networks without a junction table but requires an additional set of identifiers. Hydrologically, these nexus identifiers represent three distinct entity types: headwater initiation nexuses, catchment interaction nexuses, and network outlet nexuses. Non-dendritic networks — where a catchment connects to more than one downstream neighbor — can be represented in any of the three patterns, but self-referencing tables require extra rows at every such divergence whereas junction and catchments-as-edges tables carry the many-to-many connectivity structurally.

3.3 Relationship of diverted networks to dendritic networks

In a dendritic network, each catchment flows to at most one downstream catchment. When a divergence occurs at a catchment's outlet, one downstream path is designated primary and the others as non-primary. Recall that a flowpath is the path through a catchment from inlet to outlet, and a flowline is any linear waterbody representation; non-primary paths downstream of a divergence may be flowlines without a local catchment rather than flowpaths. Some applications refine this distinction by

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

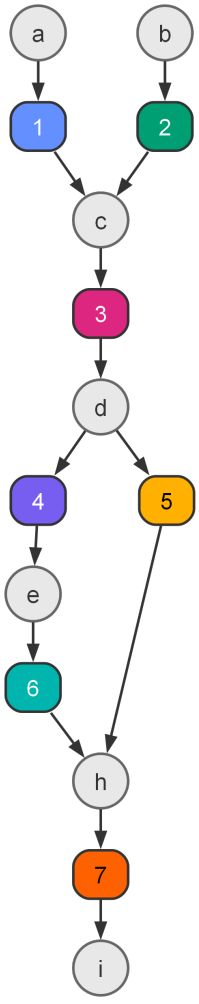

Figure 3. Bipartite graph representation of the same network. Circles are nexuses (the nodes listed in Table 5); rounded rectangles are catchments.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

assigning divergence fractions to the primary path and each diversion. Identifying a single primary downstream path matters because certain hydrologic network algorithms require one primary path through systems of diversions.

4 Algorithms for Non-dendritic Networks

The algorithms presented here operate on the graph structure exposed by any of the table patterns introduced above. Implementations may convert between representations as a preprocessing step (for example, via the conversion utilities in `hydroloom`), but the algorithms themselves are representation-agnostic: each depends only on upstream and downstream adjacency, primary-path flags where directionality is required, and, for total accumulation, undirected graph connectivity. This section treats four operations central to non-dendritic hydrologic analysis: network navigation, divergence-aware stream order, accumulation (apportioned, unapportioned, and total), and catchment aggregation. The bridge-based total-accumulation algorithm receives extended treatment as the principal algorithmic contribution of this paper. The `hydroloom` package (Blodgett, 2023) for the R statistical computing environment, version 4.6.0 (R Core Team, 2026), serves as the reference implementation throughout: it realizes the patterns of Section 3 and the algorithms developed below, and produces the empirical results of Section 5. Per-algorithm implementation details (function names, internal data structures) are deferred to Section 6; the algorithm descriptions in this section stand independent of any specific implementation.

4.1 Notation and complexity

Because some readers may be more familiar with hydrologic terminology than with graph theory, this section briefly recalls the standard terms and complexity notation used in this paper (Bondy and Murty, 1976; Cormen et al., 2022).

A *graph* consists of *nodes* and *edges* connecting pairs of nodes. We use V to denote the set of nodes and E the set of edges; the same symbols denote their counts when the meaning is clear from context. In a *directed* graph, each edge has a direction from a *tail* node (origin) to a *head* node (destination); these are graph-theory terms, unrelated to the hydrologic *headwater* and *tailwater*. In an *undirected* graph, edges have no direction. A flow network is a directed graph in which flowlines are edges and nexuses (or implicit endpoints) are nodes; its *undirected projection* is the same graph with edge directions ignored.

A *path* is a sequence of edges connecting two nodes through intermediate nodes; a *cycle* is a path that returns to its starting node. A graph is *connected* if a path exists between every pair of nodes.

A *cut edge* (synonymously called a *bridge*) is a flowline whose removal splits the network into two pieces with no remaining flow path between them — every drop of water reaching one side of the cut had to pass through the cut flowline. (In graph terms: removing the edge increases the number of *connected components*; equivalently, the edge lies on no cycle (Bondy and Murty, 1976, Theorem 2.3).) We use "bridge" hereafter, following the convention in graph algorithms and hydrologic software (see also Table 1). Figure 4 illustrates a bridge on a primitive graph.

Depth-first search (DFS) is a standard graph traversal that visits each node by following edges as far as possible before backtracking; DFS visits every reachable node and edge exactly once and runs in time proportional to $V + E$ (Cormen et al., 2022; Tarjan, 1972). A *topological sort* of a directed acyclic graph is an ordering of its nodes such that every edge points

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

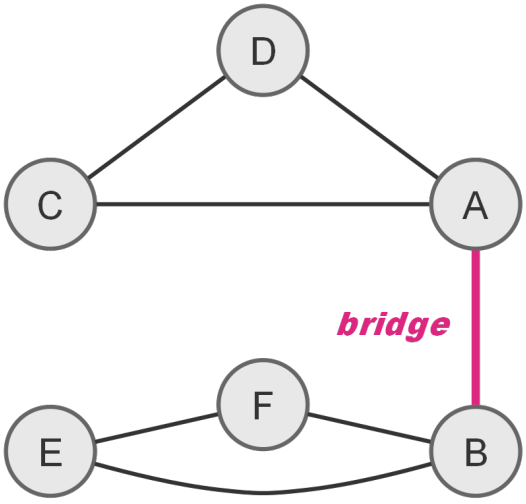

Figure 4. Primitive graph illustrating a cut edge (bridge). The red edge is a cut edge — removing it splits the graph into two disconnected pieces (connected components). The grey edges lie on cycles and are not cut edges; removing any grey edge leaves the graph connected via parallel paths.

forward in the order; it can also be produced in time proportional to $V + E$ (Cormen et al., 2022). Ordering edges by the position of their tail nodes gives the corresponding flowline order used in Section 4.4.3.

Algorithm running times are reported in *big-O notation* (Cormen et al., 2022), written $O(f(V, E))$. This expresses how the worst-case running time grows as the network grows: $O(V + E)$ means the time grows linearly with total network size, and $O(V^2)$ means running time grows with the square of the node count. The distinction matters in practice: on a continental network with millions of flowlines, an $O(V^2)$ algorithm is impractical where an $O(V + E)$ algorithm runs in minutes.

4.2 Network navigation in diverted networks

Network navigation starts from a network entity, follows a specified mode, and returns all features encountered along the walk. Four modes are commonly supported: upstream primary, upstream with tributaries, downstream primary, and downstream with diversions. Graph algorithms such as depth-first search, augmented to optionally skip tributary or diversion features, implement these modes. NHDPlus also supports navigation through database queries against pre-computed identifier attributes that encode upstream and downstream relationships within its data model, an alternative to graph traversal that works without custom code; the details fall beyond this paper's scope. All four modes use a single depth-first traversal that consults a primary-path flag (`upmain` or `downmain` in 3DHP terms) when deciding whether to descend a non-primary branch. Figure 5 illustrates the four modes from a single start point on the New Hope River (Chatham County, North Carolina) network: 1) upstream with tributaries traverses all contributing flowlines, 2) upstream main follows only the primary path, 3) downstream with diversions

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

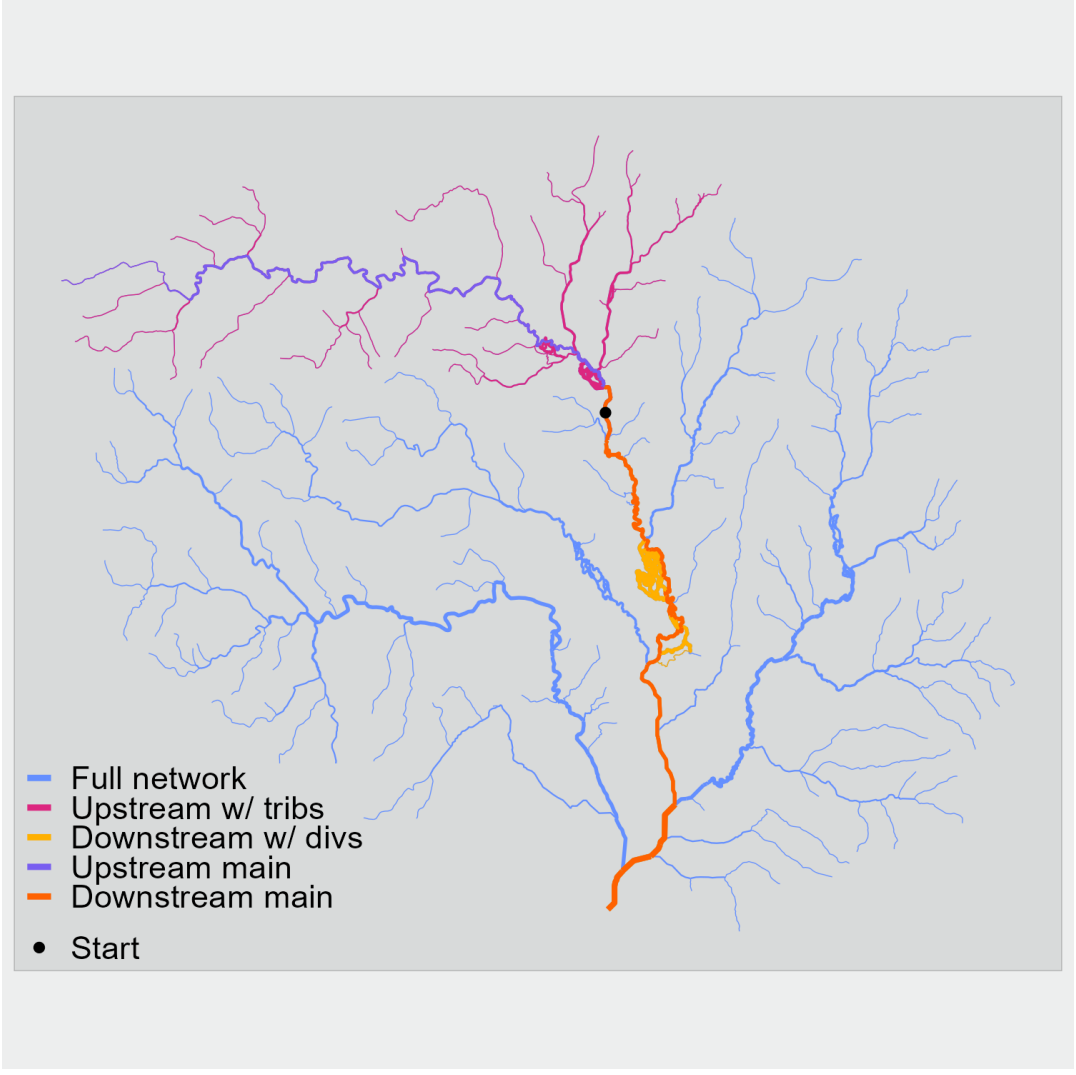

Figure 5. Four network navigation modes from a single start point on the New Hope River (Chatham County, North Carolina) network (NHDPlusV2 data (McKay et al., 2012)). Upstream with tributaries (tribs; magenta) traverses all contributing flowlines. Downstream with diversions (divs; amber) follows all downstream paths. Upstream main (purple) and downstream main (orange) follow only primary connections. Line widths are scaled by stream order.

follows all downstream paths including non-primary branches, and 4) downstream main follows only the primary downstream path.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

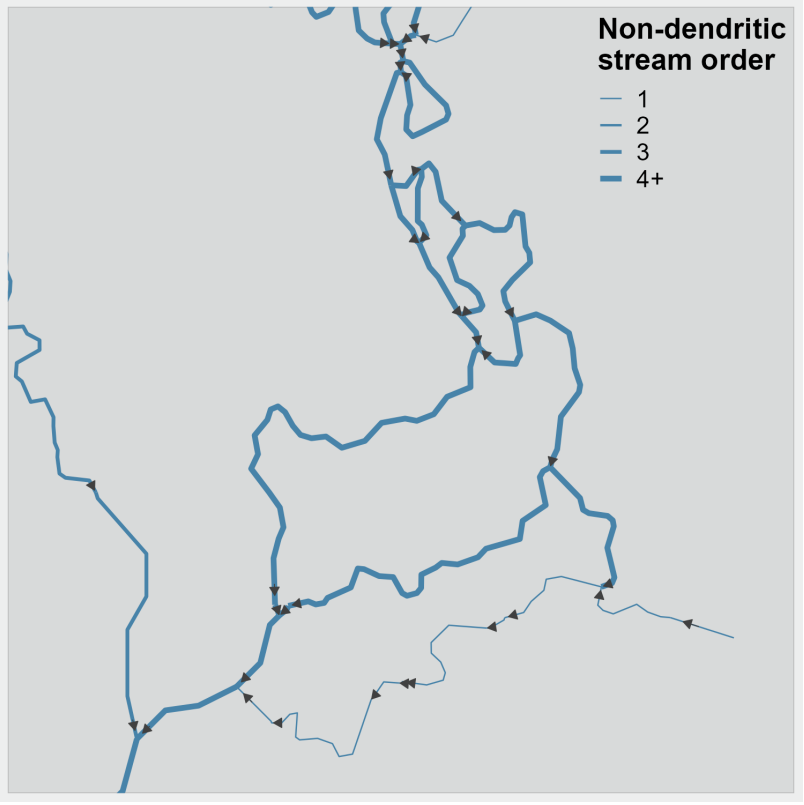

Figure 6. Non-dendritic stream order on a New Hope River subset (NHDPlusV2 data (McKay et al., 2012)). Line widths are scaled by the divergence-aware stream order, which inherits order through diversions and increments only where two distinct primary paths converge. Arrows indicate flow direction.

4.3 Stream order in diverted networks

Classical stream order (Strahler, 1957) assigns order 1 to headwaters and increments by one at every confluence of equal-order streams; this rule does not translate cleanly to diverted networks. NHDPlus (Bondelid et al., 2010) adopted a modified algorithm: diversions inherit stream order from upstream, and order increments only when two equal-order streams that do not share an upstream parent flowline converge. This requires tracking primary paths uniquely and incrementing order only where two distinct primary paths join. Figure 6 illustrates the effect: non-dendritic stream order reflects the full network topology including diversions, yielding order assignments that differ from classical Strahler values wherever secondary flow paths rejoin the primary channel.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

4.4 Accumulation in diverted networks

Drainage area and other catchment quantities accumulate downstream through a network in three forms: divergence-apportioned, unapportioned, and total. Divergence-apportioned accumulation splits flow downstream of a divergence according to known fractions (which must sum to one at each divergence). Unapportioned accumulation is a special case in which the entire value is assigned to the primary connection and nothing to diversions. Total accumulation, the focus of this section, passes the full upstream value to every downstream path regardless of flow split, and therefore must account explicitly for paths that reconverge in order to avoid double counting.

4.4.1 Apportioned and unapportioned accumulation

When divergence fractions are known and sum to one at every divergence, apportionment happens in a single upstream-to-downstream traversal. At each flowline, incoming values are summed and multiplied by the local divergence fraction. Because fractions below a divergence sum to one, upstream quantities apportion without duplication. Flow apportioned to a non-primary branch is preserved on that branch even when the branch terminates in a closed basin. Apportionment conserves flow wherever fractions sum to one, regardless of whether downstream branches share an outlet. Unapportioned accumulation is implemented as a special case by setting the non-primary fraction to zero. Both variants run in $O(V + E)$ time on a topologically sorted network (one constant-time update per node and one per edge (Cormen et al., 2022, ch. 20)), and require no special treatment of divergences or reconvergences beyond the fractions themselves.

4.4.2 Total accumulation and the double-counting problem

Total accumulation answers a different scientific question from apportioned accumulation. Where apportioned methods compute the *fraction* of an upstream value that reaches each downstream path under a known flow split, total accumulation computes the *complete set of upstream catchments topologically connected* to each downstream point, irrespective of how flow actually partitions at intervening divergences. The distinction is more than a modeling choice. Many hydrologic computations are intrinsically topologic and demand the total form: drainage area at any point on the network, upstream contributing area for water-quality and pollutant-source attribution (Brakebill et al., 2020), accumulated land cover and land use for spatial statistics (Hill et al., 2016), accumulated catchment-level climatologies for downscaling and forcing (Wieferich et al., 2021), and basin-nesting indices used to integrate datasets across providers. For each of these quantities, every catchment in the upstream basin is part of every downstream point's contributing area, regardless of how water actually splits at divergences. For these topologic quantities, apportioned accumulation gives a fractionated answer that does not match the question being asked, because contributing-area membership is binary: an upstream catchment either drains to a downstream point or it does not, and a catchment cannot drain only fractionally to a point. Apportioning contributions across multiple downstream paths therefore produces fractional values that do not describe any actual subset of the upstream contributing catchments. Apportioned accumulation remains the appropriate algorithm when the quantity itself partitions with flow, for example steady-state water or solute flux given known divergence fractions, or upstream-derived flux proxies for water-quality and contaminant mass (Schwarz et al.,

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

2006). Unapportioned accumulation gives the dendritic-only answer, equivalent to discarding the non-dendritic part of the network. Total accumulation is therefore the appropriate algorithm whenever the question is which catchments contribute to a point on the network. The remainder of this section addresses the algorithmic challenge that this correctness brings.

Total accumulation passes the full upstream value through every downstream link. When a diversion recombines, as occurs wherever two paths in a diversion system share a downstream confluence, a naive upstream-sum adds the shared upstream contribution once per path, producing a value greater than the true drainage area. Avoiding this error requires distinguishing flowlines whose upstream basin is uniquely defined by a tree from flowlines whose upstream basin is shared across multiple paths in a diversion system.

A brute-force approach resolves this by performing an upstream-with-tributaries traversal for each flowline in the full network and summing distinct contributing catchments through set operations. This is correct but scales poorly as a single upstream depth-first search visits up to V nodes and E edges, taking $O(V + E)$ time (Cormen et al., 2022, ch. 20), and repeating this once per flowline yields a total cost $O(V \cdot (V + E))$. This is impractical for continental-scale hydrography with millions of flowlines.

4.4.3 Bridge-based total accumulation

A more efficient strategy exploits the structure of diversion systems. Recall from Section 4.1 that a bridge is an edge whose removal increases the number of connected components of the undirected projection of the graph; equivalently, a bridge lies on no cycle (Bondy and Murty, 1976, Theorem 2.3). In a hydrologic network, bridges form the dendritic backbone (flowlines whose removal would separate the network into disconnected pieces), whereas non-bridge edges lie within diversion systems, where two or more parallel paths connect the same upstream and downstream points.

This distinction matters for accumulation. Because a bridge is the only undirected connection between its upstream and downstream sides, no parallel path exists along which a shared upstream contribution could inflate the bridge's total. Non-bridge flowlines, in contrast, lie within diversion groups where upstream basins overlap across paths. The non-bridge flowlines downstream of a divergence and upstream of the recombining confluence(s) form a system of diversions: a group of flowlines connected to one another by parallel flow paths, bounded by the bridges where flow enters and leaves. Together with those bounding bridges, this is a *diversion group* (see Table 1). In graph terms, the non-bridge flowlines of one diversion group form a *connected component* of the non-bridge subgraph. A bridge incident on a diversion group has exactly one endpoint in the component's node set; if it had both endpoints located there, it would lie on a cycle within the component and would not be a bridge. This formal definition refines the diversion-group concept introduced informally by Brakebill et al. (2020).

The same partition has a spectral counterpart: Tejedor et al. (2015a), building on the null-space characterization of Caughman and Veerman (2006) (which uses the bridge equivalence of Bondy and Murty (1976), Theorem 2.3), identifies for each outlet the "exclusive part" of nodes that drain only to that outlet versus the "common part" of nodes that drain to multiple outlets, via eigenvectors in the null-space of the weighted graph Laplacian. The exclusive part corresponds to nodes downstream of the bridge through which flow leaves the last diversion group on the way to the outlet, together with any tributaries joining below it, and the common part corresponds to nodes in and upstream of the diversion groups that feed more than one outlet.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

The two approaches solve different problems and operate at different scales. The spectral approach is well suited to apportioned (weighted) flux through delta channel networks, where eigenvector-valued fractions $\gamma_i(v)$ directly express the proportion of flow at v that drains to outlet i; its empirical demonstrations span tens to hundreds of links (the Wax Lake delta with 59 links and the Niger delta with 181 links in Tejedor et al. (2015a)), where eigendecomposition of an $N \times N$ matrix is fully tractable. The bridge-based combinatorial approach below targets the different problem of total drainage accumulation, where every downstream path receives the full upstream value irrespective of flow split, and runs in time linear in $V + E$ when most flowlines are bridges, extending the same conceptual partition to hydrofabric scale where networks contain millions of flowlines. The bridge partition recovers the topological structure of the spectral exclusive/common decomposition but does not directly produce the eigenvector-valued apportionment fractions. For problems requiring those fractions (apportioned flux through deltas, vulnerability mapping), the spectral approach remains the appropriate tool.

Let $w(e)$ denote the per-flowline value being accumulated (drainage area, runoff, or any additive catchment quantity), and let $T(e)$ denote the total accumulated value at flowline e. The bridge-based procedure has three steps:

1. **Identify bridges.** Let G_u denote the undirected projection of G (the same graph with directions removed; see Section 4.1). Compute the set $B \subseteq E$ of bridges of G_u using Tarjan's algorithm (Tarjan, 1974). The algorithm runs in $O(V + E)$ time and performs one depth-first traversal of G_u, marking each edge as a bridge if and only if it does not lie on any cycle. Networks with parallel flowlines between the same node pair require the algorithm to track the parent edge identifier rather than the parent node when computing low-link values, so that two parallel edges form a 2-cycle and neither is misclassified as a bridge. Building G_u from the directed network introduces connections that the directed graph lacks — the two downstream branches of a divergence have no directed path between them, but share an undirected path through their divergence node in G_u. This is the topology that bridge-finding requires, but an implementation that performs the conversion implicitly can confuse undirected connectivity with directed reachability.

2. **Order flowlines from upstream to downstream.** Compute a topological ordering of E in which every directed edge points forward, so that every flowline appears after all its upstream contributors. This step requires G to be a directed acyclic graph, which is the standard assumption for hydrologic flow networks (water does not flow uphill, so flow paths never form directed cycles even when the undirected projection does). The ordering itself is produced in a single linear-time pass (Cormen et al., 2022, ch. 20).

3. **Accumulate.** For each flowline e in topological order, perform an upstream depth-first search starting at the tail node of e. The search follows incoming directed edges only (i.e., it moves against the flow direction) and is restricted to non-bridge edges, marking nodes visited to avoid revisits. The search halts whenever a bridge is encountered. Let $U = \{e\} \cup \{f \in E \setminus B \mid f \text{ is traversed during the search}\}$ denote the visited flowline set, and let B_U denote the set of bridges encountered as halt points. Set

$$T(e) = \sum_{e' \in U} w(e') + \sum_{b \in B_U} T(b). \tag{1}$$

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Here e' indexes the flowlines of the visited set U. Because e is processed in topological order, the values $T(b)$ for $b \in B_U$ have already been computed in earlier iterations.

For flowlines on dendritic segments (where the immediate upstream connections are themselves bridges already processed), the depth-first search in step 3 visits no further flowlines, U reduces to $\{e\}$, and eq. 1 collapses to $T(e) = w(e) + \sum_b T(b)$ over the immediate upstream bridges b. For non-bridge flowlines and bridges at the downstream boundary of a diversion group, the search explores the local non-bridge subgraph until it halts at the bridges where flow enters the group.

Depth-first search is required for bridge identification in step 1 (Tarjan's algorithm depends on DFS low-link tracking (Tarjan, 1974)) but is incidental for the local upstream traversal in step 3, which could equivalently be breadth-first with no change in run time. We now establish that this procedure is correct and quantify its running time.

Theorem 1 (Correctness). Let $G = (V, E)$ be a directed acyclic flow network with undirected projection G_u, and let $B \subseteq E$ be the set of bridges of G_u. For every flowline $e \in E$, define the *upstream basin* of e as

$$R(e) = \{e\} \cup \{f \in E \mid \text{there is a directed path in } G \text{ from } \mathrm{head}(f) \text{ to } \mathrm{tail}(e)\}. \tag{2}$$

The procedure computes

$$T(e) = \sum_{e' \in R(e)} w(e'). \tag{3}$$

Proof sketch. By induction on the upstream-to-downstream topological order of e. The base case (headwater flowlines) gives $T(e) = w(e)$ directly. The inductive step partitions $R(e)$ into the visited non-bridge flowlines U and the upstream basins $R(b)$ of bridges $b \in B_U$ at which the local search halts. Disjointness of U and the $R(b)$ follows from the bridge property and DAG acyclicity (see Appendix A and Lemma 1).

Lemma 1 (Disjoint upstream basins of bridges sharing a non-bridge component). Let $b_1, b_2 \in B$ be distinct bridges of G_u such that $\mathrm{head}(b_1)$ and $\mathrm{head}(b_2)$ lie in the same connected component C of the non-bridge subgraph $G_u \setminus B$ (allowing the trivial component consisting of a single node not incident to any non-bridge edge, in which case $\mathrm{head}(b_1) = \mathrm{head}(b_2)$). Then $R(b_1) \cap R(b_2) = \emptyset$.

Proof sketch. If f lay in both $R(b_1)$ and $R(b_2)$, a directed path from $\mathrm{head}(f)$ to each of $\mathrm{tail}(b_1)$ and $\mathrm{tail}(b_2)$ combined with a non-bridge walk between their heads inside the shared component would produce an undirected $\mathrm{tail}(b_1)$–$\mathrm{head}(b_1)$ walk avoiding b_1, contradicting the bridge property of b_1. See Appendix A.

Complexity. Bridge identification in step 1 runs in $O(V+E)$ (Tarjan, 1974). The topological sort in step 2 runs in $O(V+E)$ (Cormen et al., 2022, ch. 20). In step 3, each iteration's depth-first search is bounded by the size of the diversion group containing e: a depth-first search confined to a subgraph of size n runs in $O(n)$ (Cormen et al., 2022; Tarjan, 1972). Let $D_1, D_2, \ldots$ denote the diversion groups, with $|D_i|$ flowlines in group D_i. The cumulative cost of local depth-first work across the $|D_i|$ flowlines in D_i is at most $O(|D_i|^2)$ when each flowline in D_i triggers an independent search over D_i, as in the current `hydroloom` implementation. For bridge flowlines on dendritic segments, where the search is trivial, the per-iteration cost is proportional to the number of immediate upstream connections, summing to $O(V + E)$ across all such flowlines. Combining

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

the contributions:

$$T_{\text{total}} = O(V+E) + O\left(\sum_i |D_i|^2\right) = O\left(V + E + \sum_i |D_i|^2\right). \quad (4)$$

When most flowlines are bridges and diversion groups are small — the typical case for NHDPlus (89% bridge fraction, median group size 3 on V2; 92% and 4 on HR; see Section 5.2) — the second term in eq. 4 is dominated by the first and the procedure runs in effectively linear time, $O(V+E)$. The brute-force approach of Section 4.4.2, by contrast, runs in $O(V \cdot (V+E))$ regardless of diversion-group structure. The asymptotic speedup approaches a factor of V on networks where the largest diversion group contains a vanishing fraction of flowlines.

Figure 7 illustrates the procedure on a New Hope River subset: panels A and B contrast dendritic and total accumulation, panel C marks bridge (blue) versus non-bridge (red) flowlines, and panel D shows the full-network strategy with bridge cache points. Figure 8 shows the distinction between bridge and non-bridge flowlines on a representative NHDPlus diversion system.

4.5 Catchment aggregation in non-dendritic networks

The HY_Features standard (Blodgett and Dornblut, 2018) includes a "dendritic catchment" class that is functionally equivalent to a bridge edge and identifies dendritic catchments as key to catchment aggregation. An aggregate catchment must itself satisfy the catchment data model: no more than one inflow nexus and one outflow nexus. To form an aggregate, the inlet and outlet must lie on a continuous upstream-downstream path, and both must be bridge catchments so that no parallel flowlines create additional inflows or outflows. The hyRefactor R package (Blodgett and Johnson, 2024) implements catchment aggregation along these lines, although at the time of writing, it only supports a dendritic form of the network and is cited here because experience in its implementation led to development of the methods reported here.

5 Empirical Analysis

The patterns and algorithms presented above are useful only to the extent that non-dendritic flow is common in real hydrologic networks and that the bridge-based procedure scales to hydrofabric size. This section quantifies both. Section 5.1 gives the first cross-fabric measurement of non-dendritic prevalence across six hydrofabrics for three continents (Australia, Europe, and North America) and a global compilation; Section 5.2 examines the within-conterminous United States (CONUS) comparison of NHDPlusV2 and NHDPlusHR as a controlled resolution example; Section 5.3 benchmarks the bridge-based procedure of Section 4.4.3 against the brute-force approach of Section 4.4.2.

5.1 Cross-fabric prevalence

The first cross-fabric measurement of non-dendritic prevalence applies the patterns and algorithms of Section 3 and Section 4 across six hydrofabrics: NHDPlusV2 (McKay et al., 2012), NHDPlusHR (Moore et al., 2025), the Australian Geofabric V3.3 (Bureau of Meteorology, 2015), Canada's National Hydro Network (NHN) (Natural Resources Canada, 2024), the pan-

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

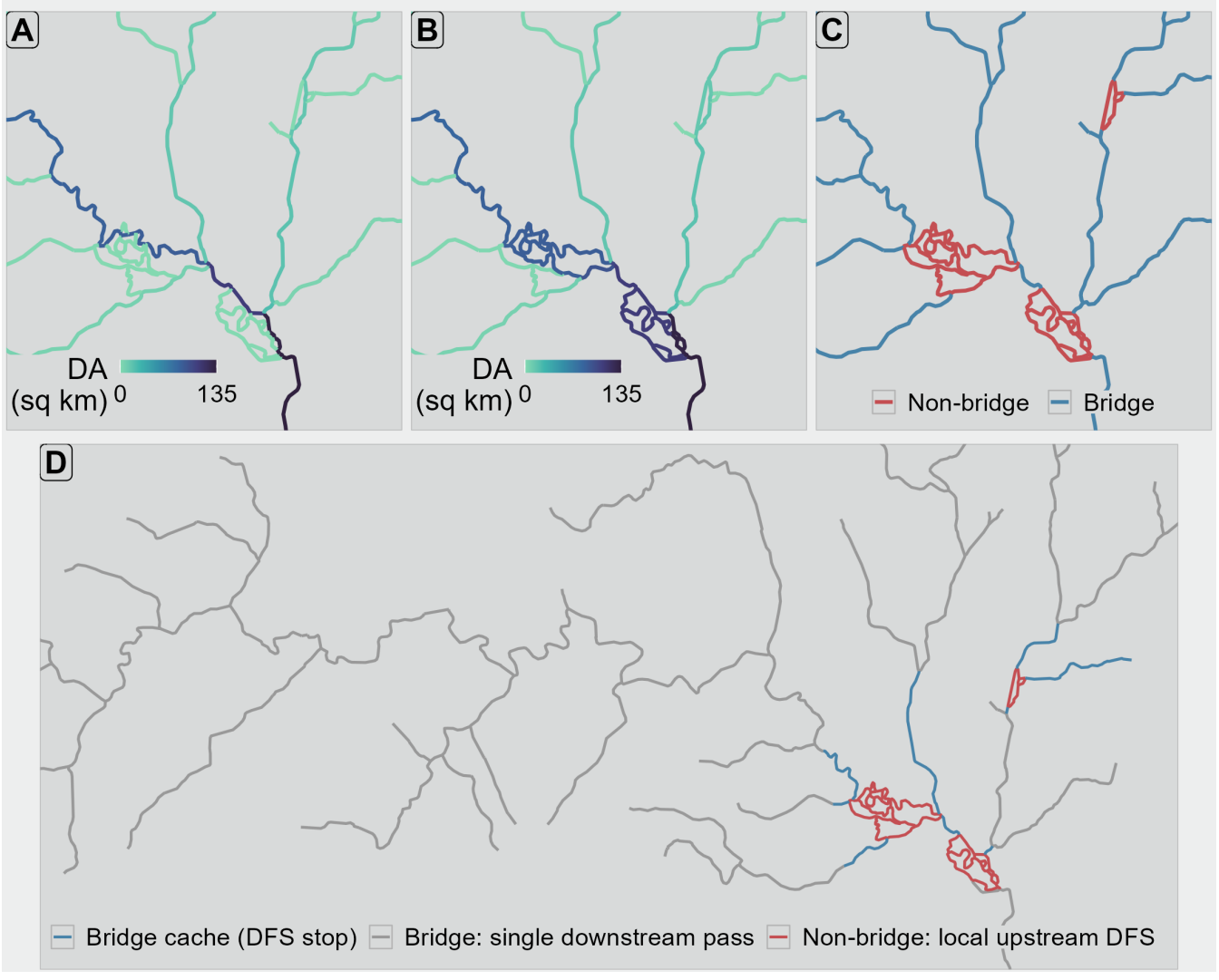

Figure 7. Accumulation walkthrough on a New Hope River subset (NHDPlusV2 data (McKay et al., 2012)). A) Dendritic accumulation treats diversions as headwaters. B) Total accumulation passes full upstream value to all downstream paths. C) Bridge (blue) vs non-bridge (red) flowlines. D) Full network showing the single-pass accumulation strategy of Section 4.4.3: flowlines are processed in upstream-to-downstream topological order; dendritic bridges (grey) accumulate directly from their upstream bridges, whereas non-bridge flowlines (red) and the bridge leaving each diversion group trigger a local upstream DFS that halts at already-processed bridges (blue). Red and blue mark every flowline upstream of the highest-accumulation non-bridge flowline by role; the search from any one flowline stops at the first blue bridge it meets.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

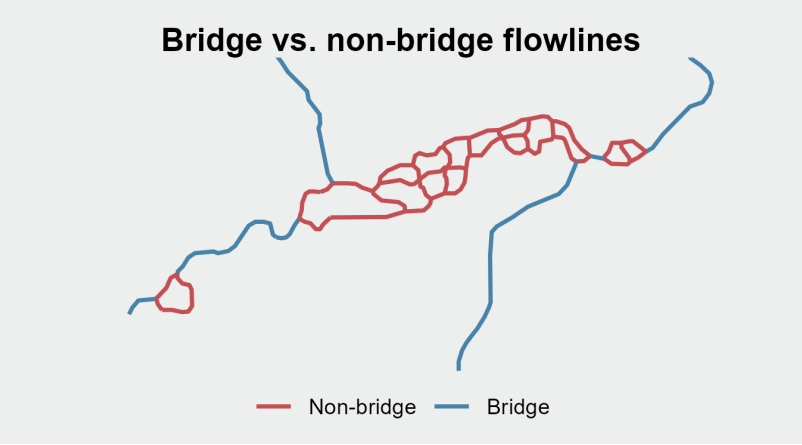

Figure 8. Bridge flowlines (blue) whose removal disconnects the graph versus non-bridge flowlines (red) within diversion systems. Data from NHDPlusV2 (McKay et al., 2012).

European EU-Hydro (Copernicus Land Monitoring Service, 2019), and the Global River Topology (Wortmann et al., 2025) — three continents and a global compilation.

5.1.1 Schemas and preprocessing

Each native flowline schema maps to `hydroloom`'s `id` / `fromnode` / `tonode` canonical attributes (see Table 6). After applying this name-mapping, a divergent node is any `fromnode` value with at least two outgoing flowlines that also appears as a `tonode`, so the per-edge divergence count reduces to a group-and-count operation on `fromnode` checked against the `tonode` set. The incoming-edge requirement restricts the count to true bifurcations and excludes nodes where two first-order channels share a headwater node; the same rule is applied to all six hydrofabrics. In NHDPlusV2 and NHDPlusHR, whose divergence attributes are hand-curated, it removes zero and eight nodes respectively.

Table 6: Mapping of native flowline-schema column names to the canonical `hydroloom` `id` / `fromnode` / `tonode` triple used by the cross-fabric pipeline. Datasets are NHDPlusV2 (McKay et al., 2012) and NHDPlusHR (Moore et al., 2025) (National Hydrography Dataset Plus, medium and high resolution), AGF (Bureau of Meteorology, 2015) (Australian Hydrological Geospatial Fabric), GRIT (Wortmann et al., 2025) (Global RIver Topology), EU-Hydro (Copernicus Land Monitoring Service, 2019) (pan-European river network database), and NHN (Natural Resources Canada, 2024) (National Hydro Network, Canada).

Dataset	`id`	`fromnode`	`tonode`
NHDPlusV2	`ComID`	`FromNode`	`ToNode`
NHDPlusHR	`NHDPlusID`	`FromNode`	`ToNode`

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Dataset	id	fromnode	tonode
AGF V3.3	HydroID	From_Node	To_Node
GRIT v1.0	global_id	upstream_node_id	downstream_node_id
EU-Hydro	OBJECT_ID	FNODE	TNODE
NHN	NID	FROM_JUNCT	TO_JUNCT

As shown in Figure 9, NHN represents flow through the interior of some waterbody polygons as a sort of "skeleton", producing divergences that are not comparable to typical non-dendritic network behavior. To account for this, divergences within lakes and other non-watercourse polygons were removed; the remaining set is termed "channel-only". Waterbody polygons of type "watercourse" (those representing major rivers) were kept because their bifurcations are real channel splits. Two additional NHN-specific filters apply. Rows with `ISOLATED = 1` (disconnected stubs) are dropped. Headwater self-loops (`FROM_JUNCT == TO_JUNCT`, NHN's "no upstream junction" encoding) are removed, because otherwise every headwater outlet would appear as a divergence. Shared headwater nodes, relatively common in NHN, are excluded by the bifurcation rule above. Among the six hydrofabrics examined, only NHN required this in-waterbody filter; the other five do not encode polygon skeletons as flowlines and so have no equivalent inflation of divergent-node counts.

5.1.2 Divergence density

Per-basin divergence density is reported normalized by km^2 of drainage area and by km of flowline within the drainage area. The latter is the more appropriate cartographic representation intensity for a point process on a linear network (Baddeley et al., 2015) and is the lead metric in the comparisons that follow. Divergences per network length is most directly a property of the cartographic representation, which is a single-resolution depiction of the landscape.

Table 7 summarizes the cross-fabric totals. Per-basin medians span an order of magnitude across datasets and another order of magnitude across basins within any single dataset. The NHN contrast between the raw count (approximately 1.7 million divergent nodes) and the channel-only count (489,619) makes explicit how strongly cartographic conventions drive the headline number. More than two-thirds of NHN divergences arise inside lake, reservoir, side-channel, and liquid-waste polygons that NHN skeletonizes into flowlines, and under a third correspond to splits along open watercourses.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

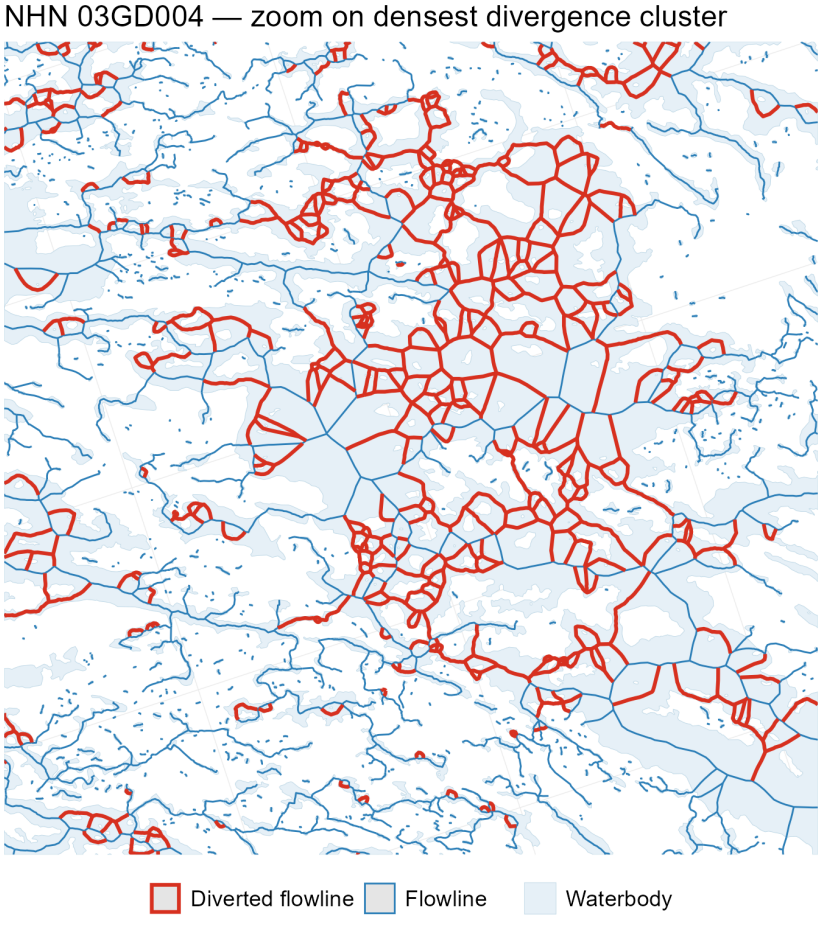

Figure 9. NHN (Natural Resources Canada, 2024) work unit 03GD004, an 8.3- × 8.3-km zoom on the densest divergence cluster. Diverted flowlines (red) trace polygon-skeletons inside waterbody features (light blue). Outside the polygons, the network reduces to single-line channels with divergences only where watercourse flow splits.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Table 7: Cross-fabric divergence summary across NHDPlusV2 (CONUS), NHDPlusHR (CONUS at finer resolution), AGF V3.3 (Australia), NHN (Canada), EU-Hydro (mainland Europe), and GRIT v1.0 (global). "Subareas" gives the count of per-dataset partition units used as the basis for the per-km min / median / max statistics; the unit is the 2-digit hydrologic unit code (HUC-2) for NHDPlusV2/HR, drainage divisions for AGF, major drainages for NHN, basins for EU-Hydro, and domains for GRIT. "Divergences" is the total divergent-node count; "Length" is the network length used as the per-km denominator; "div/km" gives the unweighted minimum, median, and maximum divergences per km across per-region densities within each dataset. NHN counts exclude divergences inside lake, reservoir, side-channel, and liquid-waste polygons (see Section 5.1).

Dataset	Subareas	Divergences	Length (10^6 km)	div/km (min / med / max)
NHDPlusV2	18	64,929	5.19	0.0029 / 0.0113 / 0.0364
NHDPlusHR	22	424,064	13.01	0.0063 / 0.0301 / 0.0857
AGF V3.3	13	112,039	3.86	0.0007 / 0.0122 / 0.0754
NHN	11	489,619	8.77	0.0149 / 0.0372 / 0.0886
EU-Hydro	32	11,713	2.69	0.0014 / 0.0038 / 0.0126
GRIT v1.0	61	67,919	18.51	0.0000 / 0.0034 / 0.0231

Figure 10 shows per-km divergence density on a single shared color scale across the four national/continental hydrofabrics (NHN channel-only, EU-Hydro, NHDPlusHR, AGF) and the global GRIT hydrofabric. The same processing pipeline (attribute name alignment, per-edge divergence-node counting on `fromnode`, and per-basin normalization by network length) produces each panel from its native flowline schema. Cross-fabric density differences reflect cartographic policy and resolution. NHN sits at the top of the per-km density distribution, though only modestly above NHDPlusHR: its channel-only median (0.0372 div/km) is about 1.2× NHDPlusHR's (0.0301). Two factors could plausibly explain why NHN ranks highest — 1) the Canadian Shield's millions of post-glacial lake-and-channel chains genuinely produce abundant in-channel divergences (Phillips et al., 2011), and 2) NHN's 1:50,000 source resolution renders splits that coarser-resolution products do not depict (Natural Resources Canada, 2024). We do not separate the two effects here. AGF has dense anabranching systems in inland floodplains (Figure 11). EU-Hydro is the sparsest of the regional fabrics, with overall drainage density ≈ 0.42 km of network per km^2 of basin area — roughly a third of NHDPlusHR's 1.15 km/km^2. GRIT has a generally coarser global resolution (≈ 0.14 km of network per km^2 of basin area). Figure 10 is the central demonstration that the patterns of Section 3 and the algorithms of Section 4 transfer across the schema and resolution variants tested here without modification beyond the name-mapping table. Within-dataset cartographic inconsistencies can also bias these per-region densities. The medium-resolution NHD inherits 1:100,000 compilation differences across its original USGS quadrangles, which remain visible in the current version of NHDPlusV2; the other hydrofabrics likely carry analogous within-product heterogeneities from their source compilations, though we have not characterized them.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

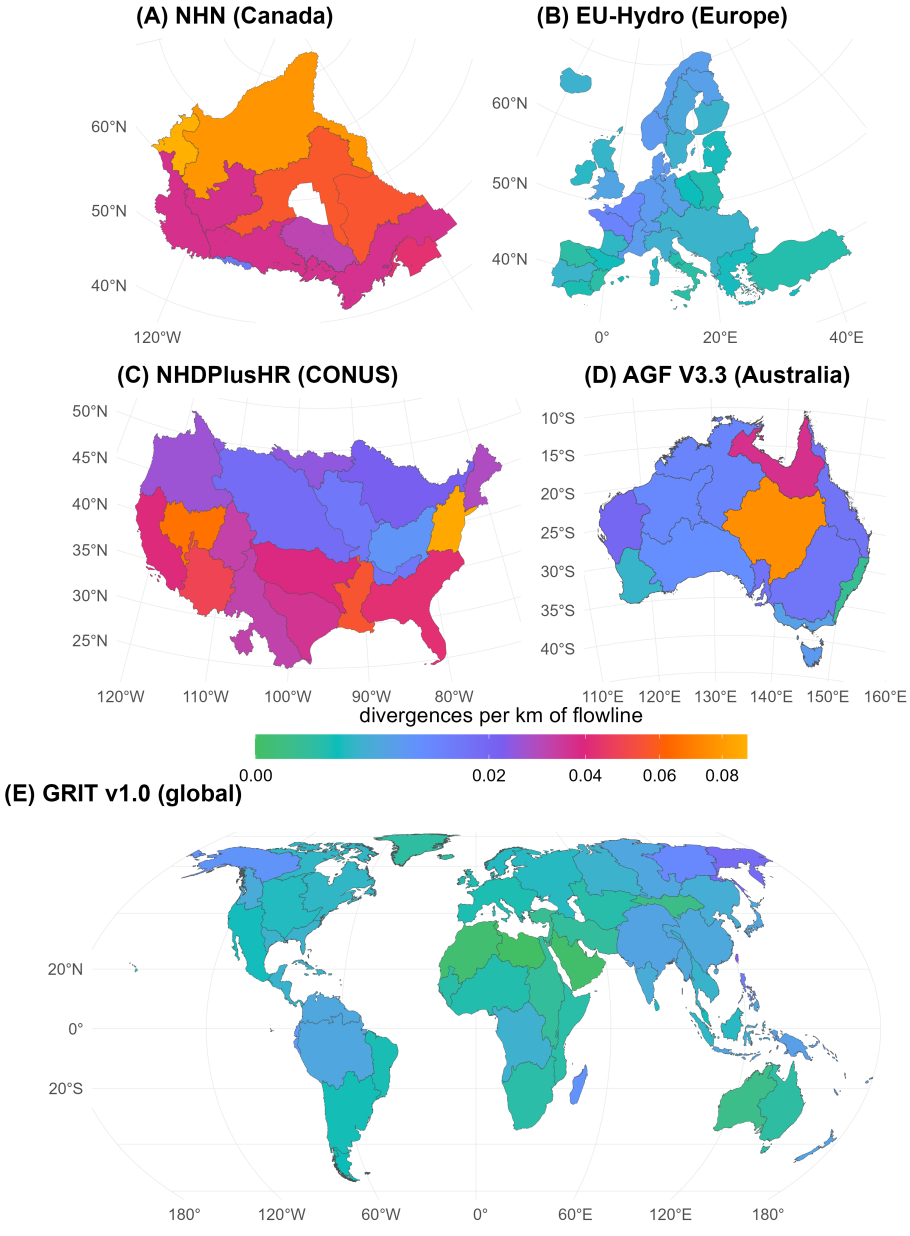

Figure 10. Per-km divergence density across five hydrofabrics rendered on a single square-root color scale. Upper grid: (A) NHN (Natural Resources Canada, 2024) (Canada, channel-only counts; in-waterbody divergence nodes excluded), (B) EU-Hydro (Copernicus Land Monitoring Service, 2019) (mainland Europe), (C) NHDPlusHR (Moore et al., 2025) (CONUS), (D) AGF V3.3 (Bureau of Meteorology, 2015) (Australia). Lower panel: (E) GRIT v1.0 (Wortmann et al., 2025) (global). The shared scale is anchored at the maximum density across the national/continental datasets; GRIT values fall below this maximum and render in the lower portion of the scale. Each panel is produced from its native flowline schema by the same `hydroloom`-based pipeline.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

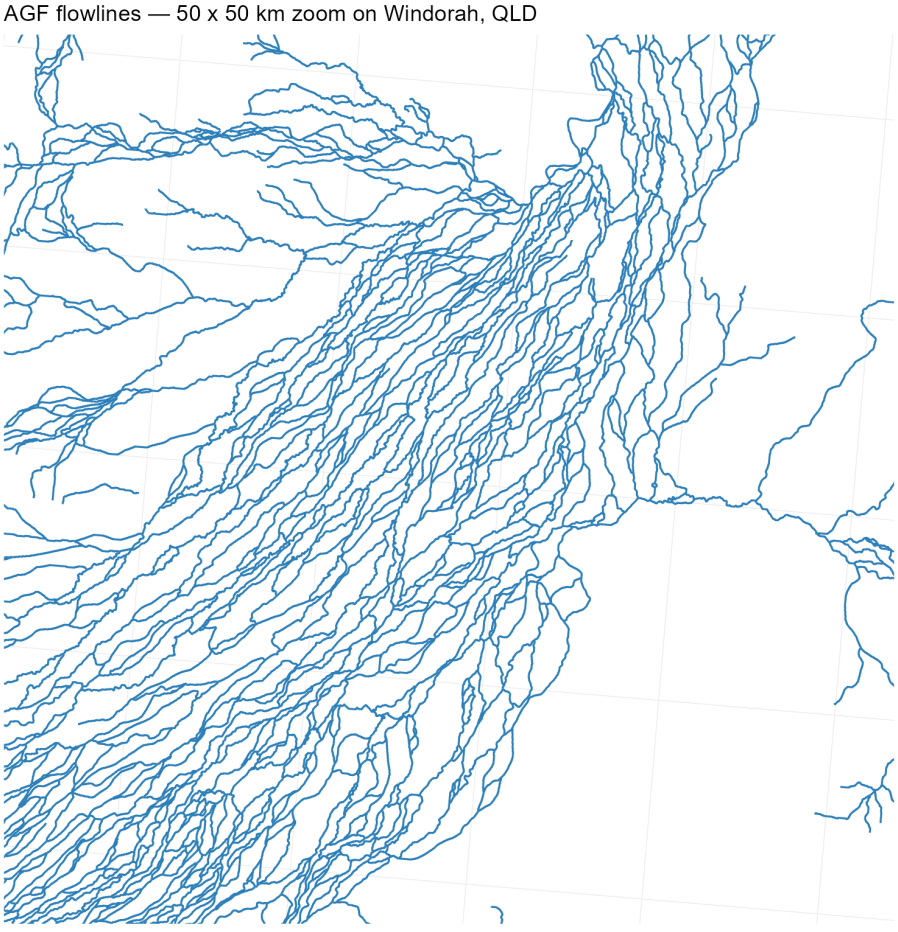

Figure 11. AGF V3.3 (Bureau of Meteorology, 2015) flowlines in a 50- × 50-km window centered on Windorah, Queensland, on the Cooper Creek inland floodplain. The dense network of interlacing channels visible here illustrates the anabranching pattern characteristic of Australia's inland drainage divisions and produces AGF's elevated divergence density in Figure 10.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

5.1.3 Diversion-group size distribution

Where divergence densities shown in Figure 10 reflect cartographic style and resolution, the size distribution of the diversion groups as described in Section 4.4.3 is a structural measurement that tests whether non-dendritic features are organized and scale comparably across hydrofabrics developed independently. Maximum-likelihood power-law fits to the upper tail of diversion-group sizes (see Figure 12 and Table 8) give $\hat{\alpha}$ exponent estimates that cluster across the six fabrics: NHDPlusV2 2.22, NHDPlusHR 2.05, AGF 2.18, EU-Hydro 2.16, GRIT 2.18, and NHN 2.37. All six fall within ± 0.17 of $\hat{\alpha} \approx 2.2$, with four (NHDPlusV2, AGF, EU-Hydro, GRIT) clustered within ± 0.07. The fabrics differ across nearly every other axis: they span an order of magnitude in flowline count, range from hand-curated authoritative networks (NHDPlusV2, NHDPlusHR) through national surveys (NHN, AGF, EU-Hydro) to a globally derived skeleton (GRIT), and cover physical settings from Canadian-Shield drainages to Australian inland anabranching corridors and mainland-European single-line channels. In short, the rate at which large diversion groups become rare follows nearly the same power law in each fabric we examined, to within a small constant. Why the scaling regime should be this consistent across that much cartographic and physical heterogeneity is left as an open question that could follow this work.

The power law (PL) fits were established with the discrete Clauset-Shalizi-Newman procedure (Clauset et al., 2009) as implemented in the R package poweRlaw (Gillespie, 2015). We estimate $\hat{\alpha}$ (slope of the power-law tail) by maximum likelihood, with $k_{\min}$ (lower bound of the fitted range) selected by minimizing the Kolmogorov-Smirnov distance between the empirical and fitted distributions. Selected $k_{\min}$ values range from 6 (NHDPlusV2, EU-Hydro) to 72 (NHDPlusHR). Sample size influences this threshold — more samples give the KS criterion finer power to detect departures from power-law at small sizes and push the optimal lower bound deeper into the upper tail — but does not determine it: NHN has more diversion groups than NHDPlusHR yet settles at $k_{\min} = 10$, indicating its distribution stays closer to a pure power-law at smaller sizes. Bootstrap resampling gives 95% confidence intervals on ($\hat{\alpha}$, $k_{\min}$) and a parametric goodness-of-fit p-value. We compare against lognormal and exponential alternatives at the same $k_{\min}$ via Vuong's test (Vuong, 1989), and against power-law-with-exponential-cutoff via a nested likelihood-ratio test. Because pure power-law lies on the boundary of the cutoff family's parameter space, the likelihood-ratio (LR) p-value is reported below with the standard one-sided boundary correction (halved). Heavy-tailedness is confirmed — exponential is rejected against power-law in every fabric (Vuong $z \geq 4$, $p \ll 0.001$). Power-law itself is not uniquely identified. Lognormal is statistically indistinguishable from power-law for all six hydrofabrics. Power-law-with-exponential-cutoff is preferred over pure power-law at the 5% level for NHDPlusHR (boundary-corrected $p = 0.035$) and EU-Hydro ($p = 0.038$), is borderline for NHN ($p = 0.044$, two-sided $p = 0.087$), and is not preferred for NHDPlusV2, AGF, or GRIT ($p \geq 0.25$); for NHDPlusHR and EU-Hydro, the cutoff term improves the upper-tail fit, indicating mild upper-tail truncation rather than a strict power-law. The CSN goodness-of-fit accepts a pure power-law for NHDPlusHR and GRIT ($p \geq 0.59$) and NHN ($p = 0.49$), marginally for EU-Hydro ($p = 0.071$), and rejects for NHDPlusV2 and AGF ($p < 0.05$). At these sample sizes the test reveals small empirical departures from a strict power-law form as significant. In practice, the data are heavy-tailed and are consistent with a power-law description over $[k_{\min}, k_{\max}]$ across all six hydrofabrics, with mild evidence for upper-tail truncation in NHDPlusHR and EU-Hydro and no single heavy-tailed family identified across all hydrofabrics. Figure 12 shows each hydro-

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

fabric's empirical CCDF (complementary cumulative distribution function, $P(K \geq k)$) with the fitted power-law CCDF (red dashed) and the fitted PL+cutoff CCDF (blue solid) overlaid above its $k_{\min}$. Table 8 summarizes the per-fabric estimates.

Table 8: Per-fabric maximum-likelihood power-law fit results. $\hat{\alpha}$ is the estimated power-law (PL) exponent; $k_{\min}$ is the KS-selected lower bound of the fitted range. CSN GoF p is the Clauset-Shalizi-Newman goodness-of-fit p-value. The "PL+cutoff vs PL" column gives the boundary-corrected nested likelihood-ratio test p-value for power-law-with-exponential-cutoff against pure power-law (bold = preferred at $\alpha = 0.05$); NHN's boundary-corrected $p = 0.044$ sits just under the threshold but its two-sided $p = 0.087$ does not, so it is left unbolded as borderline. Vuong comparisons against lognormal are indistinguishable ($|z| < 2$) for every fabric and are not shown.

Fabric	$\hat{\alpha}$	$k_{\min}$	CSN GoF p	PL+cutoff vs PL p
NHDPlusV2	2.22	6	< 0.05	≥ 0.25
NHDPlusHR	2.05	72	≥ 0.59	**0.035**
AGF V3.3	2.18	7	< 0.05	≥ 0.25
EU-Hydro	2.16	6	0.071	**0.038**
NHN	2.37	10	0.49	0.044
GRIT v1.0	2.18	10	≥ 0.59	≥ 0.25

The `hydroloom` name-mapping table (Table 6), the shared bifurcation rule of Section 5.1.1, and NHN's in-waterbody filter normalize six native flowline schemas, and the same processing pipeline was run on each — the patterns and algorithms of Section 3 and Section 4 transfer across the schema and resolution variants tested here without modification beyond the name-mapping step. In that common pipeline, per-km divergence density varies across fabrics with cartographic policy and resolution, but the regional pattern within any one fabric is reliable. The size distribution of diversion groups is remarkably consistent. Maximum-likelihood power-law exponents fall in a tight $\hat{\alpha}$ band, despite the cartographic and physical heterogeneity that the per-km density envelopes reflect.

5.2 Within-CONUS resolution example

CONUS appears twice in Table 7 — the medium-resolution NHDPlusV2 (McKay et al., 2012) and the high-resolution NHDPlusHR (Moore et al., 2025) — which share a data model, cartographic conventions, and the underlying landscape. Their side-by-side comparison isolates a resolution effect that is otherwise confounded with cartographic policy and continental-scale geography across the datasets described in Section 5.1. Table 9 summarizes the NHDPlusV2/NHDPlusHR comparison along four dimensions; the regional discussion that follows details the structure behind the totals.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

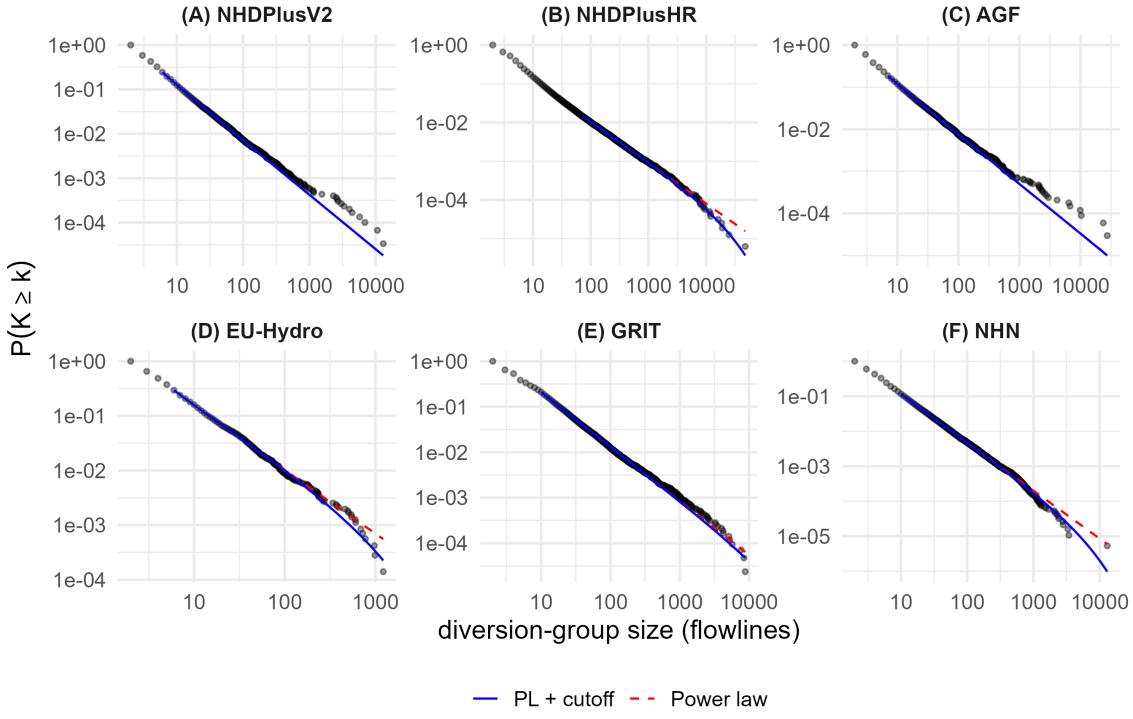

Figure 12. Empirical complementary CDF $P(K \geq k)$ of diversion-group sizes by hydrofabric — (A) NHDPlusV2, (B) NHDPlusHR, (C) AGF V3.3, (D) EU-Hydro, (E) GRIT v1.0, (F) NHN — with two fitted complementary cumulative distribution functions (CCDFs) overlaid above each fabric's KS-selected $k_{\min}$: the discrete Clauset-Shalizi-Newman power-law (PL) (red dashed) and the discrete power-law-with-exponential-cutoff (blue solid). Both axes are logarithmic. Where the cutoff fit reduces to power-law (κ effectively unbounded; NHDPlusV2, AGF, GRIT), the blue and red lines coincide. Per-fabric $\hat{\alpha}$ and $k_{\min}$ are reported in Section 5.1; full per-fabric numbers including bootstrap confidence intervals, goodness-of-fit p-values, and PL+cutoff parameters are provided in the analysis pipeline archived with this paper (see Code Availability, Section A2).

Table 9: NHDPlusV2 vs NHDPlusHR within-CONUS comparison along the four characterization dimensions. Res, resolution.

Metric	NHDPlusV2 (medium-res)	NHDPlusHR (high-res)
Flowlines	2,691,339	~24.9 million
Divergences	64,929	424,064
Diverted-flowline fraction	2.43%	1.72%
Diversion groups (count)	29,879	158,758
Median diversion group size (flowlines)	3	4
Maximum diversion group size	13,071	47,536

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Metric	NHDPlusV2 (medium-res)	NHDPlusHR (high-res)
Bridge fraction	89.4%	92.4%

The smaller diverted-flowline share in NHDPlusHR despite its finer resolution reflects the many more first-order flowlines NHDPlusHR resolves at headwaters. Diversion groups are correspondingly more numerous and somewhat larger on average (mean 12.0 flowlines versus 9.5 in NHDPlusV2), with per-HUC-2 divergent-node densities broadly paralleling NHDPlusV2.

The bridge-based decomposition of Section 4.4.3 partitions the network into bridge edges (the dendritic backbone) and non-bridge edges (interior to diversion groups). Regionally, the NHDPlusV2 bridge fraction varies from roughly 70% in HUC-2 08 (Lower Mississippi) to 96% in HUC-2 17 (Pacific Northwest); NHDPlusHR shows a comparable spread from 79% in HUC-2 16 (Great Basin) to 98% in HUC-2 05 (Ohio), running a few percentage points higher overall consistent with the extra first-order flowlines its finer resolution introduces.

The enhanced NHDPlusV2 of Brakebill et al. (2020) reassigns primary-path designations for 8,877 flowlines relative to the original NHDPlusV2 (0.33% of flowlines), of which 1,043 are at divergent nodes (about 0.80% of flowlines incident to divergent nodes). These corrections affect non-dendritic algorithms primarily through the upstream-main and downstream-main path identifiers, not through the bridge structure, which depends only on undirected connectivity and is therefore unchanged by primary-path reassignment.

Geographically, divergences concentrate in low-relief coastal-plain and deltaic regions (HUC-2 08 Lower Mississippi, HUC-2 03 South Atlantic-Gulf) and irrigation- and reservoir-dominated basins (HUC-2 16 Great Basin, HUC-2 06 Tennessee) and, in NHDPlusV2, are least common in steep mountainous terrain such as HUC-2 17 (Pacific Northwest). Figure 13 shows divergence density across CONUS aggregated to HUC-2 watersheds.

Most diversion groups in NHDPlusV2 and NHDPlusHR contain only a few flowlines, but a small number of much larger groups exists where flow re-routes through extended distributary or anastomosing channel systems (NHDPlusV2 and NHDPlusHR panels of Figure 12). These groups correspond to coastal delta and estuarine networks (Mississippi River, Chesapeake Bay, the Everglades and south Florida) and to heavily engineered irrigation and floodplain corridors (Lower Colorado River, the Missouri River floodplain). NHDPlusHR has five times NHDPlusV2's group count and a heavier tail (maximum 47,536 versus NHDPlusV2's 13,071), consistent with finer resolution resolving more non-dendritic structure.

5.3 Algorithm benchmarks

To verify that the bridge-based procedure of Section 4.4.3 delivers the asymptotic improvement over the brute-force approach of Section 4.4.2 predicted by complexity analysis, we benchmarked both procedures on NHDPlusV2 subsets of increasing size and extrapolated the fitted scaling to the full V2 CONUS network. NHDPlusHR's bridge-fraction profile is similar (92.4% versus NHDPlusV2's 89.4%, see Section 5.2), so the same scaling regime is expected, but NHDPlusHR was not directly benchmarked. Both implementations share the same preprocessing — adjacency-list build, topological sort, and Tarjan's bridge identification — and the brute-force variant differs only in disabling the bridge classification so the per-edge upstream search

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

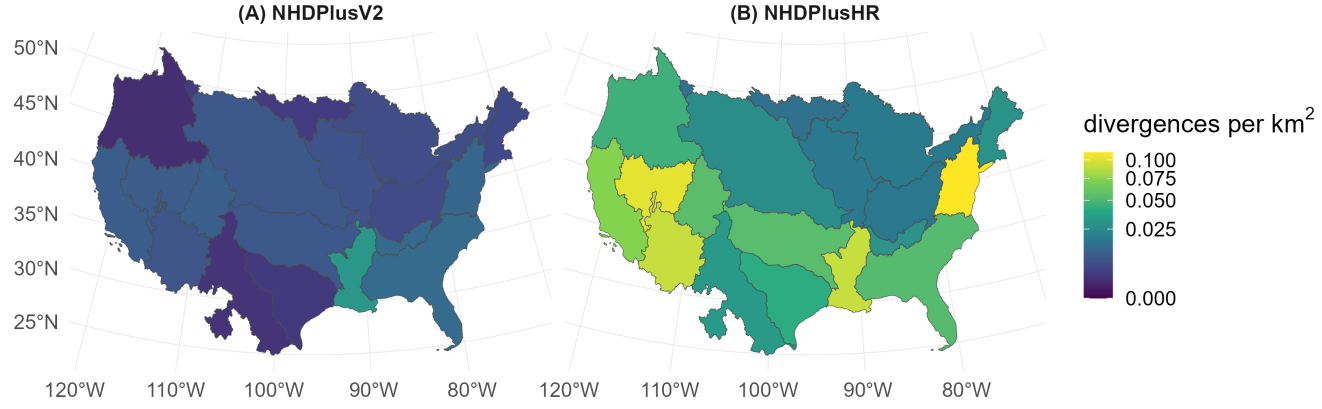

Figure 13. Divergence density across CONUS, aggregated to HUC-2 watersheds, for (A) NHDPlusV2 (McKay et al., 2012) (left) and (B) NHDPlusHR (Moore et al., 2025) (right). Both panels share the same color scale; NHDPlusHR values are systematically higher because its finer resolution resolves channel splits not represented in NHDPlusV2, but the regional pattern (high density in the Lower Mississippi River and irrigation-dominated basins) is broadly preserved across datasets; the Pacific Northwest (HUC-2 17) is the least dense region in NHDPlusV2 but mid-range in NHDPlusHR. Denominators are WBD HUC-2 polygon areas.

traverses the full basin. Hardware and software-version details are provided in the analysis pipeline archived with this paper (see Code Availability, Section A2).

Figure 14 shows the wall-clock running time result. The bridge-based procedure scales approximately as $O(V^{1.02})$, consistent with the $O(V+E+\sum_i |D_i|^2)$ bound of Section 4.4.3 collapsing to near-linear behavior when most flowlines are bridges. The brute-force approach scales approximately as $O(V^{1.41})$, in line with its $O(V \cdot (V+E))$ worst case. On the Susquehanna River basin (~30,700 flowlines), bridge-based accumulation completes in 5.3 s (median of 5 iterations) whereas the brute-force approach requires 13.2 s — a ~2.5× speedup. Across a 30-basin HUC-4 (4-digit HUC) sample, the median speedup is 1.82×, with a maximum of 4.6× on a ~14,000-flowline basin; scaling from the 1.02 slope to the full 2.69-million-flowline CONUS network implies single-digit minutes for bridge-based accumulation. Applying the same scaling to the brute-force fit puts its CONUS run time at approximately two hours, a roughly 24× speedup at continental scale — substantially larger than the 1.82× median observed at HUC-4 scale because the gap between $V^{1.02}$ and $V^{1.41}$ widens with V.

These results confirm that the bridge-based procedure of Section 4.4.3 is an appropriate choice for total accumulation on continental-scale hydrography.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

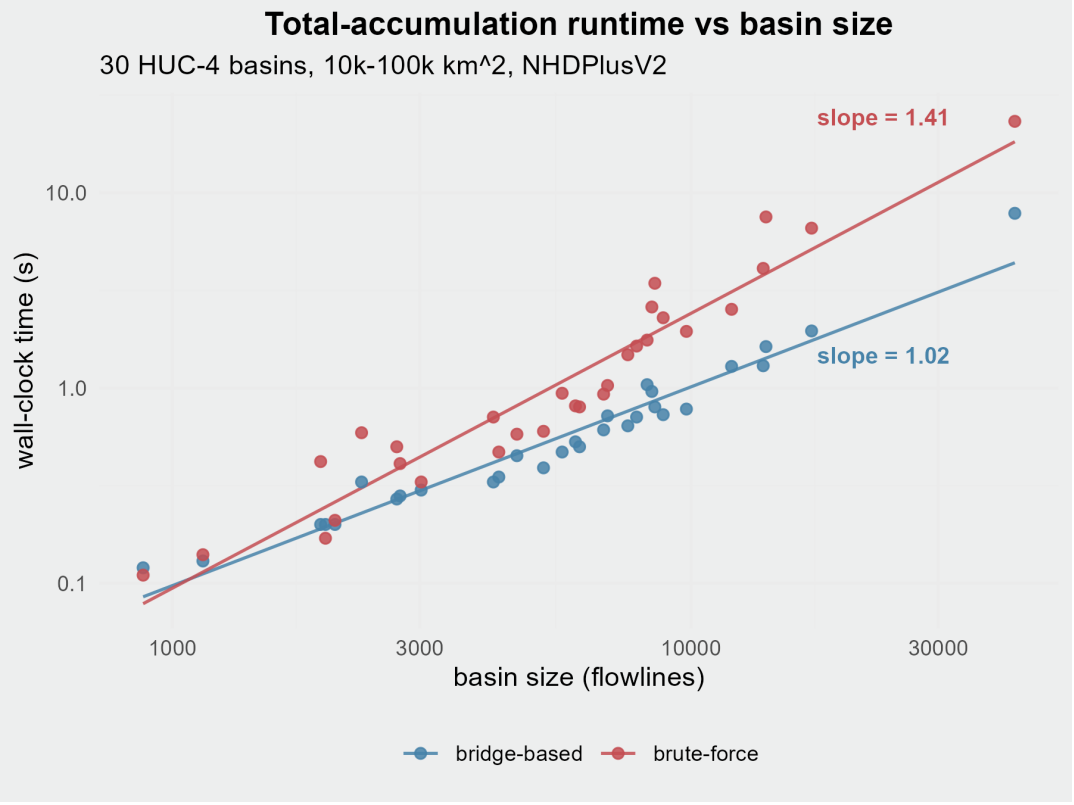

Figure 14. Wall-clock running time, in seconds (s), of brute-force versus bridge-based total accumulation on NHDPlus subsets of increasing size. Both axes are logarithmic.

6 Application

Two implementations demonstrate how the patterns of Section 3 and the algorithms of Section 4 are realized in practice. The 3DHP data model selects the junction-table pattern with primary-path flags to encode national-scale, non-dendritic connectivity. The `hydroloom` R package provides representation-conversion utilities and implements the algorithms of Section 4 across multiple table patterns. Neither set of choices is canonical; each reflects priorities specific to its setting, but together they show that the patterns and algorithms support national-scale representation and practical implementation.

6.1 3DHP data model

The 3DHP data model (Blodgett, 2026b; U.S. Geological Survey, 2023) and its initial dataset (Blodgett et al., 2024; U.S. Geological Survey, 2025) recast the high-resolution NHD (U.S. Geological Survey, National Geospatial Program, 2024) using HY_Features while retaining NHDPlus data model functionality. 3DHP illustrates one consistent set of choices among the

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

database pattern building blocks of Section 3, made with graph representation in mind even though graph operations themselves remain external to the data model.

The 3DHP flow network stores connectivity as a junction table with one row per upstream–downstream flowline connection, using from-identifier and to-identifier columns (`fromid3dhp` and `toid3dhp`). Boolean `upmain` and `downmain` attributes flag primary connections in each direction. This junction table encodes a many-to-many edge list between flowline nodes: edges with `upmain = TRUE` mark the primary upstream path and edges with `downmain = TRUE` mark the primary downstream path. To avoid duplicating topology and additional identifiers, 3DHP omits the NHDPlus `fromnode`/`tonode` construct; node identifiers can be created as needed following the algorithm described in Blodgett (2026b).

In practice, catchment size and data volume constraints prevent assigning a local catchment to every flowline (U.S. Geological Survey, 2023). Flowlines with nothing upstream and no channelization (termed "drainageways" in 3DHP) do not get their own catchment, and headwater catchments do not have a flowpath. A single catchment's flowpath may therefore comprise several flowlines. Each flowline carries a catchment-identifier column (`catchmentid3dhp`) indicating which catchment contains that flowline and a flowpath-identifier column (`flowpathid3dhp`) indicating if it is along the catchment's flowpath. Because every flowline resides in exactly one catchment, the flow network, which applies to all flowlines, supports construction of a catchment network. Figure 15 illustrates these relationships for a subset of the 3DHP network, showing flowpath flowlines (blue), non-flowpath flowlines (red), catchments with and without flowpaths, and hydrolocation features that mark confluences, headwaters, and catchment outlets.

This arrangement stores a fully resolved edge list in a many-to-many junction table that, combined with catchment/flowpath membership, can be transformed into a bipartite graph of catchments and nexuses. In the Figure 15 graph, headwater catchments flow to nexuses, catchment-junction nexuses connect catchments, and outlet catchments flow to terminal nexuses. Source nodes belong to the catchment partition and sink nodes to the nexus partition. The 3DHP data model does not store this bipartite structure directly, but a well-defined transformation can produce it. Figure 16 shows the entity-relationship structure of this design.

6.2 `hydroloom` R package

The `hydroloom` R package (Blodgett, 2023) is an open-source R implementation of the patterns of Section 3 and the algorithms of Section 4; it produced the cross-fabric measurements of Section 5. Conversion utilities transform among the self-referencing, junction, and catchments-as-edges representations, so a single set of algorithm implementations can serve people who have chosen any of the underlying patterns. For fast iteration, `hydroloom` stores the network as an adjacency list in a fixed-width matrix; geometry data volume far exceeds topology in practical hydrography use cases, so the memory cost is negligible. `hydroloom` is conceptually parallel to the `RivGraph` Python package (Schwenk and Hariharan, 2021), which extracts graph topology from remote-sensing imagery of river and delta channel networks. The two packages are complementary, with `RivGraph` operating from imagery to graph and `hydroloom` operating from existing hydrofabric topology.

The `hydroloom` R package implements all four navigation modes of Section 4 (upstream/downstream main, upstream-with-tributaries, downstream-with-diversions), divergence-aware stream order (Section 4.3), and the three accumulation modes

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

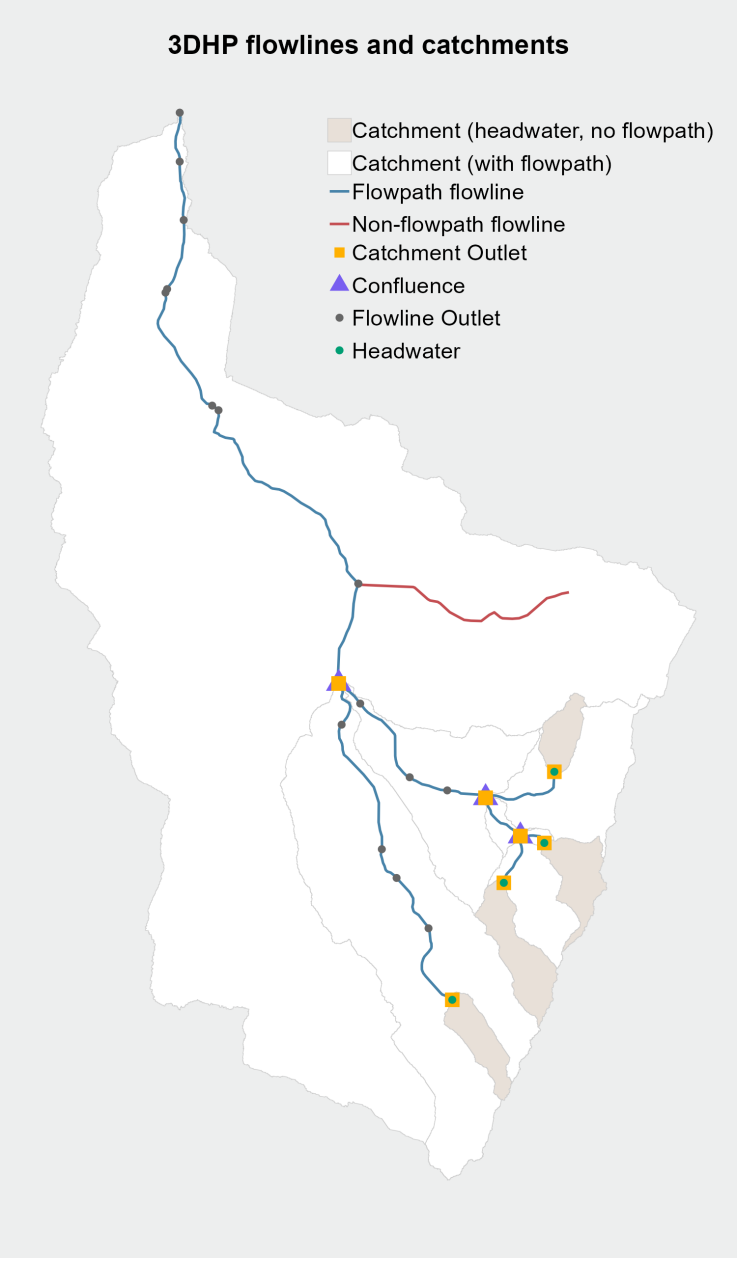

Figure 15. 3DHP flowlines and catchments for a network subset, illustrating the flowpath / non-flowpath flowline distinction and the four hydrolocation types.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

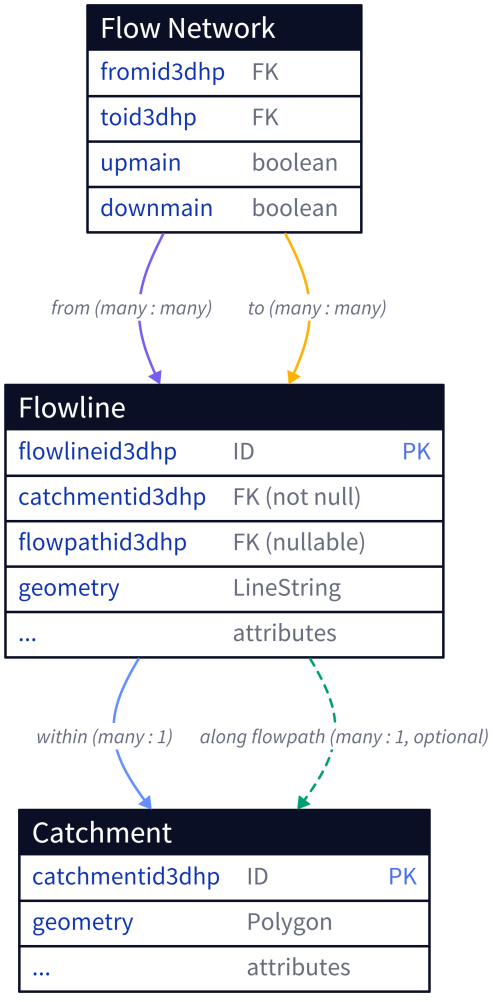

Figure 16. 3DHP data model entity-relationship diagram. The flow network junction table stores many-to-many flowline connections with primary-path flags. Each flowline belongs to exactly one catchment; flowlines along a catchment's flowpath carry an optional flowpath identifier. Combined with catchment/flowpath membership, this structure is transformable to a bipartite graph of catchments and nexuses.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

defined in Section 4.4 (divergence-apportioned, unapportioned, and total). Total accumulation uses the bridge-based procedure of Section 4.4.3: bridges identified once via Tarjan's algorithm, then a bounded upstream depth-first search per flowline that halts at already-processed bridges. Observed run-time behavior matches the complexity analysis of Section 4.4.3 — effectively linear scaling on networks dominated by dendritic structure, with the additional cost of repeated local searches showing up only in large diversion groups. Table 10 lists the `hydroloom` functions that implement each algorithm.

Table 10: `hydroloom` functions implementing the algorithms of Section 4.

Algorithm	`hydroloom` function
Primary-path assignment	`add_divergence()`
Network navigation (all four modes)	`navigate_network_dfs()`
Divergence-aware stream order	`add_streamorder()`
Total accumulation (bridge-based)	`accumulate_downstream()`
Bridge identification	`get_bridge_flowlines()`

7 Discussion

The empirical analysis in Section 5 makes two observations that ground this paper's contribution. First, non-dendritic flow is not an edge case in continental-scale hydrography. The "medium resolution" NHDPlusV2 records roughly 65,000 divergent nodes affecting 2.4% of CONUS flowlines, concentrated in low-relief coastal-plain, deltaic, and irrigation-dominated regions. The cross-hydrofabric measurements of Section 5.1 extend the same finding to two additional continents (Australia, Europe) and to the global scale, with per-basin densities spanning roughly an order of magnitude within most datasets and the same morphologies (anastomosing inland river systems, deltas, low-relief flat coasts, permafrost-thaw lake networks) recurring as divergence-dense across them. A hydrofabric data model or software that ignores these features may lose the non-dendritic structure reported here. Second, the algorithmic choice matters in practice: the bridge-based procedure of Section 4.4.3 runs in effectively linear time on NHDPlus, whereas the brute-force alternative of Section 4.4.2 becomes impractical at continental scale. These findings also clarify when each algorithm is appropriate. The choice between total and apportioned accumulation turns on what is being computed. For topologic quantities such as upstream contributing area, accumulated land use and land cover, source-attribution extent, and accumulated catchment-level climatologies, total accumulation is the correct algorithm; apportioned methods produce fractionated answers for quantities that do not partition with flow, and unapportioned methods discard the non-dendritic portion of the network entirely. For dynamic quantities such as steady-state water or sediment flux given known divergence fractions, apportioned accumulation is the right choice and runs in $O(V + E)$ on any topologically sortable network. The bridge-based procedure is therefore the appropriate default whenever the implementation must support topologic accumulation on networks that may contain diversions; the brute-force fallback remains acceptable only for small subsets where its quadratic cost is tolerable.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

The cross-fabric numbers presented in Section 5 measure how each hydrofabric represents non-dendritic structure, not how non-dendritic the underlying landscape is or what the nature and diversity of non-dendritic behavior is. A divergence count inherits the resolution and cartographic policy of the source product, regardless of how it is normalized. We read the empirical analysis presented here as a baseline characterization of digital hydrography rather than as an estimate of an idealized landscape property. Measuring the landscape directly — independent of representation — requires data fit for purpose. Representation-independent measurement of landscape divergence remains an open problem.

International hydrography datasets reflect two deliberate paradigms. Some datasets are non-dendritic-aware and encode divergences in their topology. Others are dendritic by construction, where omitting non-dendritic structure is a deliberate choice for the use cases targeted by each dataset. The patterns and algorithms offered here apply to the first paradigm. The Global River Topology of Wortmann et al. (2025) is a recent and complementary instance of the first paradigm — a bifurcating global product on which the patterns and algorithms presented operate without modification.

The patterns and algorithms presented here draw on observations from practice and relate them to design patterns from graph theory and database design. They are deliberately limited in scope as they do not interpret which implementation choices fit which use cases, nor do they prescribe conventions, standards, or operating procedures. Such normative work requires scientific interpretation and engineering that layers on top of the foundation the patterns provide. Two areas remain especially open: 1) more nuanced conventions for the type and behavior of divergent flows, and 2) more specific practices for representing hydrologic entities with bipartite graph structures.

7.1 Diversion classification and dynamics

The patterns and algorithms describe network topology, not flow dynamics. Within that scope, they accommodate any classification of diversion types: diversions within a river's bank-full channel versus those that diverge into a different corridor; diversions that stay within a drainage basin versus those that leave it; and diversions that terminate in closed basins versus those that rejoin and eventually reach the ocean. The choice of which classification to apply is a domain question outside this paper's scope, but the graph and database patterns do not constrain it.

The patterns and algorithms address the network structure and topology underlying divergent-flow dynamics without prescribing the dynamics themselves. A static "divergence fraction" attribute — specifying how much flow to apportion to each flowline downstream of a divergence — has been used in source-attribution modeling for several decades (e.g., Schwarz et al., 2006) and is published as a flowline-level attribute in NHDPlusV2 by Brakebill et al. (2020). Operational models such as RiverWare (Zagona et al., 2001) employ multiple non-dendritic object types (bifurcations, diversion objects, and distribution canals), each with distinct methods for computing flow splits that vary dynamically based on demand, physical capacity, and policy rules at each simulation timestep. Choosing among these treatments is implementation-specific and outside this paper's scope.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

7.2 Applications of bipartite structures

The catchments-as-edges / bipartite pattern identified in this paper is applied in NHDPlus with feature-identifier and from-node/to-node columns (`ComID`, `FromNode`, `ToNode` in NHDPlusV2) where each flowline connects through identified nodes. NHDPlus nodes serve as a non-dendritic topology utility, not as a consistent representation of physical entities. The 3DHP approach (Section 6) stores all topology as an edge list, avoiding explicit bipartite graph structures and leaving decisions about entity representation to users of the data.

Internationally, Australia's Geofabric stores `From_Node` / `To_Node` and an edge list; Canada's National Hydro Network uses `FROM_JUNCT` / `TO_JUNCT` junction-UUID columns with `DIRECTION` and `ISOLATED` flags, and its HY_Features-aligned successor, the Canadian Hydrospatial Network, replicates the NHDPlus `fromnode`/`tonode` pattern with divergence codes; and INSPIRE provides a node-edge graph that acknowledges divergent flow but prescribes no divergence codes or primary-flow designation. In each case the bipartite structure is adopted in its own way or left implicit. That these independent efforts converge on similar structures, without a shared vocabulary to recognize the convergence, reinforces the value of naming and formalizing the underlying patterns, as this paper aims to do.

7.3 Cross-hydrofabric questions

The empirical analysis presented here leaves several cross-hydrofabric questions open that could be further investigated. Whether the diversion-group size distribution behaves consistently across cartographic and physical heterogeneity — and, if so, whether the consistency reflects a true scaling regime in landscape branching, a shared cartographic convention propagating through derivative products, or coincidence at the sample sizes considered — is unresolved by the data presented. Continental-relief scaling of the bridge fraction is sampled here by one within-CONUS resolution pair plus the global GRIT skeleton; a systematic grain-effect treatment across more pairs could be its own analysis. Same-geography cross-fabric disagreement — GRIT North America against NHDPlus on the same continent — remains unexamined here and could directly test representation against landscape. The patterns, algorithms, and aggregates released here offer a starting point for additional work for any of these and other topics.

8 Conclusion

This paper has presented a small set of graph and database patterns for non-dendritic hydrologic networks. A set of reusable algorithms that operate on those patterns were developed and applied to the first cross-hydrofabric measurement of non-dendritic prevalence for six hydrofabrics on three continents and a global compilation. This evaluation found that diverted flow is a pervasive feature of digital hydrography rather than a regional curiosity. The bridge-based total-accumulation procedure scales to continental hydrography with formal correctness and complexity guarantees, where the brute-force alternative does not.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Just as Strahler stream order (Strahler, 1957) enabled decades of cross-basin comparison in dendritic hydrology, naming the patterns and algorithms underlying non-dendritic representation offers common ground for comparing data and software across the broader hydroscience community. Conventions, standards, and use-case-specific logic can layer on top of the foundation provided here.

Appendix A: Proofs of Theorem 1 and Lemma 1

A1 Proof of Theorem 1 (Correctness)

By induction on the upstream-to-downstream order of e.

Base case. If e is a headwater flowline (no incoming edges), then $R(e) = \{e\}$. The depth-first search in step 3 finds no further flowlines, so $U = \{e\}$ and $B_U = \emptyset$, and eq. 1 gives $T(e) = w(e)$, which matches eq. 3.

Inductive step. Assume $T(e') = \sum_{e'' \in R(e')} w(e'')$ for all e' processed before e. Every flowline $f \in R(e)$ with $f \neq e$ admits a directed path π in G from $\mathrm{head}(f)$ to $\mathrm{tail}(e)$ (by definition of $R(e)$); the reverse of π is the path the upstream search would attempt to follow from $\mathrm{tail}(e)$ toward $\mathrm{head}(f)$. Three cases arise:

- All edges of π are non-bridges and f itself is a non-bridge. The search reaches $\mathrm{head}(f)$ along the reverse of π and traverses f as an incoming non-bridge edge of $\mathrm{head}(f)$, so $f \in U$.
- All edges of π are non-bridges but f itself is a bridge. The search reaches $\mathrm{head}(f)$ as above and halts at f because f is a bridge, so $f \in B_U$ and $f \in R(f)$ trivially.
- π contains at least one bridge. Let b be the bridge on π closest to $\mathrm{tail}(e)$. The search proceeds through non-bridges from $\mathrm{tail}(e)$ to $\mathrm{head}(b)$ and halts at b, so $b \in B_U$, and f lies upstream of b on π, hence $f \in R(b)$.

So $R(e) \subseteq U \cup \left(\bigcup_{b \in B_U} R(b)\right)$, and the reverse inclusion is immediate. It remains to show that this union is disjoint, so that eq. 1 does not double-count any $w(f)$.

- For any $b \in B_U$, $U \cap R(b) = \emptyset$: U contains only e together with non-bridge flowlines whose heads are reachable from $\mathrm{tail}(e)$ via non-bridge upstream traversal, whereas $R(b)$ consists of b and flowlines strictly upstream of b. The bridge b itself is in B, hence not in U. Any non-bridge flowline upstream of b would require — together with the non-bridge path from $\mathrm{tail}(e)$ that reaches $\mathrm{head}(b)$ — an undirected cycle through b in G_u (the undirected projection of G, as in Theorem 1), contradicting the bridge property of b. Finally, e cannot be in $R(b)$ in a directed acyclic graph: b is upstream of e, so e being upstream of b would close a directed cycle.
- For distinct $b_1, b_2 \in B_U$, $R(b_1) \cap R(b_2) = \emptyset$ by Lemma 1 (proved below).

Therefore $R(e)$ is the disjoint union of U and $\{R(b) : b \in B_U\}$, and eq. 1 sums each $w(f)$ for $f \in R(e)$ exactly once. Each $b \in B_U$ is reached during the upstream search from $\mathrm{tail}(e)$, so $\mathrm{head}(b)$ is upstream of $\mathrm{tail}(e)$ in G; hence b is upstream of e in G and appears before e in the topological order, so $T(b)$ has already been computed and the inductive hypothesis applies. Combining yields $T(e) = \sum_{f \in R(e)} w(f)$. ■

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

A2 Proof of Lemma 1

Because b_1 is a bridge with $\mathrm{head}(b_1) \in V(C)$, $\mathrm{tail}(b_1) \notin V(C)$ (otherwise b_1 and C's non-bridge path between its endpoints would form a cycle through b_1, contradicting the bridge property). The same holds for b_2. Within C there is an undirected walk Q from $\mathrm{head}(b_2)$ to $\mathrm{head}(b_1)$ using only non-bridge edges (degenerating to the empty walk when $\mathrm{head}(b_1) = \mathrm{head}(b_2)$).

Suppose for contradiction that $f \in R(b_1) \cap R(b_2)$. Then there are directed paths π_1 from $\mathrm{head}(f)$ to $\mathrm{tail}(b_1)$ and π_2 from $\mathrm{head}(f)$ to $\mathrm{tail}(b_2)$ in G. Since π_1 ends at $\mathrm{tail}(b_1)$ it does not traverse b_1, and symmetrically π_2 does not traverse b_2. We claim π_1 also does not traverse b_2: traversing b_2 in flow direction would carry π_1 into $V(C)$, and π_1 would then have to exit $V(C)$ in flow direction through a bridge b' with $\mathrm{tail}(b') \in V(C)$ (exit through a non-bridge would keep both endpoints in $V(C)$) to reach $\mathrm{tail}(b_1) \notin V(C)$. Here $b' \neq b_1$ because $\mathrm{tail}(b_1) \notin V(C)$, and the segment of π_1 from $\mathrm{head}(b')$ to $\mathrm{tail}(b_1)$ does not traverse b_1 (it ends at $\mathrm{tail}(b_1)$, and traversing b_1 first would close a directed cycle). Because C is connected, there is a non-bridge walk Q' in C from $\mathrm{head}(b_1)$ to $\mathrm{tail}(b')$. Concatenating Q', b', and that segment of π_1 gives a walk in G_u from $\mathrm{head}(b_1)$ to $\mathrm{tail}(b_1)$ that avoids b_1, contradicting the bridge property of b_1. The symmetric argument shows π_2 does not traverse b_1.

The reverse of π_1 followed by π_2 is therefore a walk P in G_u from $\mathrm{tail}(b_1)$ to $\mathrm{tail}(b_2)$ that avoids both b_1 and b_2. Concatenating P, b_2 (traversed undirected from $\mathrm{tail}(b_2)$ to $\mathrm{head}(b_2)$), and Q yields a walk in G_u from $\mathrm{tail}(b_1)$ to $\mathrm{head}(b_1)$ that avoids b_1. Hence, removing b_1 does not disconnect $\mathrm{tail}(b_1)$ from $\mathrm{head}(b_1)$, contradicting the bridge property of b_1 (Bondy and Murty, 1976, Theorem 2.3). ■

For any pair of distinct bridges in B_U produced by step 3 of the bridge-based accumulation procedure (Section 4.4.3), the heads lie in the connected component of $G_u \setminus B$ containing $\mathrm{tail}(e)$ (or, if $\mathrm{tail}(e)$ has no incident non-bridge edges, all such heads equal $\mathrm{tail}(e)$ itself). The hypothesis of Lemma 1 is therefore satisfied for every pair of distinct bridges in B_U, completing the disjointness argument used in Theorem 1.

Code availability

The `hydroloom` R package (Blodgett, 2023), which implements the algorithms presented in this paper, is available at https://doi.org/10.5066/P9AQCUY0. The analysis pipeline and processing scripts used to generate the tables and figures of this paper are available as a U.S. Geological Survey software release (Blodgett, 2026a) at https://doi.org/10.5066/P13T7FHN.

Data availability

The 3DHP downloadable datasets (U.S. Geological Survey, 2025) are available at https://doi.org/10.5066/P148NT7B, and the 3DHP network dataset (Blodgett et al., 2024) is available at https://doi.org/10.5066/P94H0DAG. Source links and access details for the international hydrofabrics analyzed in Section 5 (NHDPlusV2, NHDPlusHR, AGF V3.3, NHN, EU-Hydro, GRIT v1.0) are all available from the cited sources.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Author contributions

D. Blodgett conceived the study, developed the graph and database patterns, implemented the algorithms in the `hydroloom` R package, performed the cross-fabric empirical analysis of Section 5, and wrote the manuscript. T. Chegini formalized the bridge-based procedure (Theorem 1, Lemma 1, and the complexity analysis of Section 4.4.3), restructured the manuscript around its algorithmic contributions, sharpened the positioning relative to the spectral graph-theoretic literature on delta channel networks, and prepared the Copernicus submission package. R. Gillespie originally suggested the use of bridges to partition the network when calculating an accumulation, reviewed the manuscript and suggested single-pass version of the non-dendritic accumulation algorithm. R. Gillespie also raised a continental-scale adjacency-matrix memory concern that motivated the quantification of bridge fraction across NHDPlus in Section 5.2. S. Kehoe was key in establishing the bridge-based accumulation method, reviewed the manuscript's framing, and contributed comments on the empirical analysis.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors thank the teams responsible for the six hydrofabrics analyzed in Section 5 for producing and openly distributing the underlying data: the U.S. Geological Survey for NHDPlusV2 and NHDPlusHR, Geoscience Australia and the Bureau of Meteorology for the Australian Geofabric, Natural Resources Canada for the National Hydro Network, the Copernicus Land Monitoring Service for EU-Hydro, and M. Wortmann and colleagues for the Global River Topology.

The authors thank USGS colleague reviewers and HESS community reviewers for their thoughtful and constructive feedback in preparing this paper.

Claude (Anthropic, Opus 4.x via the Claude Code CLI) was used during manuscript preparation to assist with editing prose, review of analysis code, and generation of diagrams and flowcharts. All AI-assisted output was reviewed, verified, and edited by the authors.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Financial support

Preparation of this manuscript was supported by USGS appropriations for the Water Resources Mission Area (Water Availability and Use Science Program) and Core Science Systems Mission Area (National Geospatial Program). Taher Chegini's contribution was made in kind.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

References

Agência Nacional de Águas: Base Hidrográfica Ottocodificada Multiescalas 2017 (BHO 2017), https://metadados.snirh.gov.br/geonetwork/srv/api/records/0c698205-6b59-48dc-8b5e-a58a5dfcc989, 2017.

Aho, K., Kriloff, C., Godsey, S. E., Ramos, R., Wheeler, C., You, Y., Warix, S., Derryberry, D., Zipper, S., Hale, R. L., Bond, C. T., and Kuehn, K. A.: Non-perennial stream networks as directed acyclic graphs: The R-package streamDAG, Environmental Modelling & Software, https://doi.org/10.1016/j.envsoft.2023.105775, 2023.

Baddeley, A., Rubak, E., and Turner, R.: Spatial Point Patterns: Methodology and Applications with R, Chapman and Hall/CRC, Boca Raton, FL, https://doi.org/10.1201/b19708, 2015.

Blodgett, D.: Non-Dendritic Hydrologic Network Analysis, https://doi.org/10.5066/P13T7FHN, 2026a.

Blodgett, D. L.: Mainstem Rivers of the Conterminous United States (ver. 3.1, October 2026), https://doi.org/10.5066/P13LNDDQ, 2022.

Blodgett, D. L.: hydroloom: Utilities to Weave Hydrologic Fabrics, https://doi.org/10.5066/P9AQCUY0, 2023.

Blodgett, D. L.: Logical data model for hydrographic data based on HY_Features concepts, Open Geospatial Consortium, http://www.opengis.net/doc/PER/hydrofabric, 2026b.

Blodgett, D. L. and Dornblut, I.: OGC WaterML 2: Part 3 — Surface Hydrology Features (HY_Features) — Conceptual Model, Open Geospatial Consortium, http://www.opengis.net/doc/IS/hy-features/1.0, 2018.

Blodgett, D. L. and Johnson, J. M.: hyRefactor: Tools for Manipulating the NHDPlus Network in Preparation for Hydrologic Modeling, https://doi.org/10.5066/P9EP9UQZ, 2024.

Blodgett, D. L., Johnson, J. M., and Bock, A. R.: Generating a reference flow network with improved connectivity to support durable data integration and reproducibility in the coterminous US, Environmental Modelling & Software, https://doi.org/10.1016/j.envsoft.2023.105726, 2023.

Blodgett, D. L., Ritmiller, C. L., Gold, B. L., Hayes, L., and Wieczorek, M. E.: Network Attributes for the High Resolution National Hydrography Dataset (NHD) based initial 3D Hydrography Program Network, https://doi.org/10.5066/P94H0DAG, 2024.

Bondelid, T., Johnston, J., McKay, L., Moore, R., and Rea, A.: NHDPlus Version 1 (NHDPlusV1) User Guide, U.S. Environmental Protection Agency, https://www.epa.gov/waterdata/nhdplusv1-documentation, last access: July 2025, 2010.

Bondy, J. A. and Murty, U. S. R.: Graph Theory with Applications, North-Holland, New York, 1976.

Brakebill, J. W., Schwarz, G. E., and Wieczorek, M. E.: An enhanced hydrologic stream network based on the NHDPlus medium resolution dataset, U.S. Geological Survey, https://doi.org/10.3133/sir20195127, 2020.

Bureau of Meteorology: Australian Hydrological Geospatial Fabric (Geofabric) Product Guide, Version 3.0, Bureau of Meteorology, Melbourne, Australia, http://www.bom.gov.au/water/geofabric/, 2015.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Caughman, J. S. and Veerman, J. J. P.: Kernels of directed graph Laplacians, Electronic Journal of Combinatorics, 13, R39, https://doi.org/10.37236/1065, 2006.

Clauset, A., Shalizi, C. R., and Newman, M. E. J.: Power-law distributions in empirical data, SIAM Review, 51, 661–703, https://doi.org/10.1137/070710111, 2009.

Copernicus Land Monitoring Service: EU-Hydro River Network Database, version 1.3, https://land.copernicus.eu/en/products/eu-hydro/eu-hydro-river-network-database, 2019.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C.: Introduction to Algorithms, 4th ed., MIT Press, Cambridge, MA, 2022.

Cosgrove, B., Gochis, D., Flowers, T., Dugger, A., and Ogden, F.: NOAA's National Water Model: Advancing operational hydrology through continental-scale modeling, Journal of the American Water Resources Association, https://doi.org/10.1111/1752-1688.13184, 2024.

Fullerton, A. H., Burnett, K. M., Steel, E. A., Flitcroft, R. L., Pess, G. R., Feist, B. E., Torgersen, C. E., Miller, D. J., and Sanderson, B. L.: Hydrological connectivity for riverine fish: measurement challenges and research opportunities, Freshwater Biology, https://doi.org/10.1111/j.1365-2427.2010.02448.x, 2010.

Gillespie, C. S.: Fitting heavy tailed distributions: the poweRlaw package, Journal of Statistical Software, 64, 1–16, https://doi.org/10.18637/jss.v064.i02, 2015.

Hill, R. A., Weber, M. H., Leibowitz, S. G., Olsen, A. R., and Thornbrugh, D. J.: The Stream-Catchment (StreamCat) Dataset: A database of watershed metrics for the conterminous United States, Journal of the American Water Resources Association, 52, 120–128, https://doi.org/10.1111/1752-1688.12372, 2016.

Horn, C. R.: Reach File Manual, U.S. Environmental Protection Agency, https://water.usgs.gov/GIS/browse/rf1_appA.HTML, 1986.

Howard, A. D.: Drainage analysis in geologic interpretation: A summation, AAPG Bulletin, 51, 2246–2259, https://doi.org/10.1306/5D25C26D-16C1-11D7-8645000102C1865D, 1967.

INSPIRE Thematic Working Group Hydrography: D2.8.I.8 Data Specification on Hydrography — Technical Guidelines, INSPIRE Maintenance and Implementation Group (MIG), https://inspire.ec.europa.eu/id/document/tg/hy, 2024.

Leopold, L. B., Wolman, M. G., and Miller, J. P.: Fluvial Processes in Geomorphology, W.H. Freeman, San Francisco, 1964.

Maidment, D. R., Ye, Z., and Dartiguenave, C.: Application of the ArcHydro schema to GIS-based hydrologic modeling, University of Texas at Austin, 2001.

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., and Rea, A.: NHDPlus version 2: User guide, U.S. Environmental Protection Agency, https://www.epa.gov/system/files/documents/2023-04/NHDPlusV2_User_Guide.pdf, 2012.

Metcalf & Eddy, Inc., University of Florida, and Water Resources Engineers, Inc.: Storm Water Management Model: Volume I—Final Report, U.S. Environmental Protection Agency, Washington, D.C., 1971.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Moore, R. B., McKay, L. D., Rea, A. H., Bondelid, T. R., Price, C. V., Dewald, T. G., and Hayes, L.: User's guide for the National Hydrography Dataset Plus High Resolution (NHDPlus HR), U.S. Geological Survey, https://doi.org/10.3133/sir20255031, 2025.

Natural Resources Canada: National Hydro Network (NHN), GeoBase Series, https://ftp.maps.canada.ca/pub/nrcan_rncan/vector/geobase_nhn_rhn/, 2024.

Natural Resources Canada: Canadian Hydrospatial Network (CHN) Data Model Edition 0.2, Natural Resources Canada, GeoBase, https://open.canada.ca/data/en/dataset/ae385105-e48c-4b54-bd0f-dfb7303301cb, 2025.

Nolan, J. V., Brakebill, J. W., Alexander, R. B., and Schwarz, G. E.: ERF1_2—Enhanced River Reach File 2.0, https://doi.org/10.5066/P9JVHVND, 2002.

O'Callaghan, J. F. and Mark, D. M.: The extraction of drainage networks from digital elevation data, Computer Vision, Graphics, and Image Processing, 28, 323–344, https://doi.org/10.1016/S0734-189X(84)80011-0, 1984.

Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M.-J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., and Wenger, S. J.: Modelling dendritic ecological networks in space: an integrated network perspective, Ecology Letters, https://doi.org/10.1111/ele.12084, 2013.

Pfafstetter, O.: Classification of hydrographic basins: coding methodology, Departamento Nacional de Obras de Saneamento, Rio de Janeiro, 1989.

Phillips, R. W., Spence, C., and Pomeroy, J. W.: Connectivity and runoff dynamics in heterogeneous basins, Hydrological Processes, 25, 3061–3075, https://doi.org/10.1002/hyp.8123, 2011.

Porta, S., Crucitti, P., and Latora, V.: The network analysis of urban streets: A dual approach, Physica A - Statistical Mechanics and its Applications, 369, 853–866, https://doi.org/10.1016/j.physa.2005.12.063, 2006a.

Porta, S., Crucitti, P., and Latora, V.: The network analysis of urban streets: A primal approach, Environment and Planning B - Planning and Design, 33, 705–725, https://doi.org/10.1068/b32045, 2006b.

R Core Team: R: A language and environment for statistical computing, version 4.6.0, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, last access: 23 October 2026, 2026.

Robertson, D. M. and Saad, D. A.: Nitrogen and phosphorus sources and delivery from the Mississippi/Atchafalaya River Basin: An update using 2012 SPARROW models, Journal of the American Water Resources Association, https://doi.org/10.1111/1752-1688.12905, 2021.

Schwarz, G. E., Hoos, A. B., Alexander, R. B., and Smith, R. A.: The SPARROW surface water-quality model: Theory, application and user documentation, U.S. Geological Survey Techniques and Methods 6-B3, U.S. Geological Survey, https://doi.org/10.3133/tm6B3, 2006.

Schwenk, J. and Hariharan, J.: RivGraph: Automatic extraction and analysis of river and delta channel network topology, Journal of Open Source Software, 6, 2952, https://doi.org/10.21105/joss.02952, 2021.

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Stein, J. L.: An enhanced Pfafstetter catchment reference system, Water Resources Research, 54, 9951–9963, https://doi.org/10.1029/2018WR023218, 2018.

Strahler, A. N.: Quantitative analysis of watershed geomorphology, Transactions of the American Geophysical Union, 38, 913–920, https://doi.org/10.1029/TR038i006p00913, 1957.

Tarboton, D. G., Bras, R. L., and Rodriguez-Iturbe, I.: The fractal nature of river networks, Water Resources Research, 24, 1317–1322, https://doi.org/10.1029/WR024i008p01317, 1988.

Tarjan, R. E.: Depth-first search and linear graph algorithms, SIAM Journal on Computing, 1, 146–160, https://doi.org/10.1137/0201010, 1972.

Tarjan, R. E.: A note on finding the bridges of a graph, Information Processing Letters, 2, 160–161, https://doi.org/10.1016/0020-0190(74)90003-9, 1974.

Tejedor, A., Longjas, A., Zaliapin, I., and Foufoula-Georgiou, E.: Delta channel networks: 1. A graph-theoretic approach for studying connectivity and steady state transport on deltaic surfaces, Water Resources Research, 51, 3998–4018, https://doi.org/10.1002/2014WR016577, 2015a.

Tejedor, A., Longjas, A., Zaliapin, I., and Foufoula-Georgiou, E.: Delta channel networks: 2. Metrics of topologic and dynamic complexity for delta comparison, physical inference, and vulnerability assessment, Water Resources Research, 51, 4019–4045, https://doi.org/10.1002/2014WR016604, 2015b.

U.S. Geological Survey: 3D Hydrography Program Products: 3DHP_all 2023 Service Specification, U.S. Geological Survey, National Geospatial Program, https://www.usgs.gov/ngp-standards-and-specifications/3d-hydrography-program-products-3dhpall-2023-s 2023.

U.S. Geological Survey: USGS 3D Hydrography Program (3DHP) Downloadable Datasets 2024, https://doi.org/10.5066/P148NT7B, 2025.

U.S. Geological Survey in cooperation with the U.S. Environmental Protection Agency: National Hydrography Dataset, USGS, Reston, VA, 1999.

U.S. Geological Survey, National Geospatial Program: USGS High Resolution National Hydrography Dataset (NHD), https://www.usgs.gov/core-science-systems/ngp/national-hydrography, 2024.

Vuong, Q. H.: Likelihood ratio tests for model selection and non-nested hypotheses, Econometrica, 57, 307–333, https://doi.org/10.2307/1912557, 1989.

Wang, W., Lu, H., Yang, D., Sothea, K., and Jiao, Y.: Modeling daily floods in the Lancang-Mekong River Basin using an improved hydrological-hydrodynamic model, Water Resources Research, https://doi.org/10.1029/2021WR029734, 2021.

Wieferich, D. J., Williams, B., Falgout, J. T., and Foks, N. L.: xstrm, https://doi.org/10.5066/P9P8P7Z0, 2021.

Wortmann, M., Slater, L., Hawker, L., Liu, Y., Neal, J., Zhang, B., Schwenk, J., Allen, G., Ashworth, P., Boothroyd, R., Cloke, H., Delorme, P., Gebrechorkos, S. H., Griffith, H., Leyland, J., McLelland, S., Nicholas, A. P., Sambrook Smith, G., Vahidi, E.,

https://doi.org/10.5194/egusphere-2026-6636
Preprint. Discussion started: 31 October 2026
Public domain. CC0 1.0.

Parsons, D., and Darby, S. E.: Global River Topology (GRIT): A Bifurcating River Hydrography, Water Resources Research, 61, e2024WR038308, https://doi.org/10.1029/2024WR038308, 2025.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resources Research, 55, 5053–5073, https://doi.org/10.1029/2019WR024873, 2019.

Yu, H., Ai, T., Yang, M., Huang, L., and Yuan, J.: A recognition method for drainage patterns using a graph convolutional network, International Journal of Applied Earth Observation and Geoinformation, 107, 102696, https://doi.org/10.1016/j.jag.2022.102696, 2022.

Zagona, E. A., Fulp, T. J., Shane, R., Magee, T., and Goranflo, H. M.: RiverWare: A Generalized Tool for Complex Reservoir System Modeling, Journal of the American Water Resources Association, 37, 913–929, https://doi.org/10.1111/j.1752-1688.2001.tb05522.x, 2001.