Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Topology-Aware Graph Neural Networks for River Network Reconstruction: Prediction and Classification

Neelima N V

nyayapativenkataneelima@uq.edu.au

The University of Queensland

Rashmita Soni

IIT Delhi

Shaurya Shriyam

IIT Delhi

Research Article

Keywords: River Network, GNN, GAT, Network reconstruction, Hetero-GAT, GraphSAGE

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11295179/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Topology-Aware Graph Neural Networks for River Network Reconstruction: Prediction and Classification

Neelima N V[1*], Rashmita Soni[2†] and Shaurya Shriyam[3]

[1*]UQ-IITD Research Academy, IIT Delhi, Hauz Khas, 110016, New Delhi, India.
[2]Department of Mechanical Engineering, IIT Delhi, Hauz Khas, 110016, New Delhi, India.

*Corresponding author(s). E-mail(s): nyayapativenkataneelima@uq.edu.au, Phone: +91 9902736887;
Contributing authors: rashmitasoni09@gmail.com;
†These authors contributed equally to this work.

Abstract

River networks contribute to the spatial organisation of water, sediment, and ecological connectivity across landscapes. As global hydrographic datasets remain incomplete and basin monitoring needs to grow, there is increasing demand for methods that can perform network reconstruction. This study evaluates graph neural networks (GNNs) for two reach-level tasks on four global river basins, Amur, Ganga, Mekong, and Mississippi: directed link prediction for network reconstruction and Strahler order classification. Four architectures were compared: GraphSAGE, graph attention network (GAT), heterogeneous GAT (Hetero-GAT), and a proposed weighted directed MixHop model. The proposed model outperforms the baselines in most link prediction settings and in all basins for Strahler classification, achieving balanced accuracy of approximately 92.8–94.8% at low sparsity (10% missing edges) for link prediction and macro-F1 of approximately 90.6–93.6% for Strahler classification across basins. Overall, the results demonstrate that topology-aware GNNs can extract hydrologically meaningful patterns from river networks and offer a promising approach to reconstructing incomplete hydrographic connectivity and supporting basin-scale stream classification.

Keywords: River Network, GNN, GAT, Network reconstruction, Hetero-GAT, GraphSAGE

1 Introduction and related work

Rivers serve as the backbone of a civilisation. These networks are the primary medium for water, sediment, and nutrient transport and serve as the largest conduit between continents, oceans, and the atmosphere [1]. Hence, the structure and topology of these complex networks directly influence ecosystems. Rinaldo et al. 2024 asserted that river networks act as corridors that sustain aquatic life and facilitate migration of species, optimising energy dissipation across landscapes due to their universal branching patterns [2]. Furthermore, river network topology provides information on network spatial characteristics that helps with flood management and infrastructure resilience. These inputs are crucial for extreme events caused by climate change. [3, 4].

In a rapidly urbanising world, these networks have been evolving due to anthropogenic stresses; hence, understanding the topological characteristics of river networks is essential for informed decision-making for river-oriented sustainable development. Although studies have accumulated information on global river basins, there is still scope to understand river networks better.

Monitoring of river hydrological attributes has evolved from in-situ measurements to comprehensive remote sensing frameworks. Satellite imagery has enabled observations of elevation and slope of water surface providing data for ungauged stations. However, data scarcity remains a significant challenge, indicating the need for integration of physical sensor-based data with remote sensing data, validation enhancing the accuracy of hydrological models [5].

Global river topology datasets, such as Global River Topology (GRIT), are being developed to represent bifurcating and multi-threaded channels, overcoming the limitations of traditional convergent-flow assumptions in digital elevation models [6]. Graph neural networks have increasingly been applied to river systems beyond topology reconstruction alone, reflecting their general suitability for the non-Euclidean, directionally structured nature of hydrographic data. GNNs utilise message-passing mechanisms to aggregate information from neighbouring nodes, respecting the directed acyclic graph (DAG) topology of river systems [7]. This architecture allows for the explicit incorporation of flow direction as a physical prior, ensuring that information propagates upstream-to-downstream consistent with mass conservation principles.Buchhorn et al. [8] applied a Graph Deviation Network to real-time water-quality sensor data from the Herbert River, demonstrating that GNN-based spatio-temporal modelling can detect persistent, variable-length anomalies in multivariate river sensor networks despite the scarcity of labelled anomaly data , a challenge that similarly motivates graph-based approaches to structural (rather than temporal) river data. Wang et al. [3] showed that the tree-like topology of river networks induces over-squashing in GNNs, where signals from distant nodes become excessively compressed during message passing, and proposed a reachability-based dense graph transformation to mitigate this for flood forecasting. Plettenberg et al. [9] identified that standard attention-based GNNs fail to respect the conservation-law structure inherent to physical flow networks (evaluated on electronic circuits and power grids), proposing a flow-attention mechanism constrained to satisfy Kirchhoff's first law. Together, these studies establish that generic GNN architectures require explicit structural modification whether through graph transformation, attention redesign, or directional

propagation, to adequately capture the physical constraints of flow-conserving, tree-like networks such as river systems, motivating the direction-aware and multi-hop design adopted in this work.

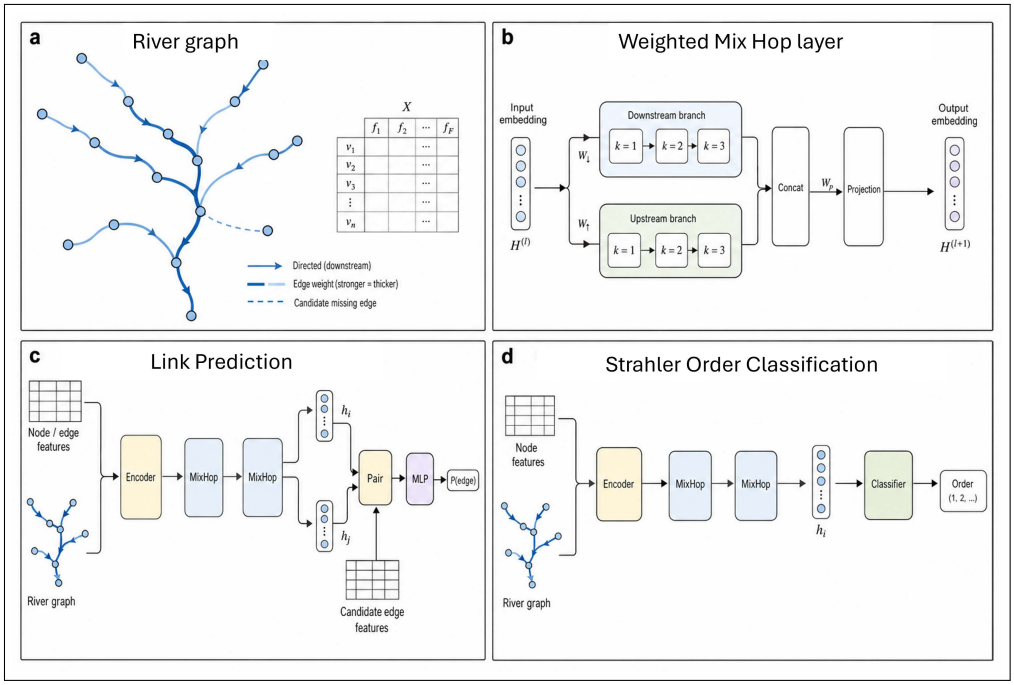


Fig. 1 Graphical Abstract

GNN-based link prediction for spatially embedded networks has been explored directly for river topology reconstruction. Fan et al. [10] proposed a Generic Multimodal Spatially Graph Convolutional Network (GMu-SGCN) that jointly encodes multimodal node and edge features for edge-existence prediction, evaluated on both a river network and a power network test bed, reporting improved accuracy over a position-only GraphSAGE baseline on the power network. This result reinforces that incorporating richer edge- and feature-level information substantially improves reconstruction over baseline neighbourhood-aggregation methods. However, there is still a gap in modelling directed networks and their performance evaluation, systematically across geographically diverse river basins. Additionally, link prediction methods for directed graphs are frequently evaluated inconsistently. This critique directly motivates the architectural comparison undertaken in this study, in which direction-naive (GraphSAGE, GAT) and direction-aware (Hetero-GAT, weighted directed MixHop) encoders are evaluated under an identical protocol. The multi-hop propagation mechanism used in the proposed architecture builds on Abu-El-Haija et al. [11], who

introduced higher-order neighbourhood mixing via sparsified multi-hop graph convolutions; this study extends that formulation to a directed, hydrologically weighted setting.

In this work, a Weighted Directed Graph Neural Network (Weighted MixHop) framework (Figure 1) for reconstructing missing connectivity in large-scale river networks has been proposed. The central idea is to learn reach-level embeddings that respect the physical direction of water flow, so that the predicted probability of a directed link between two reaches reflects both their learned hydrological compatibility and their spatial and hierarchical continuity along the drainage network. Unlike many existing GNN-based approaches applied to environmental networks, this approach explicitly investigates which hydrological and geometric signals drive successful reconstruction, thereby providing an interpretable account of what the model learns. The main contributions of this work are:

1. This study introduces a weighted directed Mix Hop architecture that combines direction-aware propagation, multi-hop diffusion, and hydrology-informed edge weighting within a single encoder, and shows that this combination consistently outperforms undirected and single-hop baselines (GraphSAGE, GAT, Hetero-GAT) for directed link reconstruction across four large, structurally distinct river basins.
2. Systematic evaluation of reconstruction performance across a wide range of network sparsity (10–90% missing edges) and, through post-hoc permutation importance, identify which hydrologic and geometric features (stream order, discharge, spatial endpoint proximity) drive successful link prediction for each architecture, showing that reconstruction success is tied to specific, physically interpretable signals rather than generic structural similarity.
3. Further, Strahler stream-order classification has been performed using the same reach-level river networks, which shows that the learned embeddings support both tasks well, with the two tasks drawing on different but hydrologically dominant features.

2 Methodology

This section describes the methodology adopted in this study for both link prediction and node classification tasks, including the model validation procedure. This study

Table 1 Summary of model characteristics.

Characteristic	GraphSAGE	GAT	Hetero-GAT	Weighted MixHop
Aggregation	Mean	Attention	Relation-wise attention	Weighted MixHop
Explicit up/downstream split	No	No	Yes	Yes
Multi-hop within layer	No	No	No	Yes ($K = 3$)
Hydrologic edge weights	No	No	Yes	Yes
Candidate-edge features in decoder	No	No	Yes	Yes
Hidden size	64	64	64	64

represents each river basin as a directed reach-level graph derived from HydroRIVERS vector data. A *reach* is a river segment between confluences (or equivalent breakpoints in the hydrographic product). The same graph representation is used for two tasks: (i) directed link prediction for network reconstruction, and (ii) Strahler-order classification of reaches. Four architectures are compared under a common protocol: GraphSAGE, GAT, Hetero-GAT, and the proposed weighted directed MixHop model. Table 1 summarises the model characteristics adopted in this study.

Table 2 maps HydroRIVERS [12] attribute names to the symbols used below. Identifiers such as reach ID and next-downstream ID are used to construct topology.

Table 2 Notation used in the study

Symbol	Description	HydroRIVER attribute name
i, u, v	River reaches (nodes)	-
id_i	Unique reach identifier	HYRIV_ID
$\mathrm{next}(i)$	Downstream neighbour of reach i	NEXT_DOWN
ℓ_i	Reach length (km)	LENGTH_KM
ρ_i	Reach sinuosity	sinuosity
d_i	Distance to outlet (km)	DIST_DN_KM
s_i	Strahler order	ORD_STRA
Q_i	Mean discharge (m^3/s)	DIS_AV_CMS
r_i	Main-river identifier	MAIN_RIV
$c_{x,i}, c_{y,i}$	Projected reach centroid	derived from geometry
g_{uv}	Endpoint gap between u and v (km)	derived
Δd_{uv}	Downstream-distance difference	$d_u - d_v$
Δs_{uv}	Strahler-order difference	$s_v - s_u$
ΔQ_{uv}	Discharge difference	$Q_v - Q_u$
$\Delta \rho_{uv}$	Sinuosity difference	$\rho_v - \rho_u$
m_{uv}	Same-main-river indicator	1 if $r_u = r_v$, else 0
$\mathbf{h}_i$	Learned embedding of reach i	-
$\hat{p}_{uv}$	Predicted probability of link $u \to v$	-

2.1 River link prediction problem

Network reconstruction is posed as directed link prediction. Let $G = (V, E)$ be the directed river graph, where $V = \{v_1, \ldots, v_N\}$ is the set of reaches and each edge $(u, v) \in E$ denotes a valid downstream connection from u to v. Topology is obtained from the hydrographic next-downstream pointer:

$$(u, v) \in E \quad \Longleftrightarrow \quad \mathrm{next}(u) = v. \tag{1}$$

Thus $\mathrm{next}(\cdot)$ defines the observed drainage network used for graph construction, but is not itself an attribute learned by the model.

The learning objective is a scoring function

$$f_\Theta : (u, v, G, \mathbf{x}_u, \mathbf{x}_v, \mathbf{q}_{uv}) \mapsto \hat{p}_{uv}, \tag{2}$$

where $\hat{p}_{uv} \in [0,1]$ is the predicted probability that the directed link $u \to v$ exists, $\mathbf{x}_u$ and $\mathbf{x}_v$ are node features of the two reaches, and $\mathbf{q}_{uv}$ contains candidate-edge descriptors of their physical compatibility.

2.2 Reach-level node and edge features

For each reach i, the node feature vector combines geometric, positional, hierarchical, and hydrologic attributes:

$$\mathbf{x}_i = [\ell_i,\ \rho_i,\ d_i,\ s_i,\ Q_i,\ c_{x,i},\ c_{y,i}]^\top . \tag{3}$$

Here ℓ_i and ρ_i describe local reach geometry; d_i locates the reach along the drainage continuum; s_i encodes stream hierarchy; Q_i is long-term average discharge; and $(c_{x,i}, c_{y,i})$ is the projected geometric centre of the reach. Highly skewed non-negative variables are transformed by $\log(1+x)$, missing values are median-imputed, and all features are standardized:

$$\tilde{x}_{if} = \frac{x_{if} - \mu_f}{\sigma_f}. \tag{4}$$

For Strahler-order classification, s_i is the prediction target and is therefore removed from $\mathbf{x}_i$.

For a candidate directed pair $u \to v$, an edge feature vector is formed to capture whether the two reaches are hydrologically compatible as an upstream–downstream connection:

$$\mathbf{q}_{uv} = [g_{uv},\ \Delta d_{uv},\ \Delta s_{uv},\ \Delta Q_{uv},\ \Delta \rho_{uv},\ m_{uv}]^\top , \tag{5}$$

with

$$\Delta d_{uv} = d_u - d_v, \Delta s_{uv} = s_v - s_u, \Delta Q_{uv} = Q_v - Q_u, \Delta \rho_{uv} = \rho_v - \rho_u, m_{uv} = \mathbb{1}\{r_u = r_v\}. \tag{6}$$

The endpoint gap g_{uv} is the Euclidean distance between the downstream endpoint of u and the upstream endpoint of v. Candidate-edge features are standardized using training-candidate statistics only. Architectures differ in whether they consume $\mathbf{q}_{uv}$: GraphSAGE and the link GAT baseline primarily use node embeddings, whereas Hetero-GAT and weighted MixHop include candidate-edge information in the decoder.

2.3 Positive and negative link construction

Observed edges from HydroRIVERS topology define the positive set. To separate message passing from supervision, E is partitioned into disjoint subsets

$$E = E_{\text{msg}} \cup E^+_{\text{train}} \cup E^+_{\text{val}} \cup E^+_{\text{test}}. \tag{7}$$

Message edges E_{msg} remain visible to the GNN during propagation. The remaining positive edges provide train, validation, and test labels. In the implemented protocol, these four subsets correspond to 70%, 10%, 10%, and 10% of edges, respectively. Missing-edge experiments further control how many true edges are removed from E_{msg} (10% to 90%), after which removed positives are allocated to supervised splits.

For each positive split, hard negative directed pairs are sampled from non-edges. Negatives include reversed true edges, spatially nearby but disconnected endpoint pairs, and same-main-river pairs that are not directly linked. This yields a supervised candidate set

$$\mathcal{D} = \{(u, v, y_{uv}, \mathbf{q}_{uv})\}, \qquad y_{uv} = \begin{cases} 1, & (u, v) \in E^+, \\ 0, & (u, v) \in E^-. \end{cases} \tag{8}$$

Training and validation use one negative per positive; testing uses ten negatives per positive, which better reflects the rarity of true links among candidate pairs.

2.4 GraphSAGE link predictor

GraphSAGE learns reach embeddings by aggregating information from neighbouring reaches on the message graph. Let $\mathbf{h}_i^{(\ell)}$ be the hidden state of reach i at layer ℓ, and let $\mathcal{N}(i)$ be its neighbours in E_{msg}. A neighbourhood message is

$$\mathbf{m}_i^{(\ell)} = \text{AGG}^{(\ell)}\left(\{\mathbf{h}_j^{(\ell)} : j \in \mathcal{N}(i)\}\right). \tag{9}$$

For mean aggregation,

$$\mathbf{m}_i^{(\ell)} = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{h}_j^{(\ell)}. \tag{10}$$

The node update concatenates the reach's own state with the neighbourhood summary:

$$\mathbf{h}_i^{(\ell+1)} = \sigma\left(W_{\text{sage}}^{(\ell)} \left[\mathbf{h}_i^{(\ell)} \,\|\, \mathbf{m}_i^{(\ell)}\right] + \mathbf{b}_{\text{sage}}^{(\ell)}\right), \tag{11}$$

where σ is ReLU. Two GraphSAGE layers map standardized features $\tilde{\mathbf{X}}$ to 64-dimensional embeddings. A candidate link $u \to v$ is scored by a bilinear decoder,

$$s_{uv} = \mathbf{h}_u^\top W_{\text{bil}} \mathbf{h}_v, \qquad \hat{p}_{uv} = \sigma(s_{uv}). \tag{12}$$

GraphSAGE therefore uses local neighbourhood aggregation, but does not explicitly separate upstream and downstream roles.

2.5 GAT link predictor

GAT replaces uniform aggregation with learned attention over neighbours. For a message edge $j \to i$ and attention head k,

$$e_{ij}^{(k)} = \text{LeakyReLU}\left(\mathbf{a}_k^\top \left[W_k \mathbf{h}_i \,\|\, W_k \mathbf{h}_j \,\|\, W_{e,k} \mathbf{r}_{ij}\right]\right), \tag{13}$$

where $\mathbf{r}_{ij}$ is an optional edge-attribute vector (omitted when edge attributes are disabled). Attention weights are

$$\alpha_{ij}^{(k)} = \frac{\exp(e_{ij}^{(k)})}{\sum_{t \in \mathcal{N}(i)} \exp(e_{it}^{(k)})}. \tag{14}$$

The head update is

$$\mathbf{h}_{i,k}^{(\ell+1)} = \sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(k)} W_k^{(\ell)} \mathbf{h}_j^{(\ell)}. \tag{15}$$

Multi-head outputs are concatenated (or averaged in the final layer) to produce a 64-dimensional embedding. Candidate links are scored by embedding similarity,

$$s_{uv} = \mathbf{h}_u^\top \mathbf{h}_v, \qquad \hat{p}_{uv} = \sigma(s_{uv}). \tag{16}$$

This baseline tests whether one-hop attention alone is sufficient for directed river-link discrimination.

2.6 Hetero-GAT link predictor

Hetero-GAT keeps a single node type (river reach) but introduces two directed relations,

$$\mathcal{R} = \{\text{flows_to},\ \text{receives_from}\}. \tag{17}$$

For each message edge $u \to v$,

$$(u, v) \in E_{\text{flows_to}}, \qquad (v, u) \in E_{\text{receives_from}}. \tag{18}$$

Thus downstream and upstream contexts are handled by separate attention operators. After an input projection,

$$\mathbf{h}_i^{(0)} = \text{Dropout}(\text{LayerNorm}(\text{ReLU}(W_x \tilde{\mathbf{x}}_i + \mathbf{b}_x))), \tag{19}$$

each layer computes

$$\mathbf{h}_{i,r}^{(\ell+1)} = \text{GAT}_r^{(\ell)} \left(\{\mathbf{h}_j^{(\ell)} : j \in \mathcal{N}_r(i)\} \right), \qquad r \in \mathcal{R}, \tag{20}$$

and combines relation-specific states by summation, followed by nonlinearity, layer normalization, and dropout. Candidate links are decoded with an edge-centred representation that includes $\mathbf{q}_{uv}$:

$$\mathbf{z}_{uv} = [\mathbf{h}_u \,\|\, \mathbf{h}_v \,\|\, \mathbf{h}_u \odot \mathbf{h}_v \,\|\, |\mathbf{h}_u - \mathbf{h}_v| \,\|\, \tilde{\mathbf{q}}_{uv}], \tag{21}$$

$$s_{uv} = \text{MLP}_\Theta(\mathbf{z}_{uv}), \qquad \hat{p}_{uv} = \sigma(s_{uv}). \tag{22}$$

Hetero-GAT therefore couples direction-aware attention with candidate-edge compatibility features.

2.7 Weighted directed MixHop link predictor

The proposed encoder combines three river-specific design choices: explicit upstream/downstream propagation, multi-hop diffusion within each layer, and

hydrology-informed edge weighting. Input features are first projected to a hidden state $\mathbf{h}_i^{(0)}$ as above. For each message edge $u \to v$, a fixed weight

$$w_{uv} = \frac{1}{g_{uv} + \varepsilon} \tag{23}$$

emphasizes spatially close endpoint pairs ($\varepsilon > 0$ for stability). Downstream and upstream aggregators are

$$\mathcal{A}_{\downarrow}(\mathbf{H})_v = \frac{\sum_{u \in \mathcal{N}_{\text{in}}(v)} w_{uv} \mathbf{h}_u}{\sum_{u \in \mathcal{N}_{\text{in}}(v)} w_{uv}}, \tag{24}$$

$$\mathcal{A}_{\uparrow}(\mathbf{H})_v = \frac{\sum_{t \in \mathcal{N}_{\text{out}}(v)} w_{vt} \mathbf{h}_t}{\sum_{t \in \mathcal{N}_{\text{out}}(v)} w_{vt}}, \tag{25}$$

where $\mathcal{N}_{\text{in}}(v)$ are reaches flowing into v and $\mathcal{N}_{\text{out}}(v)$ are reaches receiving flow from v.

Within a MixHop layer, K hop states are computed in both directions with root retention $\beta \in [0, 1]$:

$$\mathbf{D}^{(0)} = \mathbf{U}^{(0)} = \mathbf{H}_{\text{in}}, \tag{26}$$

$$\mathbf{D}^{(k)} = \beta \mathbf{H}_{\text{in}} + (1 - \beta) \mathcal{A}_{\downarrow}(\mathbf{D}^{(k-1)}), \tag{27}$$

$$\mathbf{U}^{(k)} = \beta \mathbf{H}_{\text{in}} + (1 - \beta) \mathcal{A}_{\uparrow}(\mathbf{U}^{(k-1)}), \tag{28}$$

for $k = 1, \ldots, K$. The layer output concatenates the local state with all hop states and projects back to the hidden dimension. Two such layers are stacked with hidden size 64, $K = 3$, and $\beta = 0.20$.

2.7.1 Decoder and training

After encoding, each reach has an embedding $\mathbf{h}_i$. For candidate $u \to v$,

$$\mathbf{z}_{uv} = [\mathbf{h}_u \,\|\, \mathbf{h}_v \,\|\, \mathbf{h}_u \odot \mathbf{h}_v \,\|\, |\mathbf{h}_u - \mathbf{h}_v| \,\|\, \tilde{\mathbf{q}}_{uv}], \tag{29}$$

$$s_{uv} = \text{MLP}_{\Theta}(\mathbf{z}_{uv}), \qquad \hat{p}_{uv} = \sigma(s_{uv}). \tag{30}$$

Training minimises binary cross-entropy on the training candidates:

$$\mathcal{L} = -\frac{1}{|\mathcal{D}_{\text{train}}|} \sum_{(u,v) \in \mathcal{D}_{\text{train}}} [y_{uv} \log \hat{p}_{uv} + (1 - y_{uv}) \log(1 - \hat{p}_{uv})]. \tag{31}$$

2.8 Model validation

Parameters are optimized with Adam (learning rate 10^{-3}, weight decay 10^{-4}). Validation loss is monitored for early stopping, and the best validation checkpoint is restored before testing. Link-prediction performance is assessed across missing-edge ratios of 10%, 30%, 50%, 70%, and 90% using ROC-AUC, PR-AUC (average precision), F1,

and the spatial-bootstrap balanced accuracy and macro-F1 reported in the Results. The same architectural family is also applied to Strahler-order classification with s_i as the target label, a stratified node split, and class-balanced cross-entropy; in that setting, s_i and Δs_{uv} are excluded from the inputs to avoid label leakage.

3 Results

This section reports two complementary experiments on reach-level river networks for the Amur, Ganga, Mekong, and Mississippi basins: (i) Directed link prediction for river-network reconstruction and (ii) Strahler-Order classification of river reaches. Four architectures are compared throughout: GraphSAGE, Graph Attention Network (GAT), Hetero-GAT, and the proposed weighted directed MixHop model. All experiments are within-basin; models are trained and evaluated separately in each basin under a common protocol.

3.1 River network dataset

The river basins selected for the study are Amur, Ganga, Mekong, and Mississippi. These basins differ in size substantially as shown in Table 3. The Mississippi basin contains the largest network (173,411 nodes and 172,958 directed edges), whereas Ganga is the smallest (65,262 nodes and 65,248 edges). All four basins exhibit near-tree-like connectivity, with average node degree close to 2.0 and mean reach length between 3.4 and 4.6 km. These structural differences provide a natural test of whether graph-learning methods generalise across basin scales and geomorphic settings.

Table 3 River network statistics for the four study basins.

Basin	Nodes	Edges	Avg. degree	Avg. edge length (km)	Basin area (km^2)
Amur	124,646	124,645	2	4.43	2,453,574
Ganga	65,262	65,248	2	3.87	944,595
Mekong	66,855	66,836	2	3.37	774,281
Mississippi	173,411	172,958	2	4.55	3,179,496

3.2 River link prediction

River-network reconstruction was formulated as a directed link prediction problem. Observed HydroRIVERS edges were partitioned into message edges (used only for GNN propagation) and supervised candidate links. Missing-edge ratios of 10%, 30%, 50%, 70%, and 90% controlled the number of true edges removed from the message graph; removed positives were allocated to the train/validation/test splits, with hard negative sampling targeting plausible but non-existent connections. Test evaluation used an imbalanced candidate pool (approximately one positive per ten hard negatives), which is closer to operational reconstruction than a balanced test set.

Performance was measured by ROC-AUC, precision–recall AUC (PR-AUC/average precision), and F1-score.

Across all basins and sparsity levels, the architectural ranking is stable (Figs. 2; Table 4). Weighted MixHop is the most robust model; Hetero-GAT and GraphSAGE form a competitive second tier; and standard GAT is a clear underperformer. At 10% missing edges, the three strong models achieve ROC-AUC values of approximately 0.96–0.98, indicating excellent global discrimination between true and false directed links. Even at 90% missing edges, ROC-AUC remains high for MixHop ($\approx$ 0.91–0.92) and for Hetero-GAT/GraphSAGE ($\approx$ 0.89–0.91), showing that river topology retains recoverable structure under severe message-graph sparsity.

PR-AUC and F1 reveal a harder problem than ROC-AUC alone. At 10% missing edges, MixHop PR-AUC reaches $\approx$ 0.71–0.77 and F1 $\approx$ 0.61–0.69 across basins; both metrics decline steadily as sparsity increases (e.g., Amur MixHop PR-AUC falls from 0.773 at 10% to 0.447 at 90%). Hetero-GAT is close to MixHop on F1 at low sparsity (and occasionally highest, as in Mississippi at 10% missing edges, F1 = 0.696), but MixHop retains a clearer advantage under extreme edge removal. GraphSAGE follows a similar degradation trajectory at slightly lower absolute levels. GAT remains near chance on precision-focused metrics (PR-AUC $\approx$ 0.09–0.11; F1 $\approx$ 0.20) despite moderate ROC-AUC ($\approx$ 0.70-0.76), indicating that one-hop attention without direction-aware multi-hop context and candidate-edge decoding is insufficient for hard river-link discrimination.

Cross-basin absolute scores differ modestly, but the ranking does not. Amur and Mississippi tend to yield slightly higher MixHop PR-AUC at low sparsity, whereas Ganga and Mekong tend to yield lower values. Given the near-tree-average degree shared by all four graphs, these differences are more likely to reflect basin size, class imbalance in candidate negatives, and local confluence geometry than a qualitative change in the learning problem.

Table 4 Link-prediction performance at 10% and 90% missing-edge ratios across basins and architectures. Results are spatial bootstrap mean ± SD (%).

Basin	Architecture	Bal. Acc. (10%)	Bal. Acc. (90%)	Macro-F1 (10%)	Macro-F1 (90%)
Amur	GAT	60.3 ± 0.6	59.1 ± 0.3	34.9 ± 0.7	36.2 ± 0.5
	GraphSAGE	92.3 ± 0.3	86.1 ± 0.3	78.7 ± 0.3	65.0 ± 0.2
	Hetero-GAT	94.3 ± 0.2	84.8 ± 0.4	81.4 ± 0.3	62.2 ± 0.2
	Weighted MixHop	94.8 ± 0.2	89.1 ± 0.3	82.0 ± 0.3	67.9 ± 0.2
Ganga	GAT	60.5 ± 0.9	58.8 ± 0.6	35.6 ± 1.2	36.0 ± 0.9
	GraphSAGE	92.0 ± 0.4	85.9 ± 0.5	77.6 ± 0.4	64.5 ± 0.4
	Hetero-GAT	93.4 ± 0.3	84.0 ± 0.7	79.9 ± 0.4	61.0 ± 0.5
	Weighted MixHop	94.2 ± 0.3	89.3 ± 0.5	80.4 ± 0.4	68.5 ± 0.3
Mekong	GAT	60.4 ± 0.7	58.8 ± 0.5	36.1 ± 1.0	36.1 ± 0.8
	GraphSAGE	91.9 ± 0.4	85.8 ± 0.4	77.2 ± 0.5	64.8 ± 0.3
	Hetero-GAT	92.6 ± 0.3	83.6 ± 0.5	77.5 ± 0.5	60.5 ± 0.4
	Weighted MixHop	92.8 ± 0.3	88.4 ± 0.4	77.1 ± 0.5	66.8 ± 0.3
Mississippi	GAT	61.3 ± 0.5	58.8 ± 0.3	36.0 ± 0.7	36.2 ± 0.5
	GraphSAGE	92.5 ± 0.2	85.8 ± 0.3	77.6 ± 0.3	64.3 ± 0.2
	Hetero-GAT	94.8 ± 0.2	83.0 ± 0.4	82.5 ± 0.3	59.4 ± 0.2
	Weighted MixHop	94.7 ± 0.2	88.7 ± 0.3	81.3 ± 0.3	67.3 ± 0.2

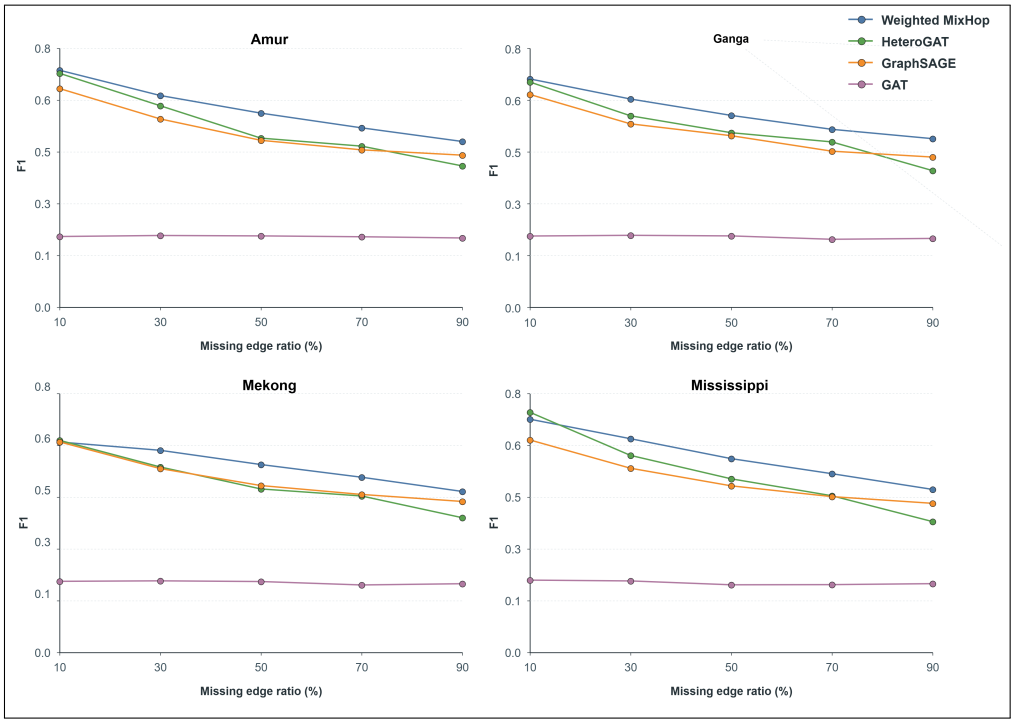


Fig. 2 F1 as a function of Missing Edge ratio across basins

3.3 Validation for network reconstruction

Link-prediction models were selected by early stopping on validation binary cross-entropy (Adam, learning rate 10^{-3}, weight decay 10^{-4}, patience 10 epochs; maximum 100 epochs), after which the best validation checkpoint was restored for testing. The held-out test set is deliberately hard: approximately ten hard negatives per positive, with negatives sampled near spatially plausible but non-existent connections. Under this protocol, ordinary accuracy is misleading, so validation and reporting emphasise balanced accuracy, macro-averaged F1, and precision–recall behaviour (Table 4; Fig. 3).

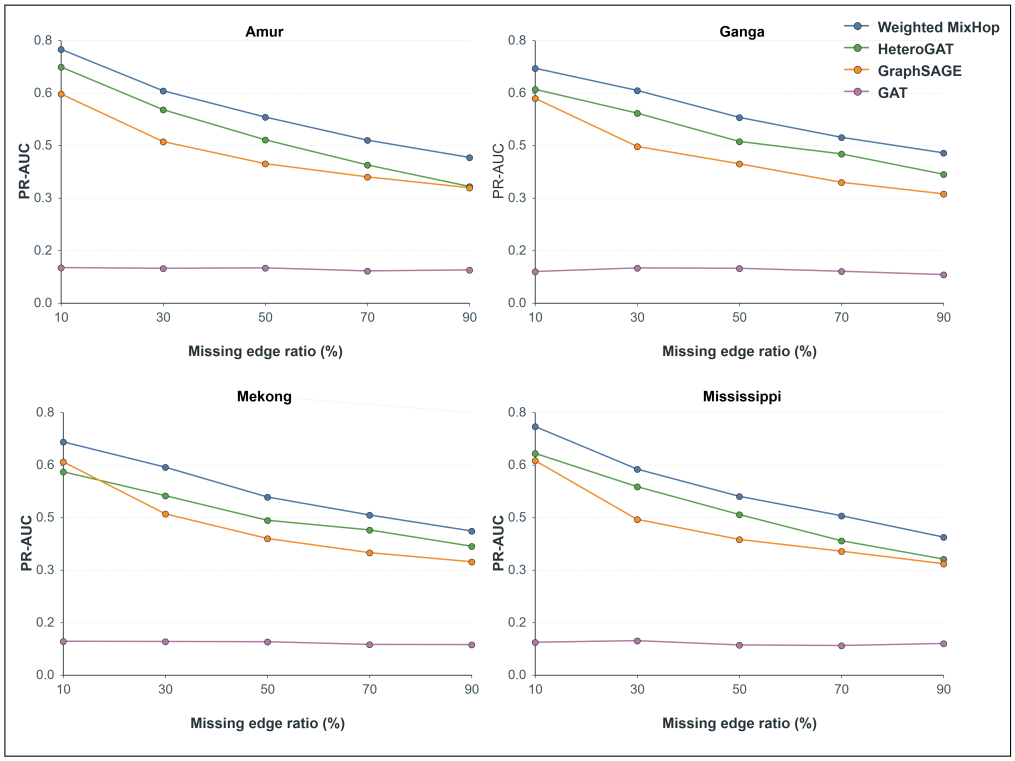


Fig. 3 PR-AUC as a function of missing-edge ratio across architectures and basins

The test results confirm a stable architecture ranking across all four basins. At 10% missing edges, Weighted MixHop and HeteroGAT reach balanced accuracies of about $92.6 - 94.8\%$ and macro-F1 of about $77.1 - 82.5\%$, with GraphSAGE close behind ($91.9 - 92.5\%$ balanced accuracy; $77.2 - 78.7\%$ macro-F1). GAT remains far weaker throughout ($\approx 58 - 61\%$ balanced accuracy; $\approx 34 - 36\%$ macro-F1), indicating that one-hop attention alone does not separate true river links from hard negatives. Bootstrap standard deviations are small relative to these architecture gaps (typically $0.2 - 0.5$ percentage points for the strong models), so the ranking is robust to spatial resampling of the test set.

Increasing the missing-edge ratio stresses the message graph and mainly degrades precision-oriented performance rather than collapsing global separability. From 10% to 90% missing edges, MixHop balanced accuracy falls only modestly (to $88.4 - 89.3\%$), whereas macro-F1 declines more clearly (to $66.8 - 68.5\%$); HeteroGAT and GraphSAGE show the same pattern at slightly lower absolute levels (Fig. 2). This is the expected signature of directed reconstruction under sparsity, true links remain distinguishable in aggregate, but recovering them among false candidates becomes harder as multi-hop neighbourhood context is removed. Gains of MixHop over single-hop baselines are therefore most visible under high missing-edge ratios, where multi-hop diffusion and direction-aware aggregation remain informative.

3.4 Feature importance for link prediction

Post-hoc permutation importance was computed on the trained link-prediction checkpoints without retraining. Each input feature was shuffled on the test candidates while model weights were held fixed (three repeats; up to 50,000 candidates per model). The drop in average precision (AP) was the primary importance score, with AUC drop as a secondary check. Importance values were summarised by architecture and basin and averaged across missing-edge ratios.

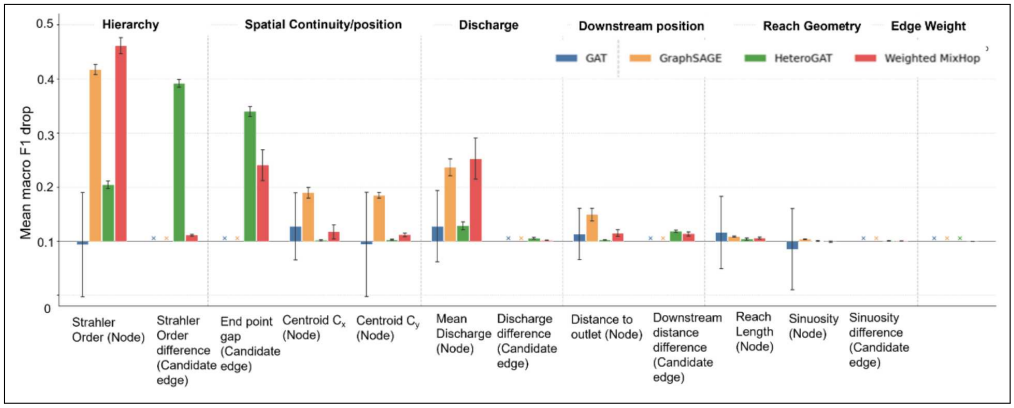


Fig. 4 Permutation feature importance for link prediction (mean macro-F1 drop).

The dominant signal depends on architecture, but the hydrologic interpretation is coherent (Fig. 4). For GraphSAGE and Weighted MixHop, the node-level Strahler order (s) is the most important feature in every basin, with mean AP drops of approximately 0.31–0.33 (GraphSAGE) and 0.35–0.38 (MixHop). Hierarchical stream order is therefore the strongest cue these models use when deciding whether two reaches should connect. Secondary features for MixHop are typically the endpoint gap (g; mean AP drop $\approx$ 0.10–0.17) and mean discharge (Q; $\approx$ 0.10–0.19), reflecting spatial continuity between endpoints and discharge compatibility along the flow path. Local geometry features such as reach length (ℓ) and sinuosity (ρ) have negligible importance.

HeteroGAT relies more on candidate-edge features than on node attributes. Across basins, the Strahler-order difference (Δs) ranks first (AP drop $\approx$ 0.28–0.30) and the endpoint gap (g) second ($\approx$ 0.23–0.25). This is hydrologically meaningful: HeteroGAT appears to score links by whether order change and spatial gap between candidate upstream–downstream reaches are consistent with valid confluence geometry. Node-level Strahler order (s) remains relevant but secondary.

Plain GAT shows weaker and less stable importance patterns. In some basins (e.g., Ganga, Mississippi), projected centroids ((c_x, c_y)) or mean discharge (Q) appear modestly useful; in others (Amur, Mekong), many AP drops are near zero or negative. Combined with GAT's persistently low balanced accuracy, macro-F1, and PR-AUC, this indicates that the baseline attention model does not form a consistent feature–decision mapping for hard river-link discrimination.

Overall, reconstruction is driven by three information families: (i) stream hierarchy, Strahler order (s) and Strahler-order difference (Δs); (ii) spatial continuity, endpoint gap (g) and projected centroids ((c_x, c_y)); and (iii) discharge — mean discharge (Q). The inverse-endpoint-gap message-edge weight ($w = 1/(g + \varepsilon)$) used inside Weighted MixHop showed essentially no additional permutation importance once other features were present, implying that most predictive power resides in learned embeddings and candidate-edge descriptors rather than in sensitivity to that scalar weight alone.

An important interpretive caution is necessary. Because Strahler order (s) and Strahler-order difference (Δs) are strong hierarchy cues, their high importance is expected and may overstate topology recovery from geometry alone. If the scientific goal is reconstruction without using order attributes, a controlled ablation removing these features should be reported separately. Under the present feature set, explainability indicates that models succeed largely by combining hierarchical compatibility with endpoint proximity, which aligns with how HydroRIVERS connectivity is physically organized.

3.5 Strahler order classification

As a complementary task, the same architectures were applied to reach-level Strahler-order classification. Here each node is assigned a discrete order class; Strahler order is the prediction target and is excluded from the input feature vector. Node attributes comprise length, sinuosity, downstream distance, discharge, and projected centroids. Class ranges differ by basin (e.g., Mekong orders $1 - 8$; Amur and Mississippi $1 - 9$), and class frequencies are strongly imbalanced toward low orders.

All four architectures achieve high overall accuracy, but balanced metrics reveal clear differences (Table 5; Fig. 5). Weighted directed MixHop consistently performs better, with accuracy of approximately $0.945 - 0.957$ across basins and macro-F1 of approximately $0.91 - 0.94$. Hetero-GAT and GraphSAGE form a strong second tier (accuracy typically $\approx 0.91 - 0.93$), while GAT remains usable but weaker on balanced accuracy and macro F1, especially in Mississippi, the largest and most class-heterogeneous basin (GAT accuracy 0.877; balanced accuracy 0.817; macro F1 0.806). These results show that direction-aware multi-hop aggregation improves hierarchical labelling as well as link reconstruction, even though the two tasks emphasise different hydrologic signals (Section 3.7).

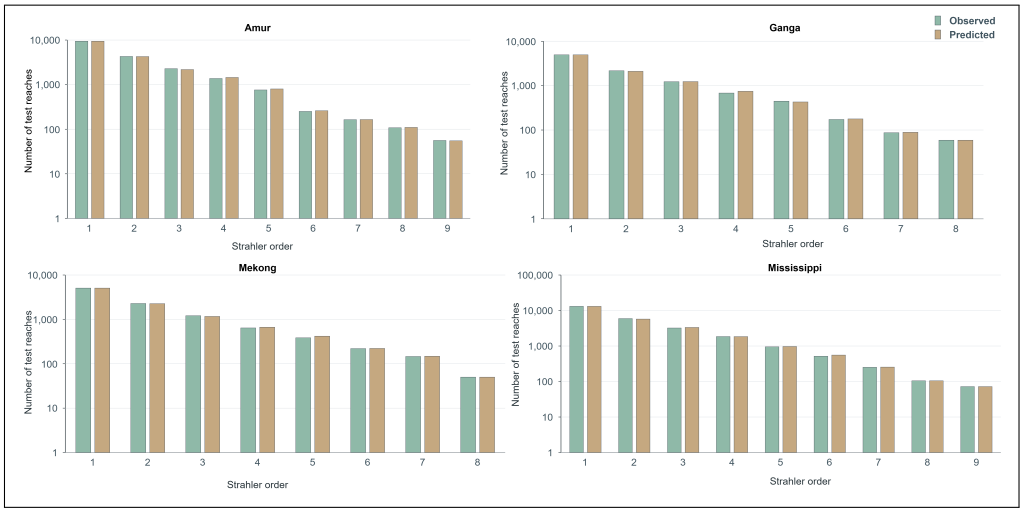


Fig. 5 Oberseved and predicted Strahler Order using the Weighted MixHop model across basins with respect to number of test reaches

Table 5 Strahler-order classification performance across basins and architectures. Results are spatial bootstrap mean ± SD (%).

	Amur		Ganga		Mekong		Mississippi	
Model	Bal. Acc.	Macro-F1	Bal. Acc.	Macro-F1	Bal. Acc.	Macro-F1	Bal. Acc.	Macro-F1
GAT	85.1 ± 1.0	85.0 ± 0.9	85.3 ± 0.8	83.8 ± 0.9	85.1 ± 1.1	84.3 ± 1.2	81.3 ± 1.2	80.2 ± 1.2
GraphSAGE	91.4 ± 0.4	90.8 ± 0.5	89.7 ± 0.5	89.2 ± 0.6	89.9 ± 0.7	89.3 ± 0.7	86.3 ± 1.0	85.0 ± 1.2
HeteroGAT	91.2 ± 0.4	90.4 ± 0.5	89.9 ± 0.6	89.7 ± 0.7	89.8 ± 0.7	88.9 ± 0.8	87.1 ± 0.7	86.6 ± 0.7
Weighted MixHop	94.0 ± 0.4	93.4 ± 0.4	94.0 ± 0.4	93.6 ± 0.5	93.9 ± 0.5	93.2 ± 0.5	91.3 ± 0.4	90.6 ± 0.4

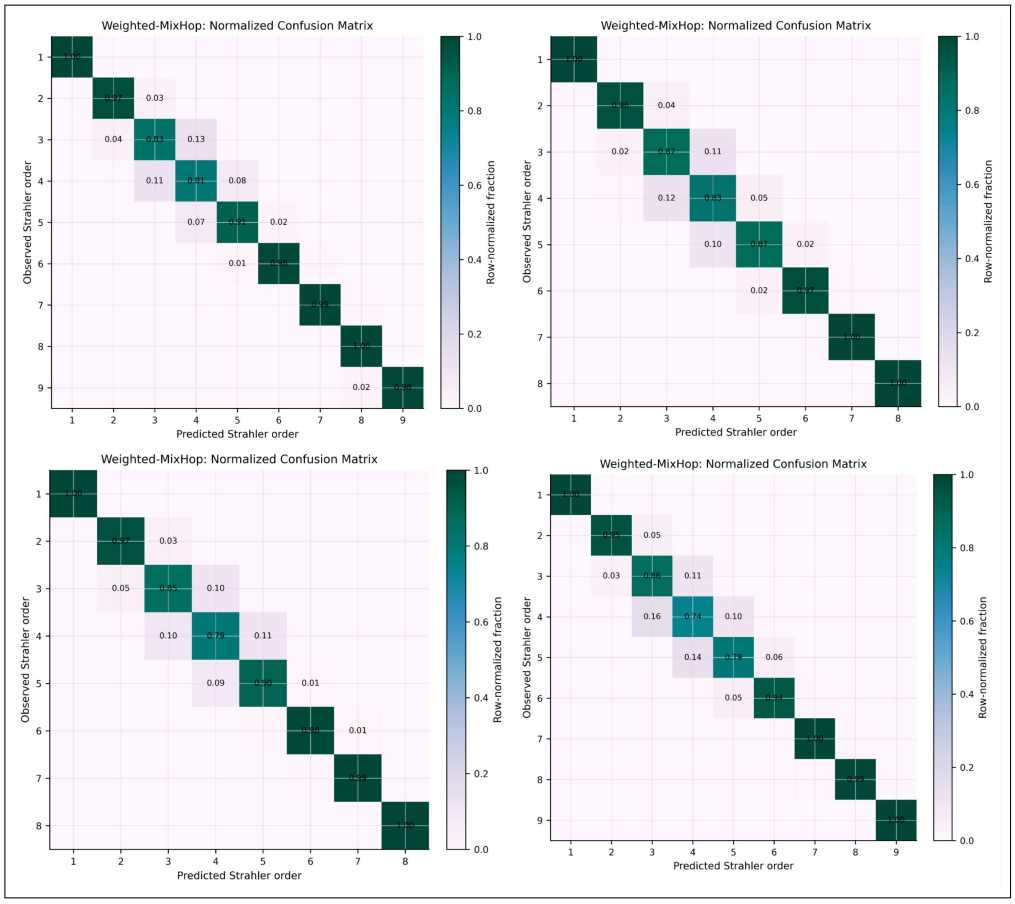


Fig. 6 Strahler Order classification performance across basins

3.6 Validation of Strahler order classification

Strahler experiments used a stratified node split (70%/15%/15% train/validation/test) so that rare high-order classes remain represented in every partition. The full directed graph remains available for message passing (transductive setting), but supervision uses only training-node labels. Training minimised class-balanced cross-entropy; early stopping monitored validation weighted loss (patience 50 epochs; minimum improvement 0.001; maximum 500 epochs), and the best checkpoint was restored for testing.

Validation supports that the Weighted Mixhop model gains are not artefacts of majority-class prediction. Balanced accuracy and macro F1 track the accuracy ranking, and Mississippi, where high-order classes are scarcest relative to network size, remains the hardest basin for every architecture. Confusion matrices and observed-versus-predicted histograms (Figs. 5 -6) corroborate this pattern: most mass lies on the

diagonal, and predicted class frequencies reproduce the heavy-tailed empirical order distribution, including rare high-order reaches.

3.7 Feature importance for Strahler order classification

Explainability analysis was conducted with the help of feature importance analysis for Strahler models, using the drop in macro-F1 after shuffling each input feature (three repeats) on held-out test nodes. It has been found that discharge dominates every architecture and every basin (Fig. 7): mean discharge (Q) ranks first throughout, with mean macro-F1 drops of approximately 0.48–0.52 for GAT, 0.63–0.69 for GraphSAGE, 0.60–0.66 for Hetero-GAT, and 0.61–0.65 for Weighted MixHop. For GAT, the message-edge discharge difference (ΔQ) is second (macro-F1 drop $\approx$ 0.29–0.39), indicating that GAT also uses how discharge changes between connected reaches. For the stronger models, secondary importance usually shifts to longitudinal and geographic position: distance to outlet (d), and projected centroids ((c_x, c_y)), while reach length (ℓ) and sinuosity (ρ) remain weak.

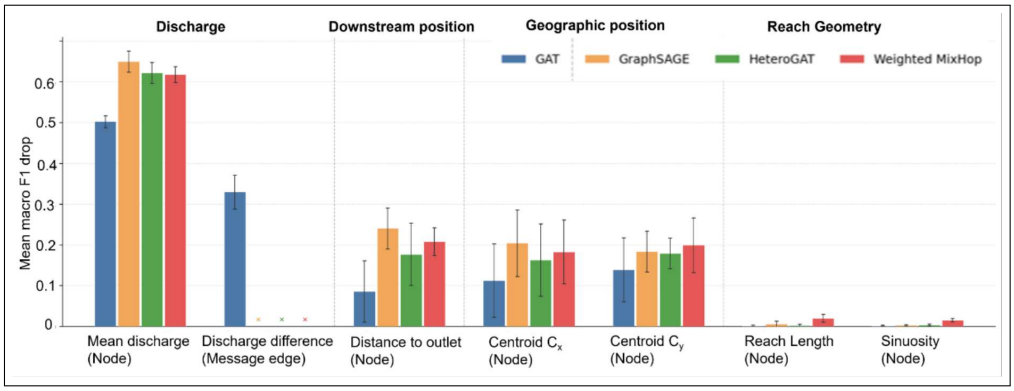


Fig. 7 Permutation feature importance for Strahler-order classification (mean macro-F1 drop).

This explainability profile contrasts cleanly with link prediction. Reconstruction depends primarily on hierarchical compatibility, Strahler order (s) and Strahler-order difference (Δs), and endpoint continuity via the endpoint gap (g) between candidate reaches; Strahler labelling depends primarily on accumulated flow (Q) and position along the drainage continuum (d, (c_x, c_y)). That contrast is hydrologically expected: stream order is tightly associated with discharge magnitude and downstream location, whereas deciding whether two reaches should connect requires matching hierarchy and spatial abutment. As in the link-prediction analysis, inverse-endpoint-gap message-edge weight in the proposed Weighted Mixhop model($w = 1/(g+\varepsilon)$) showed negligible measured importance relative to node-level hydrologic and spatial features.

4 Discussion

4.1 Does robustness to missing data depend on how a model propagates information?

Weighted MixHop's ROC-AUC degraded least as the message graph sparsified, falling by only 0.055 from 10% to 90% missing edges, averaged across basins, compared to $0.067 - 0.068$ for GraphSAGE and Hetero-GAT. This pattern held in all four basins. The plausible explanation lies in what distinguishes propagation mechanism of Mixhop: its three-hop ($K = 3$), root-retained (β) diffusion allows a node to draw on hydrological context several connections away. Hence, the removal of any single local edge is less likely to sever it from relevant upstream or downstream information entirely. This interpretation is consistent with Kirschstein and Sun [4], who identified that the tree-like topology of river networks specifically induces over-squashing in GNNs, the progressive compression of distant-node information during message passing and with Mosaffa et al. [7], who found that routing information explicitly improves rainfall-runoff predictions. Notably, this robustness advantage was specific to ranking quality, PR-AUC of the proposed weighted mixhop model degraded at a rate comparable to the other architectures ($0.25 - 0.34$ across basins), indicating that multi-hop propagation improves the ability of model to order candidate links under sparsity correctly, but does not equally protect its ability to separate true links from hard negatives.

4.2 Why does standard attention fail on river-reach graphs specifically?

GAT was the weakest architecture throughout, with PR-AUC confined to $0.09 - 0.11$ regardless of basin or sparsity level, far below the $0.34 - 0.77$ range achieved by the other three models. This can be explained with permutation importance results for GAT that internally inconsistent across basins as compared to other architectures. Strahler order (s), for instance, carried the second-highest importance in Ganga (0.118) but a negative importance in Amur and Mekong (-0.075 and -0.090 respectively), and in every basin the standard deviation of importance estimates of GAT exceeded their mean. This indicates the model did not converge on a stable feature-to-decision mapping at all. A structural explanation follows directly from the basin statistics: average node degree across all four basins is 2.0(3), which means majority of reaches have exactly one upstream and one downstream neighbour. Under this near-tree topology, an attention mechanism whose primary value lies in differentially weighting multiple competing neighbours does not have anything to select between. This reading is consistent with Zhang et al. [13], who argue that treating flow direction as an explicit physical prior, rather than leaving a generic attention mechanism to discover it, is necessary for graph networks operating on directed hydrological networks; and with Sun et al. [14], whose basin-scale river network learning explicitly incorporated physics-based connectivity rather than relying on attention alone.

4.3 Does basin geometry shape which hydrological signal the model relies on?

The secondary feature reliance of the proposed weighted Mixhop model beyond its dominant use of Strahler order (s tracked average reach length (ℓ) in a way that was consistent across all four basins. In basins with longer average reaches (Mississippi, 4.55 km; Amur, 4.43km), endpoint gap (g) was the stronger secondary predictor. In

basins with shorter average reaches (Ganga, 3.87 km; Mekong, 3.37 km), mean discharge (Q) became the stronger secondary predictor, and the gap between the two widened as reach length decreased. Because hard negatives were sampled within a fixed 5 km radius of candidate endpoints, basins with shorter reaches place more candidate reaches within that radius by construction, plausibly making spatial gap alone a weaker discriminator and pushing the model toward discharge continuity instead. With only four basins, this pattern should be treated as an observed, hypothesis-generating trend rather than an established relationship; testing it directly would require either a larger basin sample or an experiment that varies the hard-negative radius independently of basin geometry.

4.4 Is Strahler order classification governed by the same physical drivers as link prediction?

Mean discharge (Q) was the single most important feature for Strahler-order (s) classification in every architecture and every basin without exception (macro-F1 drop $0.48 - 0.69$), an order of magnitude larger than reach length (ℓ) or sinuosity (ρ). This is a materially different pattern from link prediction, where hierarchical compatibility Strahler order (s) or its candidate-edge difference (Δs) dominated instead. The distinction is hydrologically coherent, Strahler order is, by definition, a measure of accumulated tributary flow, so a model recovering it from discharge is recovering a physically grounded signal. This also clarifies an interpretive concern raised for the link-prediction task that reconstruction performance partly reflects access to an already-hierarchical input feature (Strahler order (s)) since no equivalent concern applies to the classification task, where the dominant feature (mean discharge, Q) is independent of the label being predicted. This complements Xu et al. [15], who used Strahler-coded structure as an input for drainage-pattern classification via graph convolution; the present results suggest that when Strahler order is instead the prediction target, discharge, not structural or geometric attributes, is what a GNN actually learns to associate with it.

5 Conclusion

This study evaluated a weighted directed MixHop architecture against three baseline graph neural networks: GraphSAGE, GAT, and Hetero-GAT, on two complementary tasks across four large river basins: reconstructing missing directed river connectivity and classifying Strahler stream order. Weighted MixHop outperformed all baselines on every metric, and river basin, and at every level of network sparsity tested. Feature-importance analysis showed that link reconstruction and stream order classification draw on different, physically coherent signals. Link prediction relies on hierarchical compatibility and spatial continuity, and Strahler order classification relies on accumulated discharge. Since all experiments were conducted within individual basins, rather than generalisation of a single trained model across basins; extending this work to cross-basin transfer would serve as a natural direction for future research. These findings indicate that modelling flow direction using a topologically aware, weighted multi-hop GNN architecture offers a reliable and interpretable basis for reconstructing and characterising large-scale river networks from incomplete topological data.

Declaration

Ethics approval and consent to participate Not applicable. This study used publicly available secondary datasets (road, river and satellite-derived vegetation data) and involved no human participants, human data or animals.

Consent for publication Not applicable. The manuscript contains no person's data, images or details.

Availability of data and materials The datasets analysed in this study are publicly available: river network and Strahler stream order from HydroATLAS (Linke et al., 2019).

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Not Applicable

Authors' contributions [NNV: conceptualisation, methodology, data curation, formal analysis, visualisation, writing – original draft.] [RS: Data analysis, visualisation] [SS: Conceptualisation, Writing - review and editing, supervision] All authors read and approved the final manuscript.

References

[1] Battin, T.J., Lauerwald, R., Bernhardt, E.S., Bertuzzo, E., Gener, L.G., Hall, R.O., Hotchkiss, E.R., Maavara, T., Pavelsky, T.M., Ran, L., Raymond, P., Rosentreter, J.A., Regnier, P.: River ecosystem metabolism and carbon biogeochemistry in a changing world. Nature **613**, 449–459 (2023) https://doi.org/10.1038/s41586-022-05500-8

[2] Rinaldo, A.: Ecohydrology amid a rapidly changing world. Nature Water **2**(3), 209–210 (2024) https://doi.org/10.1038/s44221-024-00205-w

[3] Wang, H., Chen, J., Zheng, Y., Song, X.: Accelerating flood warnings by 10 hours: the power of river network topology in ai-enhanced flood forecasting. Npj Natural Hazards. **2**(1) (2025) https://doi.org/10.1038/s44304-025-00083-6

[4] Kirschstein, N., Sun, Y.: The merit of river network topology for neural flood forecasting. arXiv.org (2024) https://doi.org/10.5555/3692070.3693060

[5] Wang, P., Li, H., Wang, L., Huang, Z.: The impact of teleconnections of built-up land on regional carbon burden under the shared socio-economic pathways. The Science of the Total Environment **906**, 167589 (2023) https://doi.org/10.1016/j.scitotenv.2023.167589

[6] 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., Parsons, D., Darby, S.E.: Global river topology (grit): a bifurcating river hydrography. Water Resources Research **61**(5) (2025) https://doi.org/10.1029/2024wr038308

[7] Mosaffa, H., Pappenberger, F., Prudhomme, C., Chantry, M., Rüdiger, C., Cloke, H.: A gnn routing module is all you need for lstm rainfall–runoff models. Hydrology and Earth System Sciences **30**(7), 2079–2092 (2026) https://doi.org/10.5194/hess-30-2079-2026

[8] Buchhorn, K., Santos-Fernandez, E., Mengersen, K., Salomone, R.: Graph neural network-based anomaly detection for river network systems. F1000Research **12**, 991 (2023) https://doi.org/10.12688/f1000research.136097.1

[9] Plettenberg, P., Köhler, D., Sick, B., Thomas, J.M.: Flow-Attentional Graph neural networks (2025). https://arxiv.org/abs/2506.06127

[10] Fan, X., Hackl, J.: Generic multimodal spatially graph network for spatially embedded network representation learning. Applied Network Science **10**(1) (2025) https://doi.org/10.1007/s41109-025-00751-6

[11] Abu-El-Haija, S., Perozzi, B., Kapoor, A., Alipourfard, N., Lerman, K., Harutyunyan, H., Steeg, G.V., Galstyan, A.: MixHop: Higher-Order Graph Convolutional Architectures via Sparsified Neighborhood Mixing (2019). https://arxiv.org/abs/1905.00067

[12] Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., Thieme, M.: Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Scientific Data **6**(1), 283 (2019) https://doi.org/10.1038/s41597-019-0300-6

[13] Zhang, X., He, Y., Zhou, Z., Li, W., Tao, L., Chen, J., Zou, H., Lin, H.: Incorporating flow direction as a physical prior in graph networks: A directed temporal graph convolutional network for enhanced water quality forecasting in lake basins. Transactions in GIS (2026) https://doi.org/10.1111/tgis.70309

[14] Sun, A.Y., Jiang, P., Yang, Z.-L., Xie, Y., Chen, X.: A graph neural network (gnn) approach to basin-scale river network learning: the role of physics-based connectivity and data fusion. Hydrology and Earth System Sciences **26**(19), 5163–5184 (2022) https://doi.org/10.5194/hess-26-5163-2022

[15] Xu, X., Liu, P., Guo, M.: Drainage pattern recognition of river network based on graph convolutional neural network. ISPRS International Journal of Geo-Information **12**(7), 253 (2023) https://doi.org/10.3390/ijgi12070253