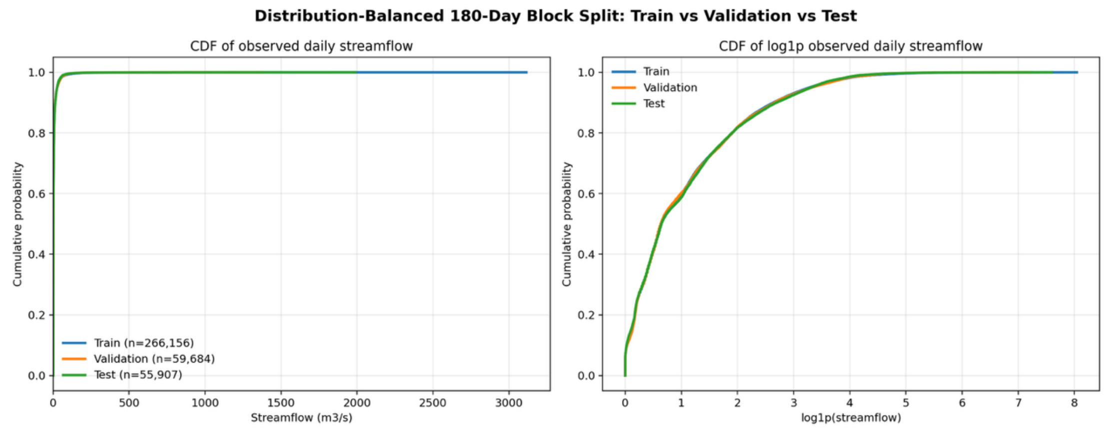

Figure S1. Cumulative distribution functions (CDFs) of observed daily streamflow and log-transformed streamflow for the training, validation, and testing datasets generated using the distribution-balanced temporal block splitting strategy. The dataset partitions were constructed using non-overlapping temporal blocks and 180-day prediction windows to preserve multi-day hydrologic structure while maintaining similar streamflow distributions across all data splits.

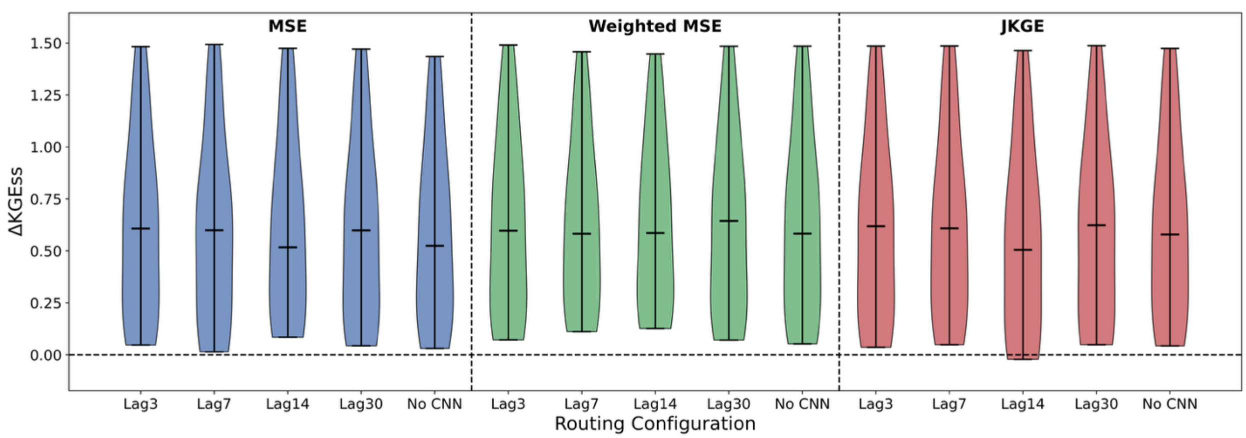

Figure S2. Violin plots of gauge-level $\mathrm{KGE}_{\{ss\}}$ improvements relative to the RAPID routing benchmark for all RiverGraphNet routing configurations. The distributions are shown for different temporal-memory configurations (3-, 7-, 14-, and 30-day runoff lags), loss formulations (MSE, weighted MSE, and JKGE), and experiments without the temporal convolutional head. Violin widths represent the density of gauge-level performance improvements. Results are consistent with the boxplots presented in Figure 6 and indicate broadly positive improvements across the gauge network, with no substantial multimodal structure or isolated clusters of performance gain.

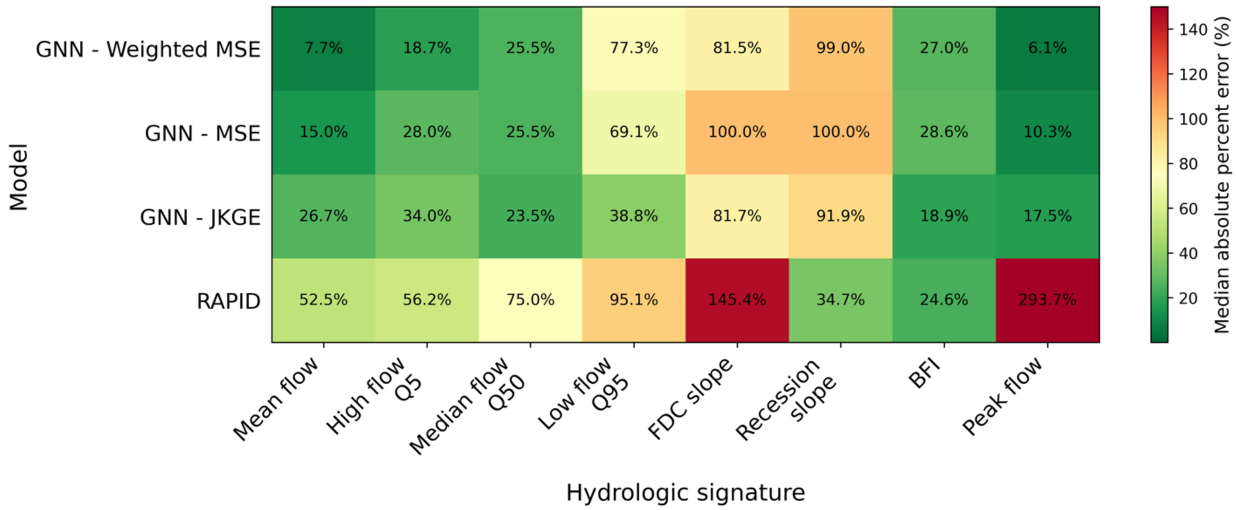

Figure S3. Hydrologic-signature errors for RiverGraphNet and RAPID. Median absolute percent errors across evaluation gauges for selected hydrologic signatures, including mean flow, high-flow magnitude (Q5), median flow (Q50), low-flow magnitude (Q95), flow-duration-curve (FDC) slope, recession slope, baseflow index (BFI), and peak flow. The results reveal clear tradeoffs among loss formulations: weighted-MSE provides the most accurate representation of overall flow magnitude and peak-flow behavior, whereas JKGE better preserves several low-flow and recession-related signatures. RAPID exhibits substantially larger errors for most signatures, particularly peak-flow magnitude.

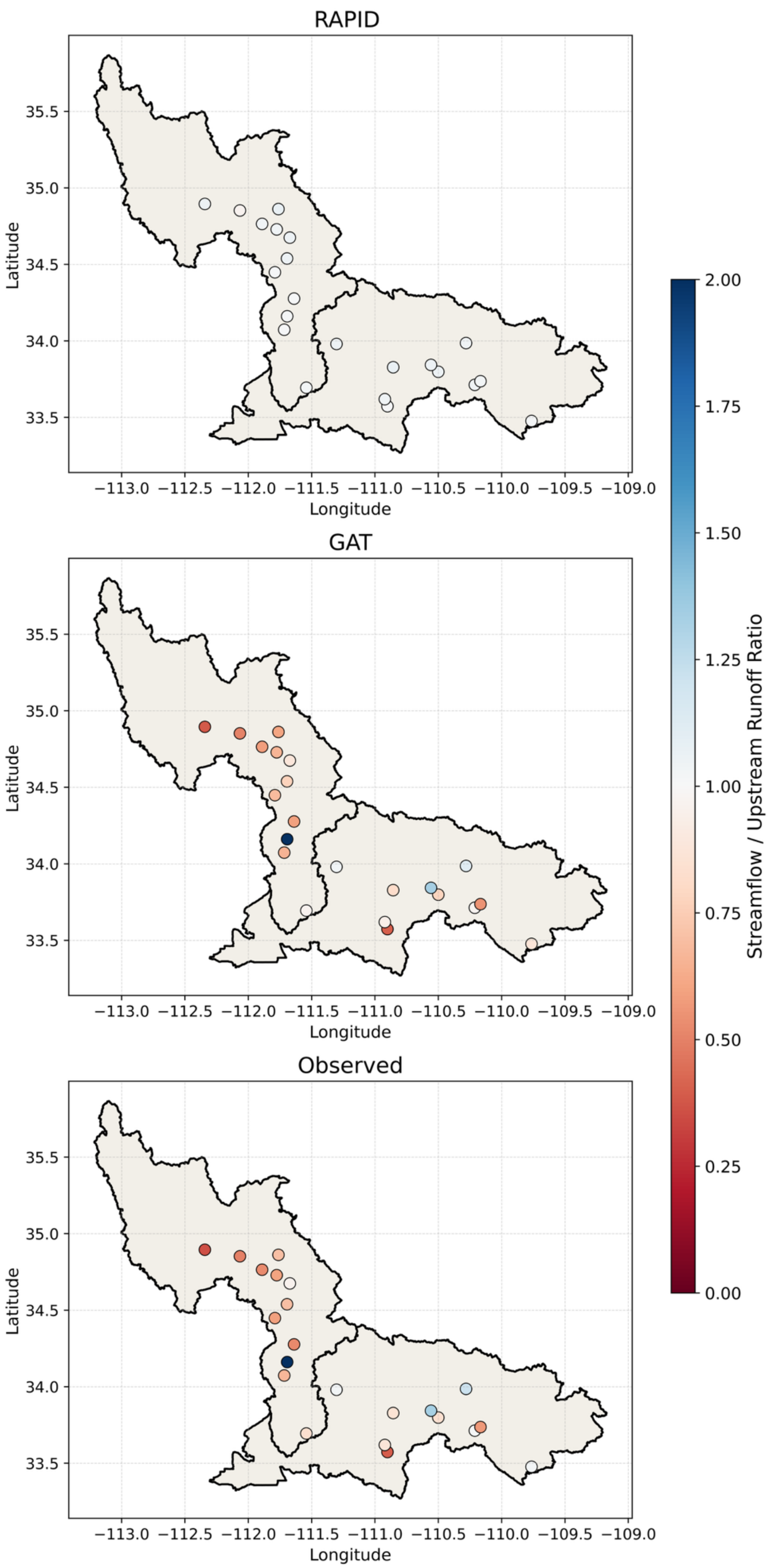
RAPID
GAT
Observed
Latitude
Longitude
35.5
35.0
34.5
34.0
33.5
−113.0 −112.5 −112.0 −111.5 −111.0 −110.5 −110.0 −109.5 −109.0
Streamflow / Upstream Runoff Ratio
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

Figure S4. Spatial comparison of streamflow-to-upstream-runoff ratios across evaluation gauges for RAPID routing, the graph-based routing framework (GAT), and observed streamflow relative to cumulative upstream Noah-MP runoff. RAPID exhibits near-conservative runoff–streamflow behavior with ratios clustered near unity, whereas the observed streamflow shows greater spatial heterogeneity reflecting integrated watershed processes and runoff-generation uncertainty. The graph-based routing framework produces runoff–streamflow relationships that more closely resemble the observed spatial variability across the Salt–Verde watershed.

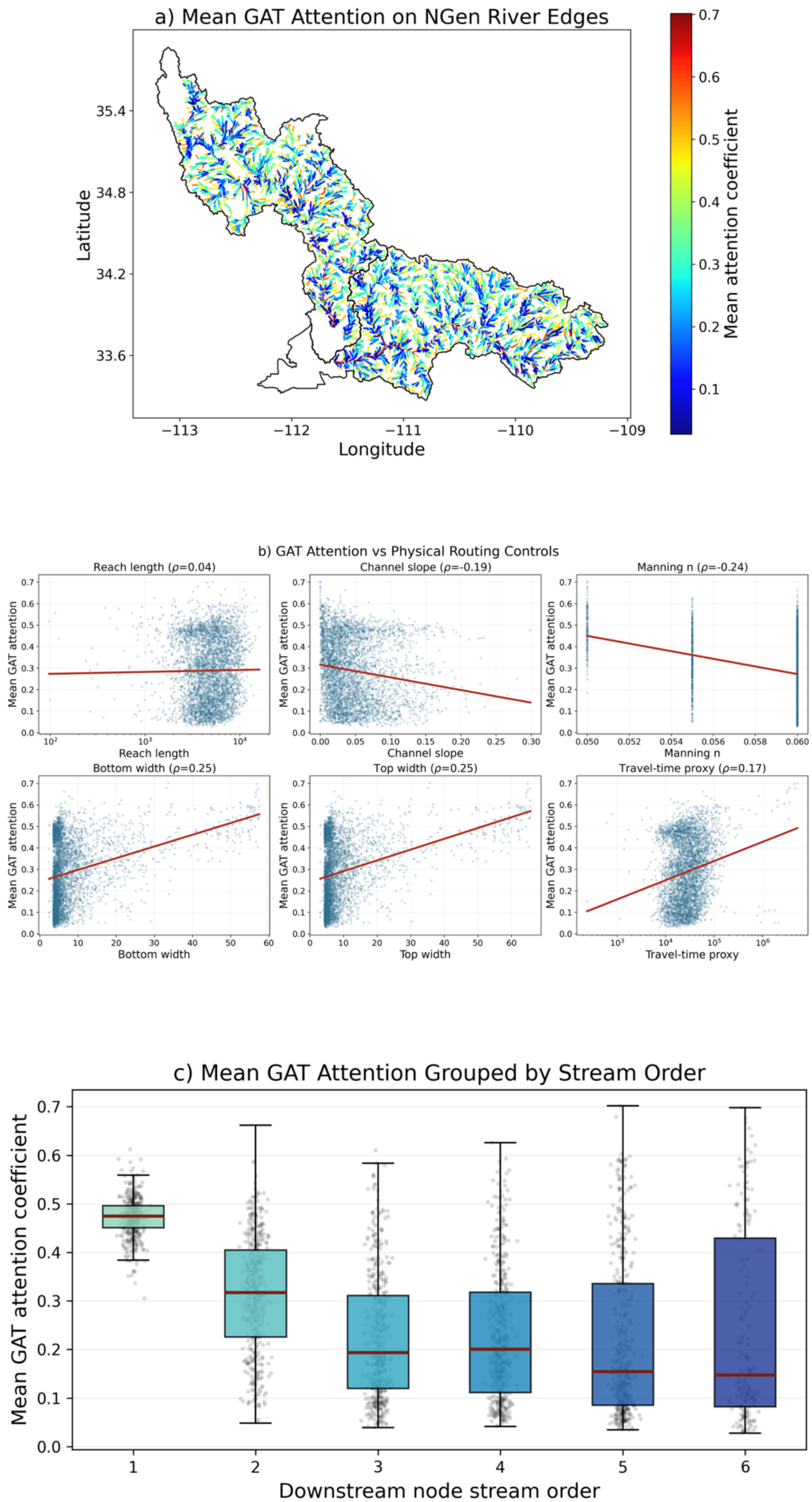
a) Mean GAT Attention on NGen River Edges
Latitude
Longitude
Mean attention coefficient
b) GAT Attention vs Physical Routing Controls
Reach length (ρ=0.04)
Channel slope (ρ=-0.19)
Manning n (ρ=-0.24)
Bottom width (ρ=0.25)
Top width (ρ=0.25)
Travel-time proxy (ρ=0.17)
Mean GAT attention
Reach length
Channel slope
Manning n
Bottom width
Top width
Travel-time proxy
c) Mean GAT Attention Grouped by Stream Order
Mean GAT attention coefficient
Downstream node stream order

Figure S5. Learned routing characteristics of RiverGraphNet. (a) Spatial distribution of mean graph-attention coefficients over the Salt–Verde river network. (b) Relationships between attention coefficients and physical channel attributes reveal that the model preferentially emphasizes wider, lower-slope, and lower-roughness reaches, which generally correspond to hydraulically efficient conveyance pathways. (c) Attention grouped by stream order shows the strongest influence in headwater reaches and progressively lower influence downstream, suggesting that accurate propagation of upstream runoff signals is a key control on routing performance. Together, these results indicate that RiverGraphNet learns spatially heterogeneous routing behavior governed by both river-network topology and channel hydraulic characteristics.

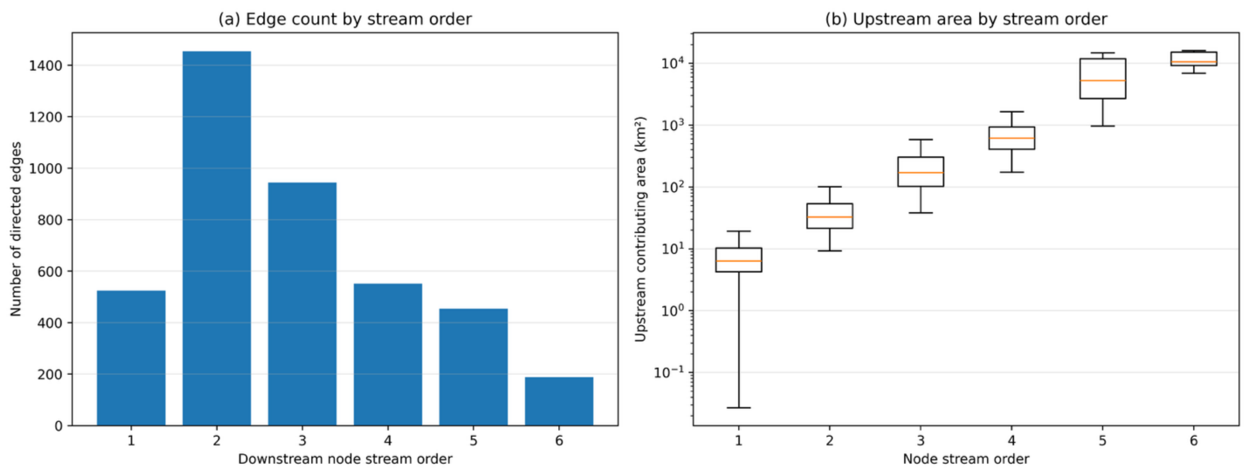

Figure S6. River-network topology summarized by stream order. (a) Number of directed graph edges associated with each downstream stream-order class. (b) Distribution of upstream contributing area for graph nodes grouped by stream order. Although contributing area increases systematically with stream order, first-order reaches receive the highest graph-attention weights (Figure 13c), indicating that RiverGraphNet preferentially emphasizes headwater runoff signals rather than simply focusing on the largest drainage areas or most abundant stream-order classes.

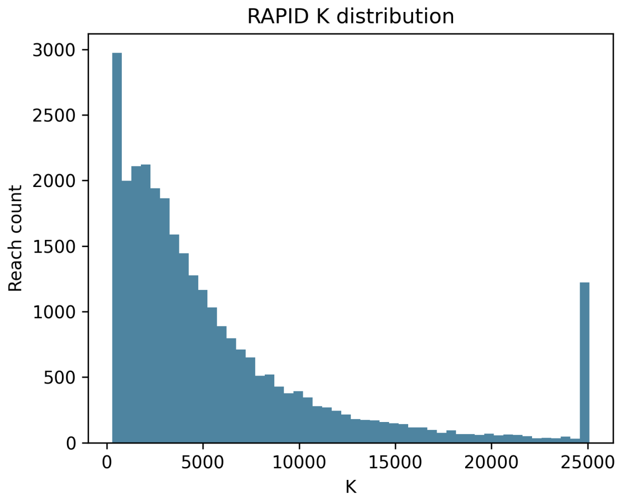

Figure S7. Distribution of RAPID Muskingum routing parameters across the Salt–Verde river network. Histograms show the spatial distributions of the reach-specific Muskingum travel-time parameter (K) and weighting parameter (X) used by RAPID. The K parameter exhibits substantial variability across river reaches, reflecting spatial differences in routing travel times, whereas X is nearly uniform throughout the network. These distributions demonstrate that RAPID already incorporates spatially varying travel-time representation through reach-specific K values.

Table S1. Detailed description of graph node and edge attributes used in RiverGraphNet.

Variable	Type	Units	Description
Length_m	Edge	m	River reach length
So	Edge	$m\ m^{-1}$	Channel bed slope
n	Edge	–	Manning roughness coefficient
nCC	Edge	–	Compound-channel Manning roughness coefficient
BtmWdth	Edge	m	Channel bottom width
TopWdth	Edge	m	Channel top width
TopWdthCC	Edge	m	Compound-channel top width
MusK	Edge	s	RAPID Muskingum travel-time parameter
MusX	Edge	–	RAPID Muskingum weighting parameter

travel_time_proxy	Edge	s	Approximate reach travel time derived from hydraulic descriptors
velocity_proxy	Edge	$m\ s^{-1}$	Approximate flow velocity estimate
conveyance_proxy	Edge	—	Proxy representing relative channel conveyance capacity
storage_proxy	Edge	—	Proxy representing relative channel storage potential
width_ratio_proxy	Edge	—	Ratio-based descriptor of channel geometry
node_dem_elevation_m	Node	m	Mean elevation of NGen catchment/divide
lengthkm	Node	km	Local flowpath length
areasqkm	Node	km^2	Incremental drainage area
tot_drainage_areasqkm	Node	km^2	Total upstream contributing area
order	Node	–	Stream order
mean.slope	Node	% or $m\ m^{-1}$	Mean catchment slope
mean.impervious	Node	%	Mean impervious fraction
mode.ISLTYP	Node	–	Dominant Noah-MP soil class
mode.IVGTYP	Node	–	Dominant Noah-MP vegetation class
distance_to_outlet_m	Node	m	Distance from node to basin outlet
travel_time_proxy_to_outlet	Node	s	Estimated travel time from node to outlet
upstream_flowpath_count	Node	–	Number of upstream contributing reaches
normalized_stream_order	Node	–	Stream order normalized by basin maximum order
is_outlet_node	Node	Boolean	Indicates basin outlet node

Table S2. Median absolute percent errors of selected hydrologic signatures across evaluation gauges. Errors are computed relative to observed streamflow. The table provides the numerical values summarized visually in Figure S3.

Model	Mean flow (%)	High flow Q5 (%)	Median flow Q50 (%)	Low flow Q95 (%)	FDC slope (%)	Recession slope (%)	BFI (%)
GNN-Weighted MSE	7.7	18.7	25.5	77.3	81.5	99.0	27.0
GNN-MSE	15.0	28.0	25.5	69.1	100.0	100.0	28.6
GNN-JKGE	26.7	34.0	23.5	38.8	81.7	91.9	18.9
RAPID	52.5	56.2	75.0	95.1	145.4	34.7	24.6