2025 MS-CC Undergraduate Research Internship

PRESENTED BY: JORDAN JOHNSON & MYA TOLBERT

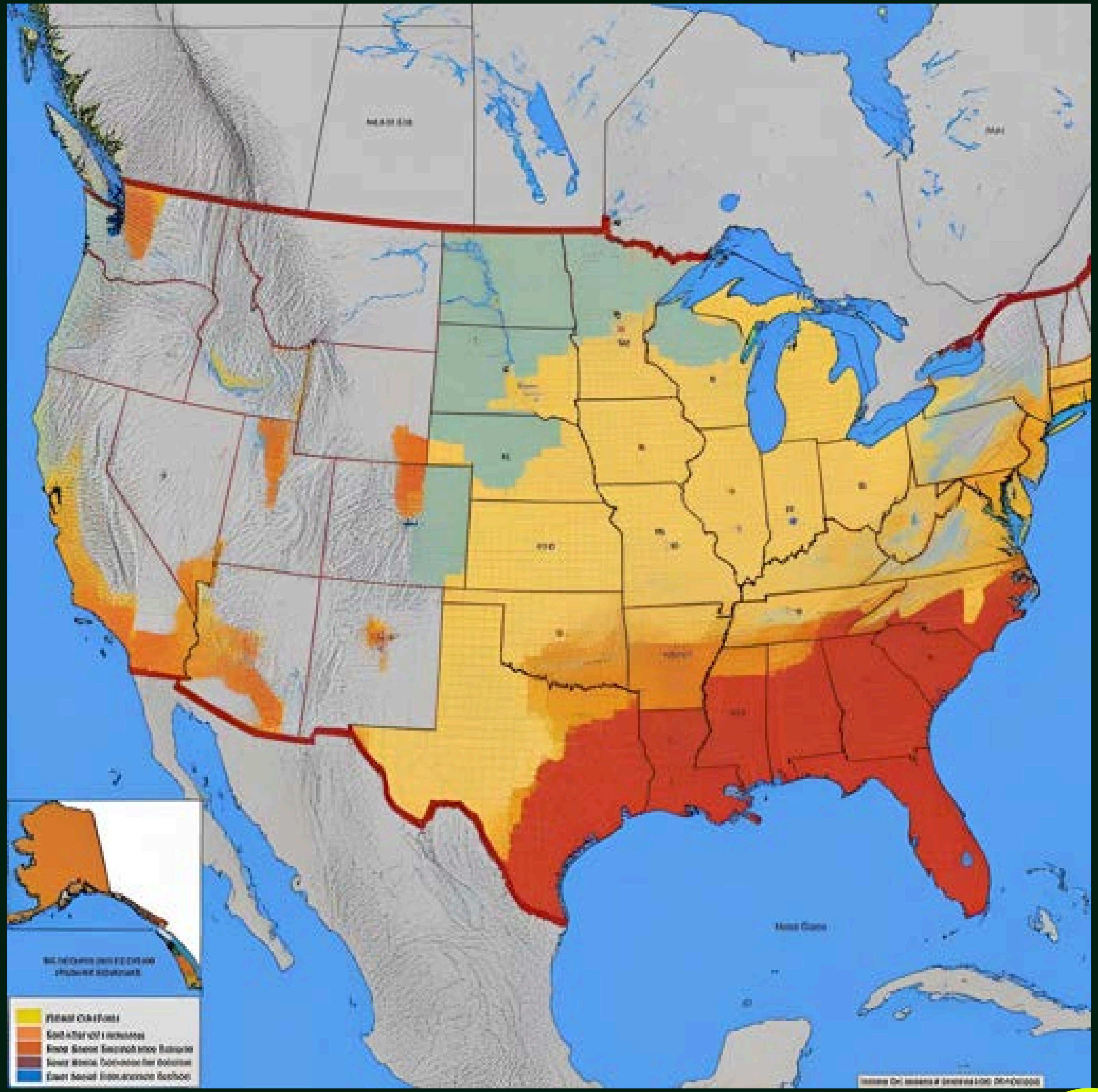
MENTOR: DR. NING ZHANG

STREAMFLOW PREDICTION USING GRAPH NEURAL NETWORKS AND CLIMATE DATA

Teaching AI to Read the Rivers

The Stakes: Why Predicting River Flow Matters

- $25B IN U.S. FLOOD DAMAGES (2023)
- IMPACTS: HOMES, FOOD SYSTEMS, INFRASTRUCTURE, EMERGENCY RESPONSE
- CLIMATE SHIFTS = CHAOTIC RAINFALL PATTERNS



Our Goal: Predict River Flow Before It Happens

- PREDICT NEXT-DAY STREAMFLOW FOR 25 U.S. BASINS
- USE 30 DAYS OF PAST CLIMATE DATA
- MODEL REGIONAL RIVER CONNECTIONS WITH GNNS

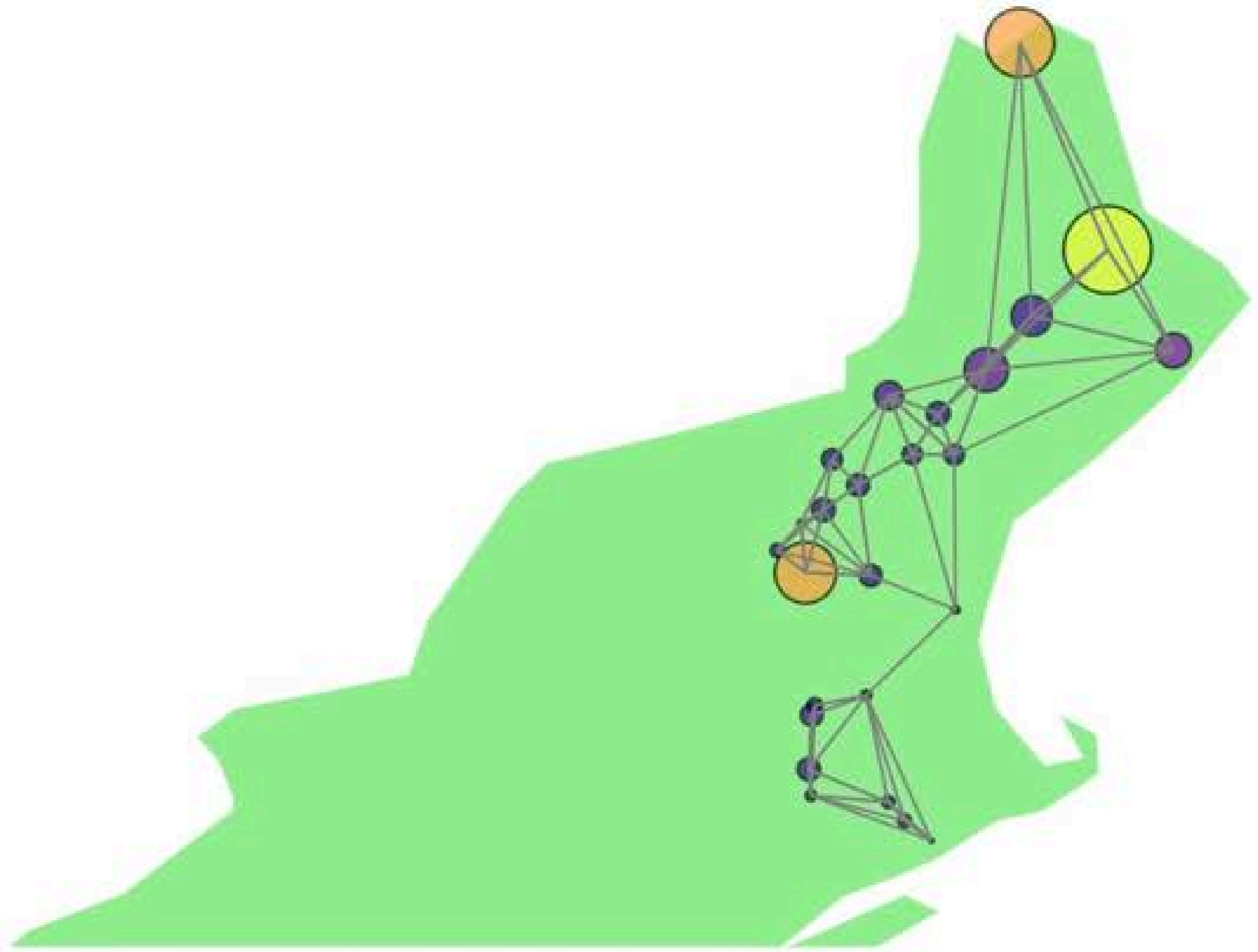
Meet CAMELS: Our Climate+Streamflow Dataset

Column	Description
Year	The calendar year (e.g., 2005)
Mnth	The calendar month (1 = January, ..., 12 = December)
Day	The day of the month
Hr	Hour of day (usually not used in daily data — might be 0 or placeholder)
dayl (s)	**Day length** in **seconds** (i.e., daylight duration from sunrise to sunset)
prcp (mm/day)	**Precipitation** in **millimeters per day**
srad (W/m²)	**Shortwave solar radiation** in **Watts per square meter**
swe (mm)	**Snow water equivalent** in **millimeters** — water content of snowpack
tmax (°C)	**Maximum daily temperature** in **degrees Celsius**
tmin (°C)	**Minimum daily temperature** in **degrees Celsius**
vp (Pa)	**Vapor pressure** in **Pascals** — indicates humidity in the atmosphere

date	prcp(mm/day)	srad(W/m2)	tmax(C)	tmin(C)	vp(Pa)
1980-01-01	0.00	153.40	-0.0654	-0.1630	171.69
1980-01-02	0.00	145.27	-0.0618	-0.1522	185.94
1980-01-03	0.00	146.96	-0.0989	-0.1886	138.39
1980-01-04	0.00	146.20	-0.1098	-0.1976	120.06
1980-01-05	0.00	170.43	-0.1129	-0.2221	117.87
...	...	...	...	...	...
2014-12-26	0.00	91.11	0.0361	-0.0050	590.58
2014-12-27	0.00	103.01	0.0215	-0.0255	508.64

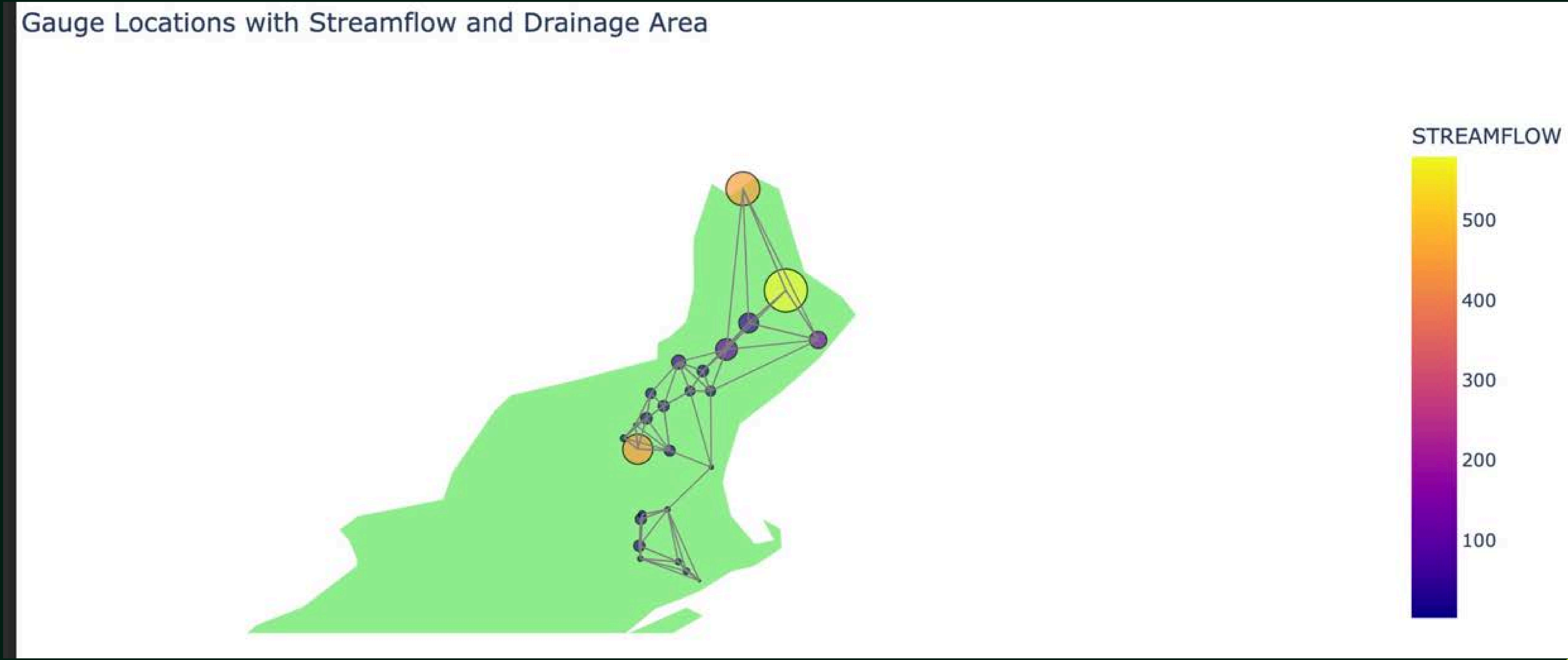
- 671 U.S. RIVER BASINS (WE USED 25 WITH COMPLETE RECORDS)
- 30+ YEARS OF DAILY DATA
- VARIABLES: PRECIPITATION, TMAX/TMIN, SOLAR RADIATION, VAPOR PRESSURE, STREAMFLOW
- 30-DAY ROLLING CLIMATE WINDOW USED AS MODEL INPUT
- TARGET: NEXT-DAY STREAMFLOW PREDICTION

Understanding GNNs: Teaching AI to Learn from Connections



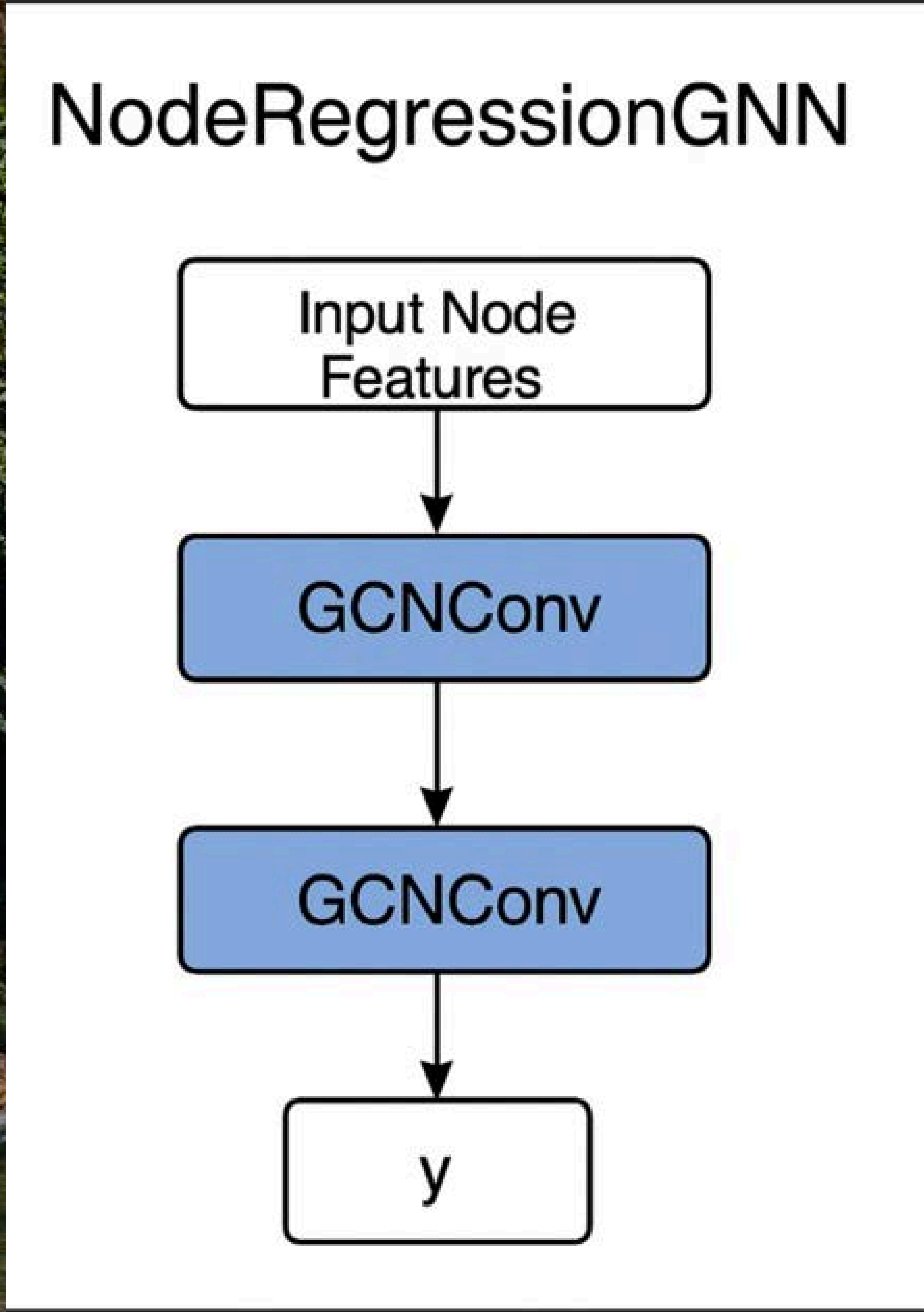
- **GNN'S LEARN FROM RELATIONSHIPS, NOT JUST INDIVIDUAL POINTS**
- **RIVERS AREN'T ISOLATED , GNN'S CAPTURE HOW ONE BASIN AFFECTS ANOTHER**

Turning Rivers into a Graph Neural Network

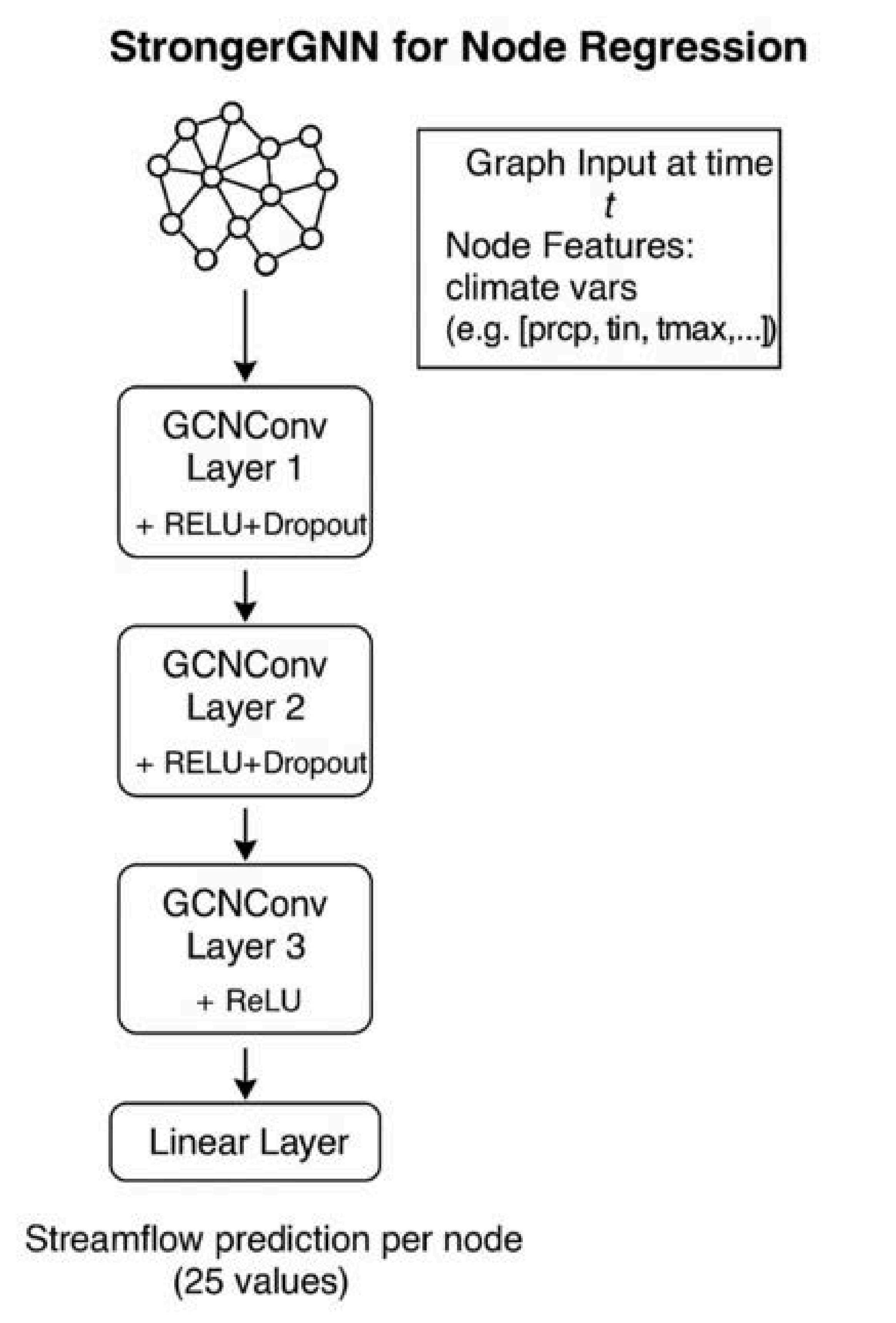


- NODES = BASINS
- EDGES = NEARBY BASINS (K-NEAREST NEIGHBORS)
- FEATURES = CLIMATE VARIABLES OVER 30 DAYS

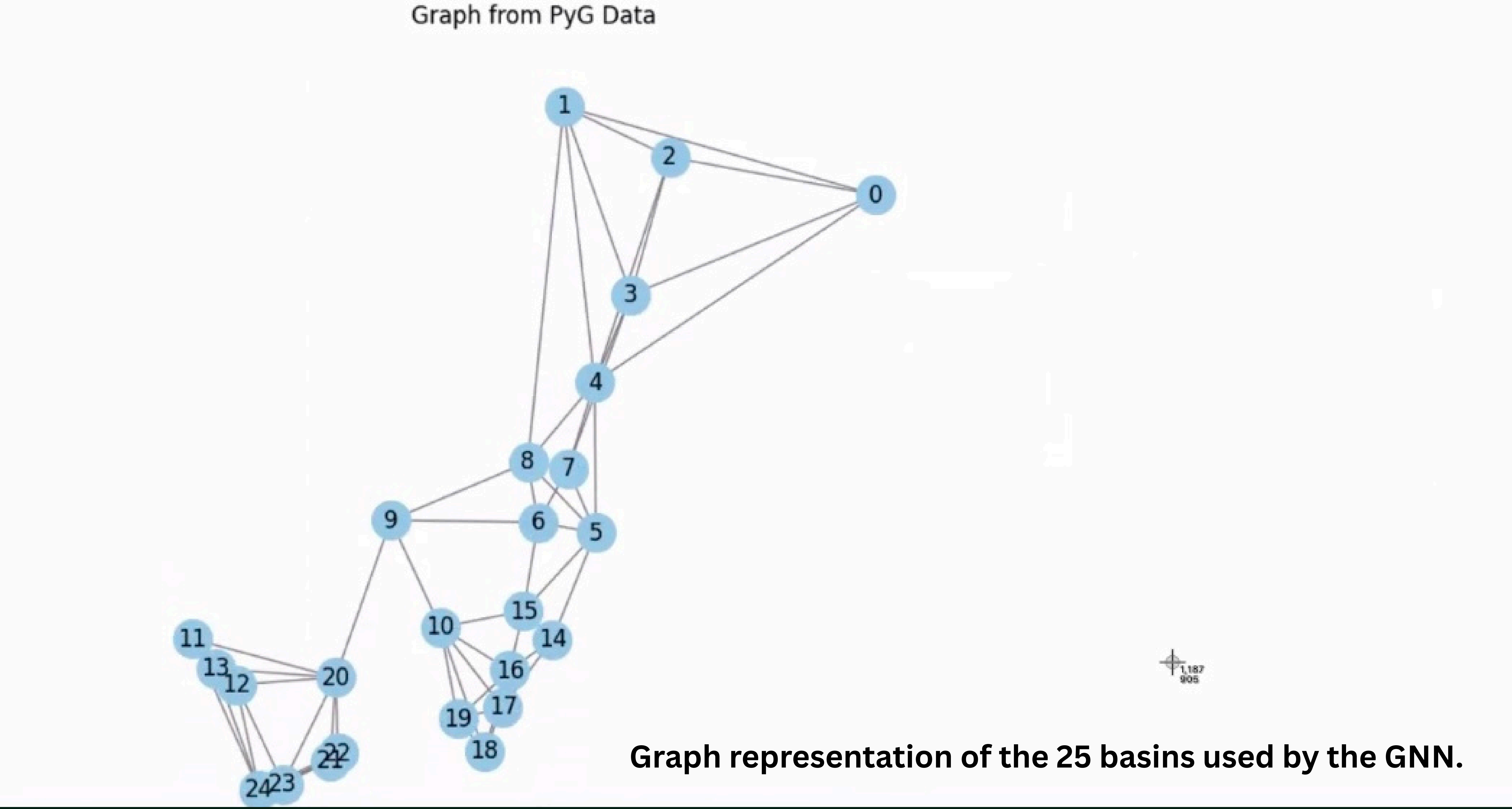
FROM DATA TO PREDICTIONS: BUILDING THE STREAMFLOW PIPELINE



- LOAD AND CLEAN 30-DAY WINDOWS OF CLIMATE DATA
- JOIN WITH STREAMFLOW LABEL (NEXT DAY)
- CREATE A DAILY GRAPH SNAPSHOT



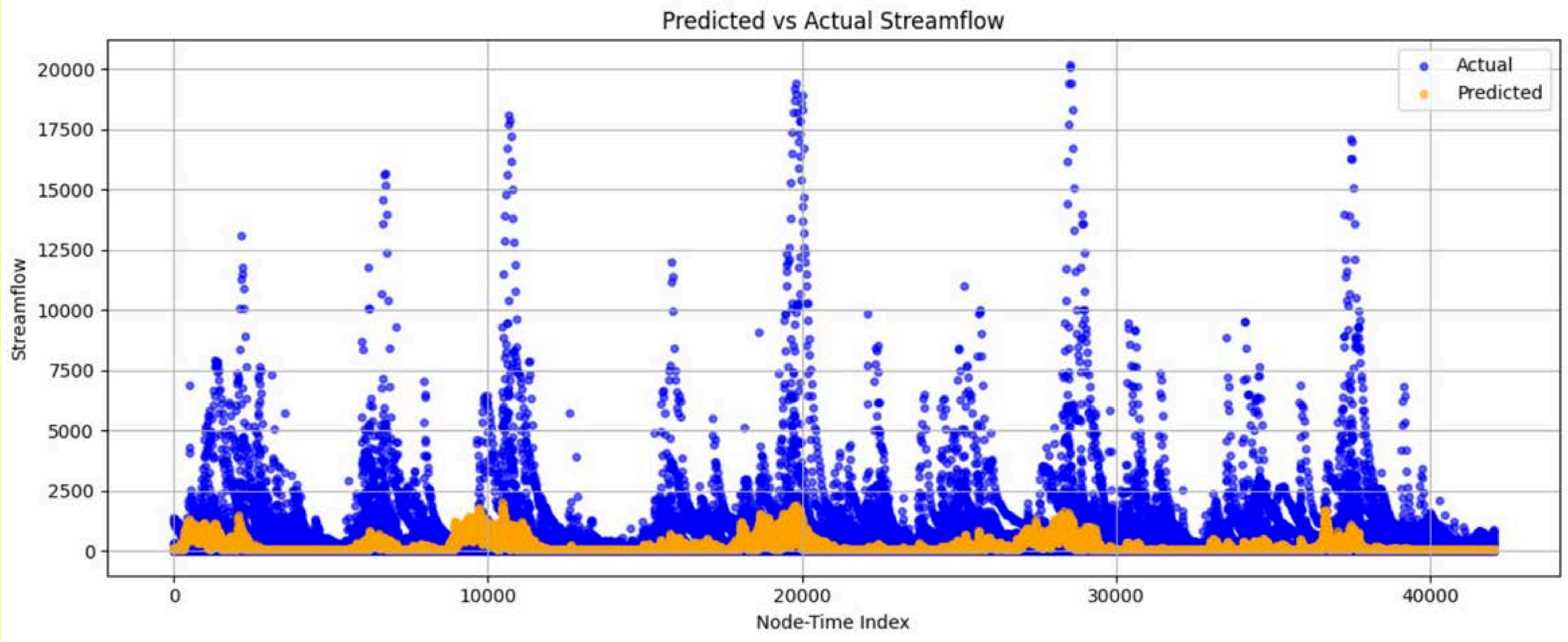
How the GNN Sees the River System



Graph representation of the 25 basins used by the GNN.

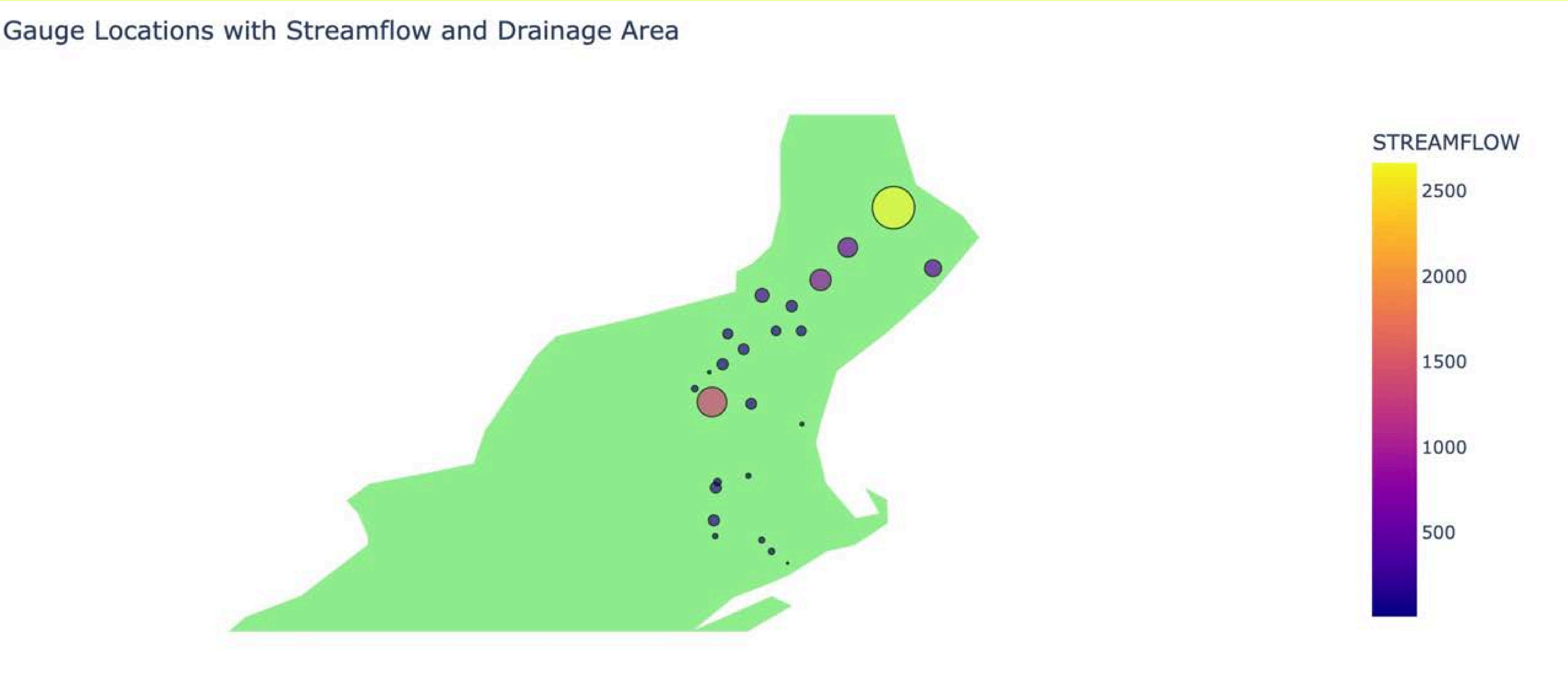
Prediction Results: Did It Work?

- TRACKED REAL STREAMFLOW PATTERNS
- CAPTURED RAINFALL SPIKES AND DRY PERIODS
- UNDERSTOOD UPSTREAM AND DOWNSTREAM EFFECTS

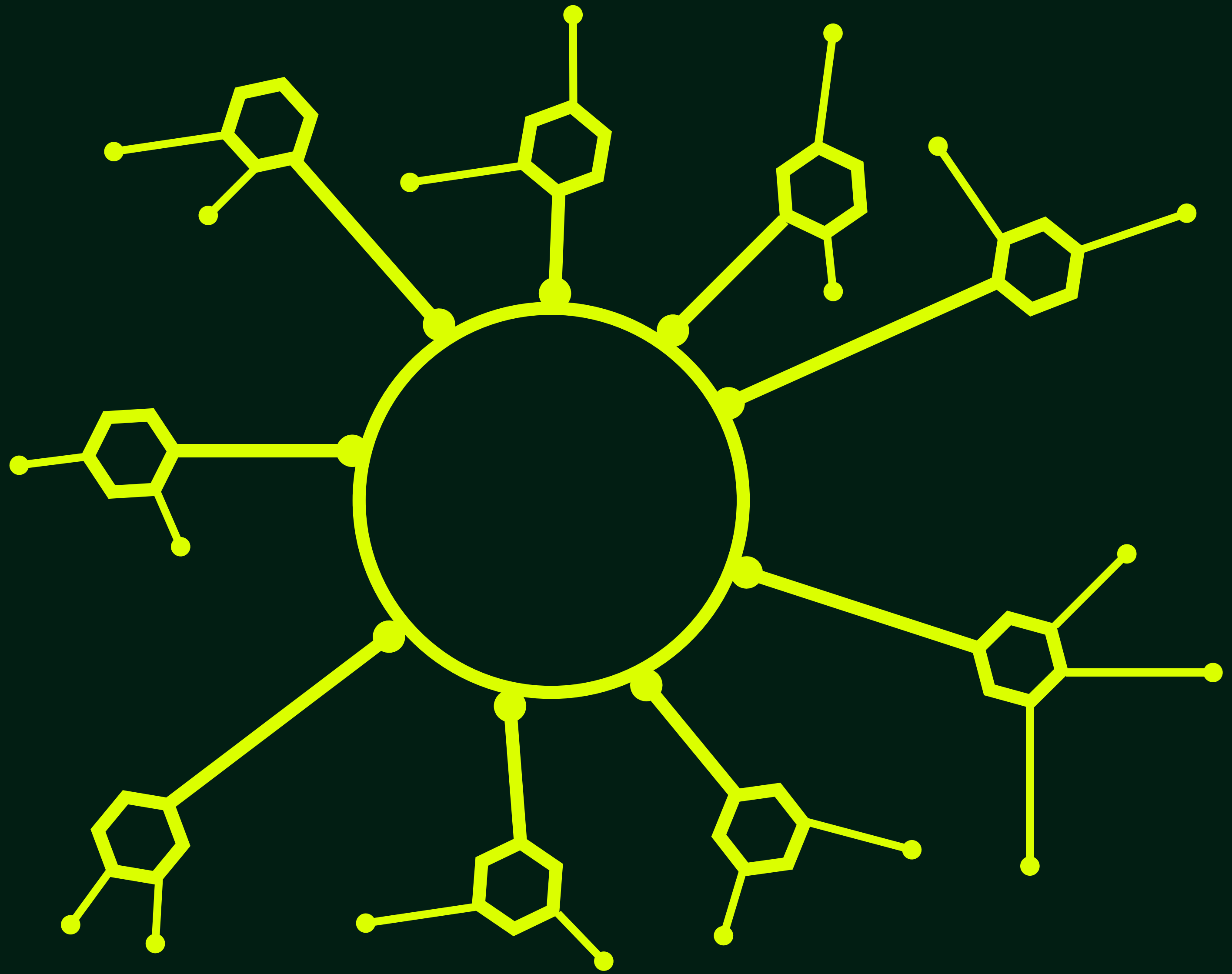


Prediction Results: Did It Work? (Pt.2)

- **Each node's color and size reflect the magnitude of streamflow.**



What We Learned from the River



- CLEANING REAL DATA IS ESSENTIAL
- GNNS ARE POWERFUL FOR SPATIAL DATA
- TEAMWORK MADE THE PROJECT POSSIBLE

The Future of Flood Forecasting

- TRY NEW GNN ARCHITECTURES (GRAPHSAGE, GAT)
- PREDICT MORE DAYS AHEAD
- EXPAND TO MORE BASINS
- REAL-TIME SYSTEM POTENTIAL

Thank you for your time! We look forward to your feedback and questions.

Jordan Johnson
Texas Southern University
Chemistry Major, Biology Minor
Rising Senior
: j.johnson2820@student.tsu.edu

Mya Tolbert
University of Arkansas at Pine Bluff
Biology/ Pre-Dentistry
Rising Senior
:tolberm5979@uapb.edu