EGU GENERAL ASSEMBLY 2026 · OSPP CANDIDATE

Session HS4.10 · Hall A · A.37 · Wed 6 May

Deep learning for monthly precipitation prediction in mountainous terrain

Manuel Ricardo Pérez Reyes

Marco J. Suárez Barón · Óscar J. García Cabrejo

UPTC · Pedagogical and Technological University of Colombia

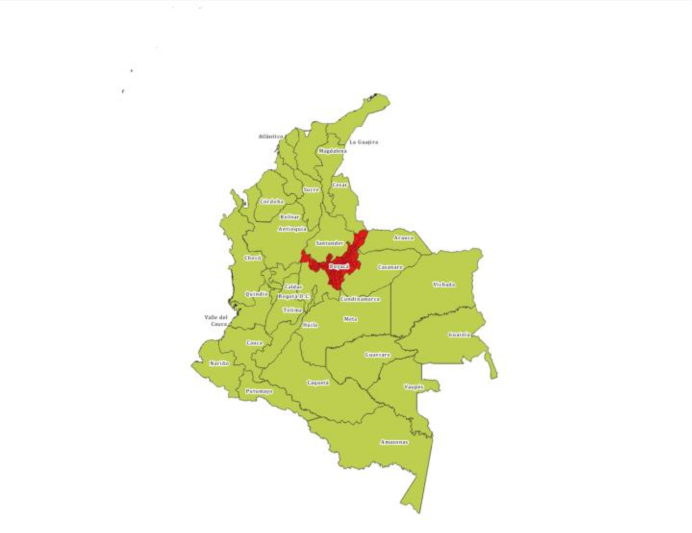
HEADLINE RESULT

$R^2 = 0.655$

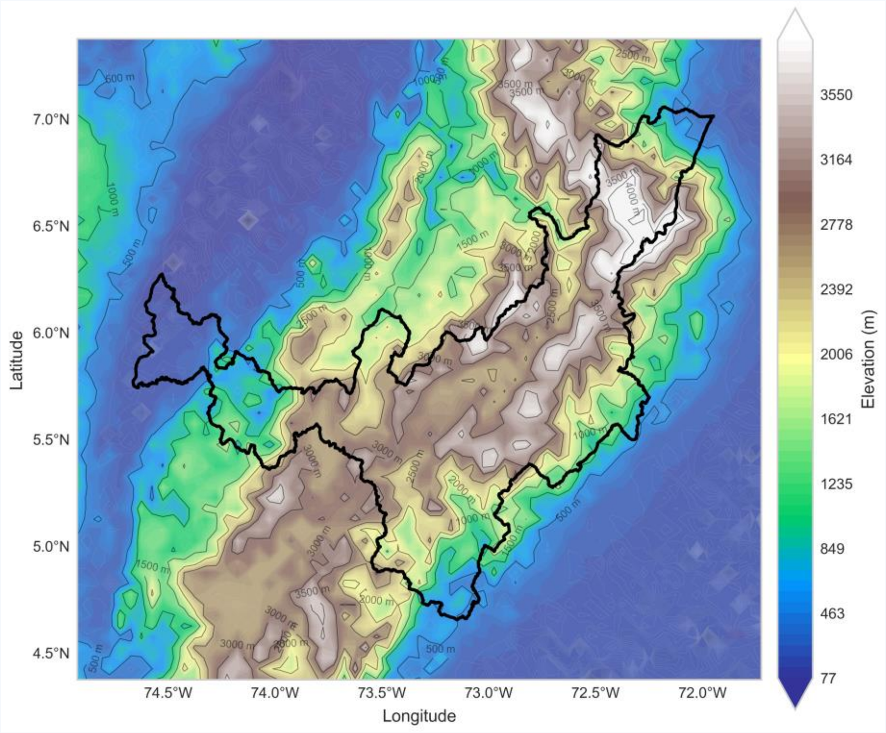
±0.018 (3-seed mean)

Late Fusion (Ridge) of two deep models beats every single architecture across all 12 forecast horizons.

Boyacá, Colombian Andes: extreme orographic gradients



Boyacá department (red) in Colombia.



SRTM elevation: 145 m (valleys) → 5,500 m (páramo).

ELEVATION

145 to 5,500 m

single watershed range

GRID

3,965 cells

61 × 65 CHIRPS lattice

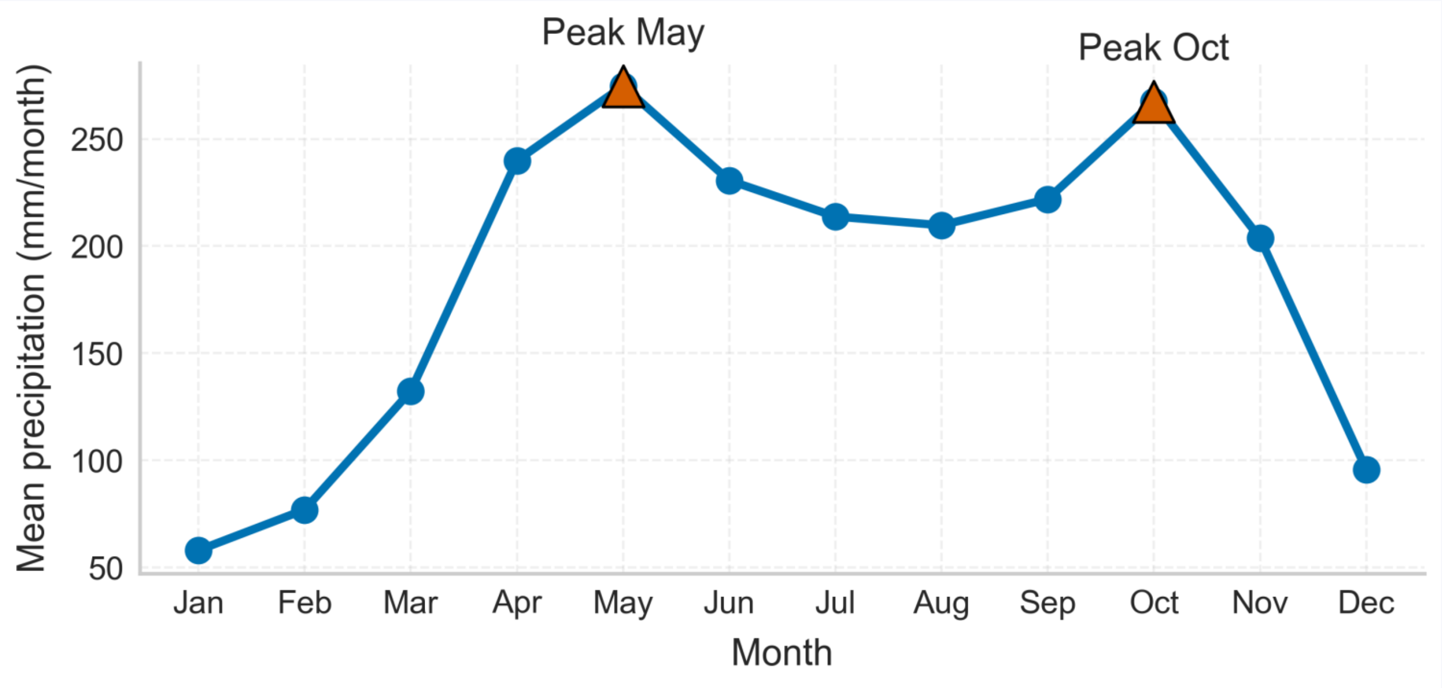
HISTORY

35 years

monthly · 1985–2022

Bimodal regime: Apr–May & Oct–Nov peaks

Two rainy seasons motivate temporal feature engineering



Mean monthly precipitation over Boyacá (1982–2025). Peaks at May (275 mm) and October (267 mm).

PEAK 1

May · 275 mm

first wet season

PEAK 2

October · 267 mm

second wet season

MOTIVATES

sin/cos + t–12

temporal features in PAFC

Why it matters: a single sinusoid cannot capture twin peaks 5 months apart. This motivates the sin/cos month + t-12 lag in the PAFC bundle.

Three deep-learning families, comparable performance

CONVLSTM

Convolutional LSTM

R^2 = 0.628 · 148K params · BASIC features

FNO

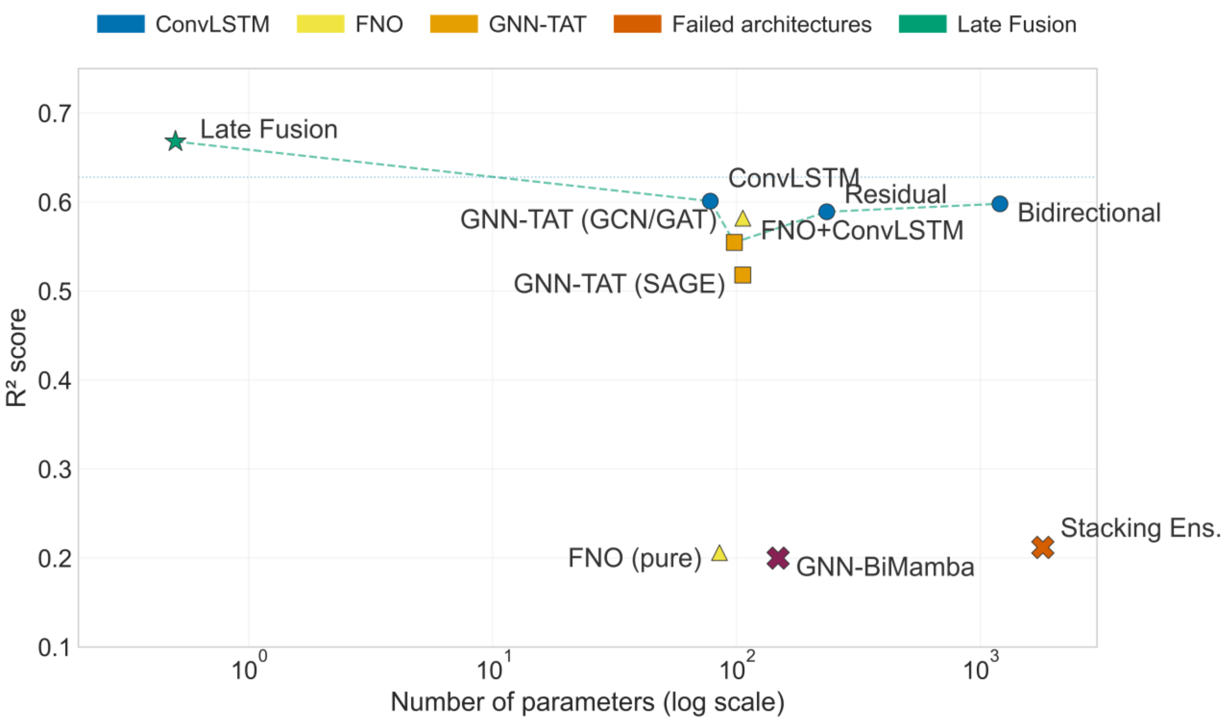
Fourier Neural Operator

R^2 = 0.582 · FNO–ConvLSTM hybrid (rescue effect)

GNN-TAT

Graph NN + Temporal Attention

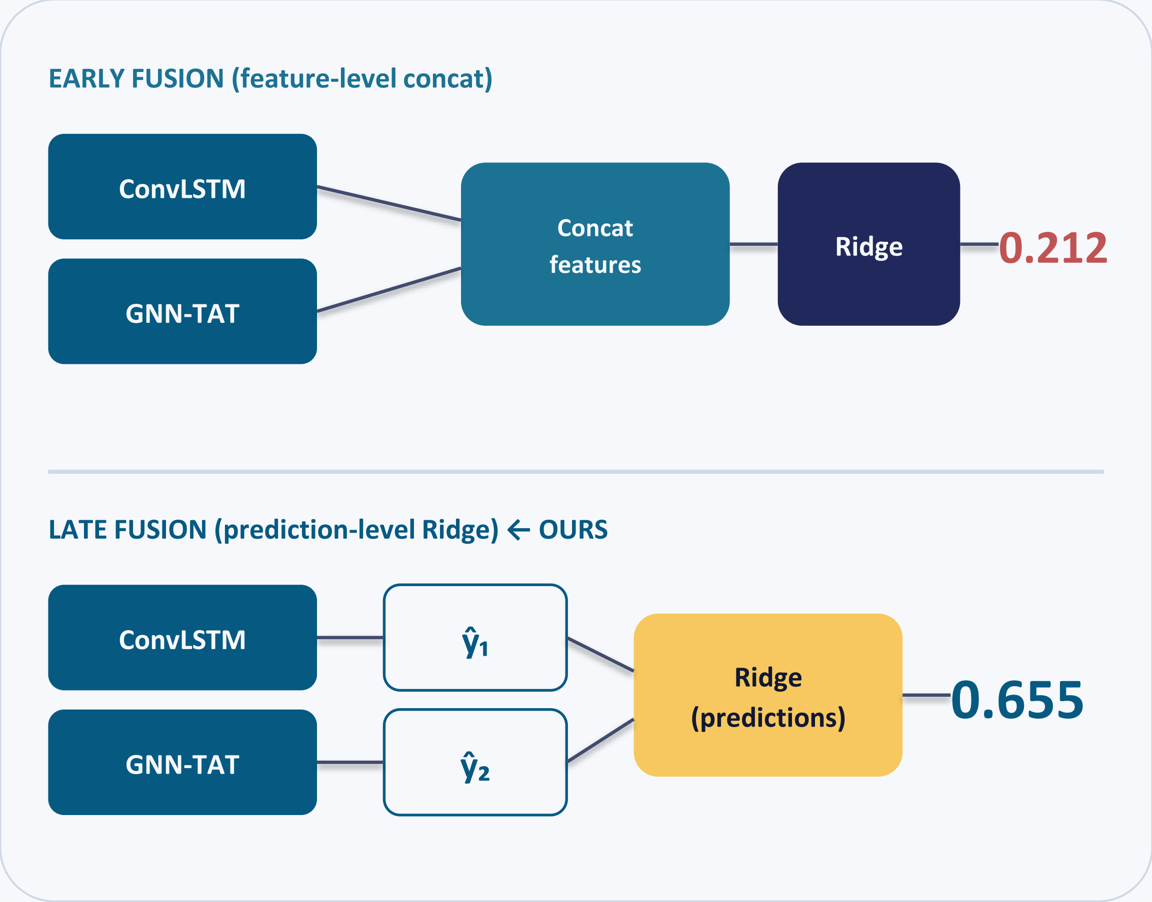
R^2 = 0.628 · 98K params · 34% fewer than ConvLSTM

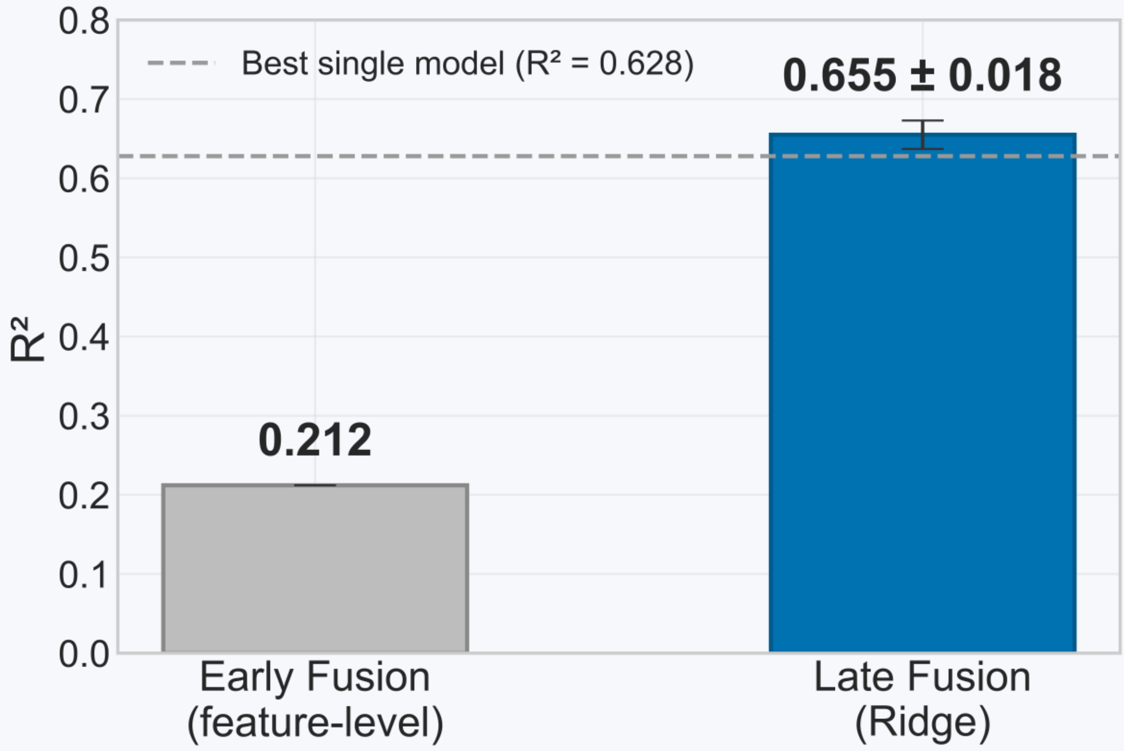


Parameters vs R^2. Late Fusion delivers the best score with the smallest footprint.

BASIC = temporals + topography · KCE = BASIC + elevation clusters · PAFC = KCE + precipitation lags

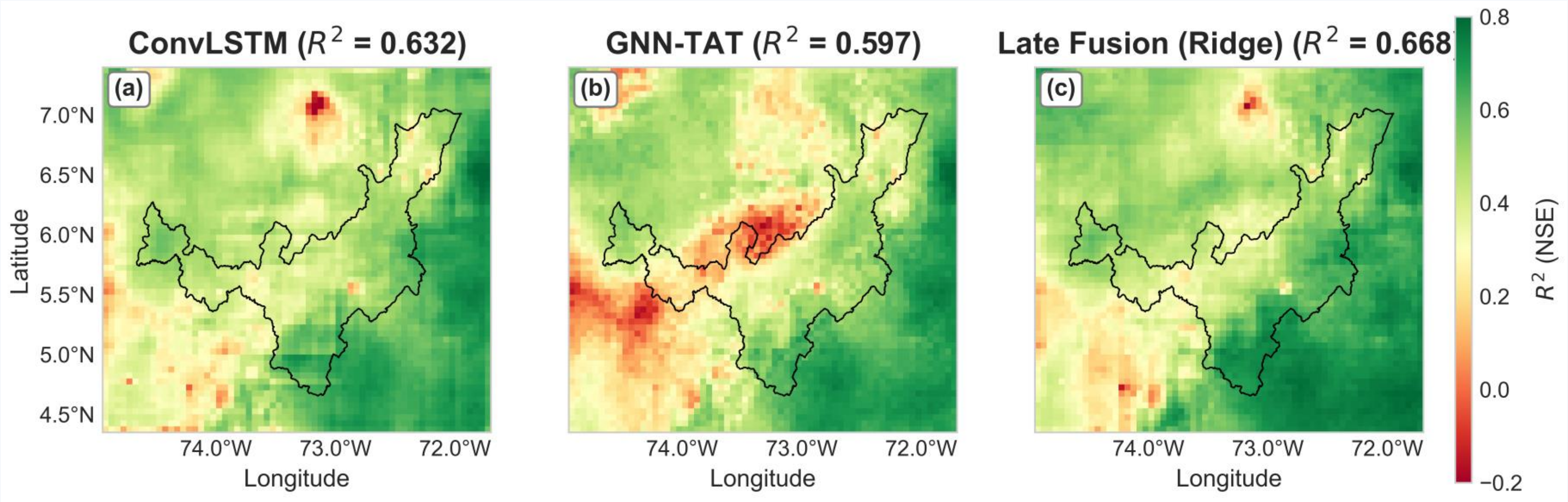
Late Fusion wins. Early Fusion fails. Same components.





Same components, two fusion strategies. Fuse predictions, not features.

Why fusion works: complementary failure modes

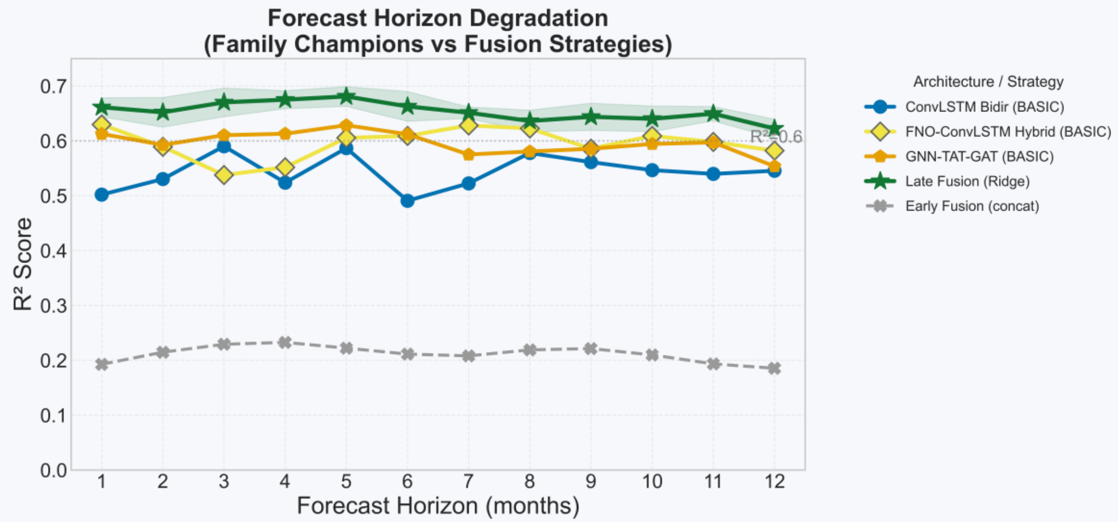


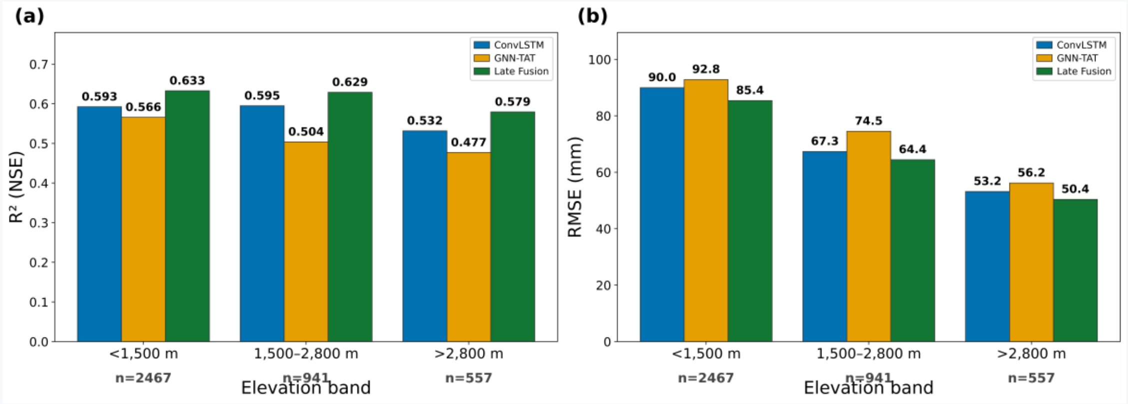
Per-cell R² maps. ConvLSTM and GNN-TAT fail in different places; Late Fusion absorbs both.

THE FUSION INTUITION

Each model corrects the other's spatial blind spots. Ridge learns the per-cell weighting that combines two flawed predictors into one reliable forecast.

Stable across horizons, KCE-aligned bands, and ENSO years





Forecast horizon (1–12 months): Late Fusion leads at every step.

Elevation bands: gain holds from valleys to páramo.

12/12 horizons

Late Fusion best at every H

16% drop

low → high (KCE bands, n=2467/941/557)

Med ≈ High

transition zone is hardest, not páramo

What this means and where we go next

01

Predictability ceiling

Three families converge to $R^2 \approx 0.580–0.630$. Limit may be data, not model.

02

Fusion beats the ceiling

+6.4% with only ~1,000 extra parameters. Cost is low, gain is real.

03

Strategy > architecture

Where you combine matters more than which models you pick.

NEXT

Wider seed ensembles · physics-guided Mixture of Experts · operational deployment with IDEAM

COMPANION OUTPUTS

Benchmark · Hydrology 2026, 13(3), 98 | Methodology · Journal of Hydrology (in review) | Code · github.com/ninja-marduk/ml_precipitation_prediction

THANK YOU

Fuse predictions, not features.

Combining ConvLSTM and GNN-TAT predictions with Ridge regression improves accuracy at every forecast horizon, by up to +6.4%.

CONTACT

Manuel R. Pérez Reyes · manuelricardo.perez@uptc.edu.co

ORCID 0009-0003-2963-1631 · Pedagogical and Technological University of Colombia

Paper · code · poster