Pure and Physics-encoded Spatiotemporal Deep Learning for Climate-Vegetation Dynamics

Qianqiu Longyang (qlongyang@ku.edu), Kansas Geological Survey, The University of Kansas

Ruijie Zeng, Arizona State University

Introduction

- Vegetation regulates land surface water and energy exchanges.
- Accurate representation of climate-vegetation feedback is crucial for reliable climate projections and water resource management.

Climate change → **Climatic conditions** $\boldsymbol{X}$

$$Y = g(\boldsymbol{X}, \boldsymbol{\alpha}, \boldsymbol{\beta})$$

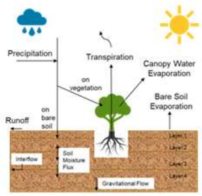
Static $\boldsymbol{\alpha}$	Time-varying $\boldsymbol{\beta}$	Fully-Dynamic
Elevation Soil type	**<u>Vegetation</u>** Albedo	Climatic/terrestrial- feedbacks

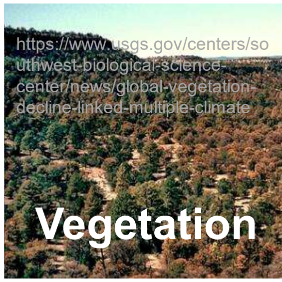
Data-driven models of vegetation dynamics (e.g., ML)
\+ Satellite-derived vegetation dynamics

- Climate($\boldsymbol{X}$)-driven vegetation dynamics modeling

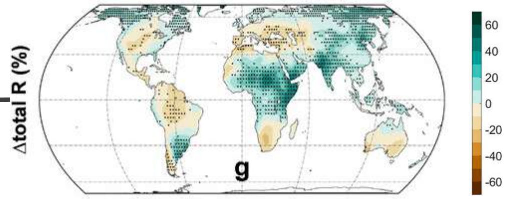
$$\boldsymbol{\beta}_t = \boldsymbol{f}(\boldsymbol{X}_t)$$

g





Hydrologic response $\boldsymbol{Y}$



Zhao, T., & Dai, A. (2022). CMIP6 model-projected hydroclimatic and drought changes and their causes in the twenty-first century. *Journal of Climate*, *35*(3), 897-921.

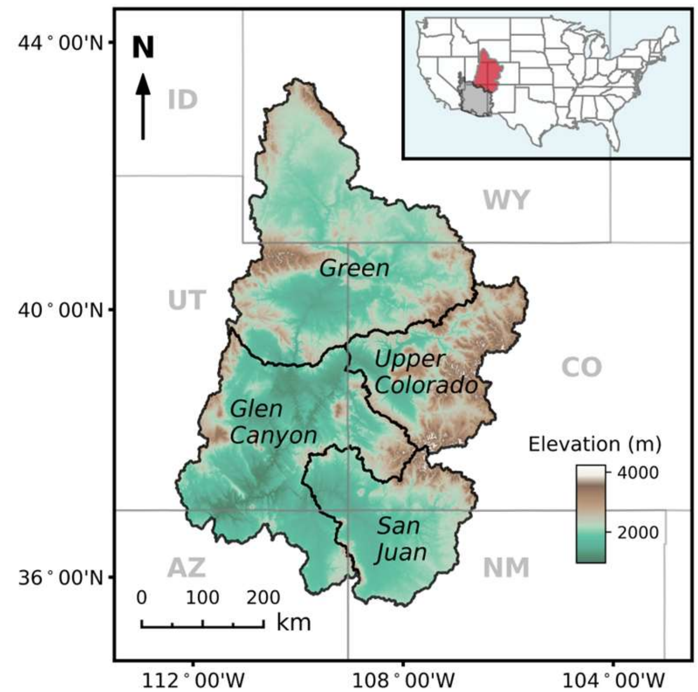
Limitations of current approaches exist

Traditional Process-Based Vegetation Models	Existing Deep Learning (DL) Applications
Oversimplified parameterizations - monthly climatology fails to capture extremes and trends; fixed plant functional types insufficiently represent real diversity	**Black-box limitations** - lacks interpretability and physical consistency for scientific understanding
Structural errors - rules derived from point-scale observations do not scale well	**Forecast-focused instead of projection-focused** - incorporates prior vegetation states, unsuitable for climate impact studies
Limited spatiotemporal interactions - prioritize vertical atmospheric feedbacks while neglecting horizontal processes	**Limited spatially distributed applications** - point-scale or flattened inputs overlook heterogeneity and risk overfitting

- We aim to develop a climate-driven, spatially distributed deep learning framework for daily-scale vegetation dynamics.

Study Area and Dataset

- Upper Colorado River Basin (UCRB) is a critical water source region with diverse hydroclimatic gradients and complex vegetation dynamics
- **Climate data**: daily precipitation, temperature, radiation, and humidity from NLDAS-2 at 1/8° resolution, providing consistent forcing inputs across CONUS
- **Vegetation data**: GLOBMAP Leaf Area Index (LAI) at 0.073° resolution, with global coverage from 1981 onward at half-monthly to 8-day intervals
- **Data split**: temporal split, training (1981-2001), validation (200-2008), testing (2009-2016)

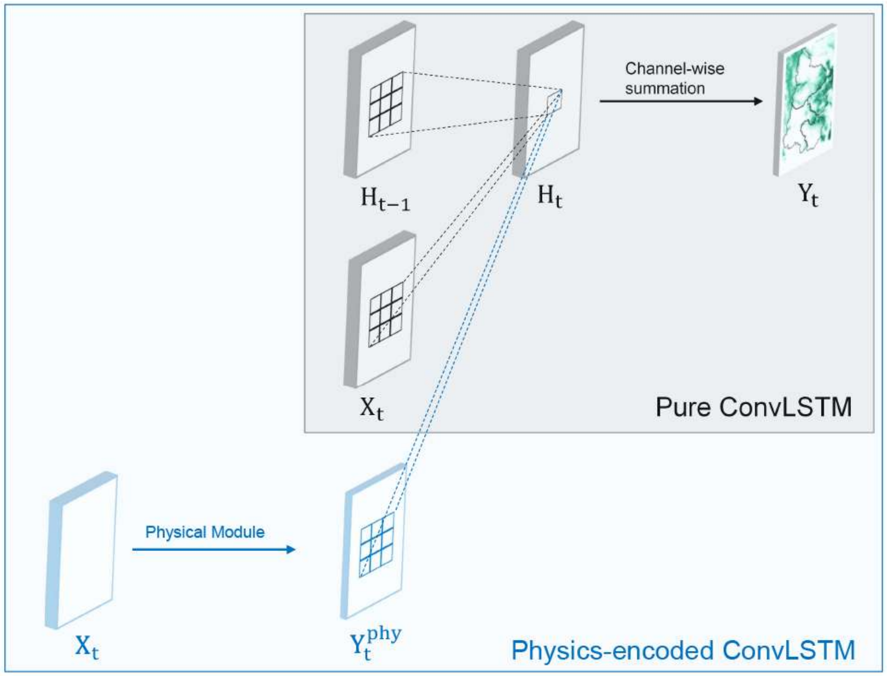


Methods

- A one-layer ConvLSTM model with 16 filters is employed, framing LAI prediction as a spatiotemporal sequence-to-sequence problem to process 2D climate inputs and output LAI maps

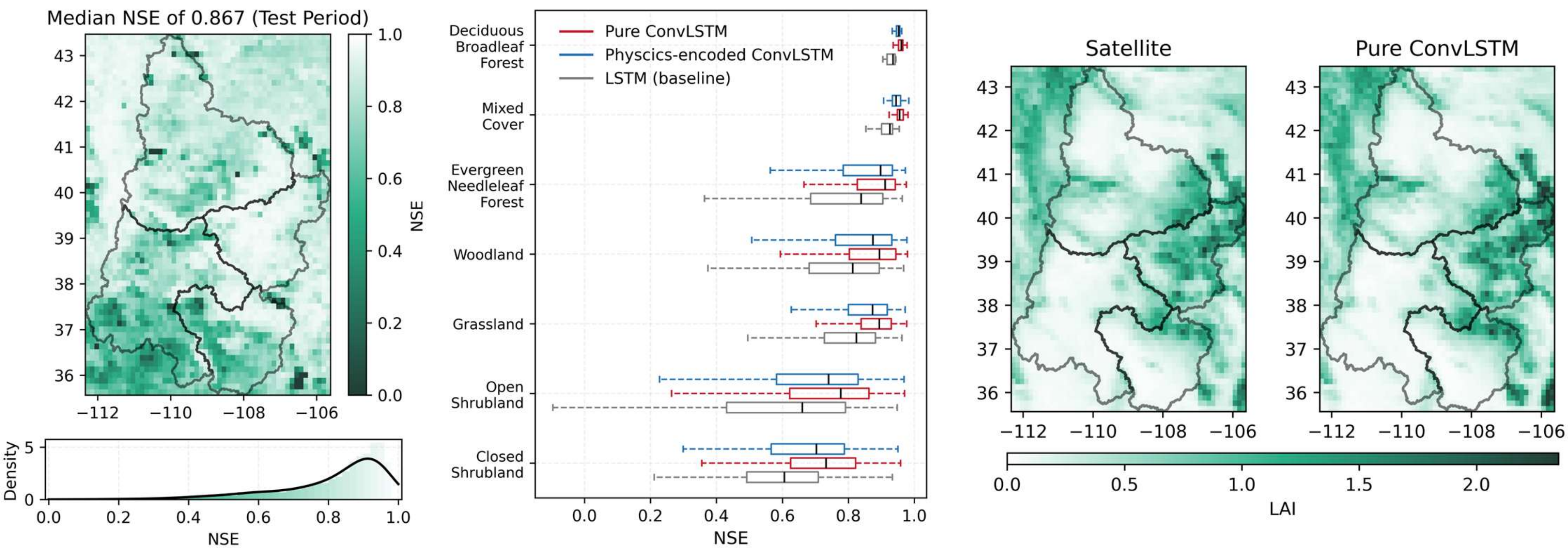
- Two variants are implemented: (1) Pure ConvLSTM using only climate forcings; (2) Physics-Encoded ConvLSTM that integrates process-based LAI simulation as an additional input, acting as a bias-corrector

- Models are trained on 63×54 spatial grids using the AdamW optimizer with early stopping to ensure robust learning and prevent overfitting



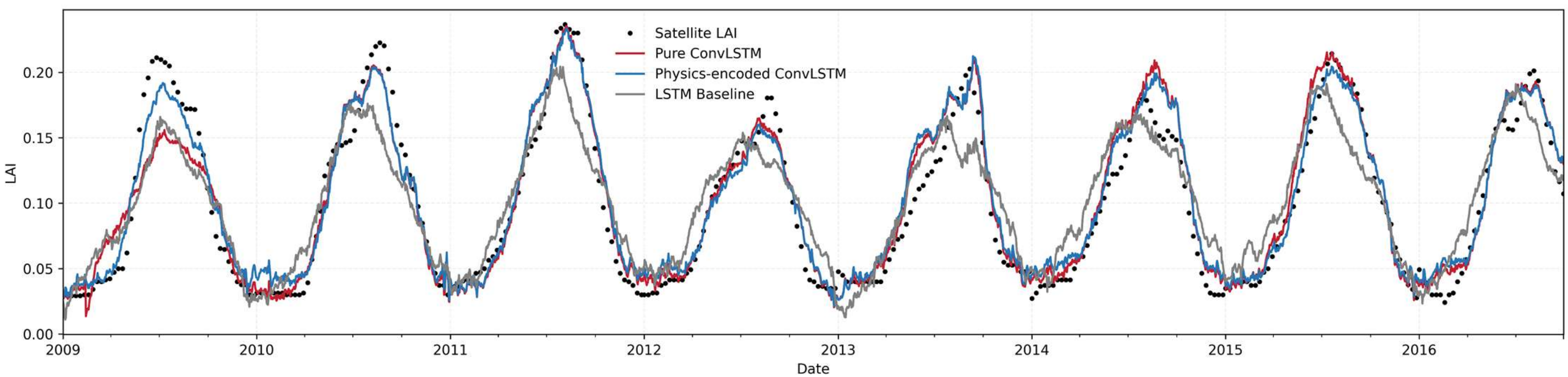
Results

- Pure ConvLSTM achieves the highest accuracy (median NSE: 0.867), followed by physics-encoded ConvLSTM (0.839), both substantially outperforming the LSTM baseline (0.787)



Results

- Pure ConvLSTM excels at capturing both spatial patterns and temporal dynamics across diverse vegetation types, demonstrating its ability to learn complex ecohydrological interactions
- The physics-encoded variant shows enhanced stability in water-limited regions and reduces unrealistic fluctuations, proving valuable for reliable extrapolation in climate projection scenarios



Future Work: Knowledge-guided DL

- This spatially distributed deep learning framework enables more reliable vegetation and water resources projections under climate change
- However, translating domain expertise to guide the design of **interpretable, trustworthy** scientific models still remains an open and nontrivial research challenge