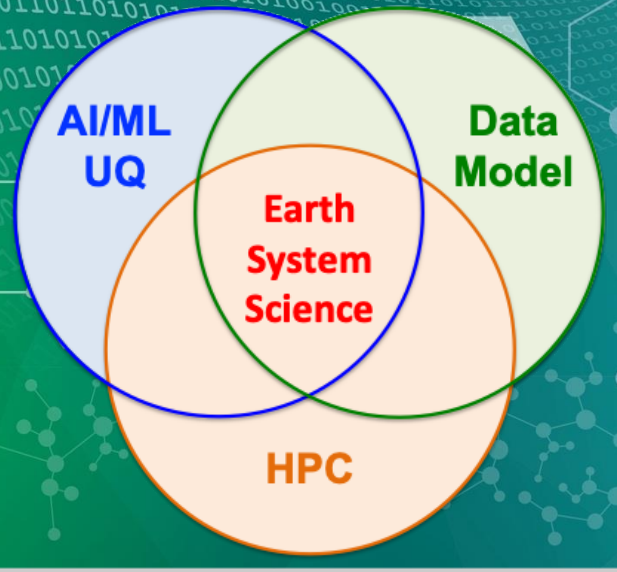
OAK RIDGE National Laboratory

2025 ML4LM Webinar Series

Advancing Land Surface Modeling through Trustworthy AI



Dan Lu (lud1@ornl.gov)

Senior Computational Earth Scientist

Oak Ridge National Laboratory

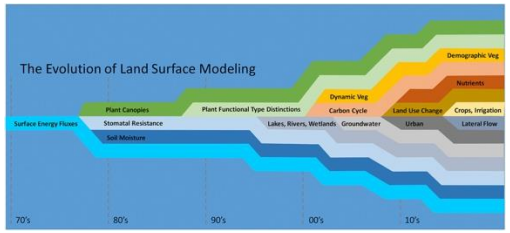
This work represents a collaboration with members of my research team and external collaborators.

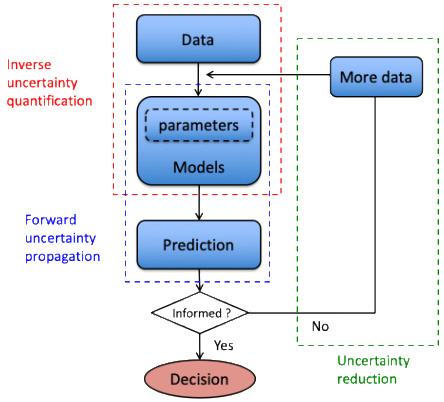
April 3, 2025

U.S. DEPARTMENT OF ENERGY

Advancing land surface modeling through data-model integration

Physical Model





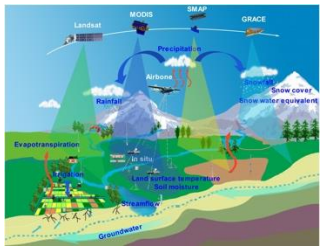
Challenges:

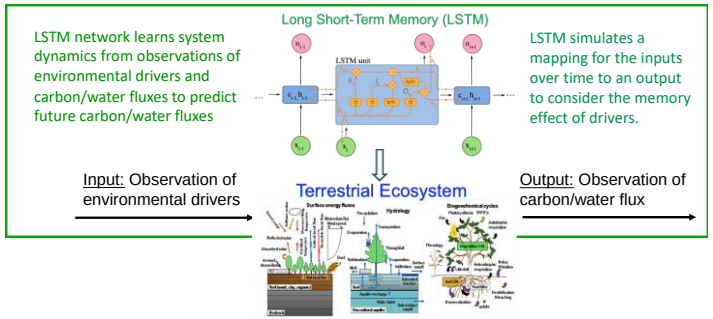
- High computational costs;
- Large parameter uncertainty;

Our study:

- Efficient emulation;
- Generative AI for UQ.

Data-Driven ML Model





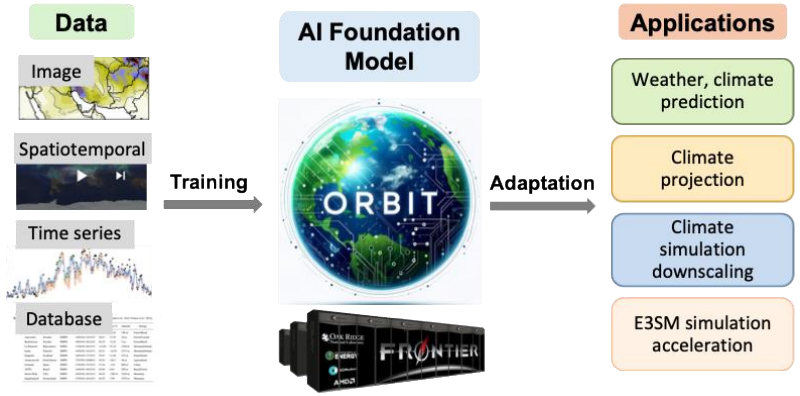
Challenges:

- Generalizing across space and time;
- Explainability, physical consistency;
- Reliability under changing conditions;

Our study:

- Advanced ML integrating diverse data;
- Interpretable AI for explainability;
- UQ to improve predictive reliability.

AI Foundation Model



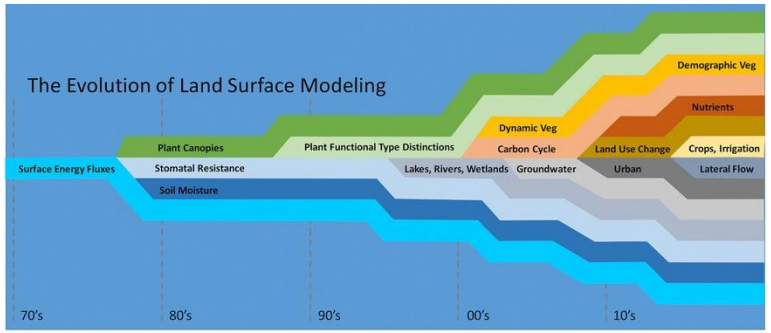
Challenges:

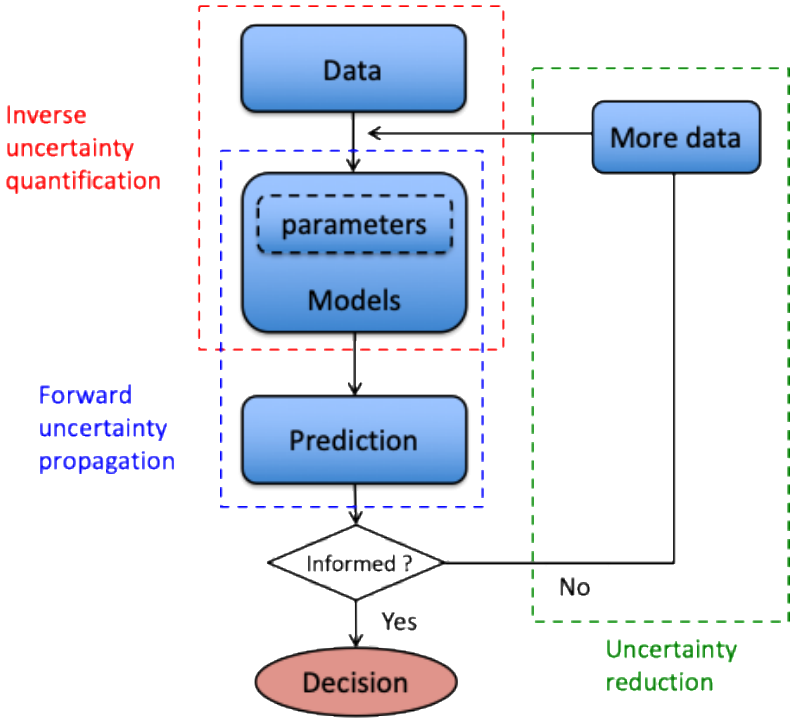
- Heterogeneous, unlabeled data;
- Diverse modeling application needs;

Our study:

- Billion-size AI foundation model trained on CMIP6 climate data;
- Adapted for weather forecasting, climate downscaling, and land model acceleration.

Physics-based land surface modeling needs model calibration





- Physics-based land surface modeling (LSM) requires model calibration and ensemble simulations for UQ.

Emulation

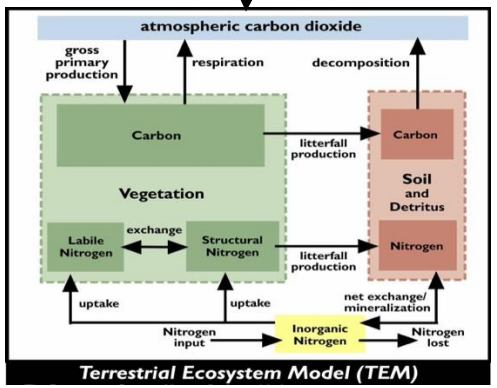
- High computational cost of LSM simulations needs improved efficiency.
- We build a fast LSM emulator from ensemble runs and evaluate it for parameter estimation and uncertainty quantification (UQ) to reduce runtime.

Generative AI

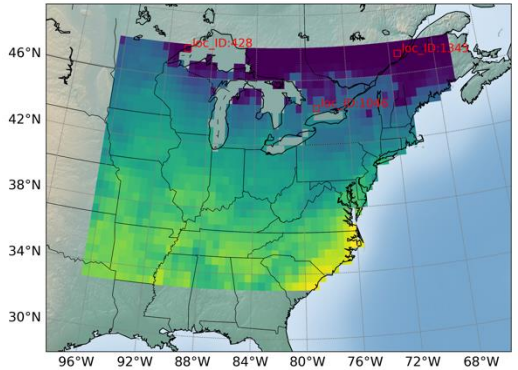
- Land surface heterogeneity requires efficient UQ for rapid, site-specific model calibration at large scales.
- We developed a diffusion model to quickly generate parameter posterior samples, enabling fast, large-scale model calibration.

Emulation to reduce computational costs of LSM

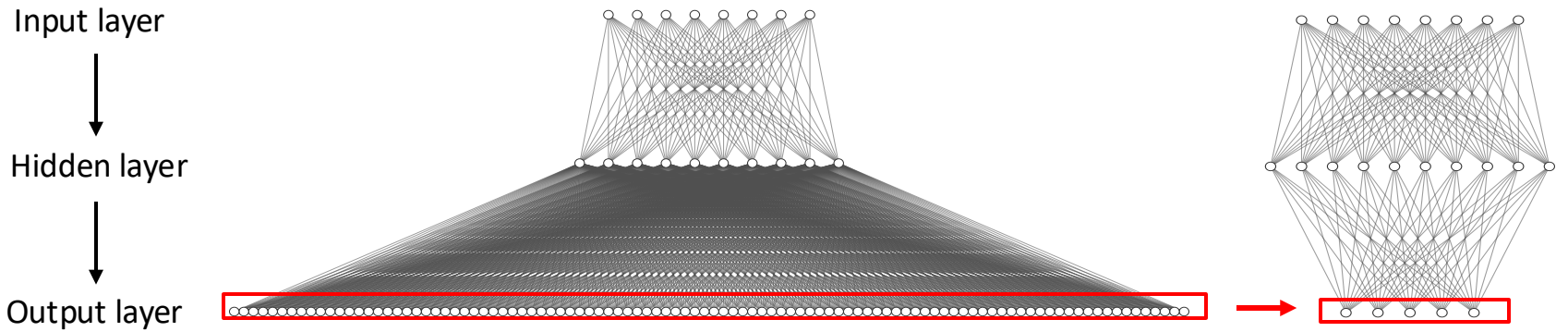
TEM parameters



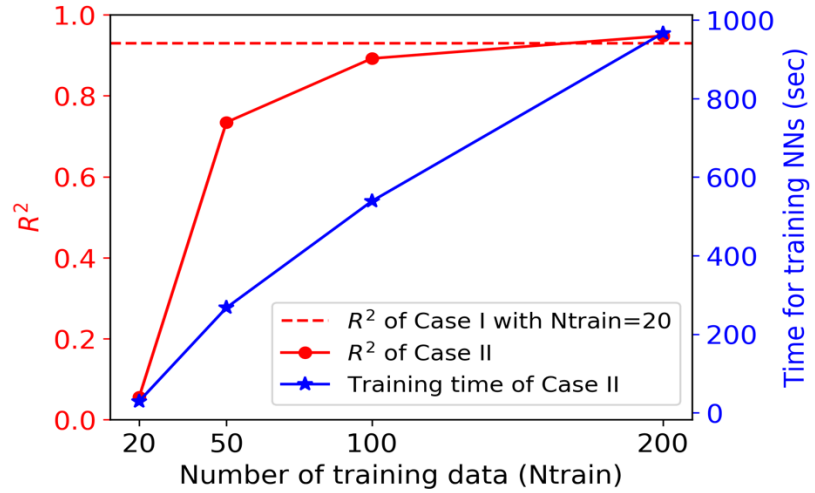
TEM outputs



Large number of spatiotemporal model outputs



Dimension reduction on output layer reduces NN parameters from 10^4 to 10



- Inputs: 8 parameters;
- Outputs: annual GPP in 1422 grid cells for 30 years, 42660 outputs;
- Time: one model run takes 24 hours.

➢ The resulted simple NN enables only 20 training data to produce accurate predictions otherwise 200 data are needed for the similar accuracy.

❖ Dimension reduction enabled an accurate NN-based emulator with fewer required samples.

- Lu, D., and Ricciuto, D., *GMD*, 2019. Lu, D., et. al., *JAMES*, 2018.

Generative AI method (DBUQ) for efficient parameter calibration

- Objective: draws samples to approximate posterior distribution of parameter X given observed y,

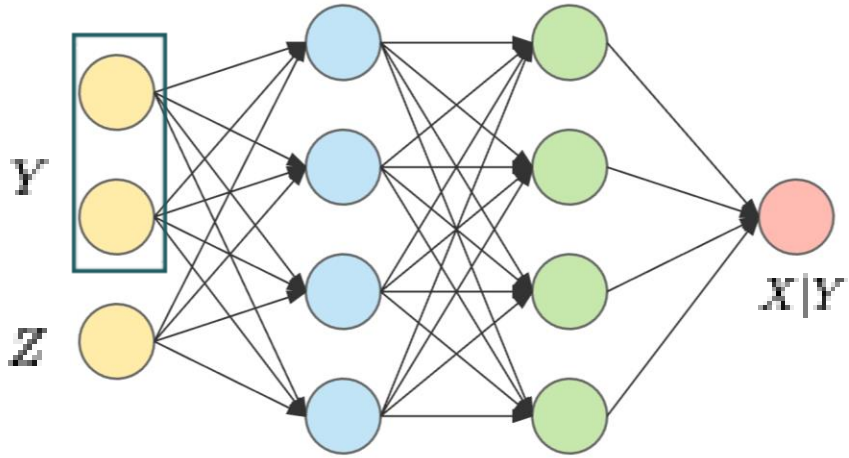
$$p(X|Y=y) \propto p(Y=y|X)p(X)$$

- Our diffusion-based UQ (DBUQ) formulates a generative model *F* to draw the target samples ,

$$X|Y \approx F(Y, Z; \theta)$$

- A neural network (NN) is trained to estimate *F*;
- After training, the NN evaluate Z to quickly generate desired parameter posterior samples

$$X|Y \text{ at } Y = y$$



❖ Use a NN to learn the relationship between [Y, Z] and X|Y;

- X|Y is the parameter of interest;
- Y is the observation variable;
- Z is the standard Gaussian variable.

❖ The generation of target samples of X|Y is computationally and memory efficient;

❖ For any given observational data, the NN generates corresponding parameter posterior samples for UQ without the need for re-training the network.

Apply DBUQ to improve LSM parameter calibration

- Problem: Use DBUQ to estimate 8 land surface model parameters;
- Observation: Annual averaged latent heat flux (LH) for 5 years at the Missouri Ozark AmeriFlux site in 2006-2010;
- Prior sample: 1000 samples from LSM simulation $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^{J}$
- Two case studies:
  - Synthetic case for method verification
  - Real observations application
- Compare DBUQ with MCMC for performance evaluation

Parameter name	Parameter range
rootb_par	[0.5, 4]
slatop	[0.01, 0.05]
flnr	[0.1, 0.4]
frootcn	[25, 60]
froot_leaf	[0.3, 1.5]
br_mr	[1.5e-6, 4e-6]
crit_dayl	[35000, 45000]
crit_onset_gdd	[600, 1000]

DBUQ

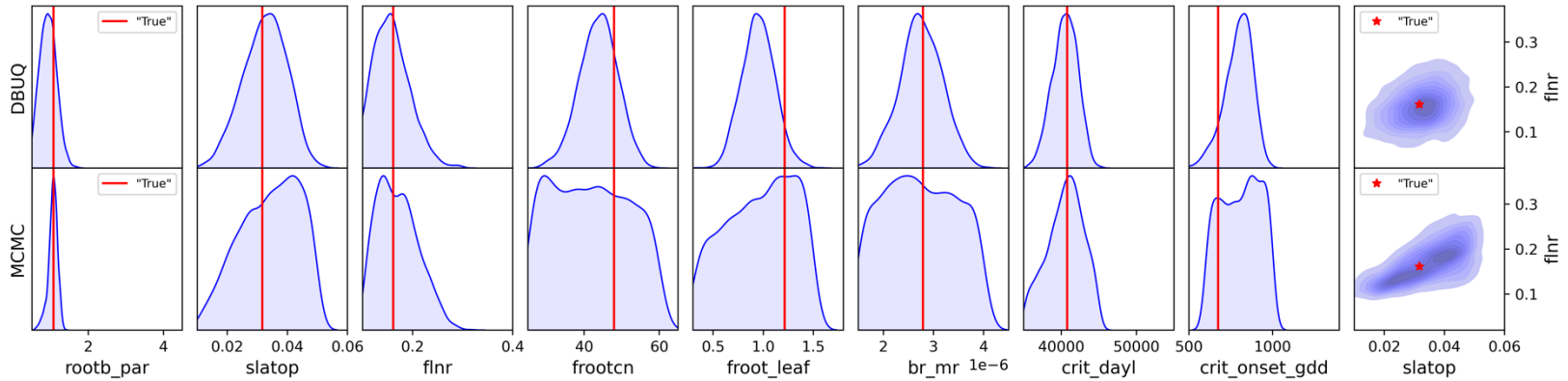
- Input: 1000 LSM samples $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^{J}$
- Output: **a trained generator** which can be quickly evaluated to generate target samples for any given observations;
- Computing time: < 10 min for solving both cases
- **Particularly suitable for site-specific LSM calibration at a global scale** due to its computational efficiency and amortized inference.

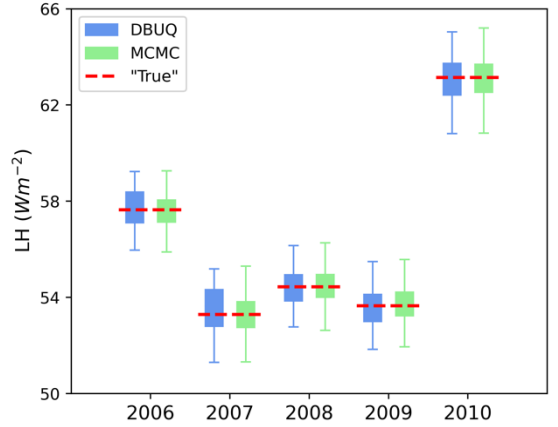
Surrogate + MCMC

- Input: 1000 LSM samples $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^{J}$
- Procedure: build an emulator on the LSM samples, and then perform MCMC simulations on the emulator;
- Output: **a set of posterior samples;** For a different observation, we need to re-run MCMC;
- Computing time: ~ 5 hours for one case to generate the same number of posterior samples as DBUQ.

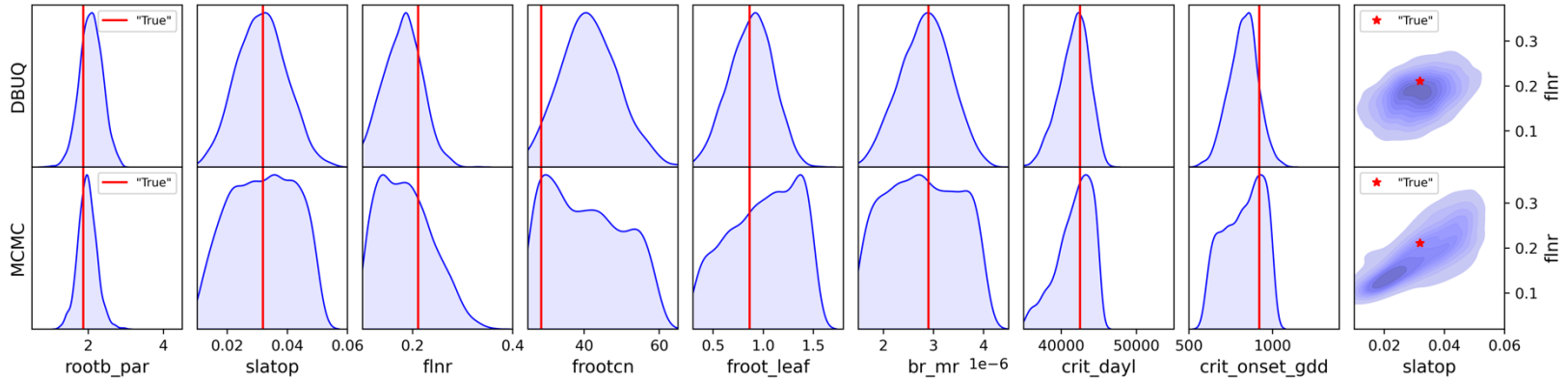
DBUQ accurately and efficiently estimated model parameters

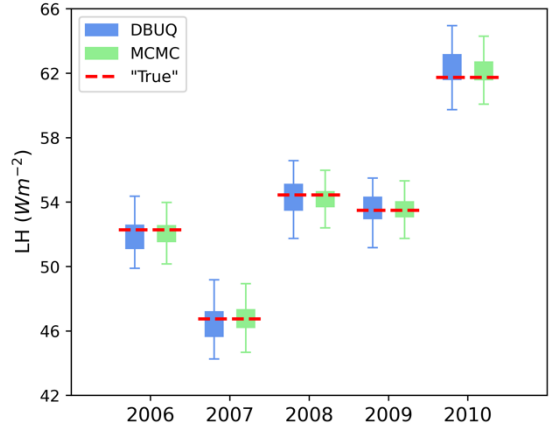
Synthetic case I





Synthetic case II

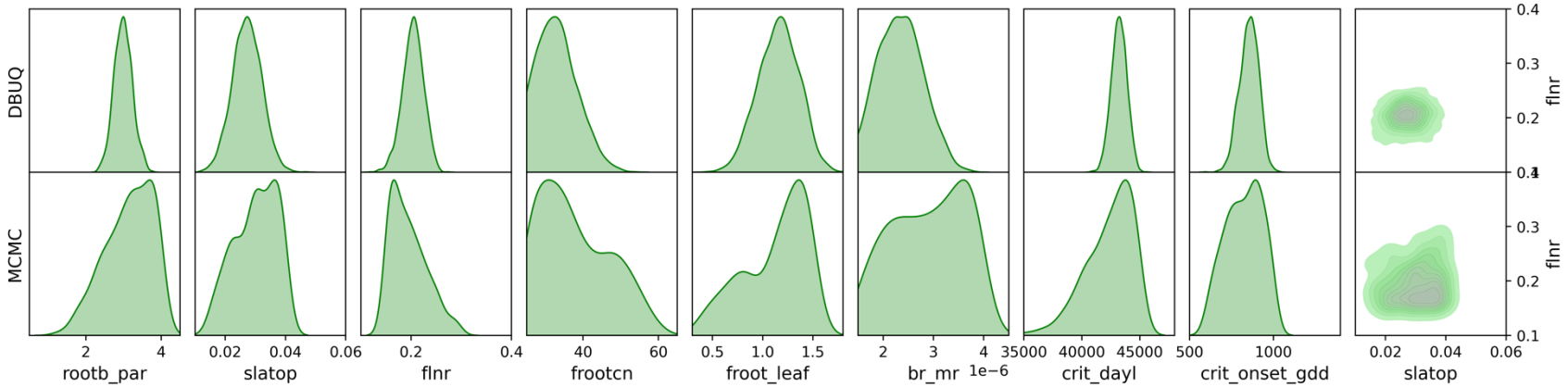


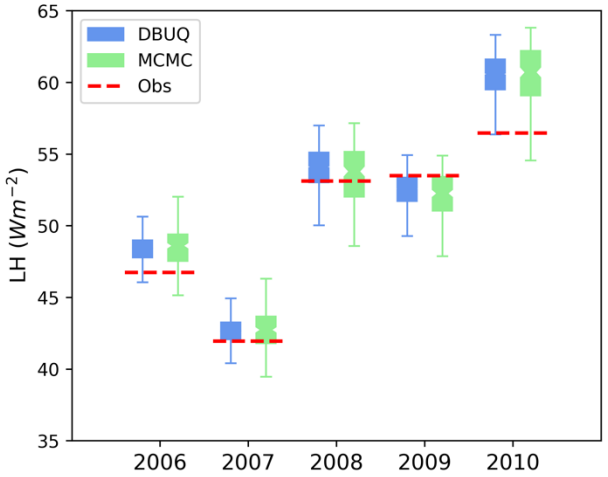


- DBUQ shows high accuracy in approximating the parameter posterior distributions.
- DBUQ demonstrates an accurate model calibration, as the prediction samples simulated from the parameter posterior samples are closely around the “true” observation.

DBUQ accurately and efficiently calibrated the land model

Real observation case

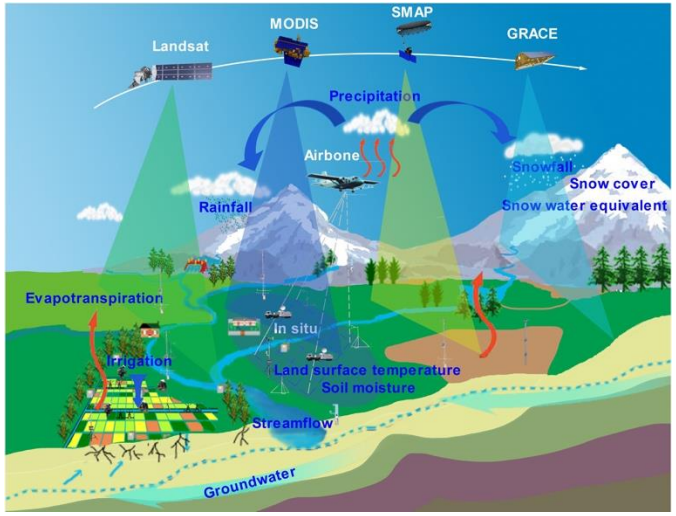




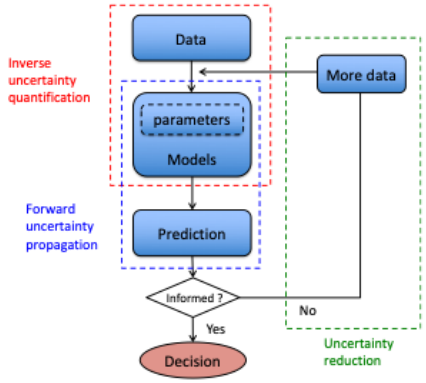
- ❖ DBUQ again shows high accuracy in approximating parameter posterior distributions.
- ❖ It showed accurate calibration, with prediction samples tightly enclosing the observations.
- ❖ DBUQ achieves comparable accuracy with MCMC with significantly less computational time.
  - DBUQ: 10 mins for all the three case studies;
  - MCMC: 5 hours for one case study;

https://github.com/patrickfan/GenAI4UQ

- Lu et. al, *JGR--Machine Learning and Computation*, 2024.



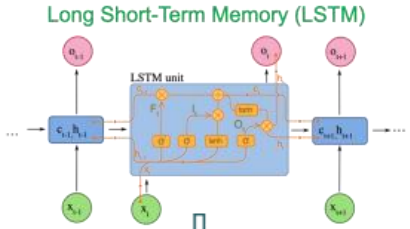
Process-based Model



- Model calibration and UQ are critical to improve prediction.
- Advanced surrogate modeling and generative AI for efficient model calibration and UQ.

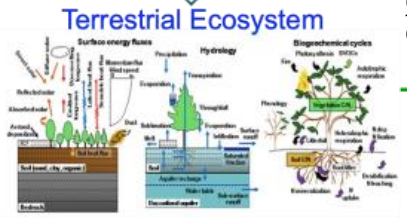
Data-driven Model

LSTM network learns system dynamics from observations of environmental drivers and carbon/water fluxes to predict future carbon/water fluxes



LSTM simulates a mapping for the inputs over time to an output to consider the memory effect of drivers.

Input: Observation of environmental drivers



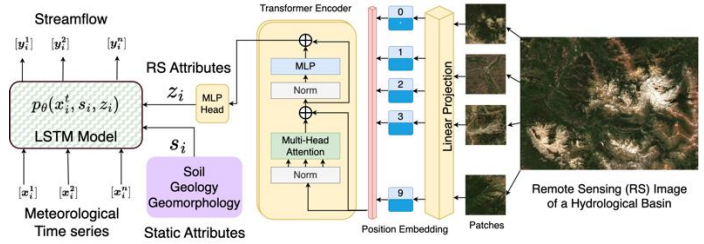
Output: Observation of carbon/water flux

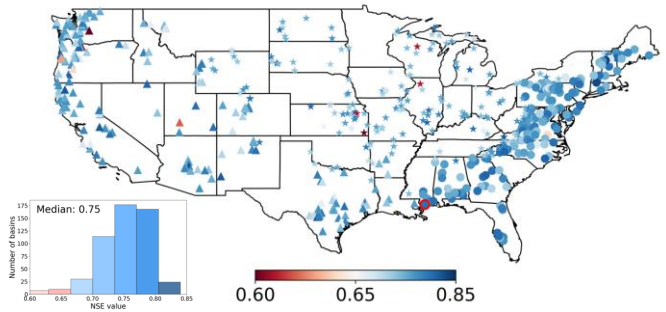
- ML model has challenges in trustworthiness.
  - How can we ensure that ML solutions generalize across space and time?
  - How do we verify that models are making good predictions for the right reasons?
  - How can we guarantee prediction reliability under changing environmental conditions?

Advanced, explainable, reliable ML for land surface modeling

Advanced ML

- Integrate diverse data from satellite and sensor networks
- Develop advanced model architectures



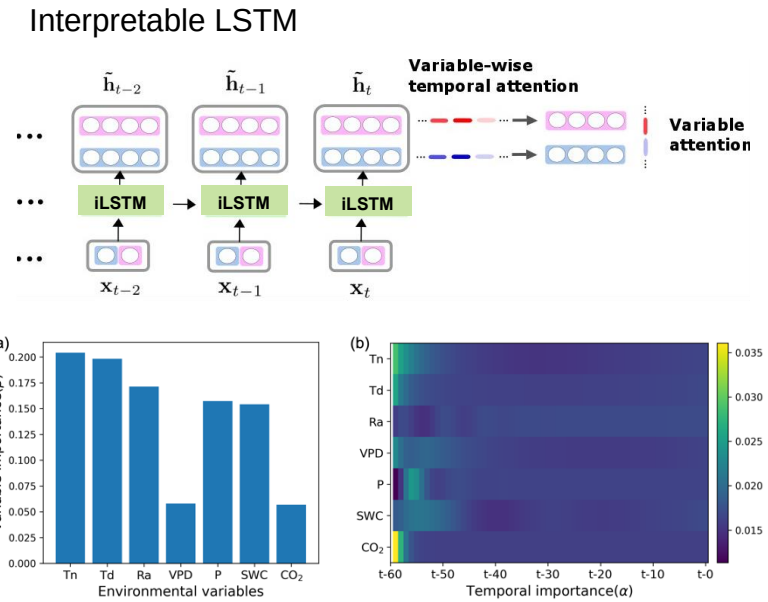


NSE value: = 1.0 means perfect fit; > 0.65 suggests good prediction

❖ Leverage diverse data and advanced ML models to improve accuracy and generalizability.

Explainable ML

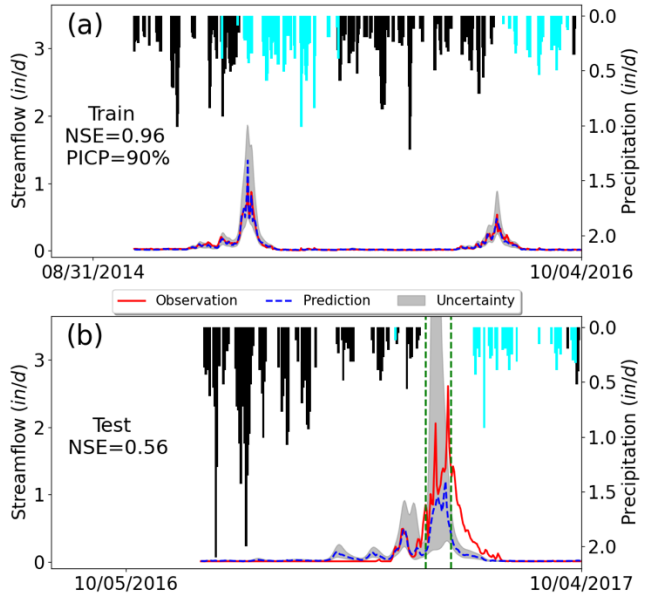
- Permutation analysis: SHAP
- Gradient-based method: IG
- Interpretable LSTM network
- Attention maps of transformer model



❖ Validate model decisions ensuring physical consistency; identify key drivers for prediction.

Reliable ML

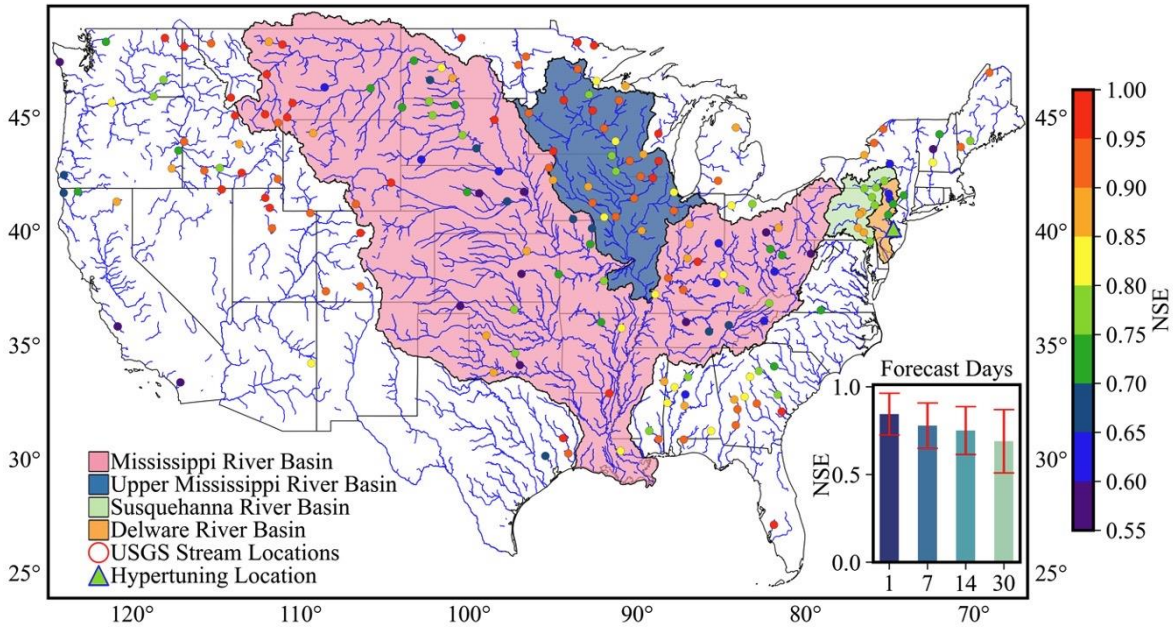
- Bayesian neural networks
- Gaussian processes
- Ensemble-based methods
- Prediction interval methods



❖ Quantify prediction uncertainty to evaluate & ensure reliability under changing conditions.

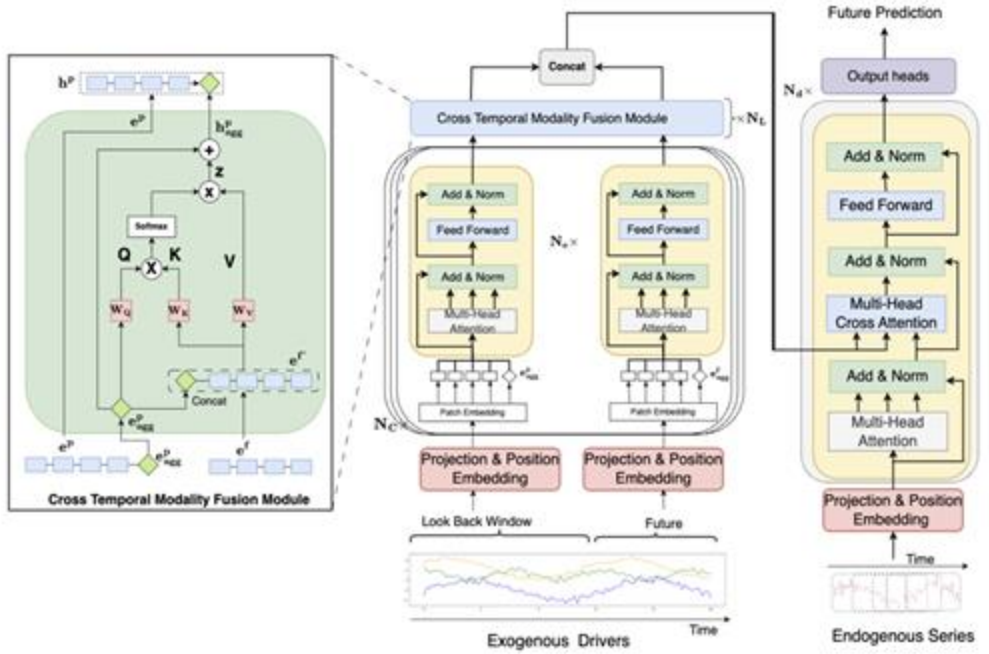
Transformer model to improve long-term streamflow prediction

- Problem: Predict daily streamflow 30 days ahead;
- Data: Past weather observation (Daymet), future weather forecast (ECMWF), and past streamflow;



❖ TST model achieved high accuracy and reliability in 30-day streamflow forecasts.

Our Temporal Sequence Transformer (TST) model

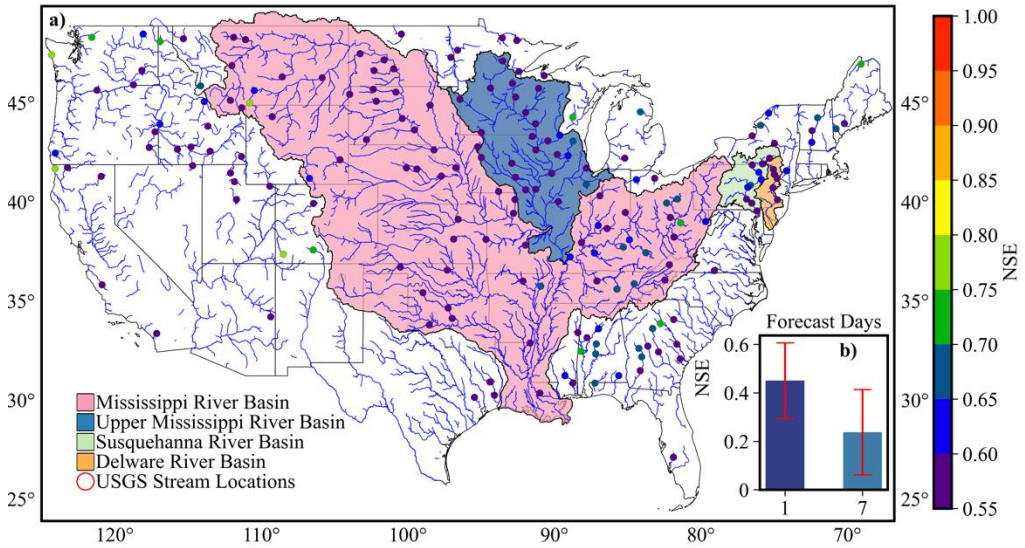


- Two encoders process past and future drivers' data, and one decode handles past streamflow data;
- A cross-temporal fusion module denoises and integrates encoder data, while the decoder uses cross-attention to combine this with past flow data for future predictions.

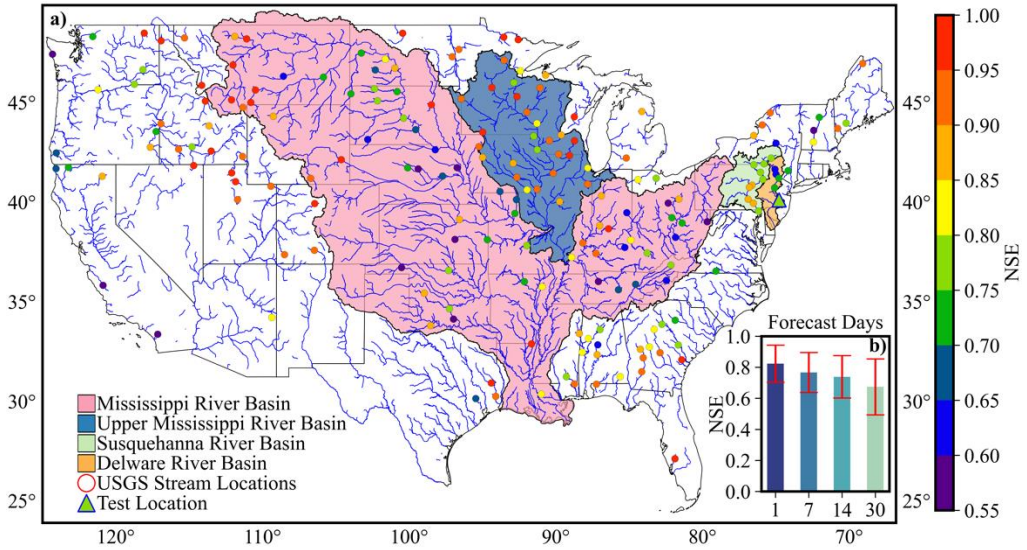
- Tayal, Renganathan, and Lu, *ICML*, 2024. Ambika, Tayal, and Lu, *GRL*, under review.

Transformer model achieved higher accuracy in long-term prediction

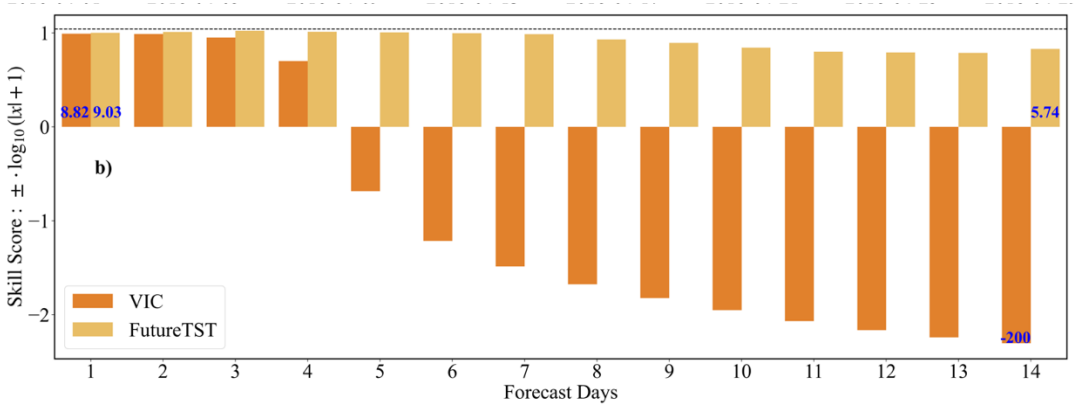
Standard LSTM model

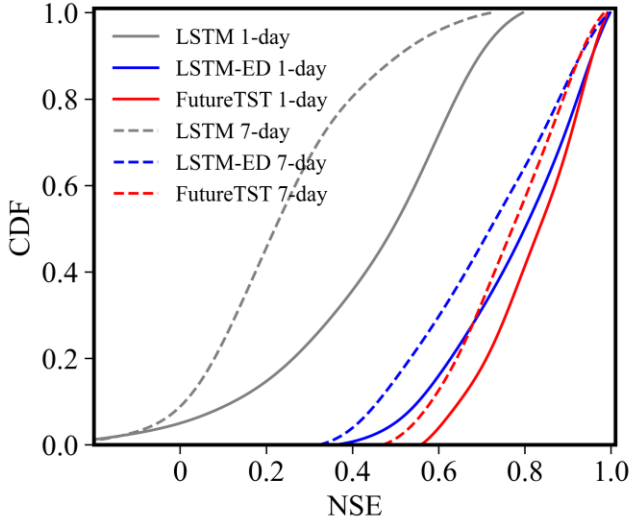


Our Transformer model



- TST model achieved higher accuracy than VIC model





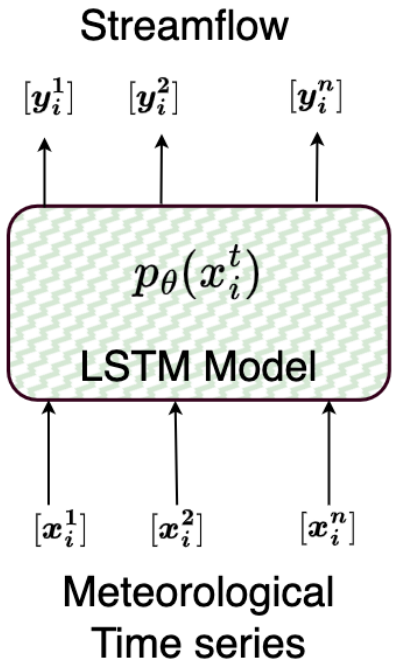
Summarized NSE over 190 gauges across US for 1- and 7-day forecast among

- LSTM
- LSTM-ED
- Our FutureTST

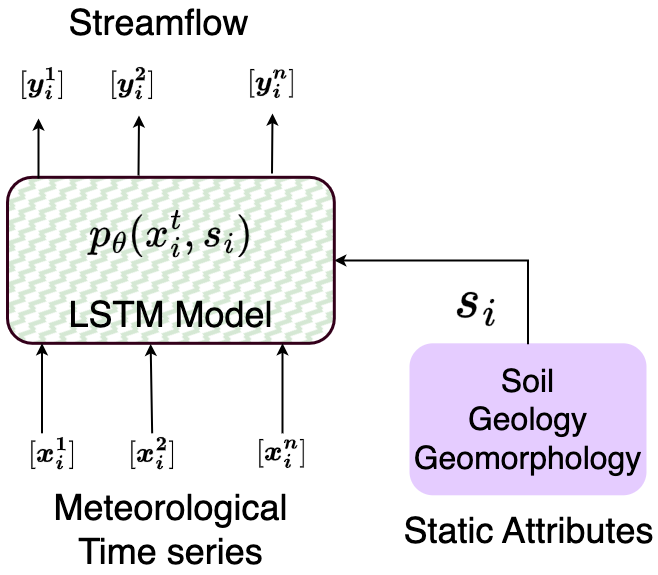
❖ TST model performed better than LSTM and VIC models, demonstrating robustness against noisy weather forecasts.

Advanced ML models to improve spatial generalizability

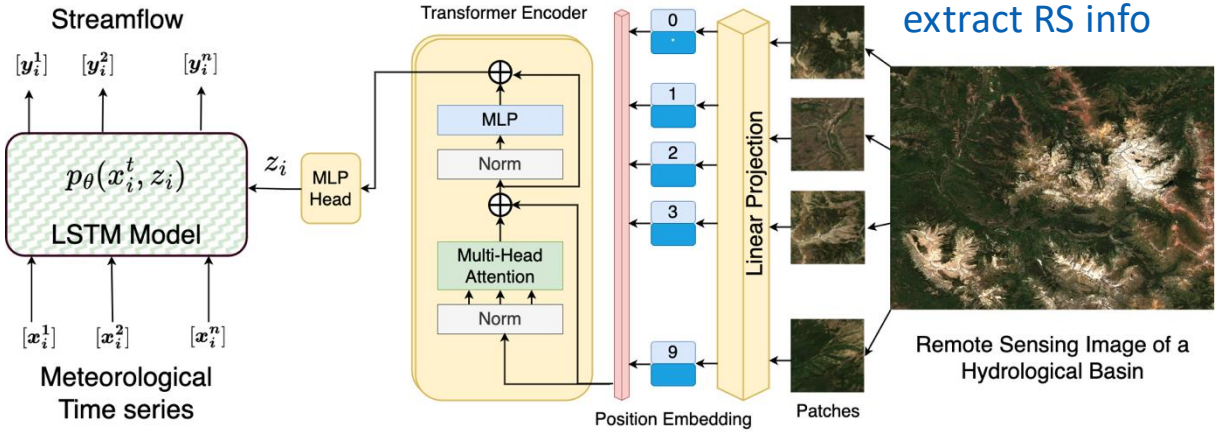
LSTM



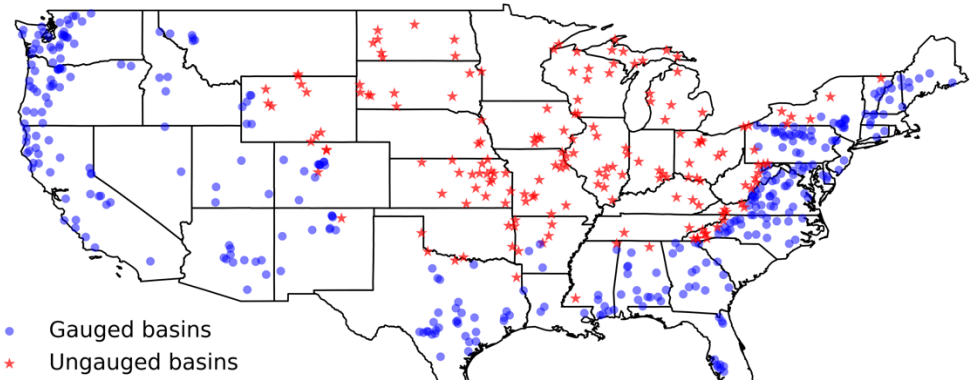
Geo-LSTM



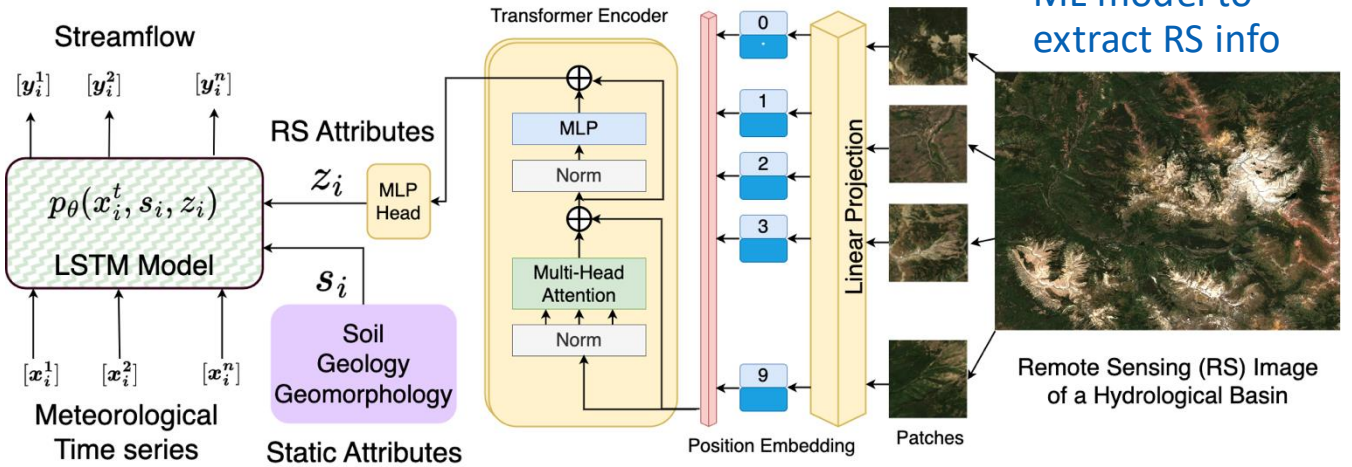
RS-LSTM



ML model to extract RS info



Geo-RS-LSTM

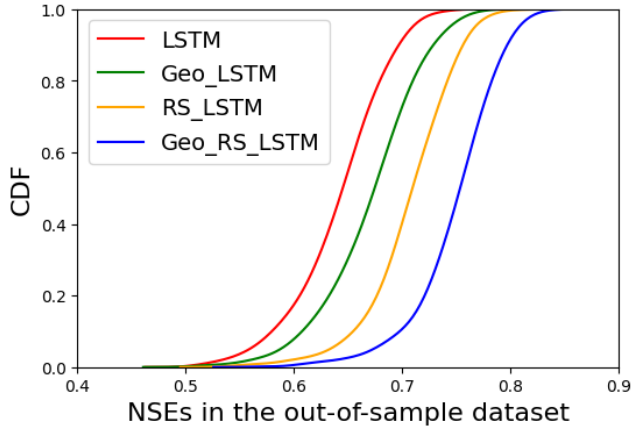


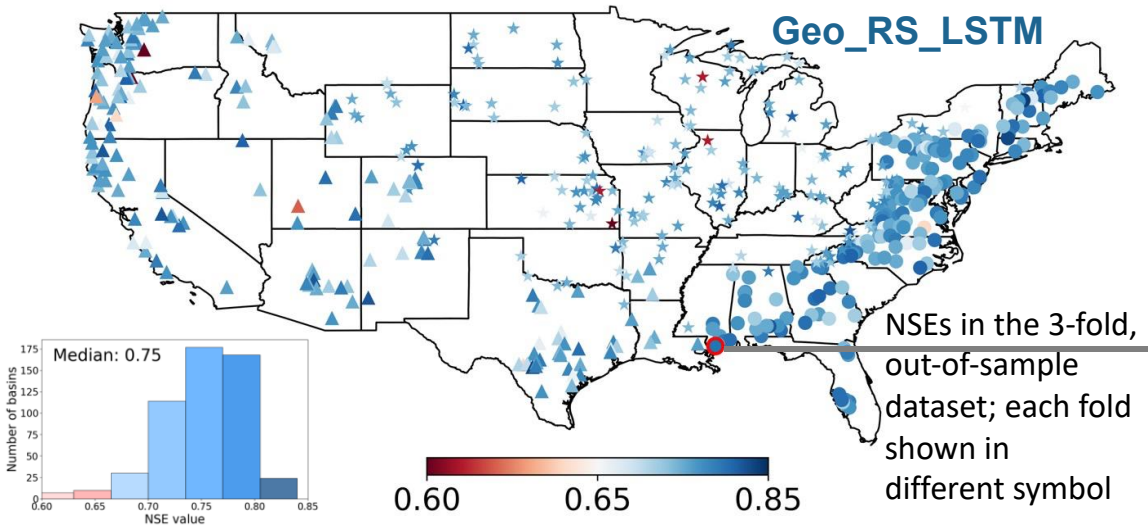
ML model to extract RS info

ML models integrating diverse data enhanced generalization

- Problem: Predict streamflow across the CONUS;
- Data: 35 years of CAMELS dataset of 531 basins and Sentinel-2 satellite images;
- Model: 4 ML models with diverse inputs;
- Evaluate: Model performance in spatiotemporal out-of-sample prediction using the NSE metric. Perform 3-fold cross-validation.

	1980-2007	2008-2014
354 basins	Training	
177 basins		Evaluation

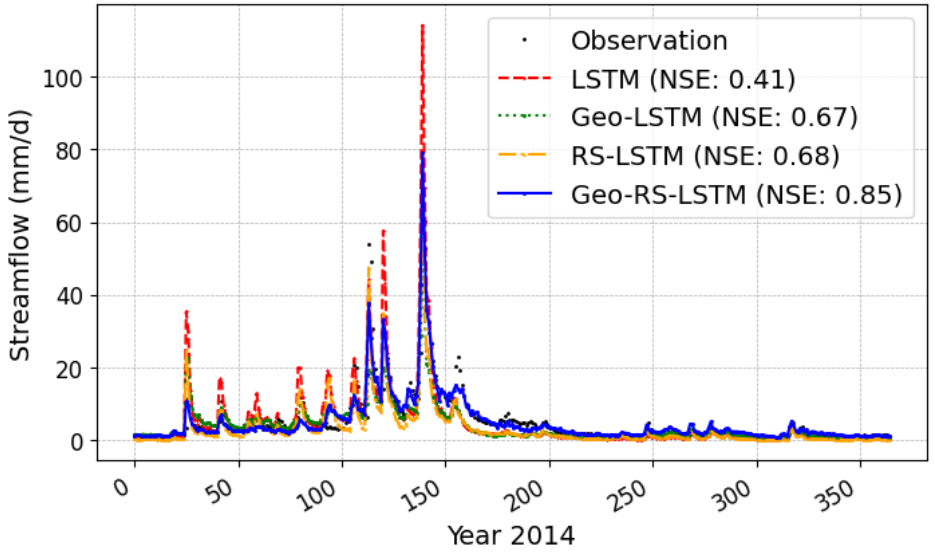




NSEs in the 3-fold, out-of-sample dataset; each fold shown in different symbol

NSE value: = 1.0 means perfect fit; > 0.65 suggests good prediction

- Geo_RS_LSTM model performs the best.



❖ The ML model, integrating diverse data from gauged basins and satellite images, excels in predicting streamflow at 'ungauged' basins under 'future' meteorological conditions.

- Tayal, Renganathan, and Lu, *ERL*, 2024.

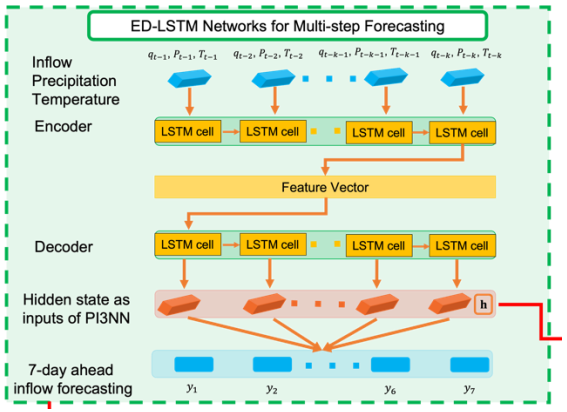
Explainable ML improved our predictive understanding

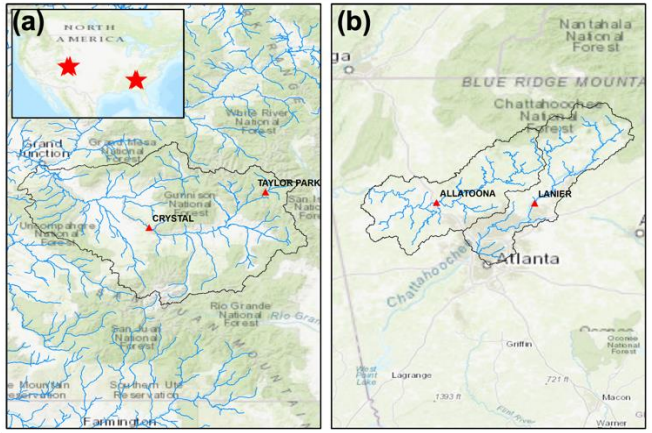
SHAP (SHapley Additive exPlanations)

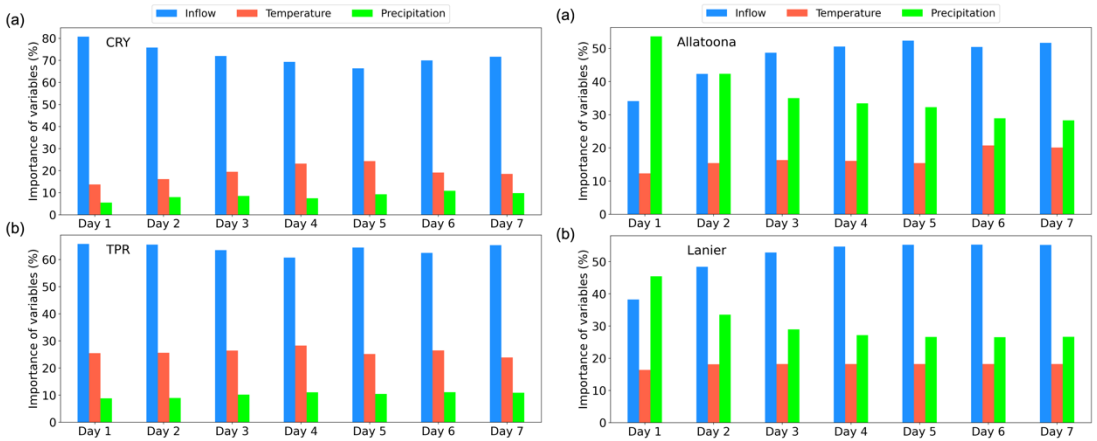
- Explains the contributions of hydrological drivers to response predictions;
- Model agnostic, flexible, and widely used;
- Computationally expensive ($\mathcal{O}(2^N)$) and unable to separate individual and interactive contributions.

IG (Integrated Gradients)

- Calculates input importance by integrating output gradients w.r.t. input along a baseline path;
- Computationally efficient; captures both individual and interactive input contributions;
- Improves understanding of multi-driver mutual impacts.







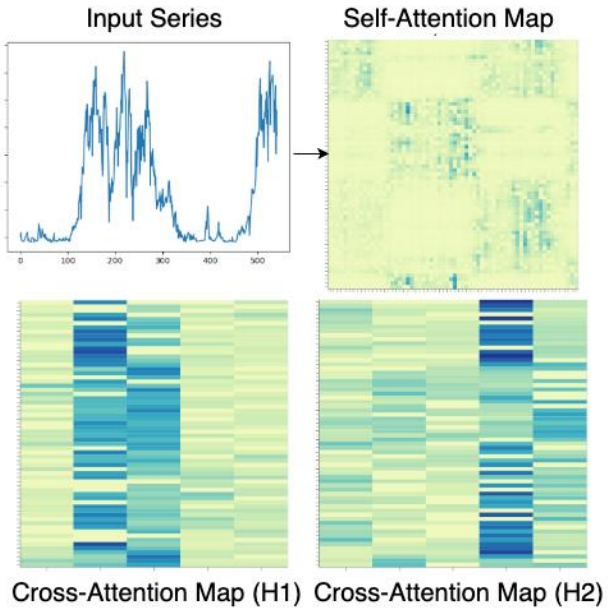
❖ SHAP and IG methods identified key drivers of reservoir inflow forecasts, improving understanding, validating predictions, and supporting hydropower operations.

- Fan, Zhang, Liu, Yang, and Lu, *Frontiers in Water*, 2023. Fan, Liu, and Lu, *JH,* 2023. Fan, Liu, Lu, Gangrade, and Kao, *EMS*, 2023.

Interpretable ML can guide process-based LSM development

Transformer-based model

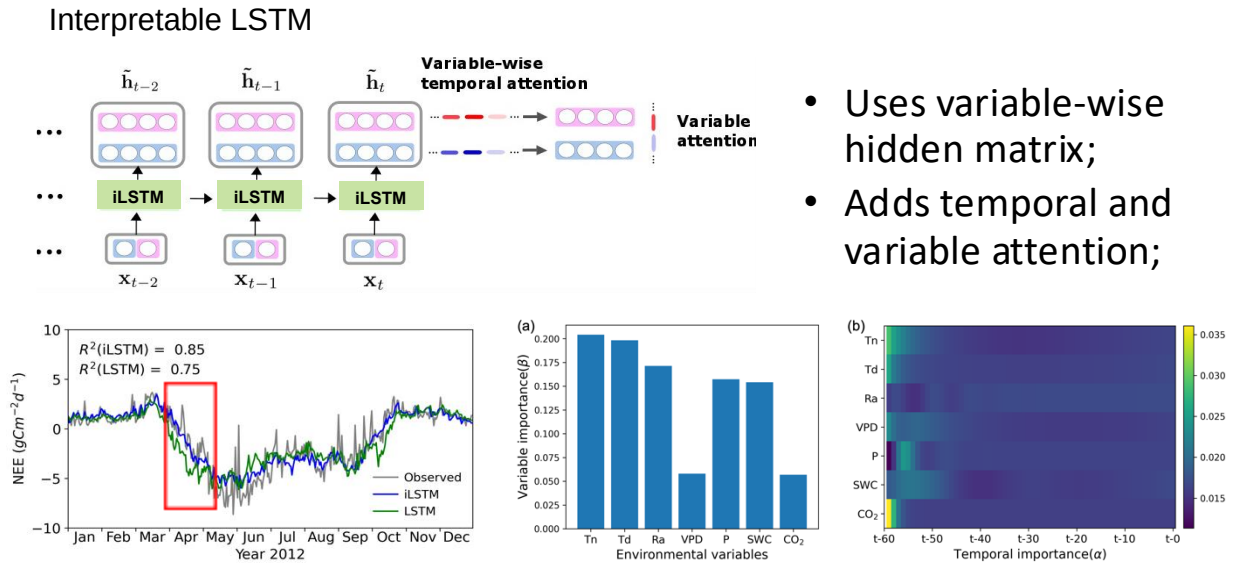
- Visualize Transformer model's learning process to improve prediction understanding.



- Self-attention identifies temporal pattern of each driver;
- Cross-attention captures relationships among drivers.

Interpretable LSTM (iLSTM)

- iLSTM explains variable and temporal importance through its advanced model architecture.



- Uses variable-wise hidden matrix;
- Adds temporal and variable attention;

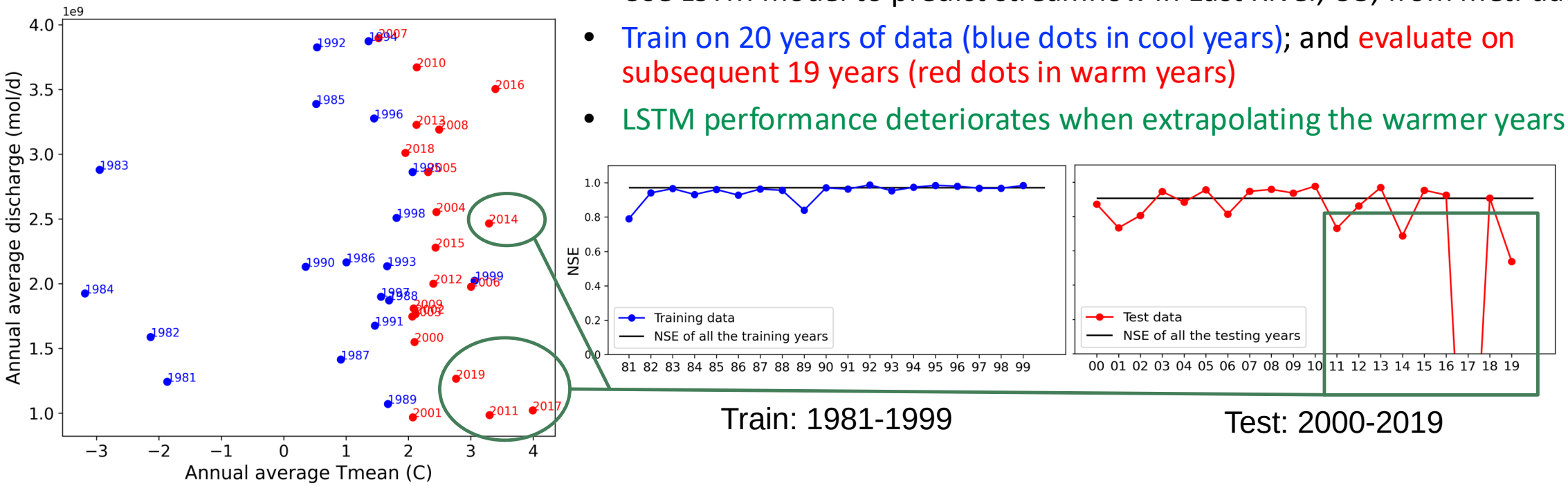
- iLSTM achieved more accurate prediction;
- iLSTM revealed new variable relationships and their temporal importance.

❖ Advanced interpretable ML models enhanced prediction accuracy, revealed learning processes, and provided insights to inform process-based model development.

- Lu, Ricciuto, and Liu, *ICLR*, 2022. Tayal, Renganathan, and Lu, *ICML*, 2024.

ML model needs UQ for trustworthy prediction under climate change

- ML model typically perform well under conditions similar to those they have been trained on but struggle with new, unseen conditions.
- Identifying the reliability of ML predictions is crucial for their effective use.
- UQ helps address the challenge of assessing ML model reliability in climate projection.

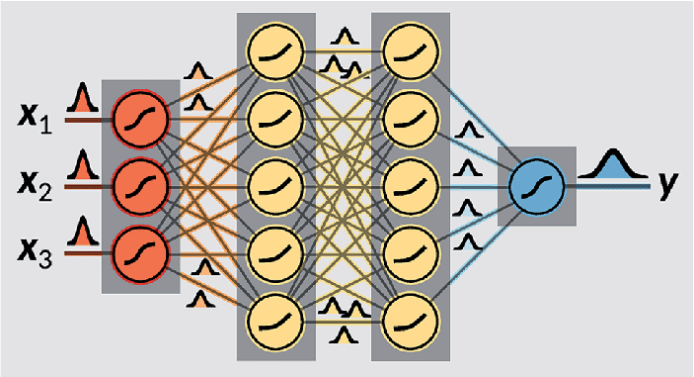


- Use LSTM model to predict streamflow in East River, CO, from met. data.
- Train on 20 years of data (blue dots in cool years); and evaluate on subsequent 19 years (red dots in warm years)
- LSTM performance deteriorates when extrapolating the warmer years.

- Topp, S., Barclay, J., Diaz, J., Sun, A., Jia, X., Lu, D., Sadler, J., and Appling A., *WRR*, 2022.

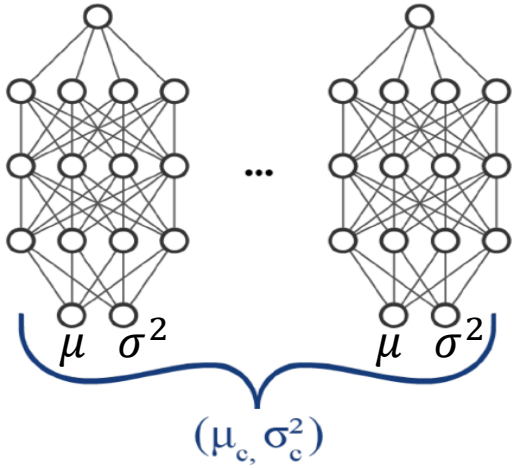
State-of-the-art UQ methods have limitations for scientific ML

Bayesian Neural Networks



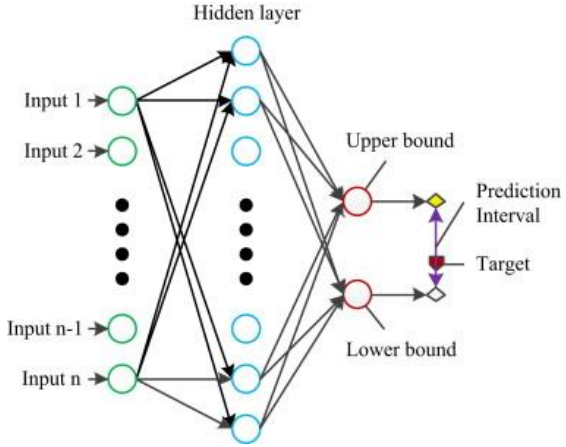
- Pros:
  - Full distribution to quantify predictive uncertainty;
- Cons:
  - Sensitive to the choice of prior distribution;
  - Overconfident results;
  - Slow to train;
  - Difficult to scale.

Deep Ensembles



- Pros:
  - Simple to implement;
  - Easy to scale;
- Cons:
  - Gaussian assumption;
  - Cost in computing and memory increases linearly with #NN in the ensemble.

Prediction Interval (PI)

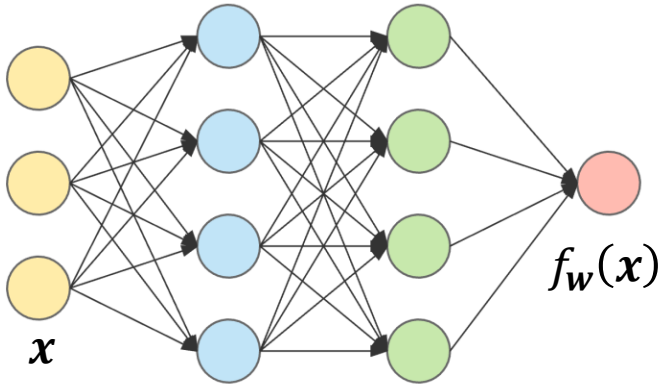


- Pros:
  - Understandable Uncertainty;
  - No distributional assumption;
- Cons:
  - No point estimates;
  - Unstable training and unreliable performance;
  - Overconfident on out-of-distribution (OOD) samples.

Our UQ method, PI3NN, for trustworthy and reliable ML prediction

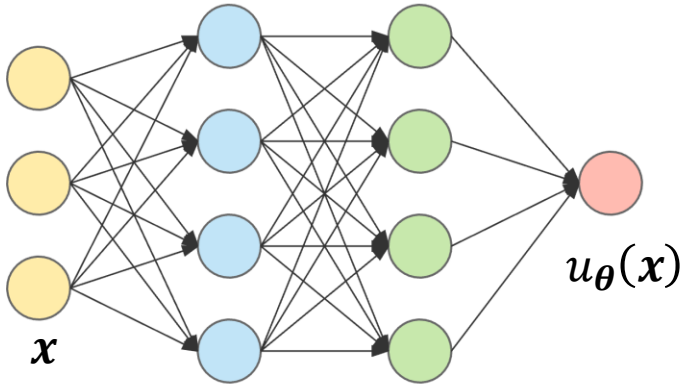
- We developed a prediction interval method from three NNs to quantify prediction uncertainty.

Step 1: Train NN $f_{\boldsymbol{w}}(\boldsymbol{x})$ to estimate y



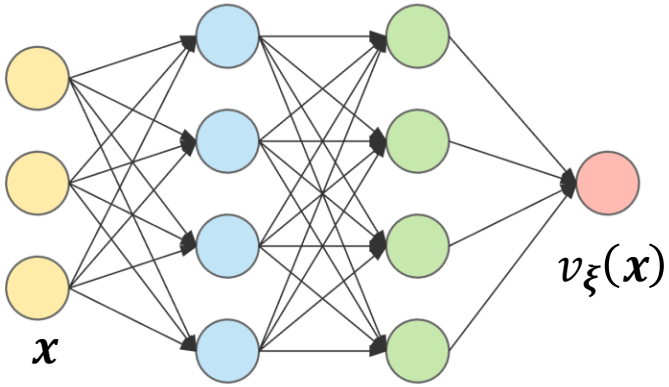
Training data: $\mathcal{D}_{train} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N}$

Step 2: Train NN $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ to learn upper bound of the interval



$\mathcal{D}_{upper} = \{(\boldsymbol{x}_i, y_i - f_{\boldsymbol{w}}(\boldsymbol{x}_i)) | y_i \geq f_{\boldsymbol{w}}(\boldsymbol{x}_i)\}$

Step 3: Train NN $v_{\boldsymbol{\xi}}(\boldsymbol{x})$ to learn lower bound of the interval



$\mathcal{D}_{lower} = \{(\boldsymbol{x}_i, f_{\boldsymbol{w}}(\boldsymbol{x}_i) - y_i) | y_i < f_{\boldsymbol{w}}(\boldsymbol{x}_i)\}$

Step 4: For a given confidence level, calculate the PI $[L(\boldsymbol{x}), U(\boldsymbol{x})]$ via root-finding to determine α and β

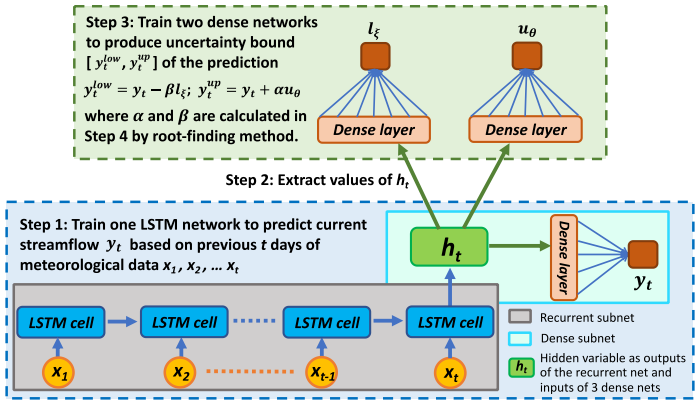
$$L(\boldsymbol{x}) = f_{\boldsymbol{w}}(\boldsymbol{x}) - \beta v_{\boldsymbol{\xi}}(\boldsymbol{x}) \qquad U(\boldsymbol{x}) = f_{\boldsymbol{w}}(\boldsymbol{x}) + \alpha u_{\boldsymbol{\theta}}(\boldsymbol{x})$$

❖ PI3NN produces accurate and reliable uncertainty bounds that precisely enclose a specified portion of data with a narrow interval width.

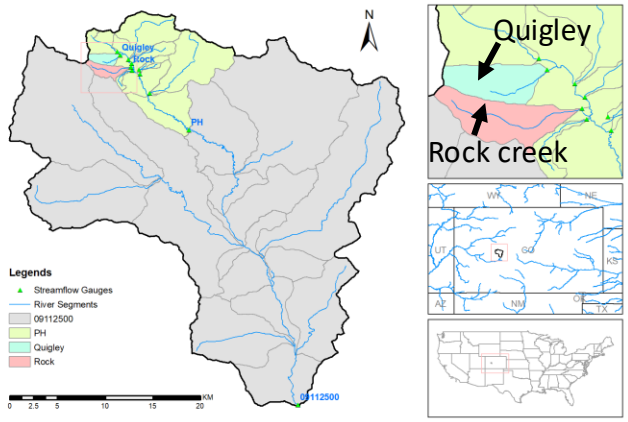
• Liu, Zhang, Lu, and Zhang, *ICLR, 2022*.

UQ ensures reliable streamflow prediction under changing conditions

Our UQ method produces prediction and its uncertainty using three NNs.

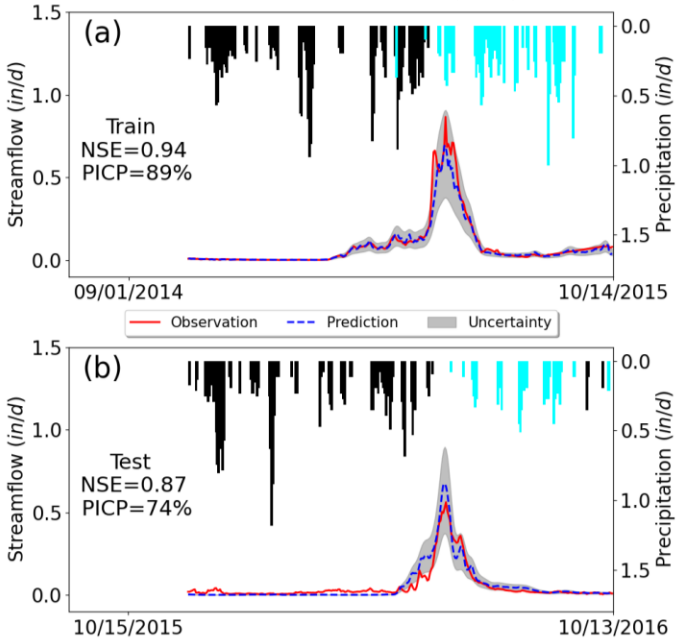


East River Watershed, CO



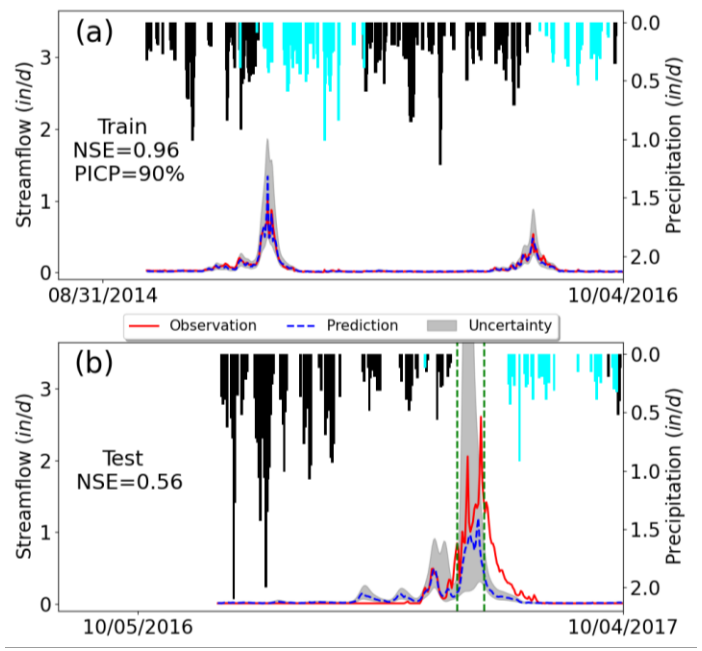
- Input: precip, max and min air T
- Output: daily streamflow
- Model: LSTM network
- UQ: calculate 90% prediction interval

Catchment Quigley



- In Quigley where test and training conditions are similar, LSTM accurately predicts the streamflow.
- Our UQ method accurately quantifies prediction uncertainty consistent with the confidence level.

Catchment Rock creek



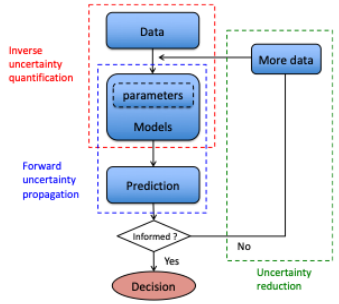
- In Rock Creek, LSTM cannot predict the test data well due to data shift and new conditions.
- Our UQ method detects this shift by producing a wider uncertainty consistent with larger errors.

❖ Our error-consistent UQ method prevents overconfidence and ensures reliable predictions under changing conditions.

- Lu et al., *JHM*, 2022; Liu, Lu, Painter, Griffiths, and Pierce, *Frontiers in Water, 2023*.

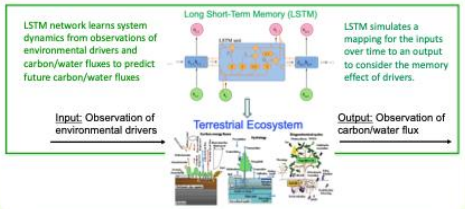
From physics-based to data-driven, now to AI foundation models

Process-based Model



- Model calibration and UQ are critical to improve prediction.
- Advanced surrogate modeling and generative AI for efficient model calibration and UQ.

Data-driven Model



- Trustworthy prediction requires generalizable, explainable, and reliable ML models.
- Advanced ML integrating diverse data to improve generalizability;
- Explainable AI to ensure physical consistency;
- UQ to support reliability under changing conditions.

Data

Training

AI Foundation Model

Adaptation

Applications

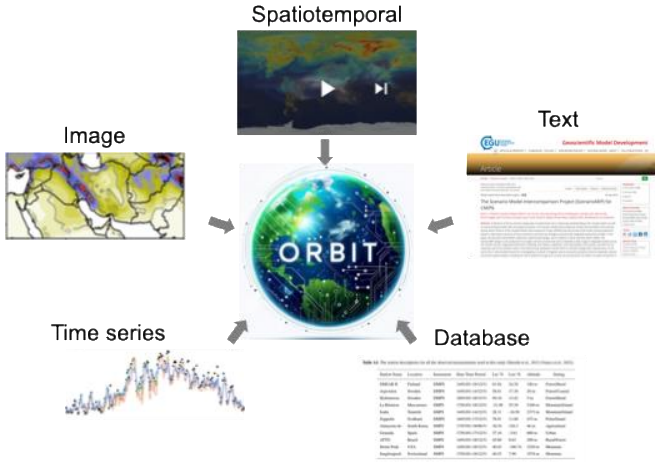
- Weather, climate prediction
- Climate projection
- Climate simulation downscaling
- E3SM simulation acceleration

❖ An AI foundation model is a large-scale neural network trained on extensive, diverse datasets and adaptable to a variety of modeling tasks.

AI foundation model can advance Earth system modeling

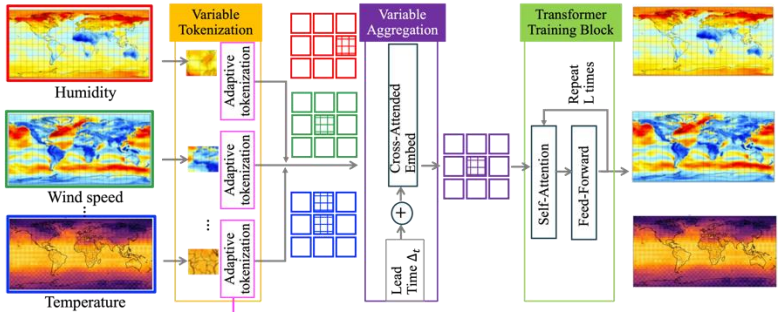
Heterogeneous Data

- Observations from lab, field, and satellite
- Model simulation data
- Data have multiple types, scales, and resolutions.
- These heterogeneous data cannot be fully integrated by numerical models and task-specific ML models.



Scalable Model

- Vision Transformer model
- Integrate heterogenous data
- Scale with data size and resolution

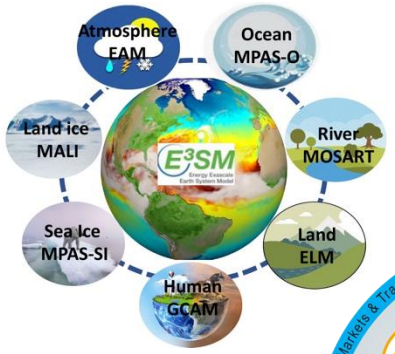


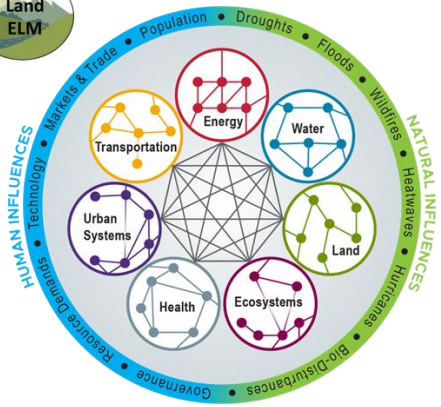
Foundation model:

- Integrate rich, multimodal data
- Reduce reliance on labeled data
- Improve accuracy, efficiency, and generalization
- Ensure high versatility

Various Applications

- Earth system is a coupled system.
- Its simulation advances various scientific applications and impacts multiple sectors.
- Foundation models can save effort, cost, and energy.





ORBIT: our AI foundation model for Earth system modeling

Pre-train on CMIP6 simulation dataset

- Simulation data from 10 CMIP6 models;
- Each model provides 65 to 100 years of data at 6h interval;
- Consider 91 variables with spatial-res of 128*256;
- 1.2 million data point and 223.6 billion tokens.

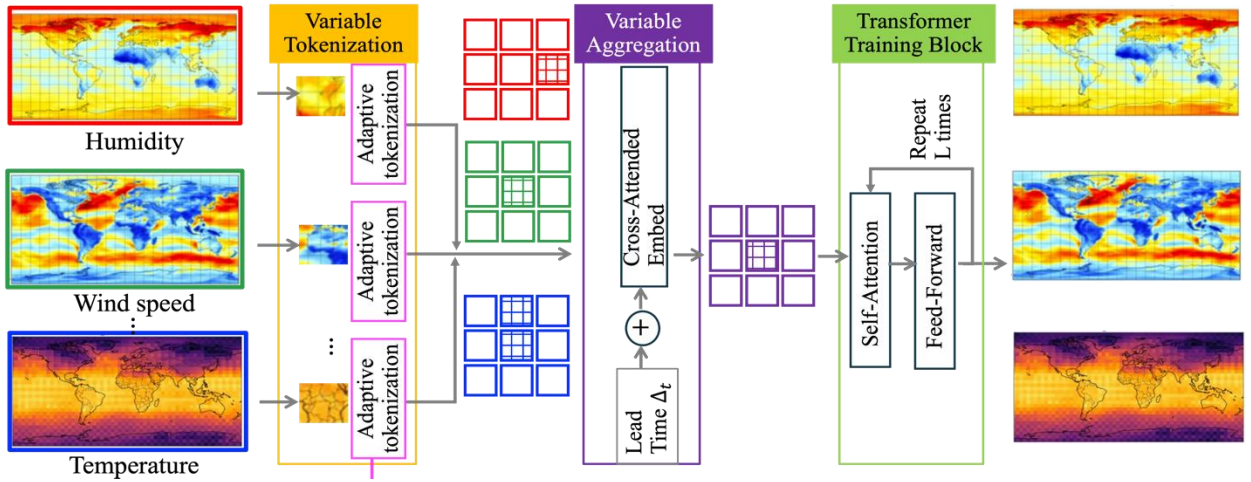
CMIP
The World Climate Research Programme's Coupled Model Intercomparison Project

ESGF
Earth System Grid Federation

PMP
PCMDI Metrics Package

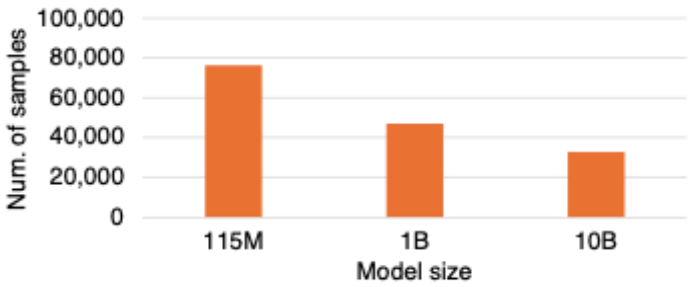
- Use ESGF to access data and PMP to select quality data.

Develop large ViT models to enable effective learning of Earth systems from extensive data



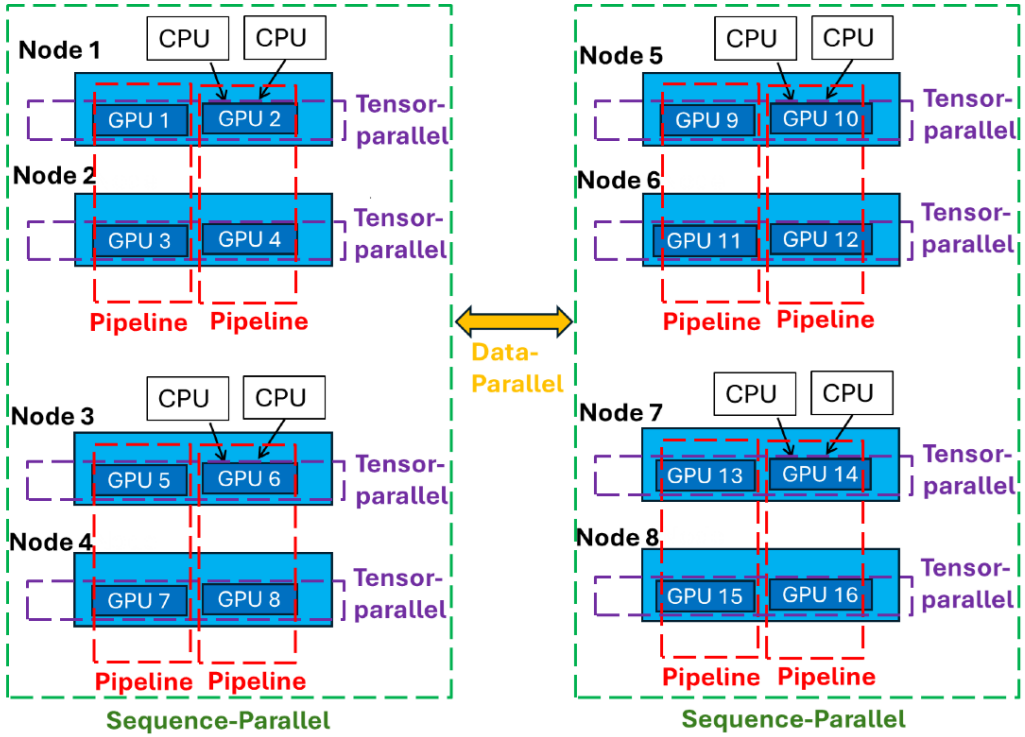
- ORBIT has four model sizes with 115M, 1B, 10B, and 113B parameters.
- It is the largest AI model for Earth system.

Larger models are more effective in Earth system modeling



- As model size increases, the required training samples decreases in Earth system modeling fine-tuning tasks;
- This data efficiency can lead to significant cost and time savings in various Earth system modeling applications.

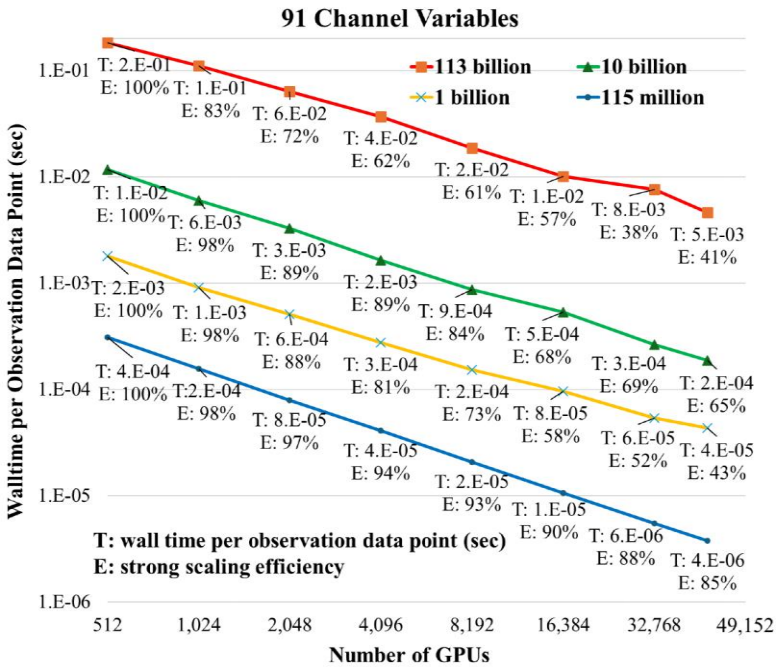
ORBIT achieved strong scaling efficiency on Frontier supercomputer



❖ We develop a novel hybrid model-data-sequence parallelism that merges

- Tensor
- Pipeline
- Data
- Sequence

parallelism orthogonally to accelerate ORBIT training.



➢ ORBIT achieves 1.6 exaflop sustained computing throughput on 6,144 Frontier nodes (49,152 GPUs), with strong scaling efficiency between 44% to 85% for model sizes of 100M, 1B, 10B, and 113B.

Collaborating with

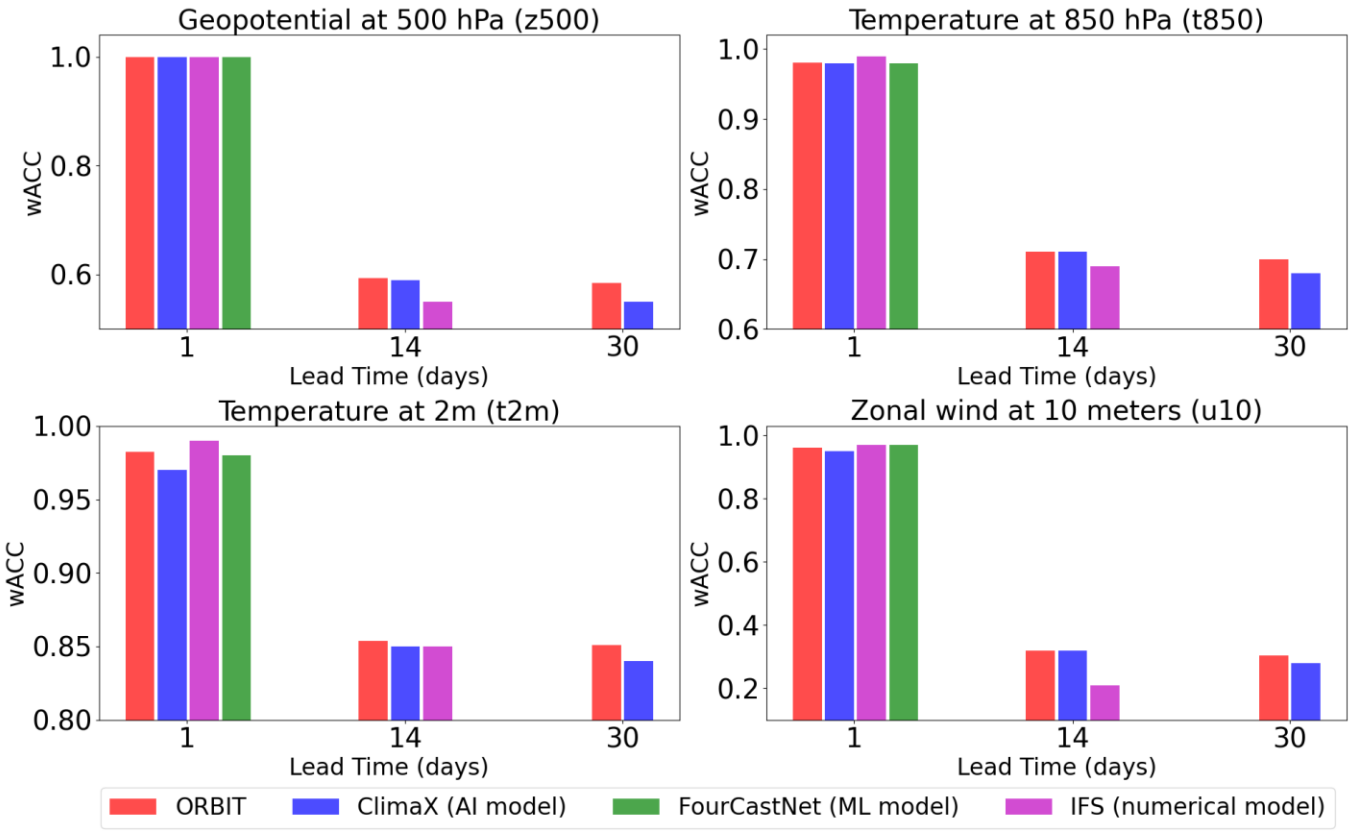
- Microsoft DeepSpeed4Science Team
- AMD Team on Frontier platforms for AI

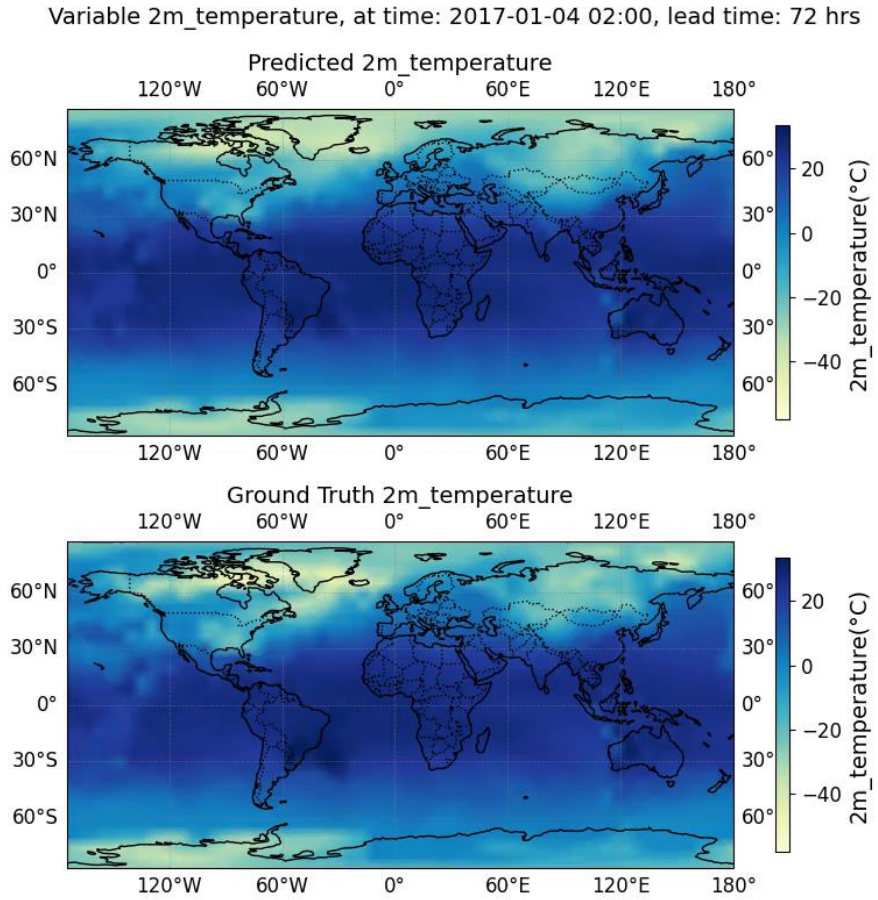
Microsoft AMD

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

ORBIT produced fast and accurate weather forecasts

- Finetune ORBIT using ERA5 data for weather forecast





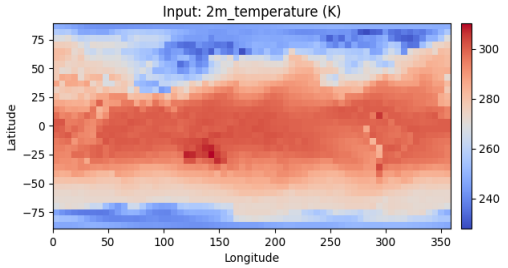
❖ ORBIT achieves competitive performance in weather forecasting, matching or surpassing state-of-the-art numerical, machine learning, and foundation models.

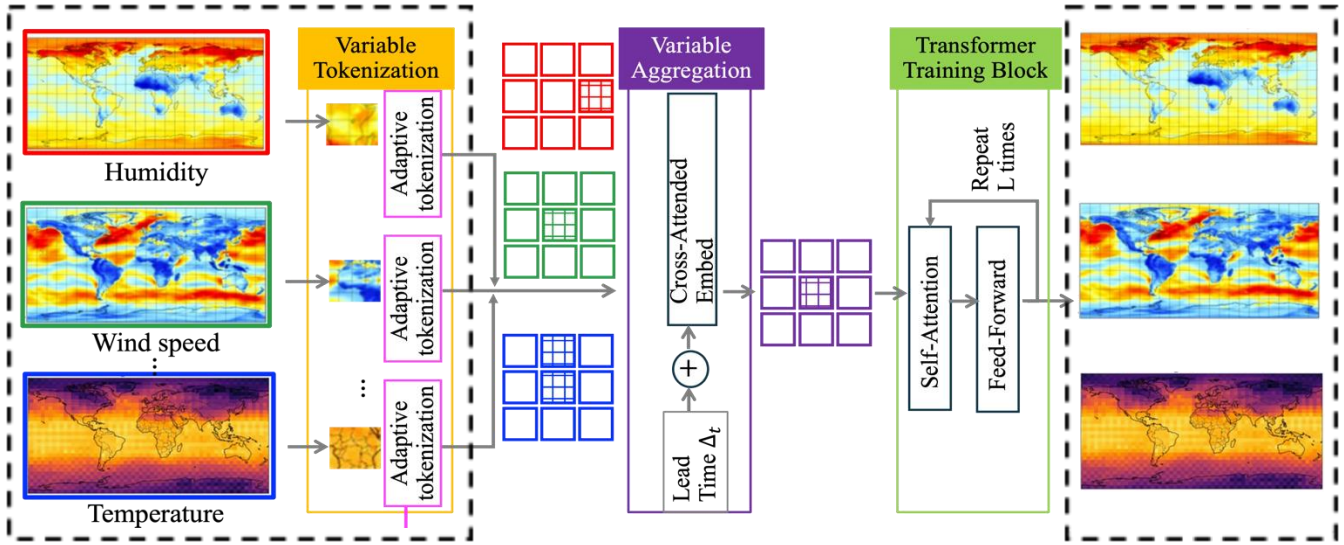
Model Size	115 million
GPUs	1 GPU
Forecast Time	0.04 sec

Fine-tuned ORBIT for weather/climate downscaling

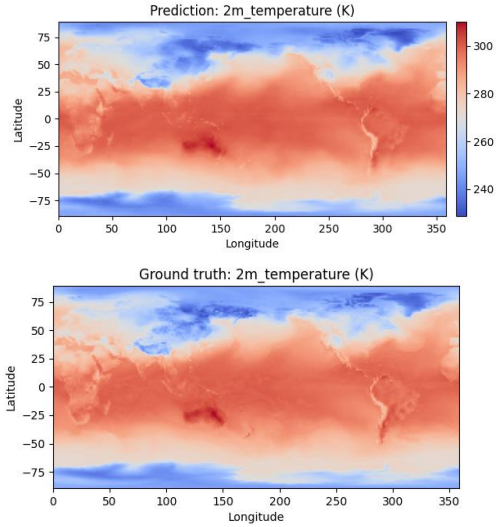
- Finetune ORBIT using pairs of low-resolution and high-resolution data for downscaling

Input: low-res data

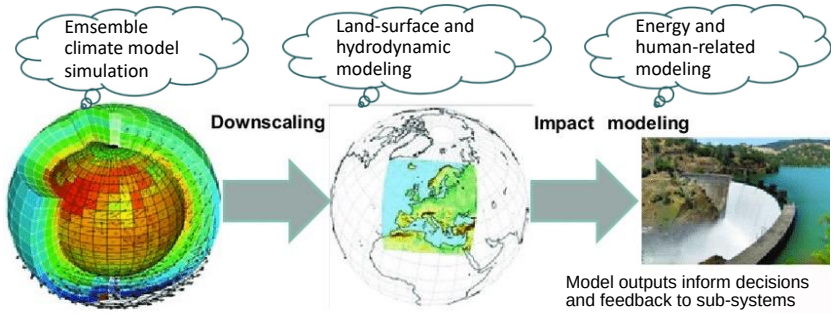




High-res data

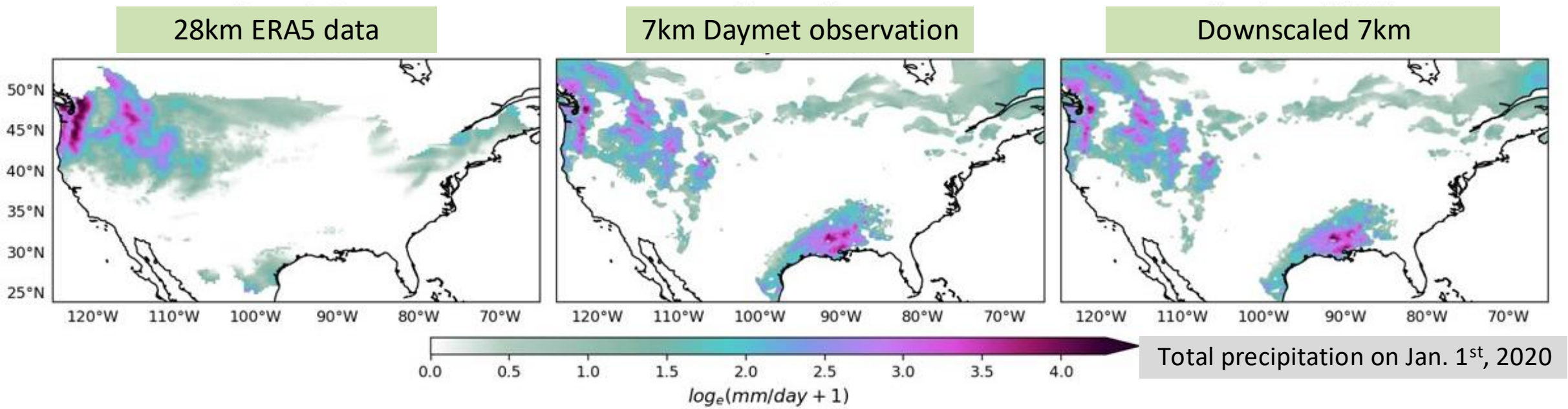


- We adapted ORBIT for weather downscaling by replacing its embedding layers and prediction heads, while retaining its attention layers and variable aggregation module.



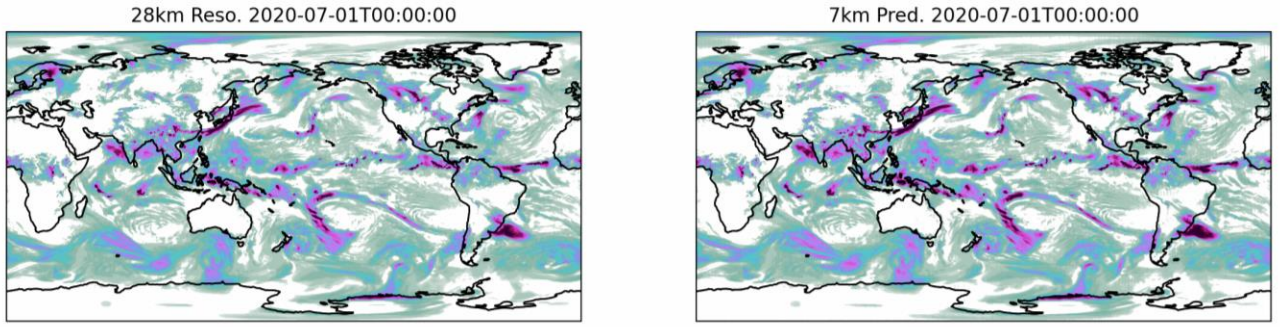
ORBIT accurately generated high-res precipitation data

Model size	corr	RMSE	RMSE σ1>68%	RMSE σ2>95%	RMSE σ3>99.7%	RMSE >99.99%	SSIM
117M	0.974	0.151	0.172	0.355	0.465	0.6439	0.924

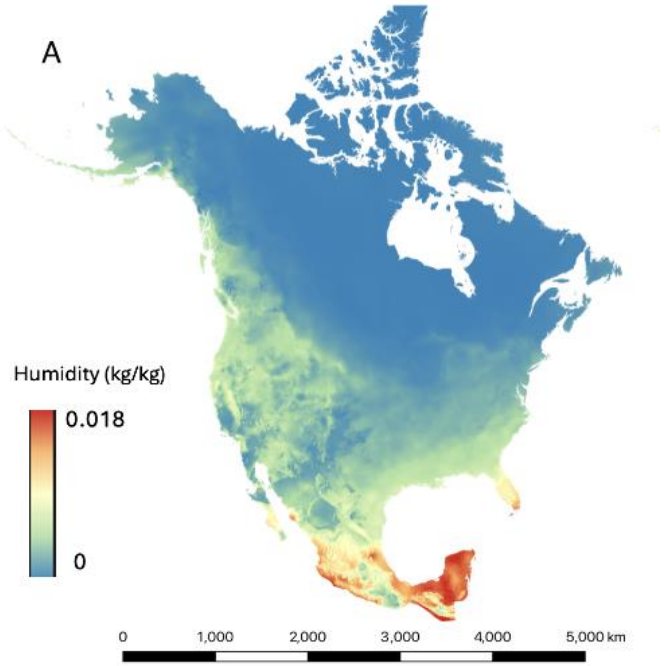


Total precipitation on Jan. 1st, 2020

- ORBIT demonstrated very accurate results in generating high-resolution precipitation from low-resolution data, even in capturing the extreme values.



Fine-tune ORBIT to accelerate land model simulation



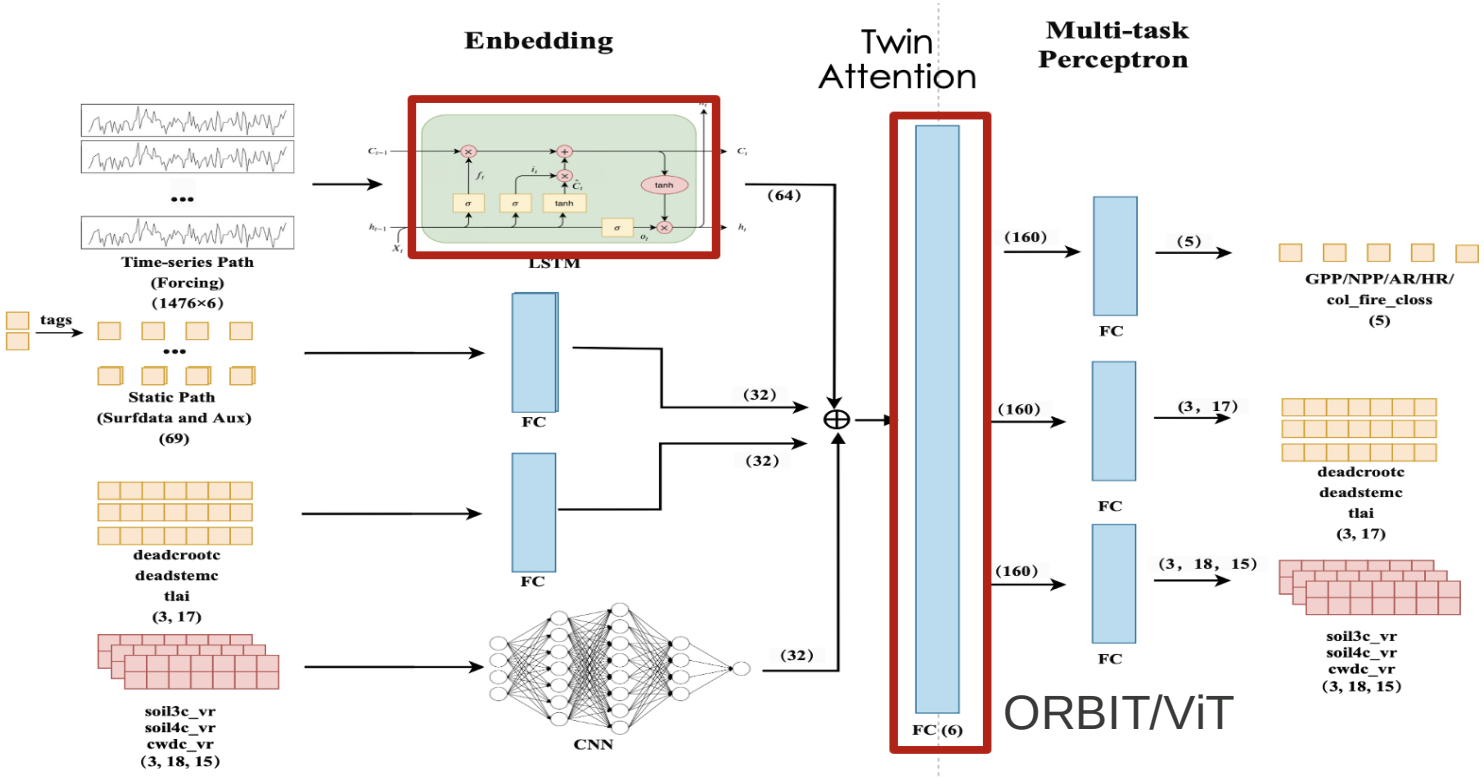
- Multimodal data including time-series, static variables, and spatial varying variables.
- Use different encoder to extract information from these multimodal data.

Current limitation:

- We developed high-res, km-scale DOE's land surface model (ELM) but it is computationally expensive, mainly due to biogeochemical (BGC) spin-up process.

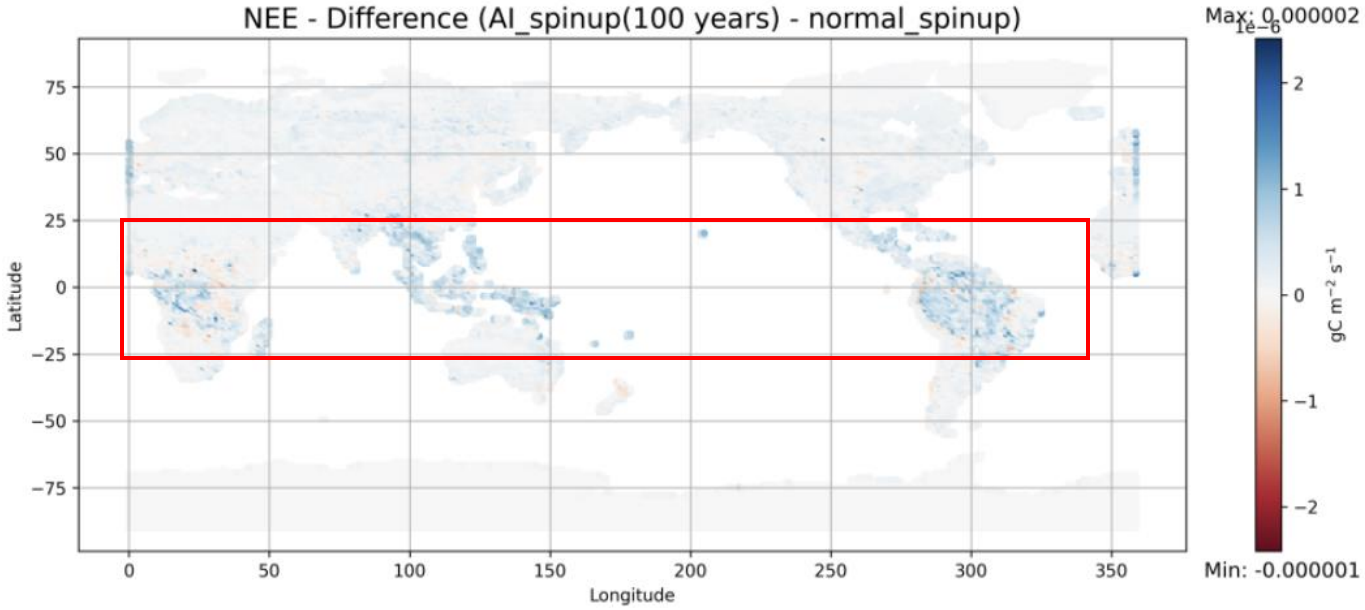
Goal:

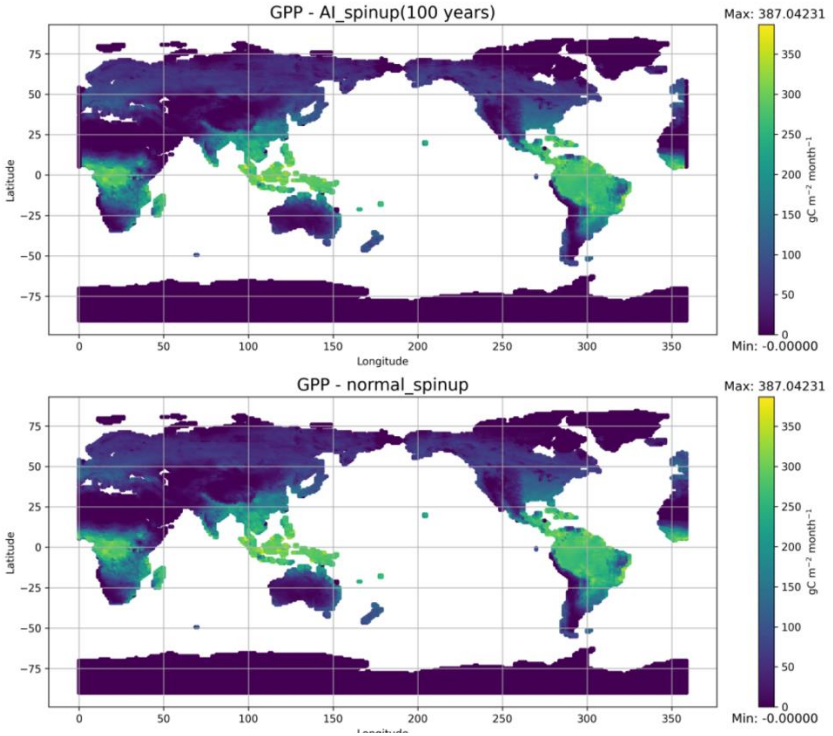
- Aim to build a fast emulator of ELM to accelerate simulation;
- first to accelerate BGC spin-up by leveraging ORBIT foundation.



ORBIT effectively emulated ELM outputs with high fidelity

- **Verification (Our model produces the results right):**
  - ML outputs closely match ELM simulations across 380 variables, with R2 > 0.97;
- **Validation (Our model produces the right results):**
  - ML model produces accurate initial conditions that lead to equilibrium, with NEE approaching zero globally.
  - Large errors occur in tropical regions, likely due to missing processes and variables not yet included in the ML model.





Acknowledgement:

- ORNL team: Wang, D., Shi, X., Ricciuto, D., Thornton, P., and Yang, X.
- North Texas University

AI foundation model has potential to transform Earth system modeling

- ORBIT has potential to advance Earth system modeling by leveraging diverse datasets and its well-trained foundation.

Fine-tuning forecasts: ORBIT brings long-range weather prediction within reach

November 13, 2024

Researchers at Oak Ridge National Laboratory used the Frontier supercomputer to train the world's largest AI model for weather prediction, paving the way for hyperlocal, ultra-accurate forecasts. This achievement earned them a finalist nomination for the prestigious Gordon Bell Prize for Climate Modeling.

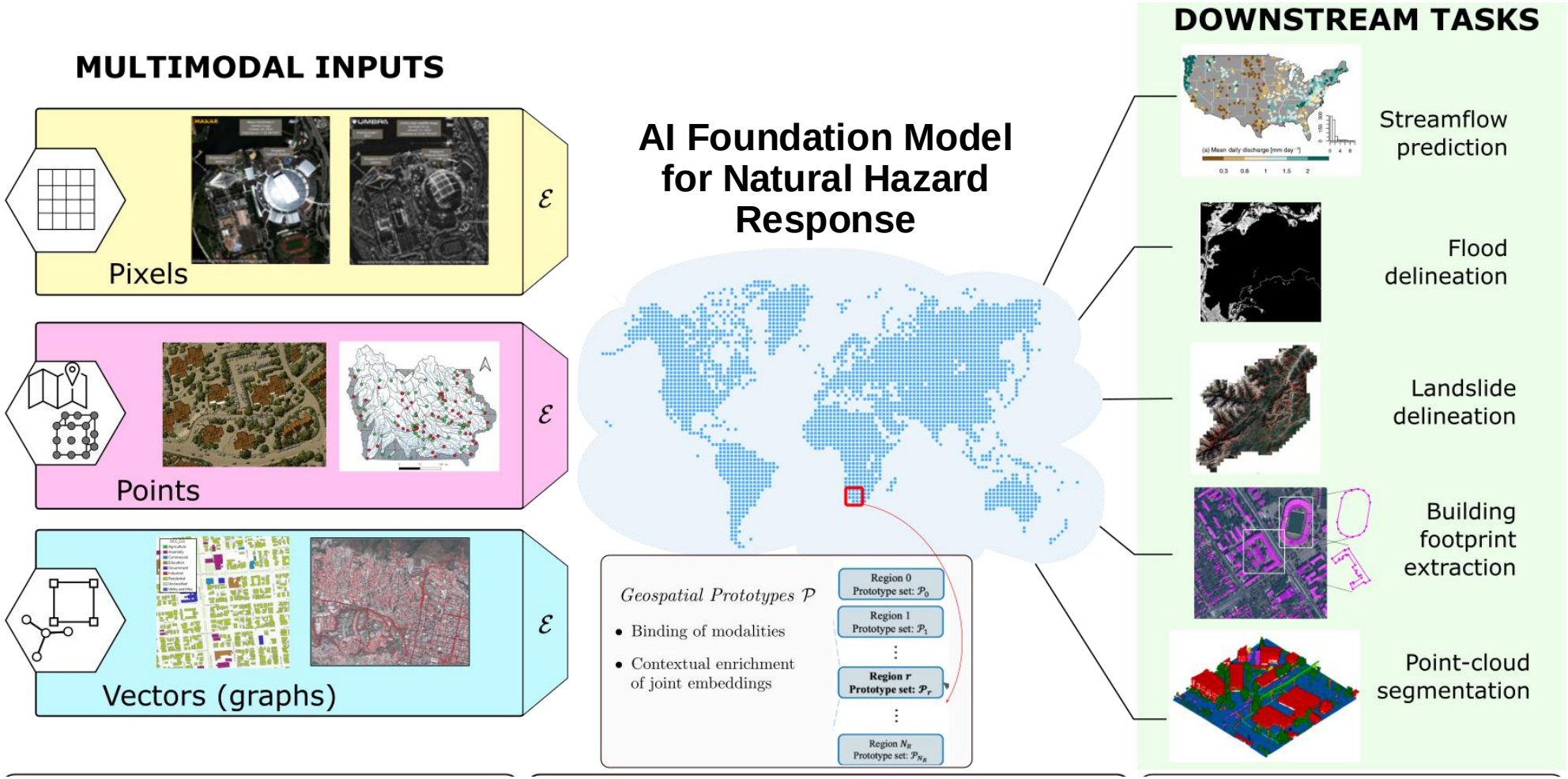
Gordon Bell Prize for Climate Modeling Finalist
Top Supercomputing Achievement Award

Oak Ridge National Laboratory receives honors in 2024 HPCwire Editors' Choice award

November 19, 2024

ORNL has been recognized in the 21st edition of the HPCwire Readers' and Editors' Choice Awards, presented at the 2024 International Conference for High Performance Computing, Networking, Storage and Analysis in Atlanta, Georgia.

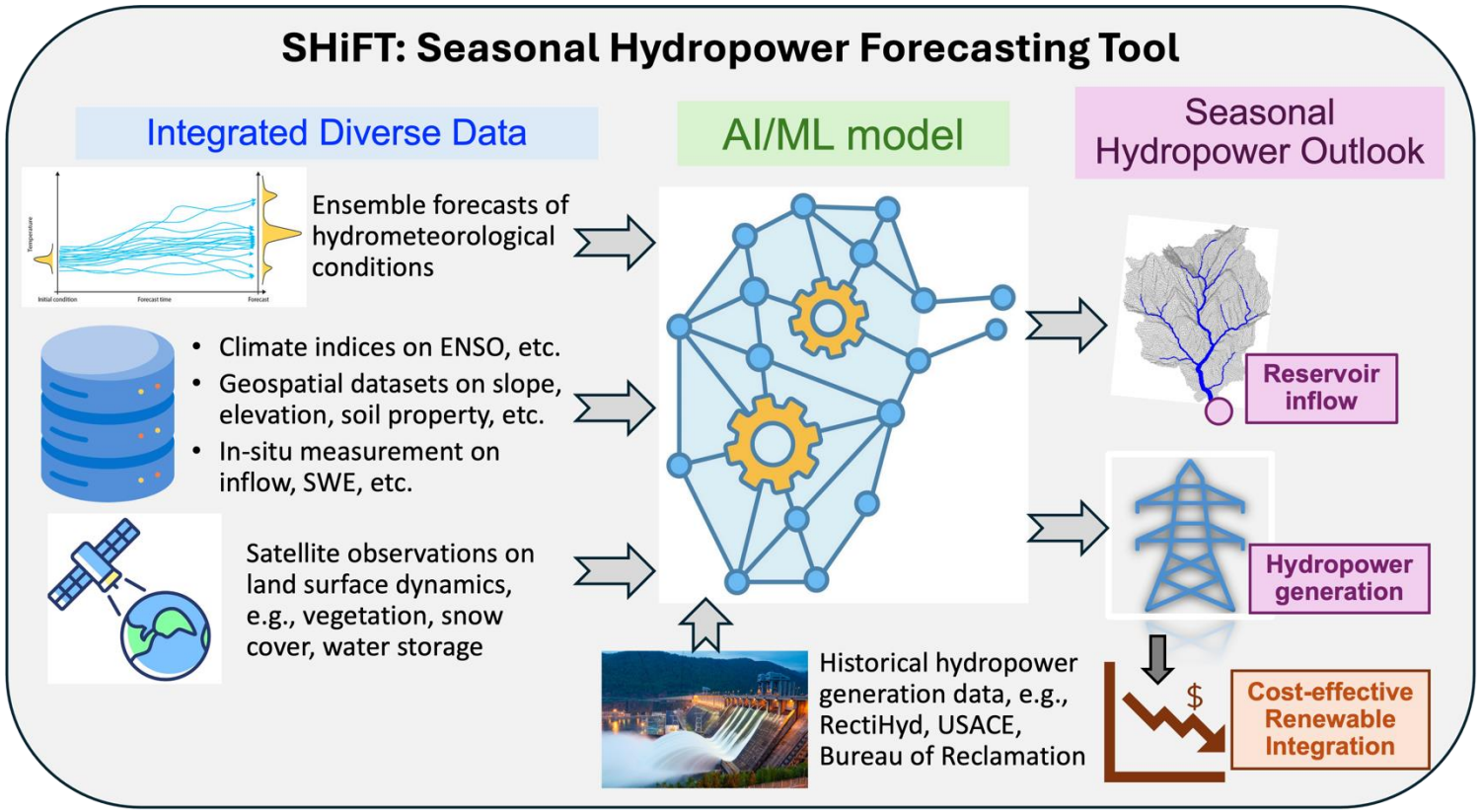
AI foundation model for natural hazard assessment and response



❖ The model integrates multimodal data to enhance disaster impact assessment and inform response strategies across multiple sectors.

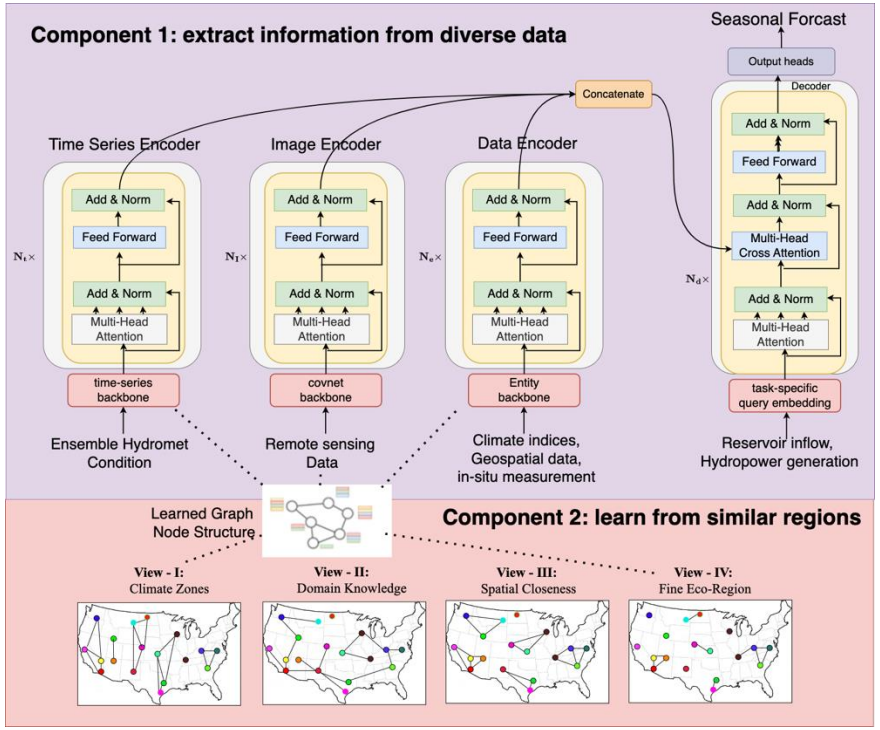
• EarthBind project, PI: Dalton Lunga

AI model for probabilistic seasonal hydropower forecasts



- The tool provides probabilistic seasonal forecasts of reservoir inflow and hydropower generation at individual plants and energy regions.

- Our ML model uses multiple encoders to extract information from various data and employs graph networks to facilitate information sharing across similar regions.

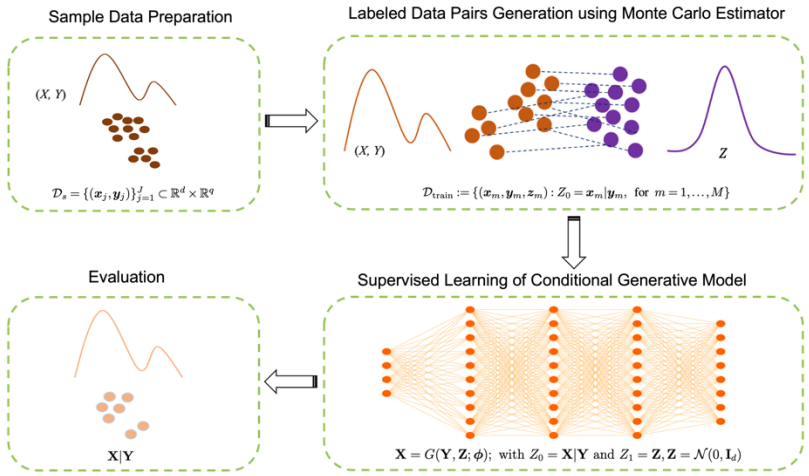


❖ The model integrates comprehensive data to provide probabilistic forecasts of inflow and hydropower generation, informing energy and water management decision.

Two open-source code packages

Generative AI for UQ

- **GenAI4UQ**: Uncertainty Quantification (UQ) package Using Conditional Generative AI

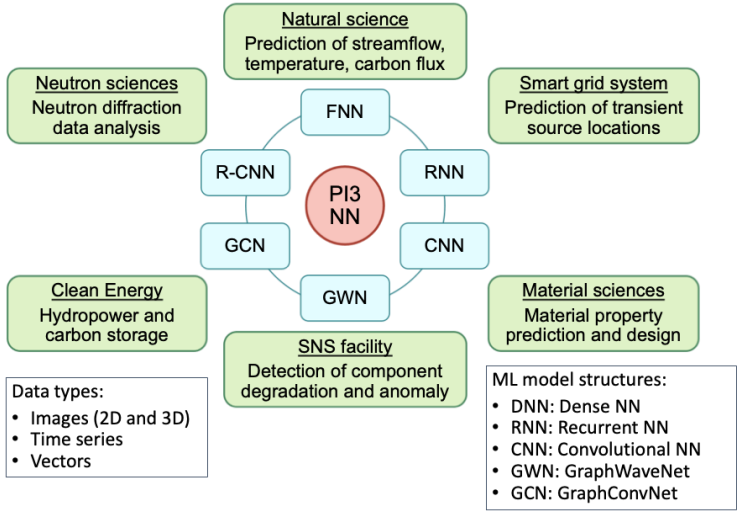


- Quantify parameter uncertainty;
- Make probabilistic forecast;
- Computationally and memory efficient;
- Perform amortized Bayesian inference;
- Enable real-time and large-scale model calibration.

https://github.com/patrickfan/GenAI4UQ

UQ for ML models

- **UQnet**: Quantify ML prediction uncertainty and identify out-of-sample regimes.

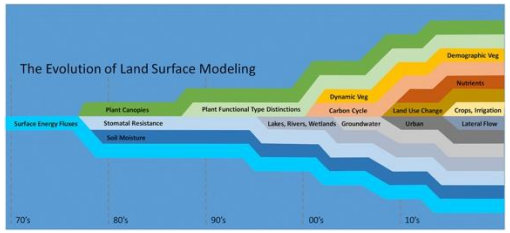


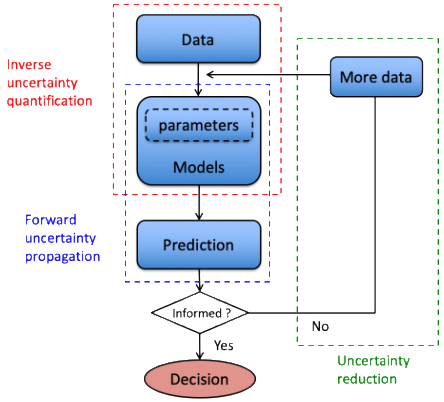
- Produce uncertainties consistent with confidence level and prediction error;
- Computationally efficient;
- Appliable to various neural network architectures;

https://github.com/liusiyan/UQnet

Advancing land surface modeling through data-model integration

Physical Model





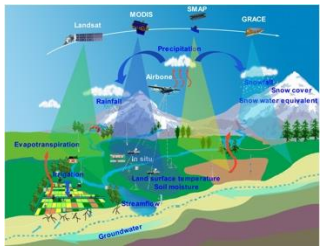
Challenges:

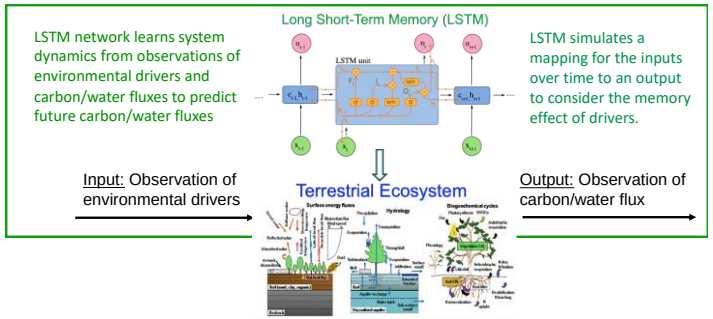
- High computational costs;
- Large parameter uncertainty;

Our study:

- Efficient emulation;
- Generative AI for UQ.

Data-Driven ML Model





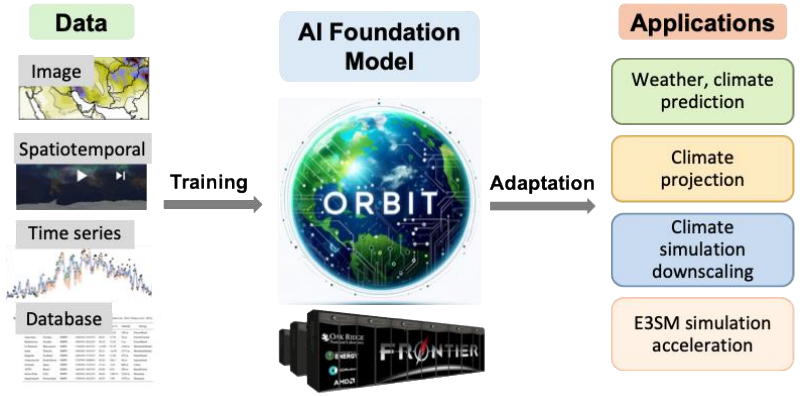
Challenges:

- Generalizing across space and time;
- Explainability, physical consistency;
- Reliability under changing conditions;

Our study:

- Advanced ML integrating diverse data;
- Interpretable AI for explainability;
- UQ to improve predictive reliability.

AI Foundation Model



Challenges:

- Heterogeneous, unlabeled data
- Diverse Earth system modeling needs

Our study:

- Develop an AI foundation model trained on CMIP6 climate data;
- Fine-tune for weather forecasting, climate downscaling, and land model acceleration.