

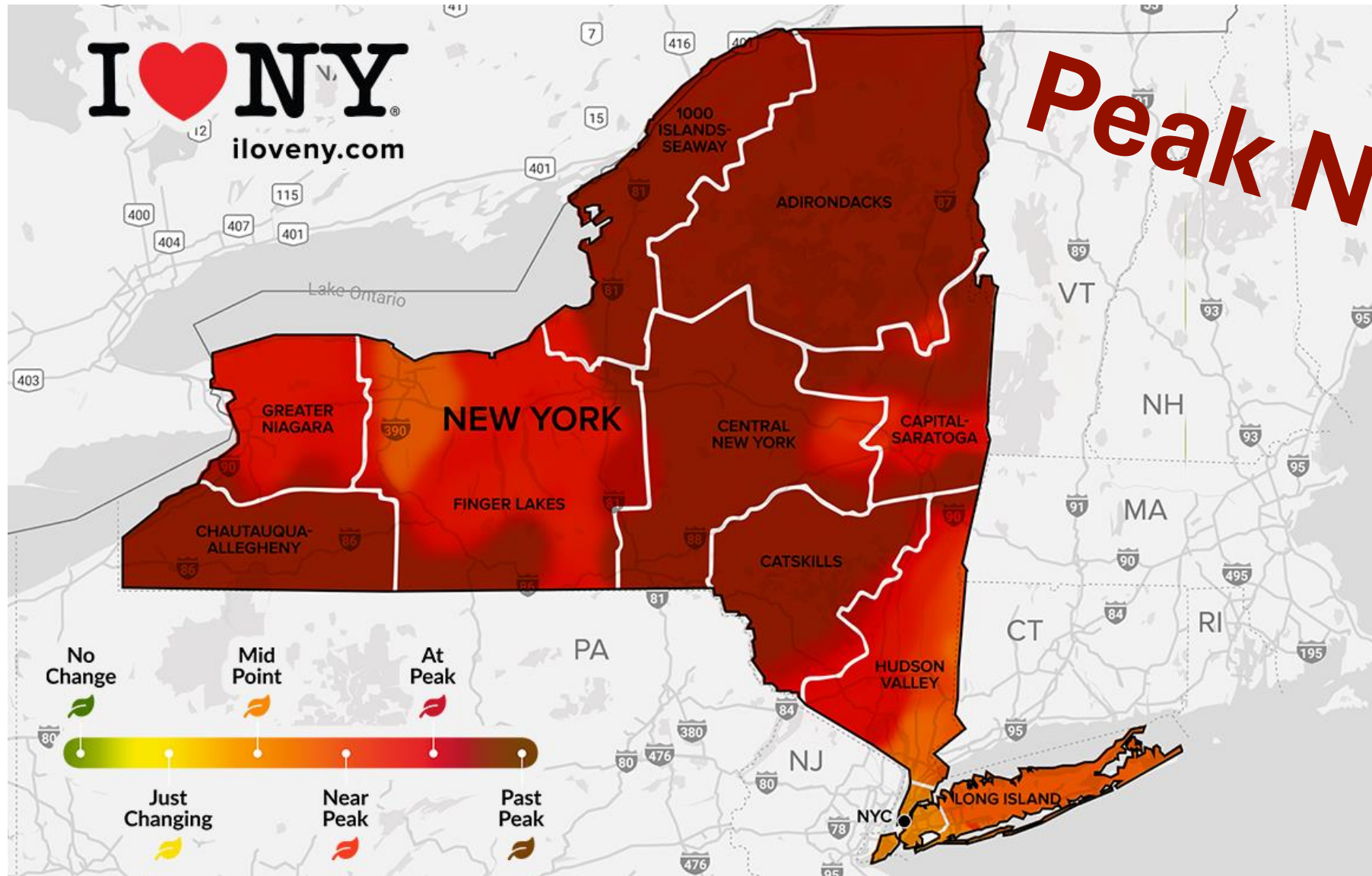


Water Shortage Attribution in the Western U.S. using eXplainable Artificial Intelligence (XAI) Emulators of Process-Based Hydrologic Models

EWRS Seminar

Yifan Luo

October 23, 2025



Peak Now!

Source: Check out New York State foliage! Iloveny.com. <https://www.iloveny.com/things-to-do/fall/foliage-report/>

Outline

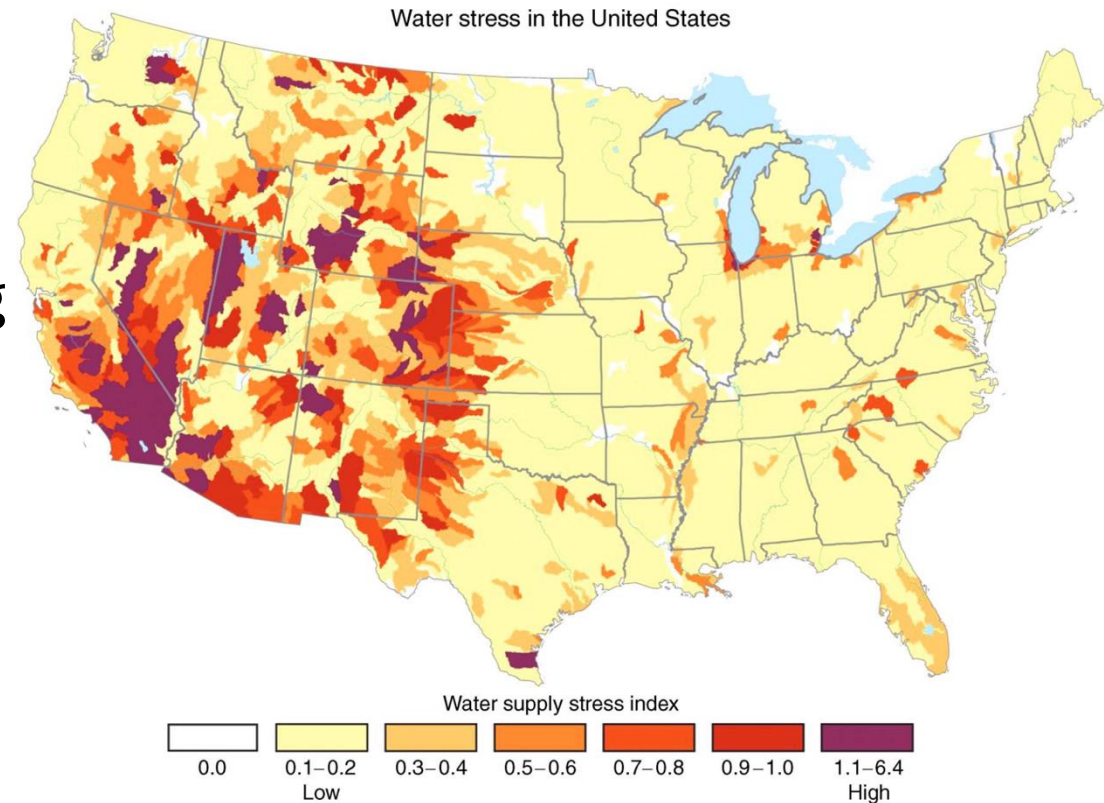
- Motivation & Background
- Research Questions
- Methods:
 - Processed-Based Model (WBM)
 - Deep Learning Emulator (ConvLSTM)
 - Explainable AI Ensemble
- Interactive Demo
- Results:
 - Emulator Performance
 - SHAP Analysis
 - Attribution of Dominant Drivers
 - Consensus Analysis
- Discussion & Conclusions



Interactive Demo

Water Scarcity in the Western U.S.

- **Definition:** Mismatch between freshwater demand and availability
- **Key Impacts:** Groundwater depletion, drought conflicts, risks to human well-being
- **Drivers:** Meteorological variability (droughts, temperature rises) + Human water use (irrigation, urban/industrial demands)
- **Regional Context:** Megadrought since 2000s, aging infrastructure, increasing aridity



Source: Averyt, K., Meldrum, J., Caldwell, P., Sun, G., McNulty, S., Huber-Lee, A., & Madden, N. (2013). Sectoral contributions to surface water stress in the coterminous United States. *Environmental Research Letters*, 8(3), 035046. <https://doi.org/10.1088/1748-9326/8/3/035046>

The Attribution Challenge

Complex Interactions:

- Hydro-climate (precipitation, temperature) modulates availability; socioeconomic use (population, agriculture) drives demand
- Nonlinearities + lags + routing → attribution isn't separable or stationary

Process-Based Hydrologic Models:

- **Strength:** Integrates hydrology + operations + human use; upstream changes propagate
- **Constraint:** High computational need; difficult to isolate driver effects

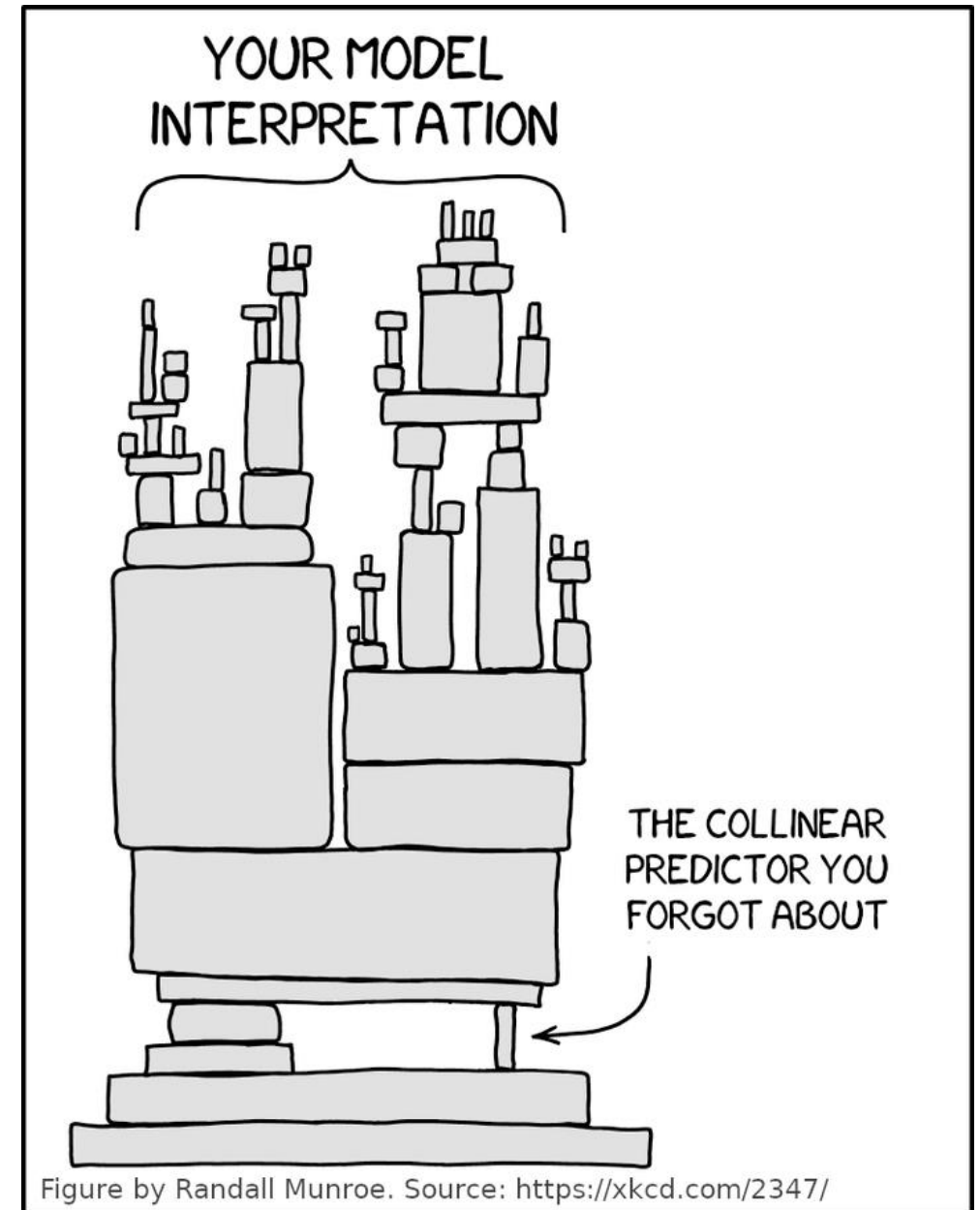


Figure by Randall Munroe. Source: <https://xkcd.com/2347/>

Source: Benito, B. M. (2023, October 29). Multicollinearity Hinders Model Interpretability. Blas M. Benito, PhD.
<https://www.blasbenito.com/post/multicollinearity-model-interpretability/#>

The Attribution Challenge – Cont.

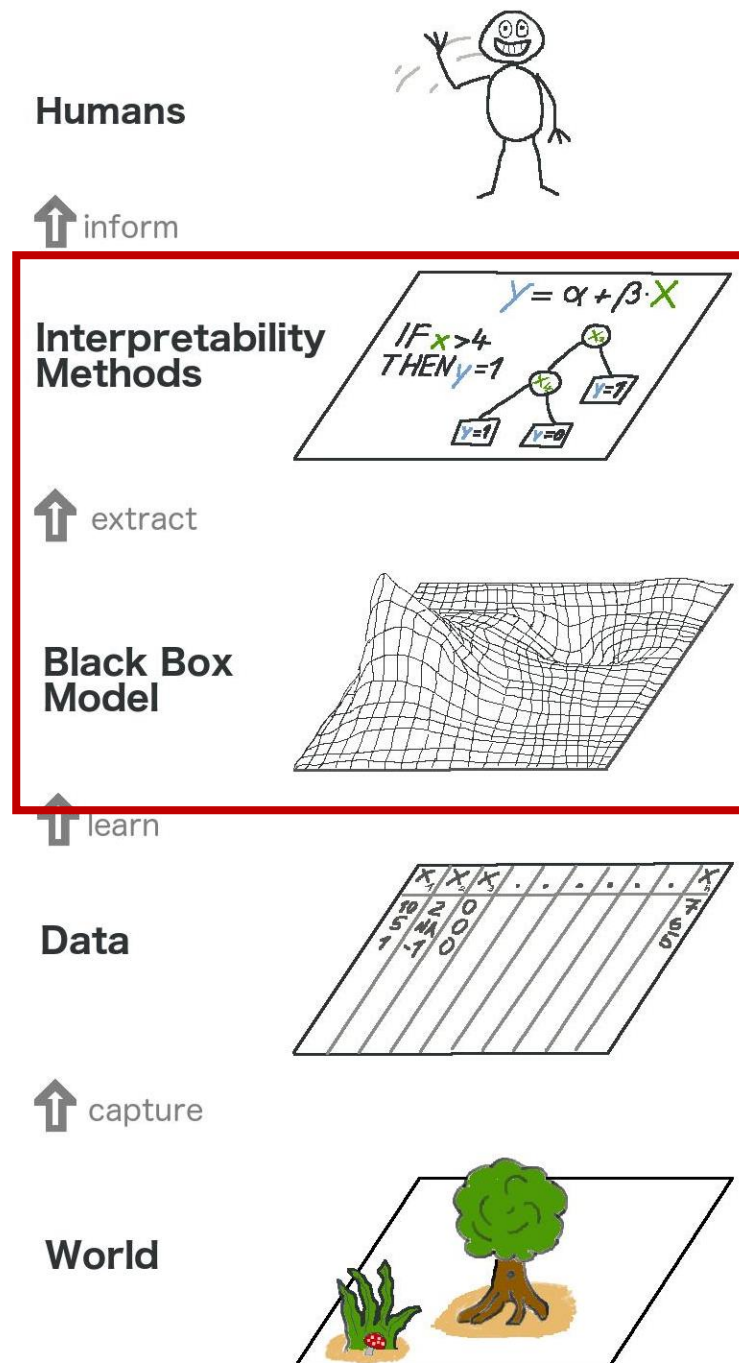
DL Emulators:

- **Strength:** Orders-of-magnitude speedup → enabling extensive analysis
- **Constraint:** Black-box behavior; lacking transparency

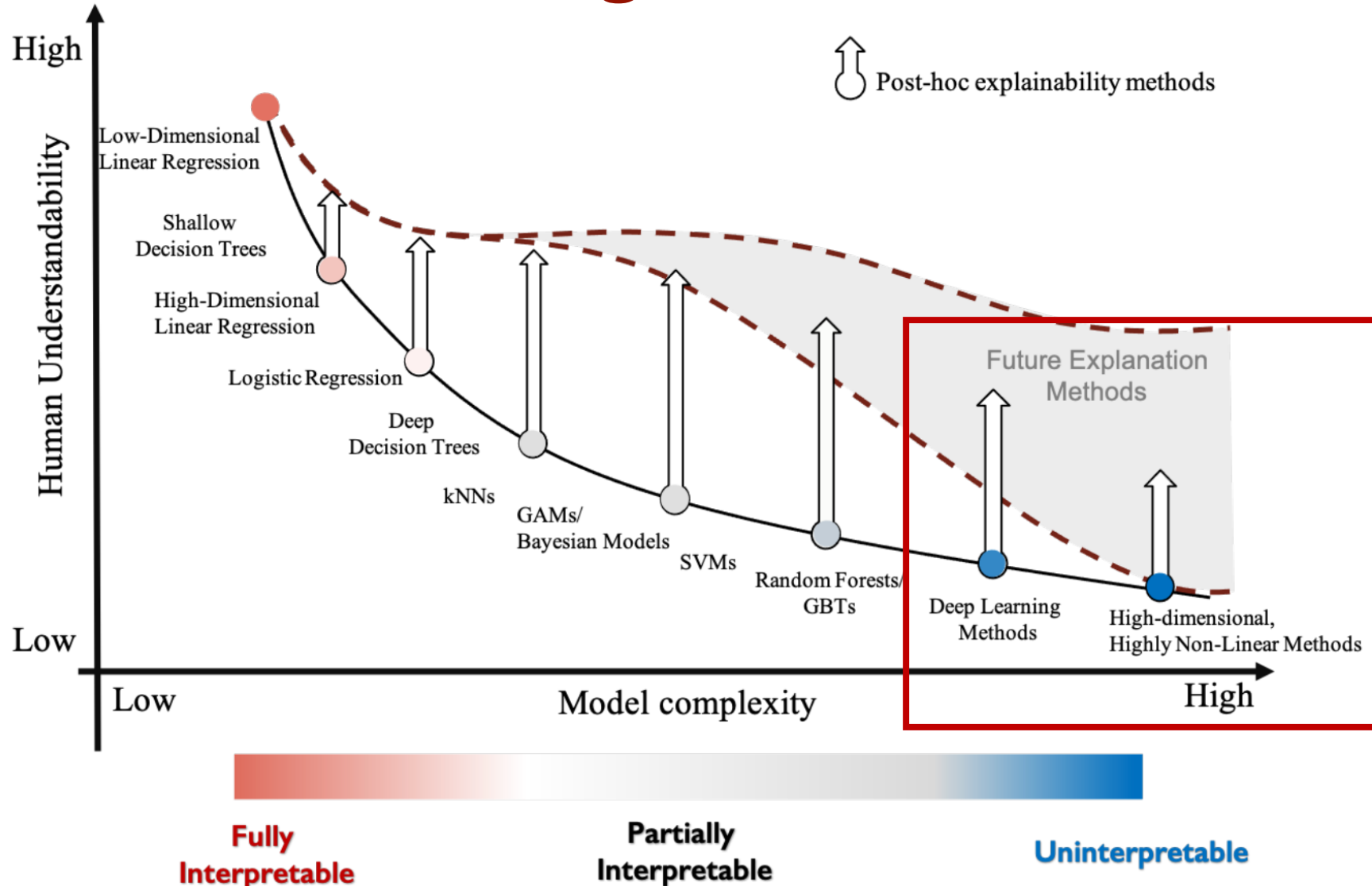
Explainable AI (XAI):

- **Strength:** Attribute emulator predictions to specific input drivers
- **Constraint:** Individual XAI methods have different assumptions and can disagree, creating explanation uncertainty

Source: Molnar, C. (2025). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (3rd ed.). <https://christophm.github.io/interpretable-ml-book>



The Attribution Challenge – Cont.



Source: Flora, M., Potvin, C., McGovern, A., & Handler, S. (2022). Comparing Explanation Methods for Traditional Machine Learning Models Part I: An Overview of Current Methods and Quantifying Their Disagreement. ArXiv.org. <https://arxiv.org/abs/2211.08943>

Outline

- Motivation & Background
- Research Questions
- Methods:
 - Processed-Based Model (WBM)
 - Deep Learning Emulator (ConvLSTM)
 - Explainable AI Ensemble
- Interactive Demo
- Results:
 - Emulator Performance
 - SHAP Analysis
 - Attribution of Dominant Drivers
 - Consensus Analysis
- Discussion & Conclusions



Interactive Demo

Research Questions

1. Can a deep learning emulator **accurately reproduce** water scarcity simulations from the process-based hydrologic model?
2. Using the emulator, what factors (climate vs socioeconomic) are **most important** for causing water shortages in different regions?
3. Do XAI methods **agree** on driver importance for the same emulator, and can we map their consensus to quantify **explanation uncertainty** spatially?

⇒ ***Moving beyond "black-box" prediction to transparent, uncertainty-aware attribution***

Outline

- Motivation & Background
- Research Questions
- **Methods:**
 - Processed-Based Model (WBM)
 - Deep Learning Emulator (ConvLSTM)
 - Explainable AI Ensemble
- Interactive Demo
- Results:
 - Emulator Performance
 - SHAP Analysis
 - Attribution of Dominant Drivers
 - Consensus Analysis
- Discussion & Conclusions

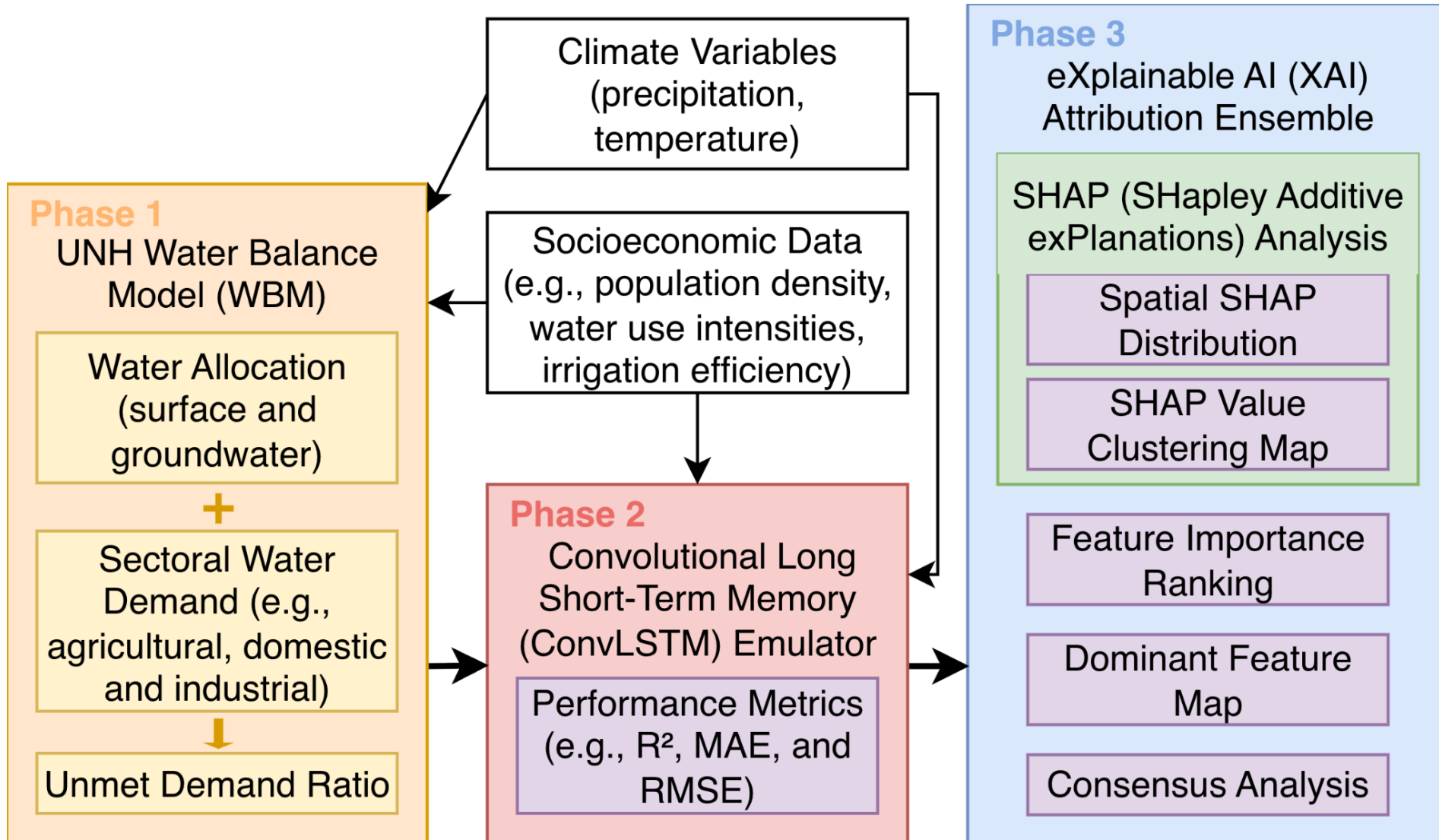


Interactive Demo

Methods: Data & Study Area

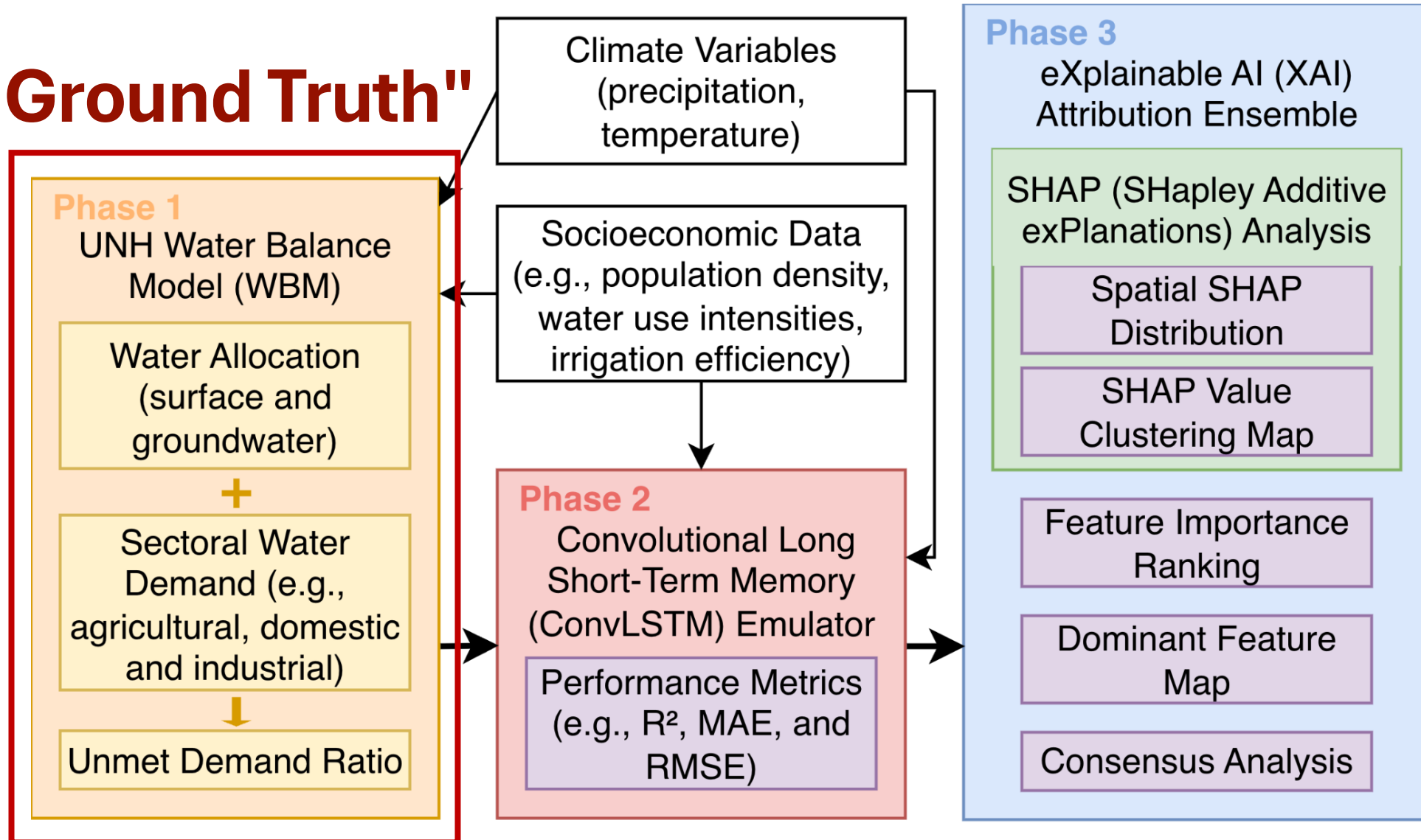
- **Study Area:** Western United States (11 states)
- **Timeframe:** 2010–2019 (10 years)
- **Data Sources:**
 - **Climate Data (Monthly):** Temperature and Precipitation
 - **Socioeconomic Indicators (Static):**
 - Population density
 - Cropland fraction
 - Irrigation efficiency
 - Sectoral water usages (domestic and industrial)
 - **Target Variable:** Unmet water demand ratio (difference between demand and allocation)
 - 0 = All demand met
 - 1 = Complete shortage

Methods: Three-Stage Framework



Methods: Three-Stage Framework

The "Ground Truth"



Methods: Processed-Based Model (WBM)

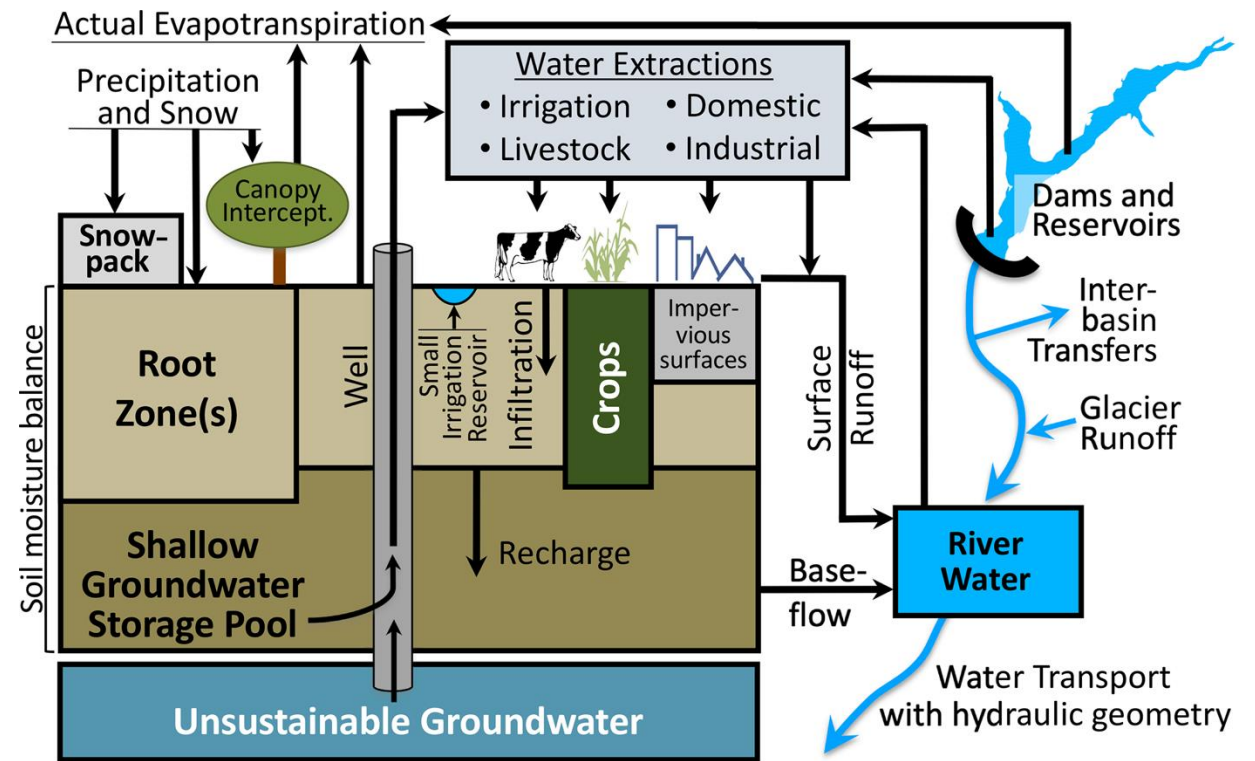
Inputs: Natural

- Climate: temperature & precipitation
- Land surface topography
- Soil properties
- Vegetation properties
- Aquifer properties
- Lake locations
- Glacier melt

Inputs: Human

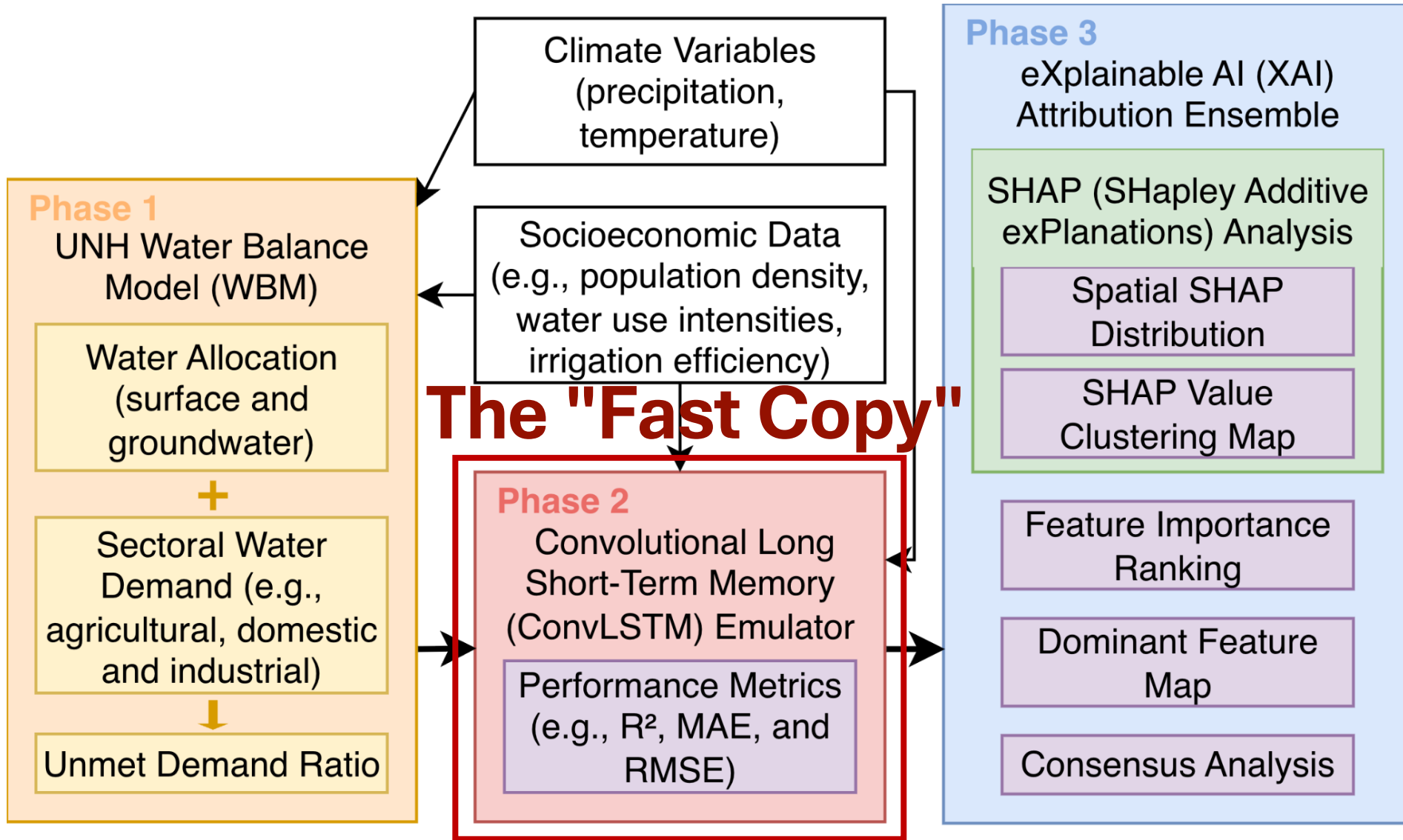
- Crops & irrigation
- Dams & reservoirs
- Population density
- Livestock
- Impervious surfaces (i.e., pavement)
- Interbasin transfers (canals across basins)

Macro-scale process-based hydrologic model with routed flows, reservoirs, allocations



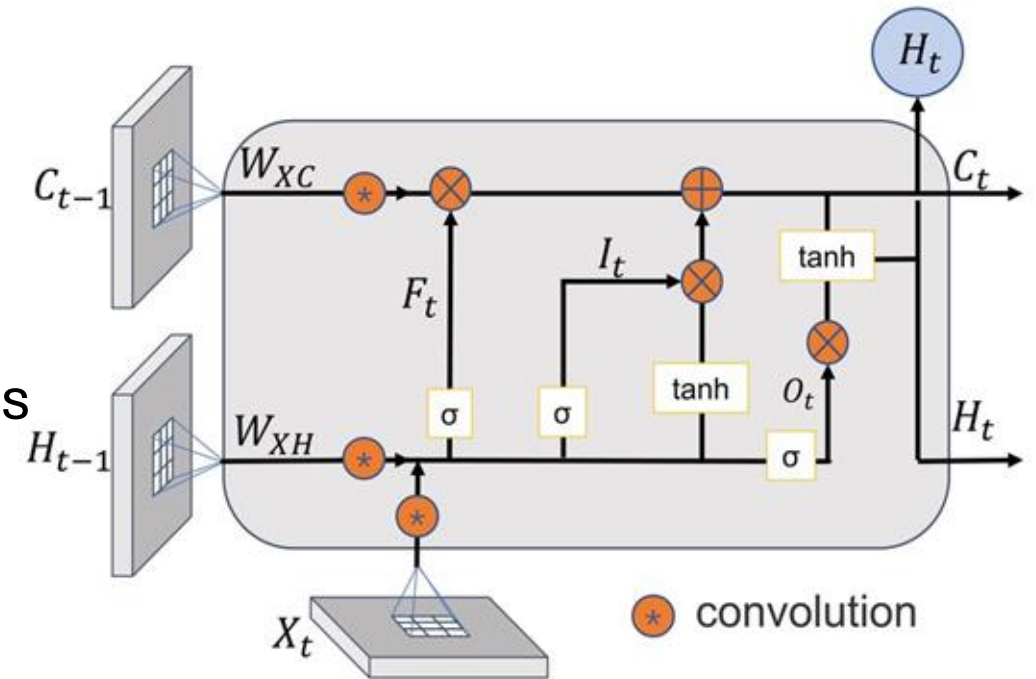
Source: Grogan, D. S., Zuidema, S., Prusevich, A., Wollheim, W. M., Glidden, S., & Lammers, R. B. (2022). Water balance model (WBM) v.1.0.0: a scalable gridded global hydrologic model with water-tracking functionality. *Geoscientific Model Development*, 15(19), 7287–7323. <https://doi.org/10.5194/gmd-15-7287-2022>

Methods: Three-Stage Framework



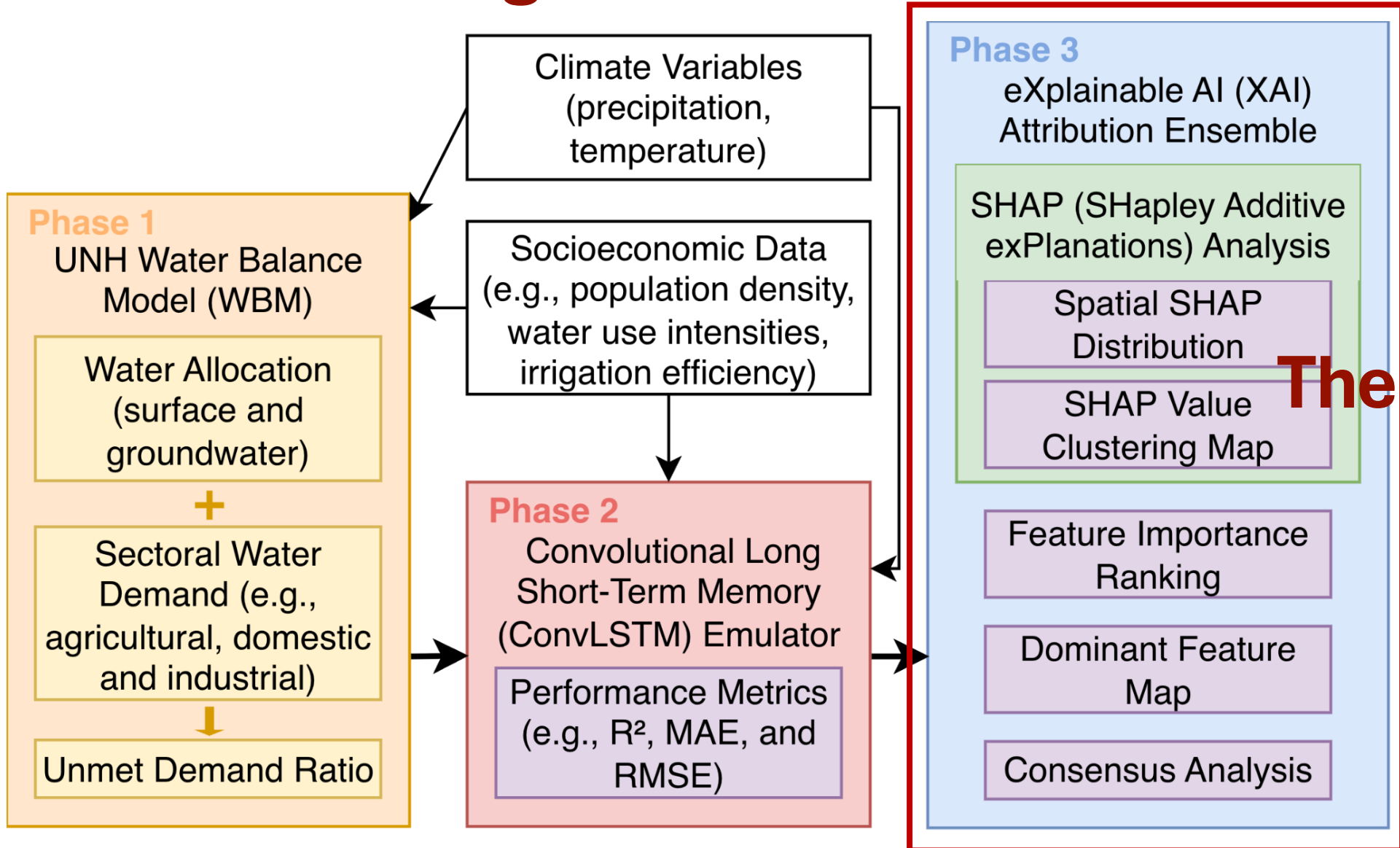
Methods: Deep Learning Emulator (ConvLSTM)

- **Goal:** Approximate WBM input-output mapping for computational efficiency.
- **Architecture:** Convolutional Long Short-Term Memory network (ConvLSTM).
 - Combines Convolutional layers (captures **spatial** patterns) with Recurrent LSTM units (captures **temporal** dynamics).
- **Applications:** Enables large-scale experiments like XAI attribution.



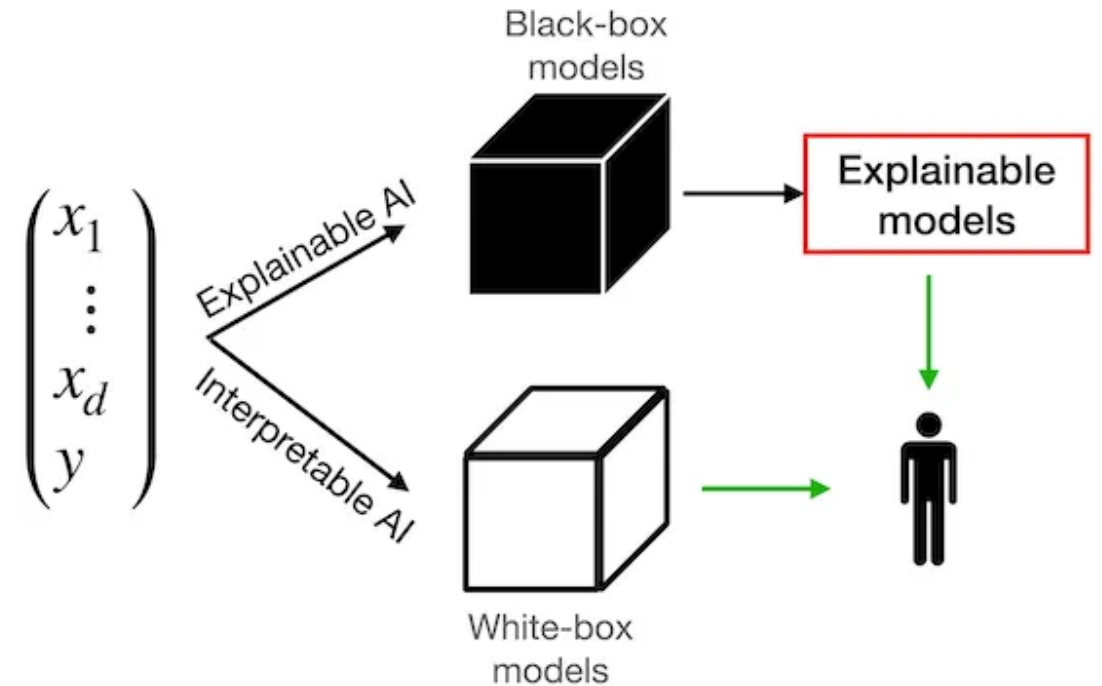
Source: Shi, C., Zhang, Z., Zhang, W., Zhang, C., & Xu, Q. (2022). Learning Multiscale Temporal–Spatial–Spectral Features via a Multipath Convolutional LSTM Neural Network for Change Detection With Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16. <https://doi.org/10.1109/tgrs.2022.3176642>

Methods: Three-Stage Framework



Methods: Explainable AI (XAI) Ensemble

- **Challenge:** Deep learning emulators are often "black boxes," limiting explainability.
- **Solution:** Apply XAI techniques to understand why the emulator makes predictions.
- **Problem:** Individual XAI methods have different assumptions and can yield different explanations.
- **Our Approach:** Use an ensemble of 6 complementary XAI methods to assess explanation consistency and uncertainty



Source: Nour Makke. (2024, May 12). AI Explainability vs. Interpretability. Medium.
<https://medium.com/@nour.makke/ai-explainability-vs-interpretability-5901a43b950d>

Methods: Explainable AI (XAI) Ensemble – Cont.

Method	Full Name	Measures	Foundation
R²	Leave-One-Covariate-Out (LOCO)	Drop in R ² when feature removed	Statistical comparison (Lei et al., 2018)
PFI	Permutation Feature Importance	Performance drop when feature shuffled (breaks correlation)	Breiman (2001); Fisher et al. (2019)
SHAP	SHapley Additive exPlanations	Feature contributions to predictions	Game theory (Shapley values) (Lundberg & Lee, 2017)
Weight	Connection Weights (Garson's)	Magnitude of learned network weights	Neural net interpretation (Garson, 1991)
Gradient	Input Gradients	Sensitivity to small input changes ($\partial\hat{y}/\partial x$)	Gradient computation (Simonyan et al., 2013)
LIME	Local Interpretable Model-agnostic Explanations	Local linear surrogate around predictions	Approximation theory (Ribeiro et al., 2016)

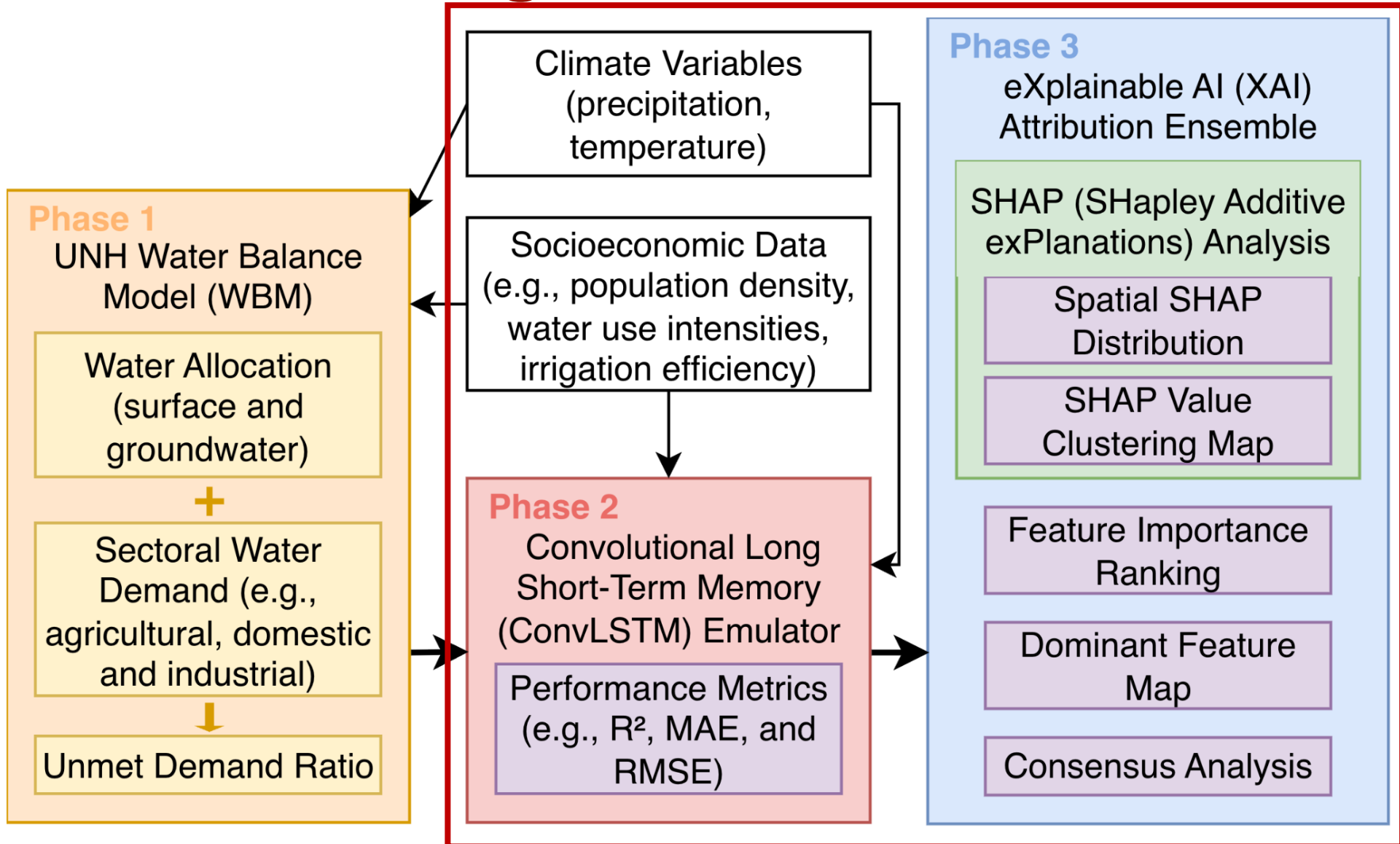
Outline

- Motivation & Background
- Research Questions
- Methods:
 - Processed-Based Model (WBM)
 - Deep Learning Emulator (ConvLSTM)
 - Explainable AI Ensemble
- **Interactive Demo**
- Results:
 - Emulator Performance
 - SHAP Analysis
 - Attribution of Dominant Drivers
 - Consensus Analysis
- Discussion & Conclusions



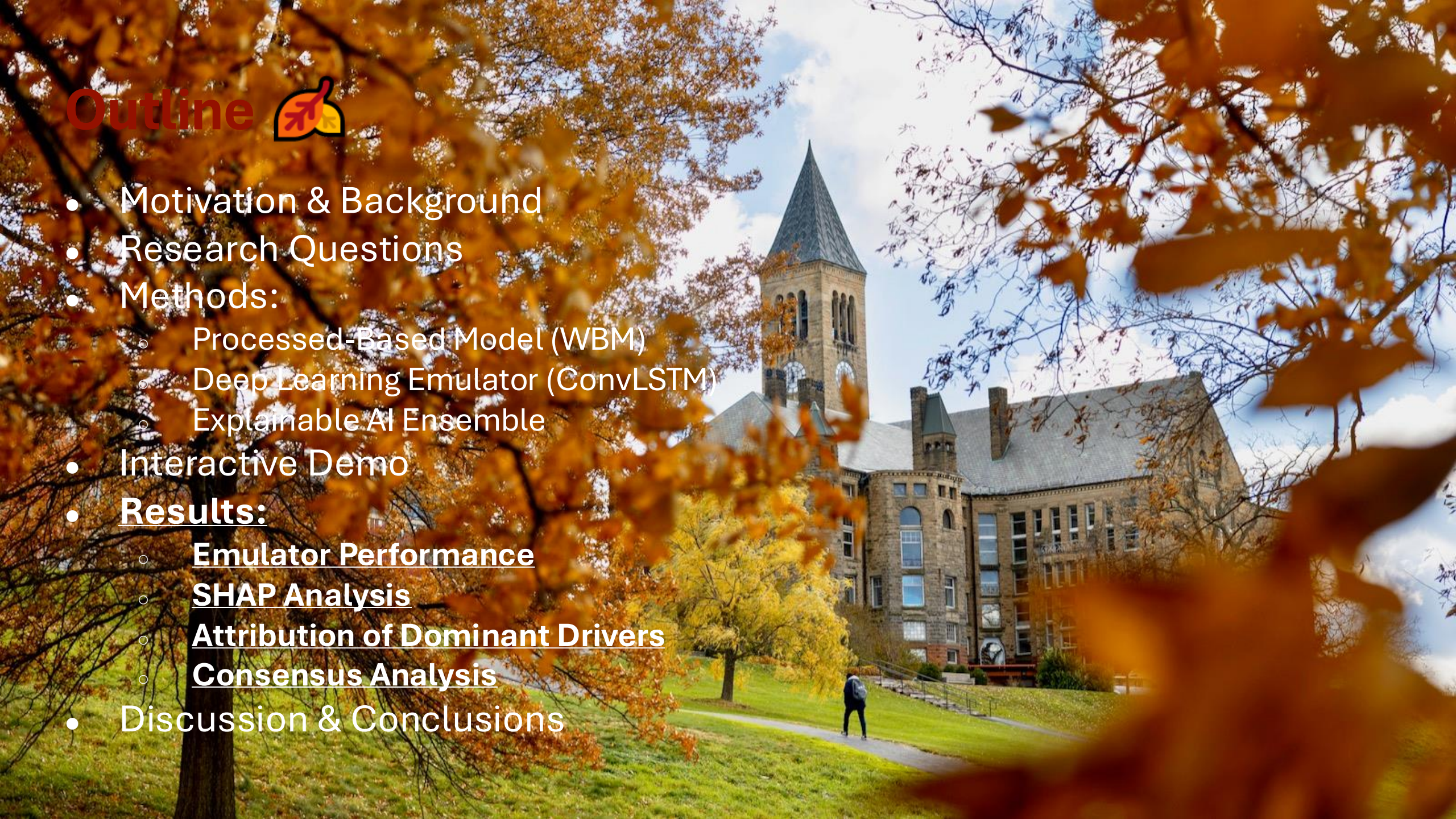
Interactive Demo

Methods: Three-Stage Framework



Outline

- Motivation & Background
- Research Questions
- Methods:
 - Processed-Based Model (WBM)
 - Deep Learning Emulator (ConvLSTM)
 - Explainable AI Ensemble
- Interactive Demo
- Results:
 - Emulator Performance
 - SHAP Analysis
 - Attribution of Dominant Drivers
 - Consensus Analysis
- Discussion & Conclusions



Research Questions

1. Can a deep learning emulator **accurately reproduce** water scarcity simulations from the process-based hydrologic model?
2. Using the emulator, what factors (climate vs socioeconomic) are **most important** for causing water shortages in different regions?
3. Do XAI methods **agree** on driver importance for the same emulator, and can we map their consensus to quantify **explanation uncertainty** spatially?

⇒ *Moving beyond "black-box" prediction to transparent, uncertainty-aware attribution*

Results: Emulator Performance

Model Performance Metrics:

- $R^2 = 0.896$
- MAE = 0.053 (mean absolute error)
- RMSE = 0.130 (root mean squared error)
- Training time = ~18 hours

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 12, 27, 27, 7)	0
conv_lstm2d (ConvLSTM2D)	(None, 27, 27, 32)	124,928
batch_normalization (BatchNormalization)	(None, 27, 27, 32)	128
flatten (Flatten)	(None, 23328)	0
dense (Dense)	(None, 64)	1,493,056
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 32)	2,080
dropout_1 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 1)	33

Total params: 1,620,225 (6.18 MB)

Trainable params: 1,620,161 (6.18 MB)

Non-trainable params: 64 (256.00 B)

Research Questions

1. Can a deep learning emulator **accurately reproduce** water scarcity simulations from the process-based hydrologic model?
2. Using the emulator, what factors (climate vs socioeconomic) are **most important** for causing water shortages in different regions?
3. Do XAI methods **agree** on driver importance for the same emulator, and can we map their consensus to quantify **explanation uncertainty** spatially?

⇒ *Moving beyond "black-box" prediction to transparent, uncertainty-aware attribution*

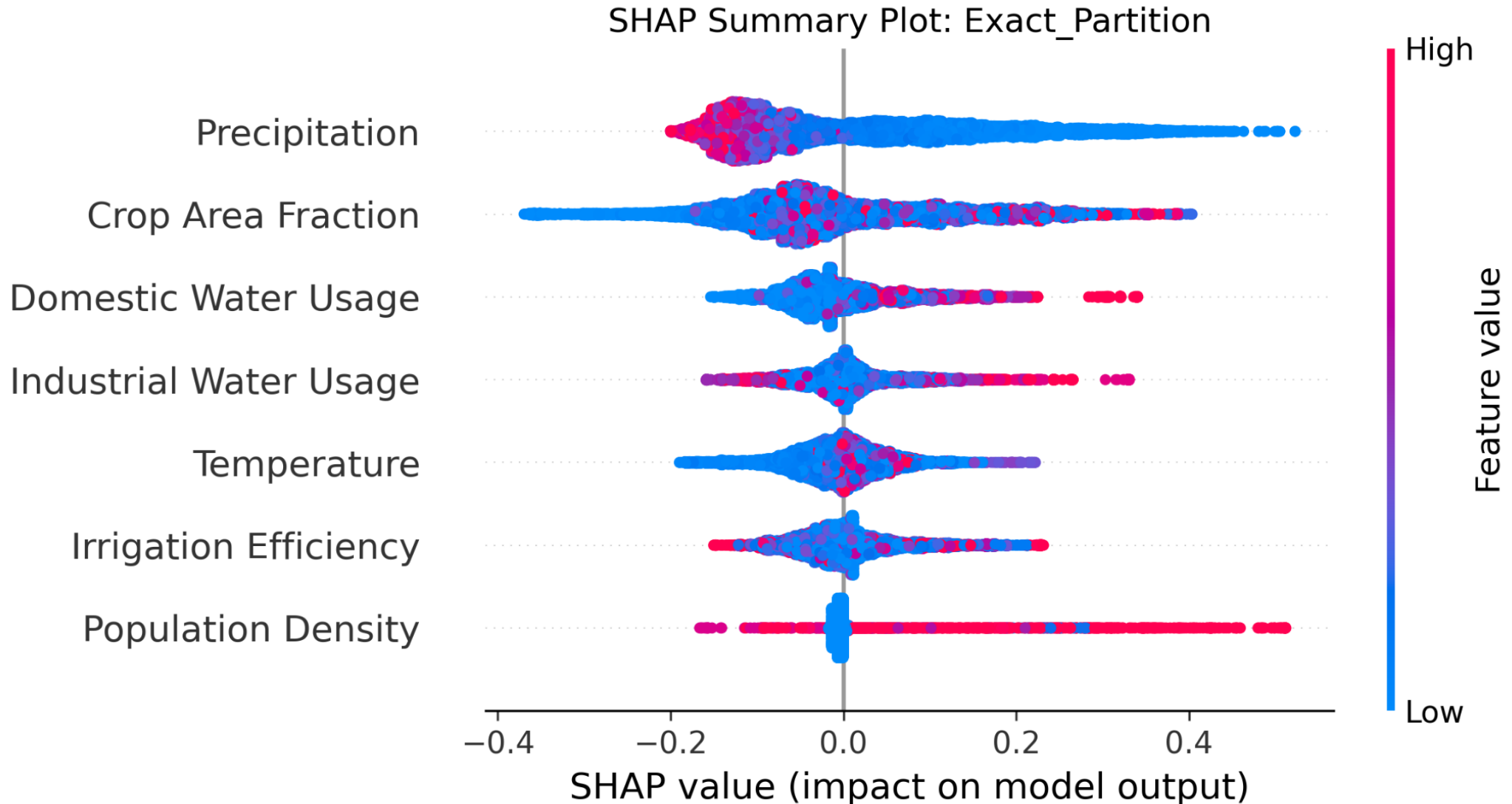
Results: SHAP Analysis

SHAP (SHapley Additive exPlanations) explains a single prediction by assigning each feature a signed *Shapley* value—its **average marginal contribution across all feature coalitions**—so the attributions **add up to the model's output relative to a chosen baseline**.

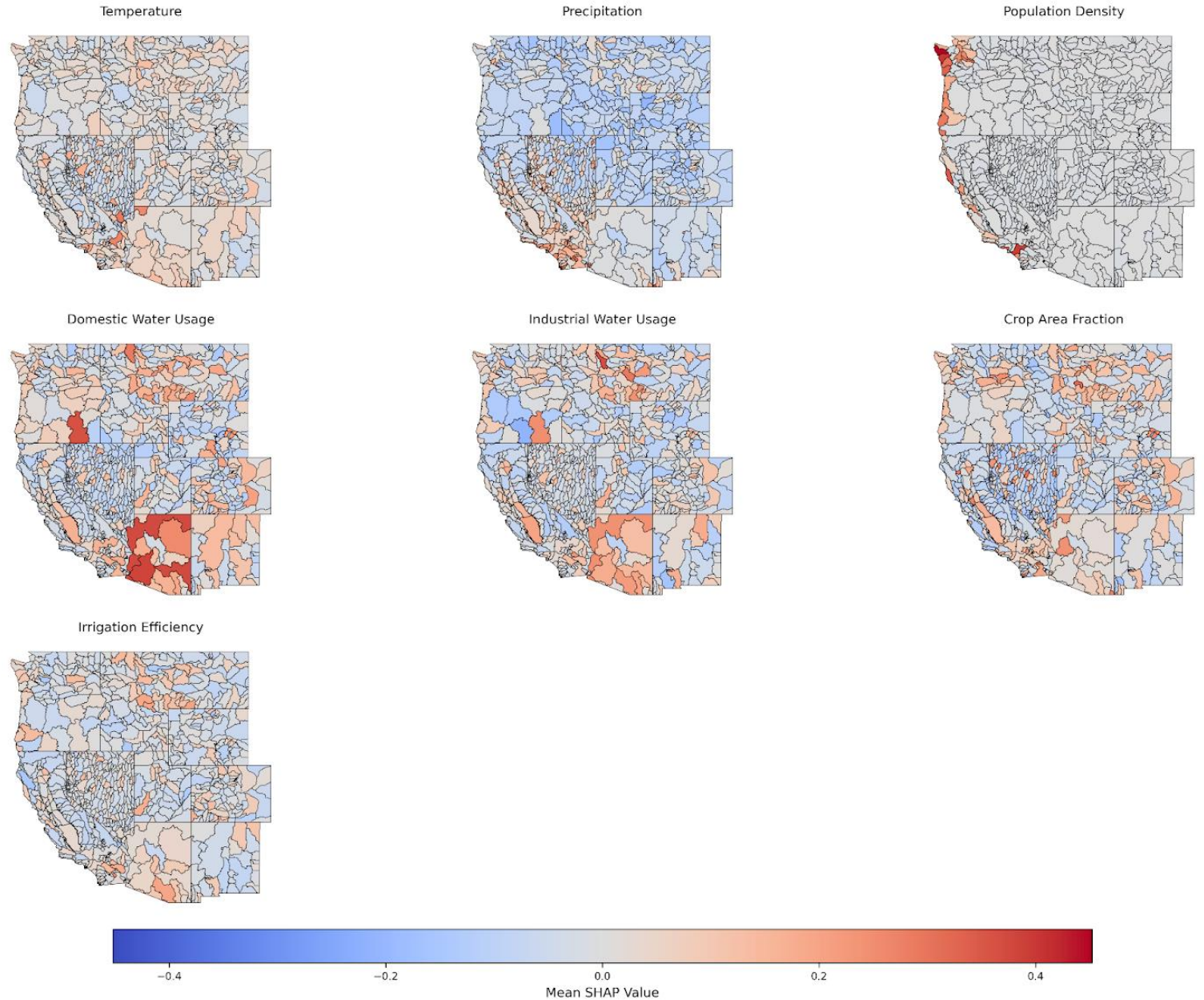


Source: SHAP: A game theoretic approach to explain the output of any machine learning model. (2025, April 17). GitHub. <https://github.com/shap/shap>

Results: SHAP Analysis – Cont.

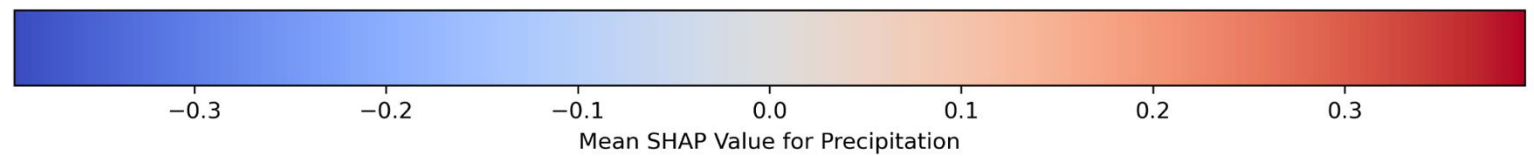
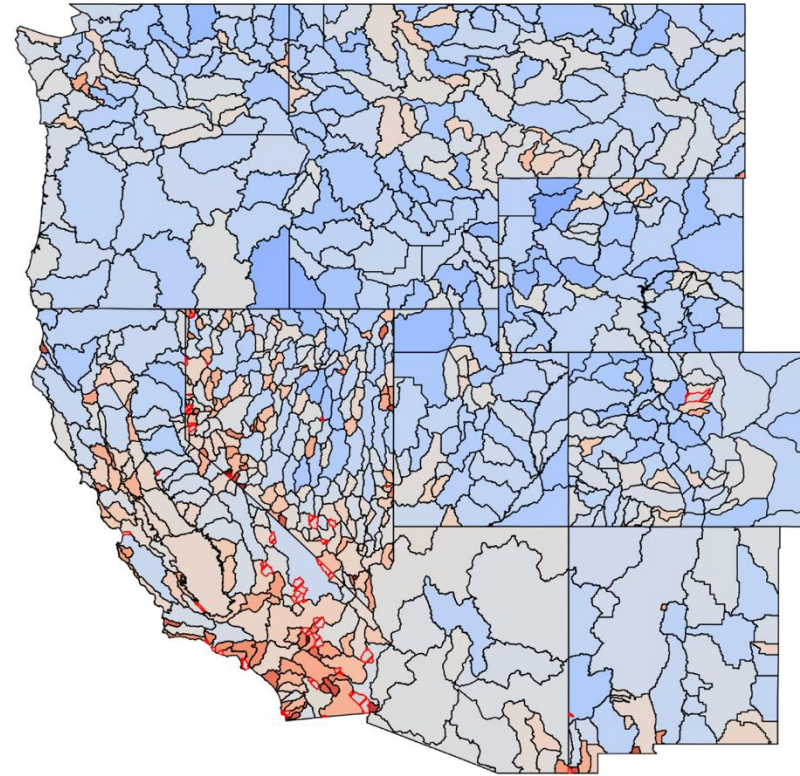


Results: SHAP Analysis – Cont.



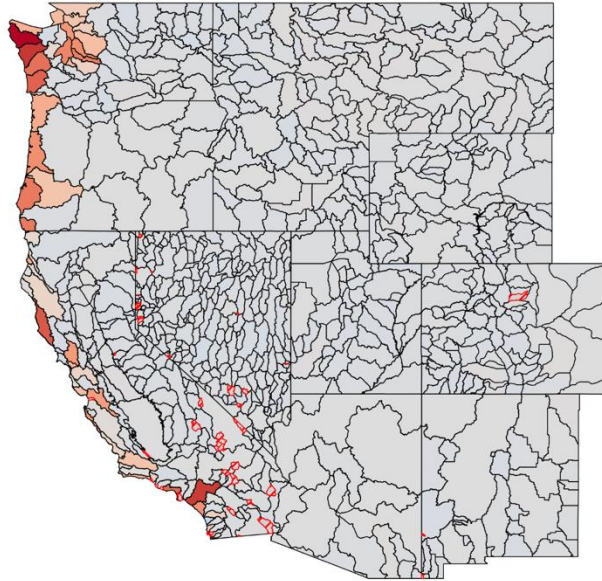
Results: SHAP Analysis – Cont.

Mean SHAP Values for Precipitation Across WMAs

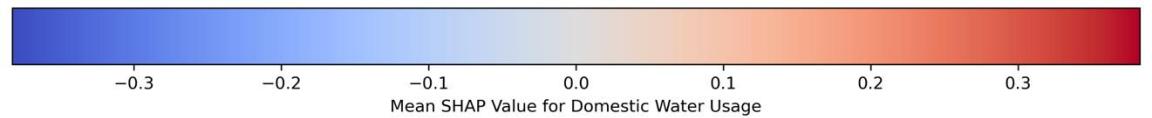
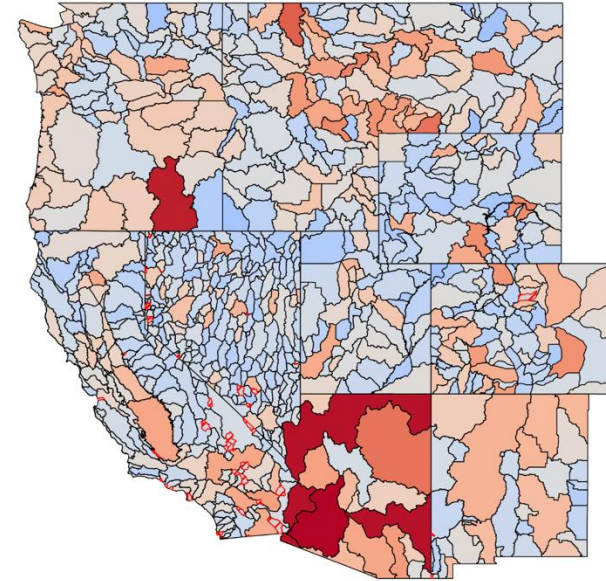


Results: SHAP Analysis – Cont.

Mean SHAP Values for Population Density Across WMAs

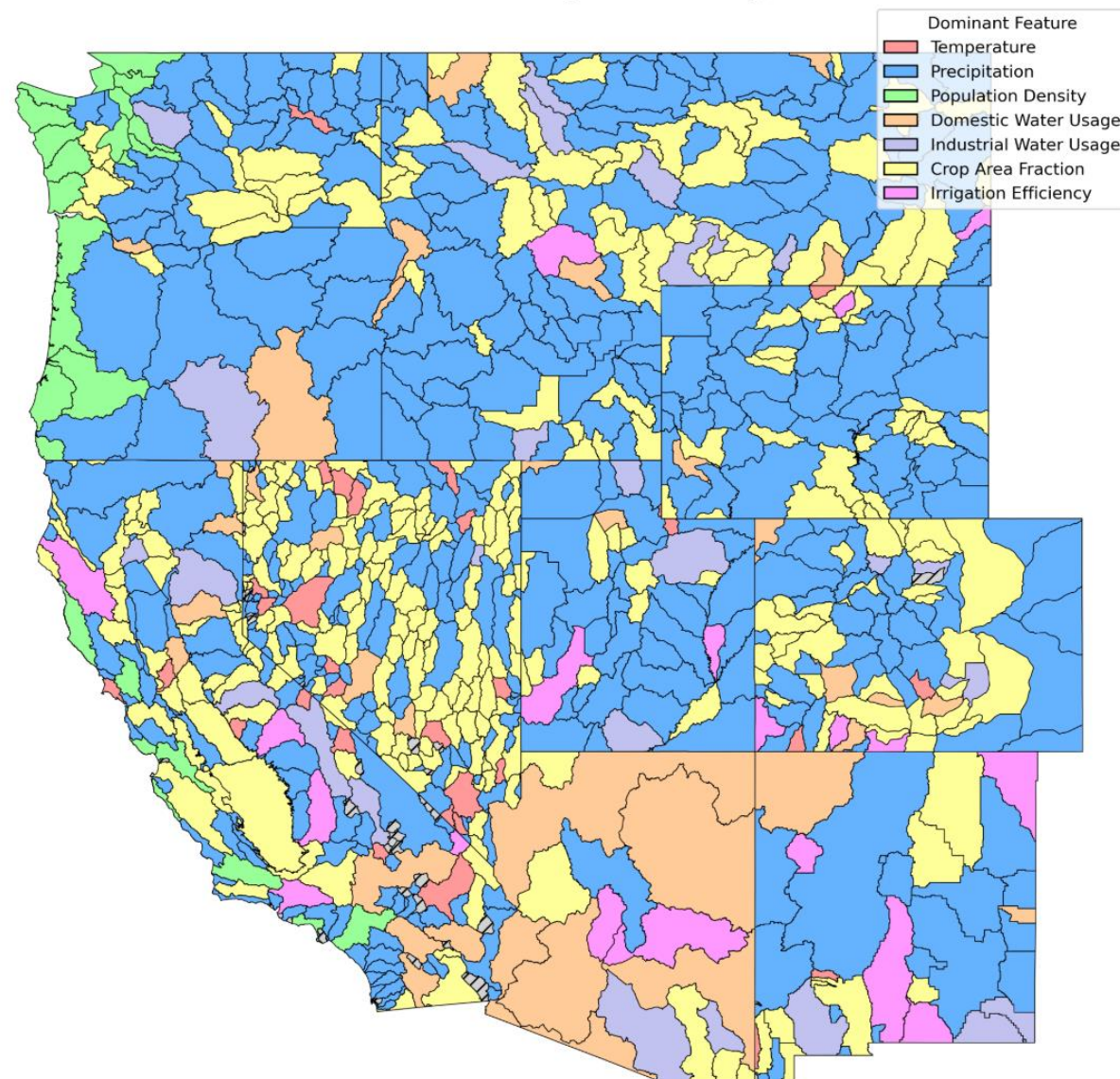


Mean SHAP Values for Domestic Water Usage Across WMAs



Results: SHAP Analysis – Cont.

SHAP Dominant Feature by WMA: Exact_Partition

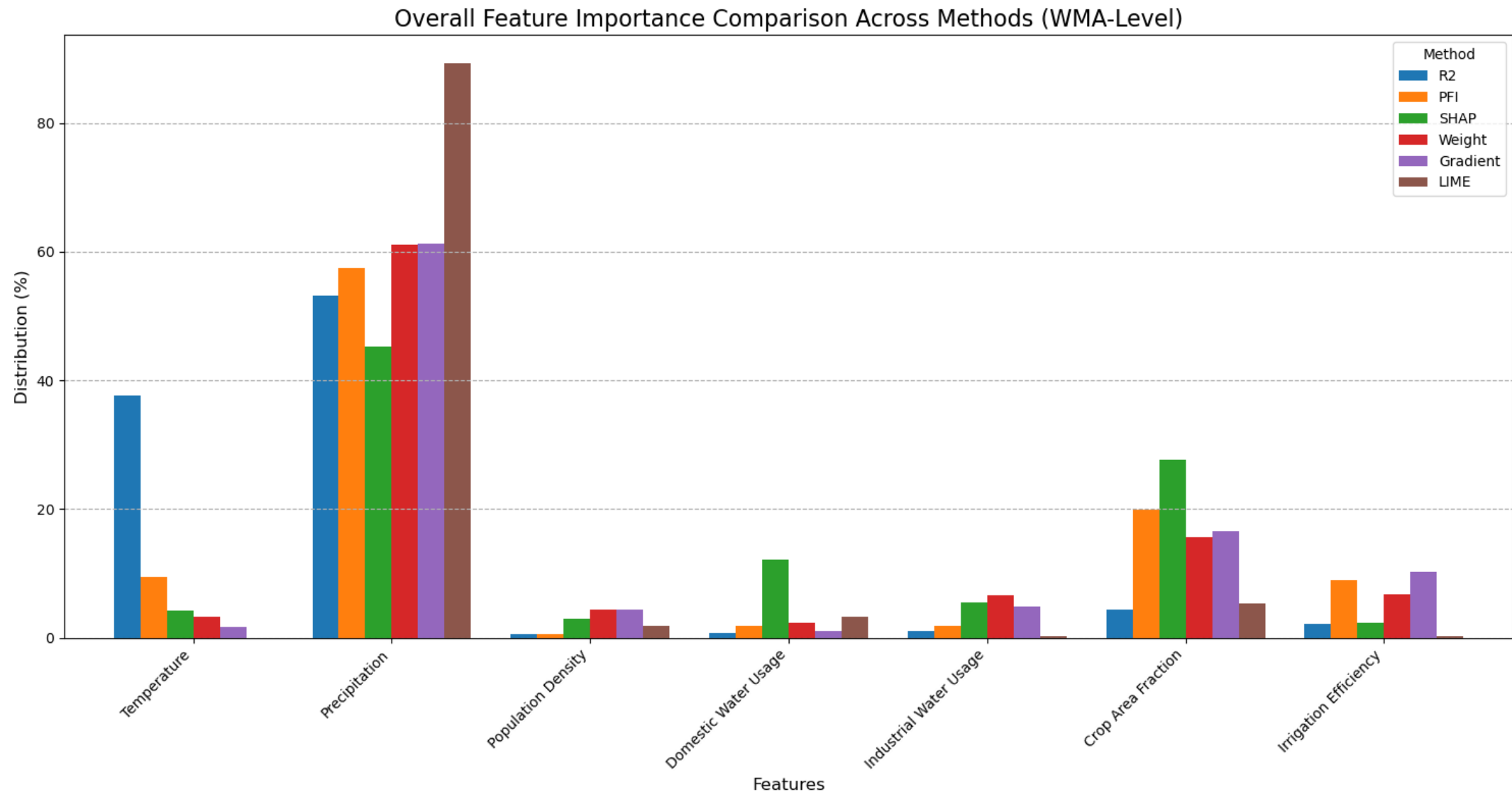


Research Questions

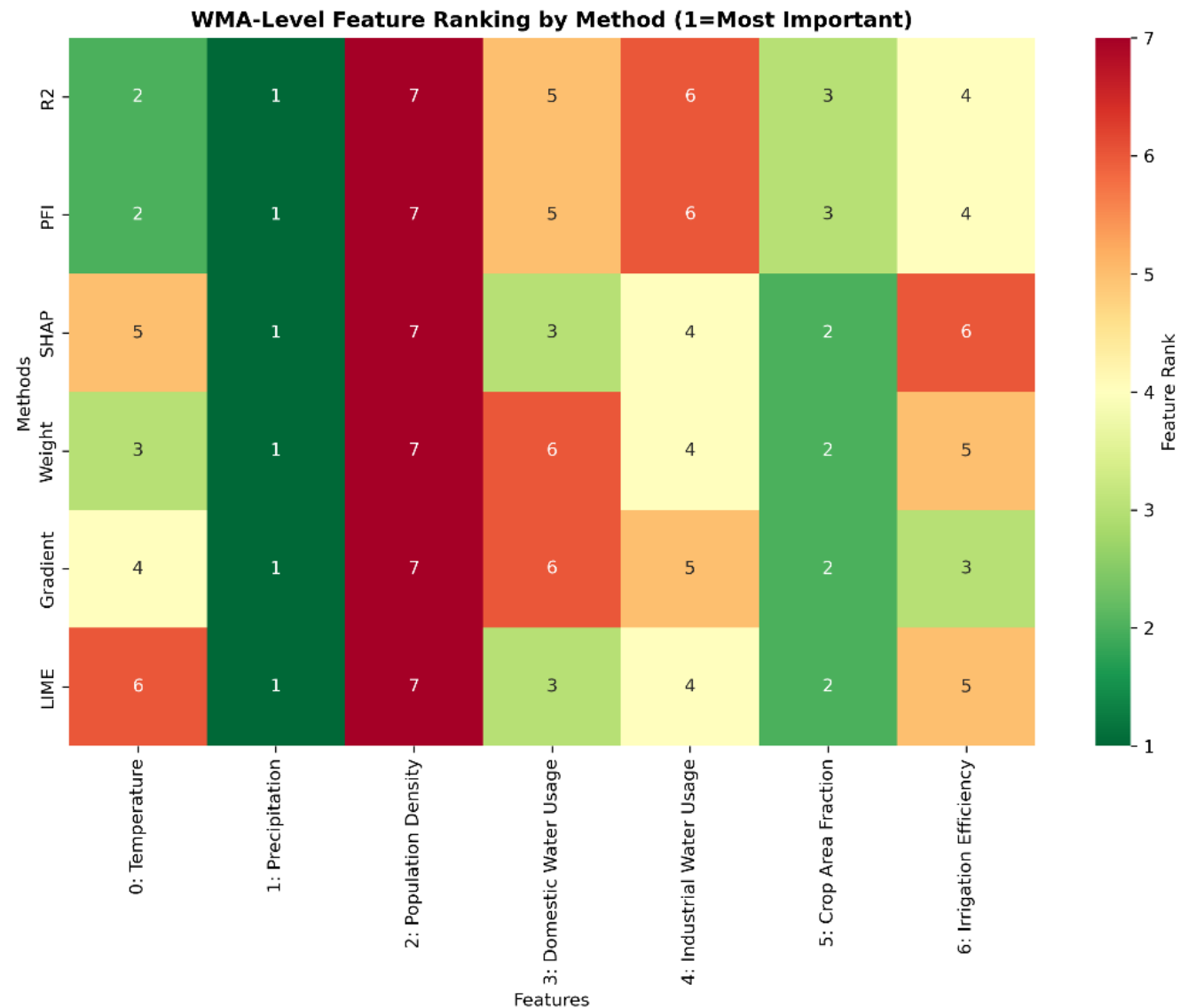
1. Can a deep learning emulator **accurately reproduce** water scarcity simulations from the process-based hydrologic model?
2. Using the emulator, what factors (climate vs socioeconomic) are **most important** for causing water shortages in different regions?
3. Do XAI methods **agree** on driver importance for the same emulator, and can we map their consensus to quantify **explanation uncertainty** spatially?

⇒ *Moving beyond "black-box" prediction to transparent, uncertainty-aware attribution*

Results: Attribution of Dominant Drivers

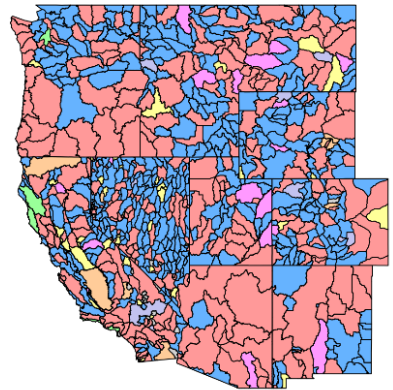


Results: Attribution of Dominant Drivers – Cont.

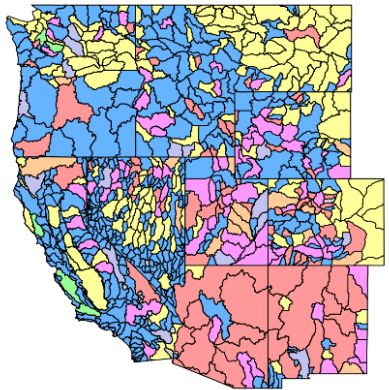


Results: Attribution of Dominant Drivers – Cont.

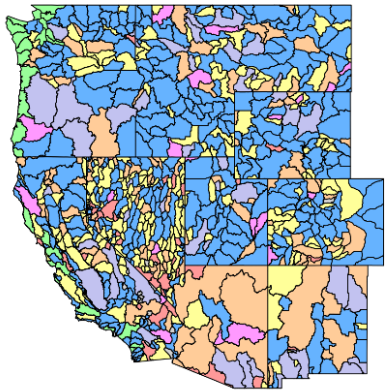
R2
7/7 features, 0.78 avg confidence



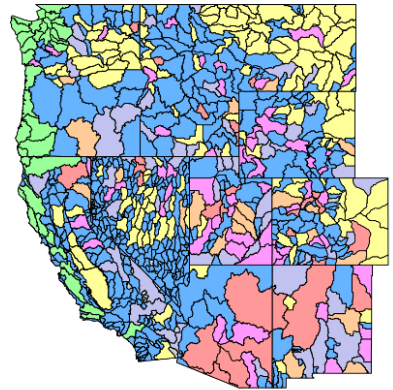
PFI
7/7 features, 0.88 avg confidence



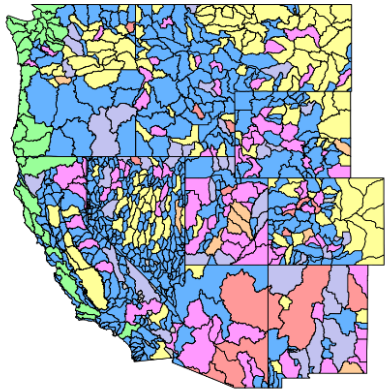
SHAP
7/7 features, 0.86 avg confidence
(Exact Partition)



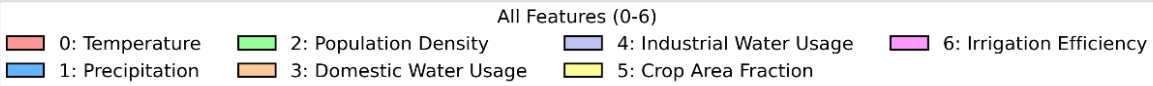
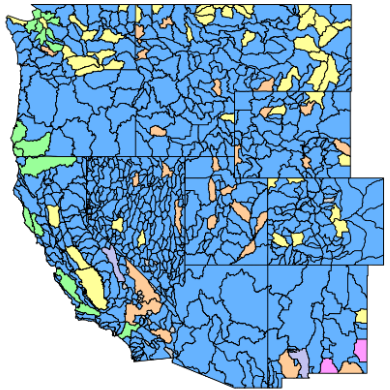
Weight
7/7 features, 0.89 avg confidence



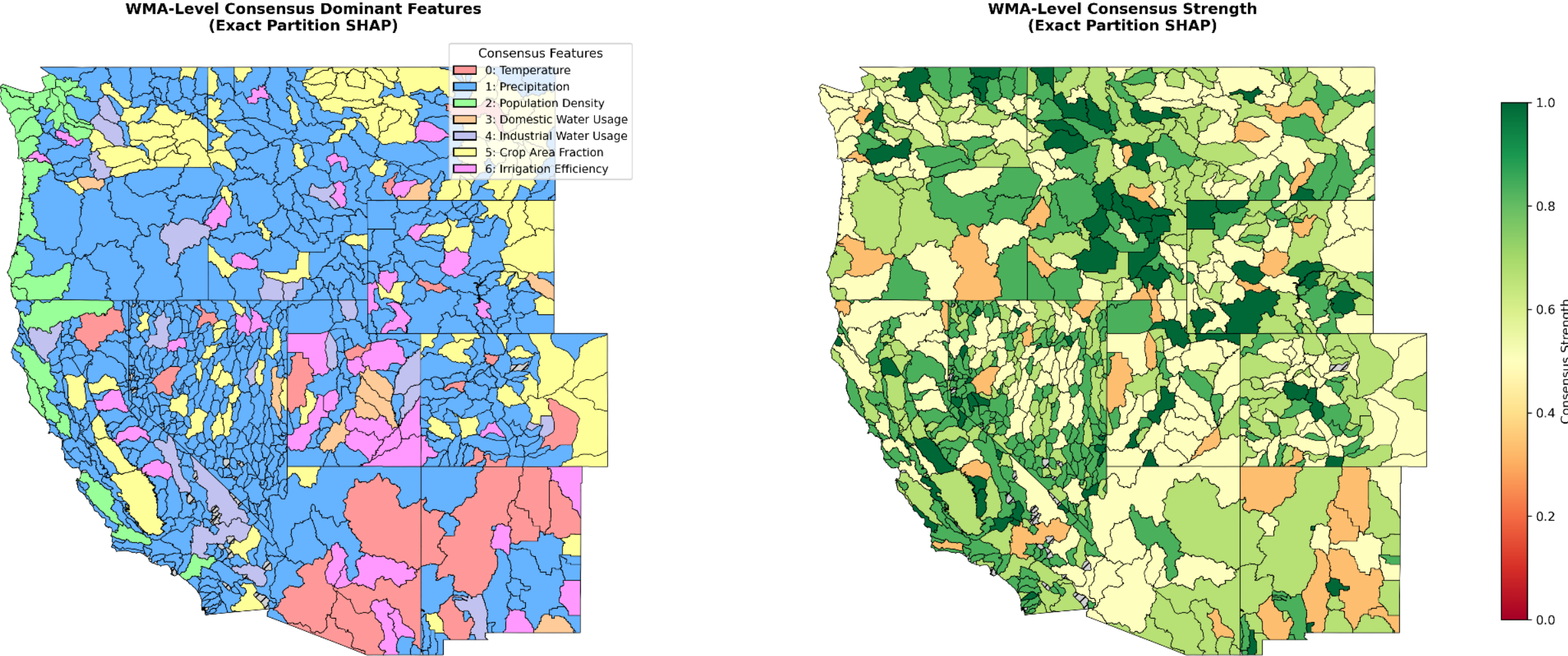
Gradient
7/7 features, 0.90 avg confidence



LIME
6/7 features, 0.91 avg confidence



Results: Consensus Analysis & Explanation Uncertainty



Outline

- Motivation & Background
- Research Questions
- Methods:
 - Processed-Based Model (WBM)
 - Deep Learning Emulator (ConvLSTM)
 - Explainable AI Ensemble
- Interactive Demo
- Results:
 - Emulator Performance
 - SHAP Analysis
 - Attribution of Dominant Drivers
 - Consensus Analysis
- **Discussion & Conclusions**

Discussion & Conclusions

- **Attribution Summary:**

Water scarcity is robustly climate-driven in mountains, influenced by agriculture in irrigated zones, but highly contested (low consensus) in complex, human-dominated basins.

- **Methodological Insight:**

XAI methods can produce divergent explanations for the same model output due to differing assumptions. Using a single XAI method is risky and may not yield reliable findings. Our ensemble/consensus approach provides a "confidence map" for attribution.

Limitations & Caveats

Model Constraints:

- Emulator inherits WBM structural assumptions
- Different process models would yield different patterns
- Socioeconomic inputs incomplete (missing water rights, groundwater pumping)

Methods:

- Post-hoc explanations are associational, not causal
- Each XAI method makes strong statistical assumptions
- No ground truth for validation in real systems

Data:

- 2010-2019 only (doesn't capture long-term trends)
- Monthly resolution (misses sub-monthly dynamics)
- Random split inflates performance metrics

Key Takeaways

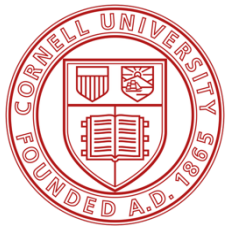
- **DL emulator enables attribution at scale:** The ConvLSTM closely matches WBM behavior and is far more efficient, making broad XAI studies practical.
- **Regional driver patterns:** Precipitation leads in the Pacific Northwest and Rockies, agriculture in the Central Valley and Plains, temperature effects concentrate in the Southwest, and population density is most important along the Pacific coast.
- **Methods disagree—map consensus to guide action:** Different XAI techniques highlight different “dominant” drivers, so building a consensus provides a necessary "confidence map" for attribution.

Acknowledgement



PCHES

Program on Coupled Human and Earth Systems



Center for
Research & Innovation



AR | C
IN

Computing



Graduate School



**ANAL &
SYSTEMS LAB**



U.S. DEPARTMENT OF
ENERGY

Office of
Science



Want to learn more 🤔

Come to my A-Exam!

Date: November 11 (Tuesday)

Time: 10:30 AM

Location: 202 Hollister Hall