Exploring continental-scale relationships between stream temperature signatures and watershed characteristics

Jeffrey Wade[1], Christa Kelleher[1], David M. Hannah[2]

(1) Department of Earth and Environmental Sciences, Syracuse University, Syracuse, NY, United States
(2) School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

Contact: jswade@syr.edu

H25X-1295

Predicting and Understanding Stream Temperature Behavior

Stream temperature is often referred to as the 'master' water quality variable, because it closely controls several instream processes, including the rate of solute processing, the viability of micro- and macrofauna habitats, and ecosystem services[1]. As changes in global climate conditions cause perturbations in energy and hydrological inputs to streams, it is likely that stream temperature behavior with be altered[2], with potential negative consequences to the overall health of river ecosystems. To better tailor thermal management solutions to expected future impacts, it is valuable to investigate how the dominant controls on stream temperature regimes in the United States change seasonally and temporally.

What are the key controls on stream temperature (ST) regimes in the conterminous US across seasonal and spatial scales?

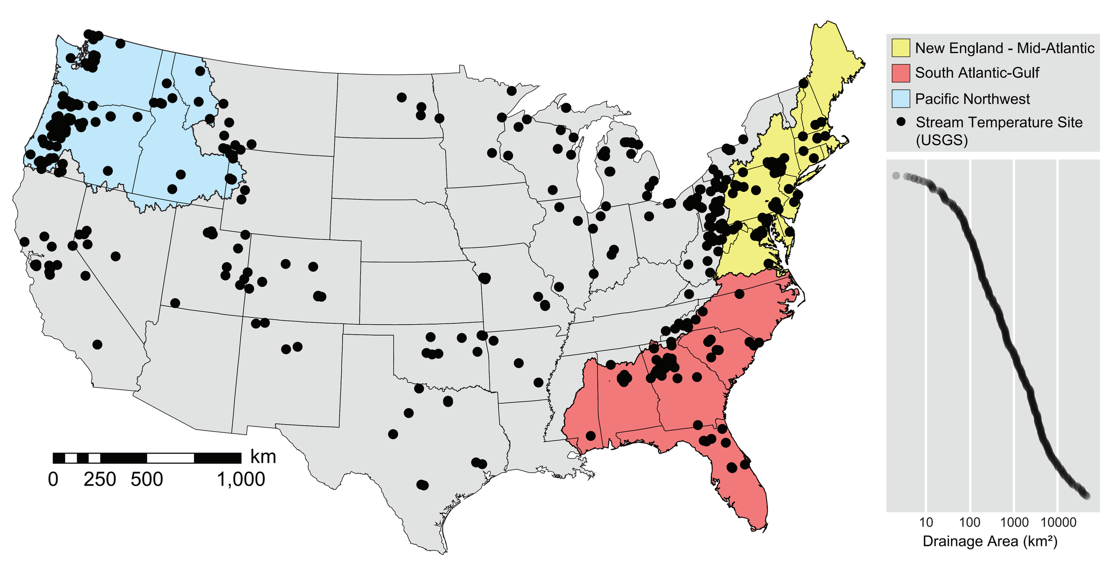
Objective 1: Use supervised machine learning methods to predict metrics of stream temperature regimes from a publicly available dataset.

Objective 2: Interpret model structure using process-based knowledge to assess drivers of stream temperature variability.

Model Inputs: Sites, Predictors, and Metrics

Sites:

- 412 USGS gaging stations in the conterminous US with relatively complete (90%) records of daily discharge and stream temperature from 2016 to 2020.
- Selected watersheds span a broad range of watershed sizes (2 - 50,000 km²), geographic settings, and degrees of human influence.

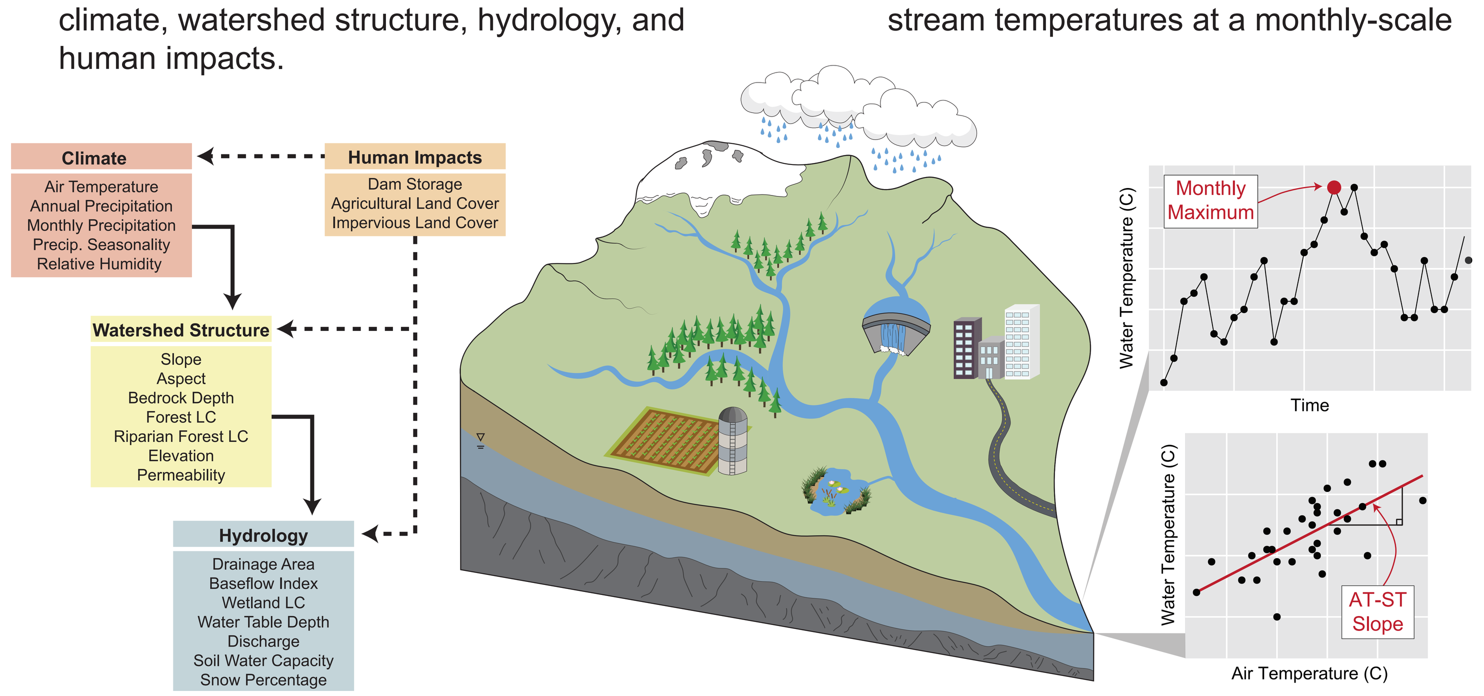


Predictors:

- 22 quantitative variables from the GAGES-II dataset[3] with a strong potential to interact with and shape stream temperature regimes.
- Predictors grouped into four categories: climate, watershed structure, hydrology, and human impacts.

Target Metrics: (Averaged across 4 years)

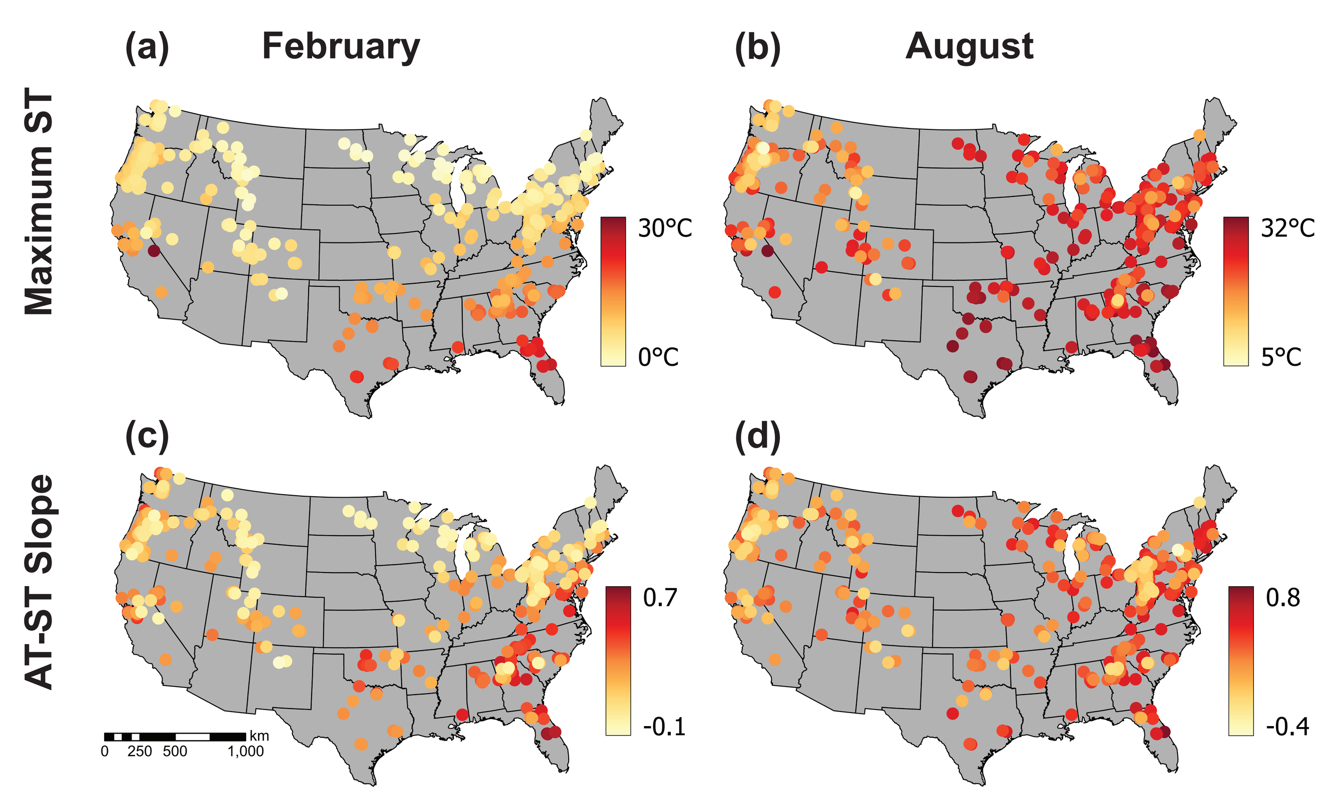
- **Maximum ST** - warmest stream temperature in each month
- **AT-ST Slope (Thermal Sensitivity)** - slope of linear regression between daily air and stream temperatures at a monthly-scale



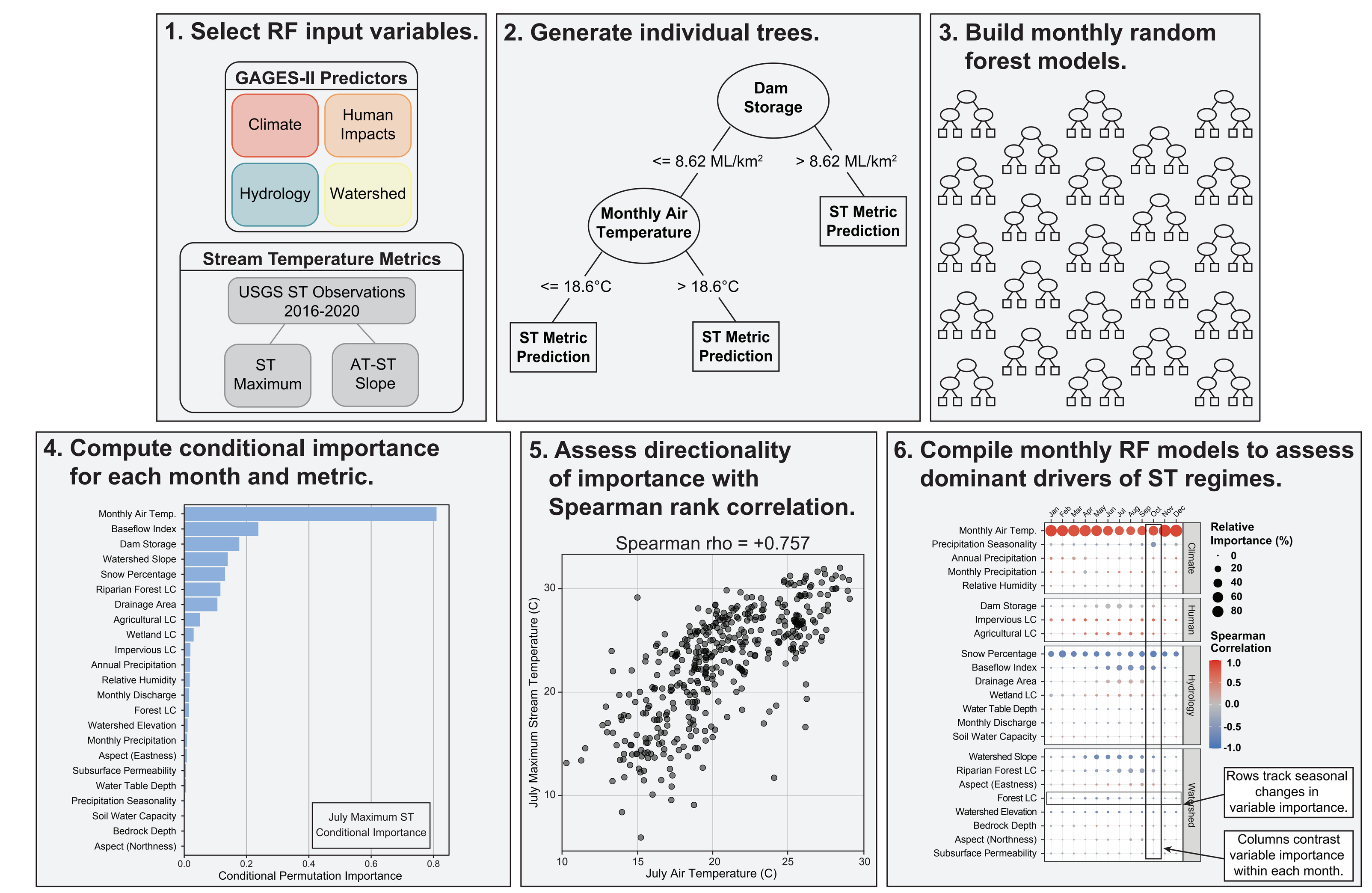
References:
[1] Caissie, D. (2006). The thermal regime of rivers: A review. Freshwater Biology, 51(8), 1389–1406. https://doi.org/10.1111/j.1365-2427.2006.01597.x
[2] Isaak, D. J., Wollrab, S., Horan, D., & Chandler, G. (2012). Climate change effects on stream and river temperatures across the northwest U.S. from 1980-2009 and implications for salmonid fishes. Climatic Change, 113(2), 499–524. https://doi.org/10.1007/s10584-011-0326-z
[3] Falcone, J. A., Carlisle, D. M., Wolock, D. M., & Meador, M. R. (2010). GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States. Ecology, 91(2), 621–621. https://doi.org/10.1890/09-0889.1

Spatiotemporal Variability in ST Metrics

During winter months, both Maximum ST and AT-ST Slope varied along a latitudinal gradient, with cool and insensitive water temperatures in the North and West and warm and atmospherically-coupled water temperatures in the South. In summer, strong regional patterns were superseded by increased basin-scale heterogeneity in both metrics.

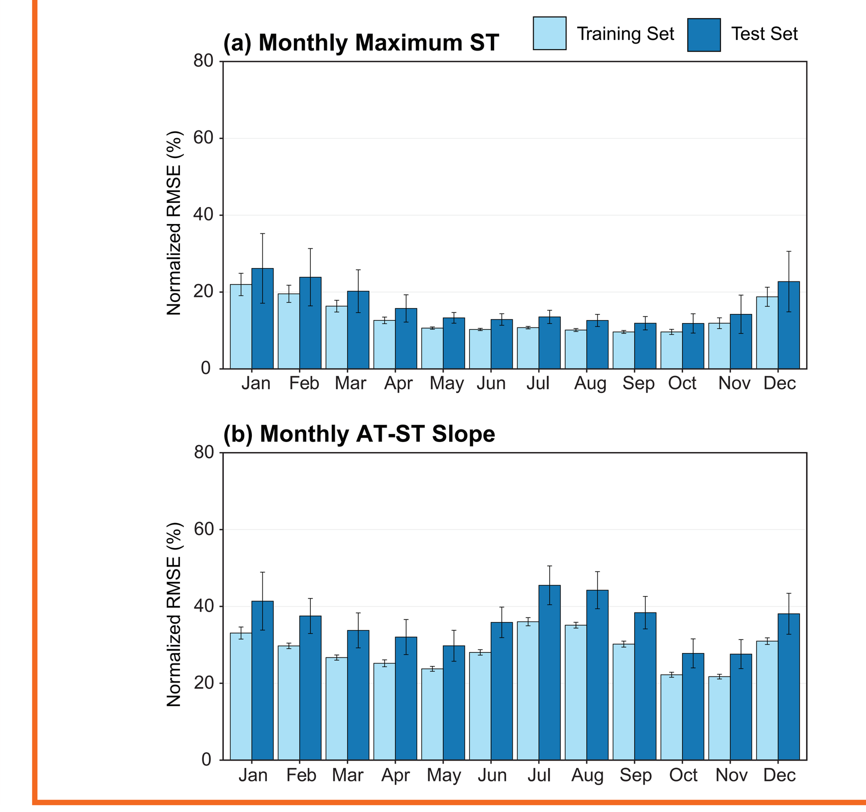


Modeling Stream Temperatures: A Random Forest Approach



Validating Random Forest Predictions

Monthly random forest models were fitted using a geographically-stratified training set (80%) and validated with a test set (20%). Five-fold cross validation was used to tune RF hyperparameters, including *mtry* and *ntree*. RF models of Maximum ST tended to have lower error than models of AT-ST Slope, particularly in the summer.



Geographic and Temporal Contrasts in Dominant Stream Temperature Controls

"Black Box" Models → Human-interpretable Results

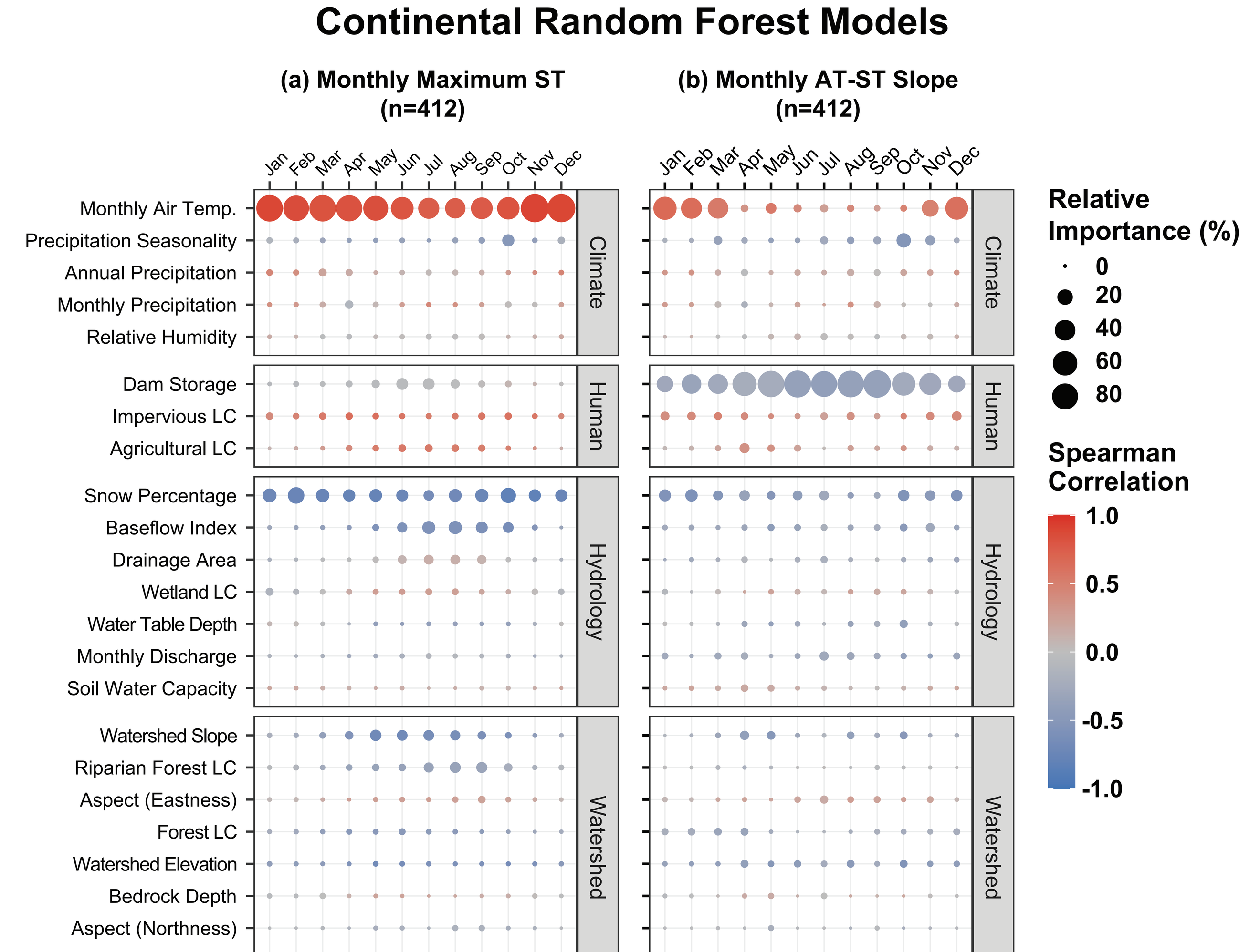
Permutation Variable Importance

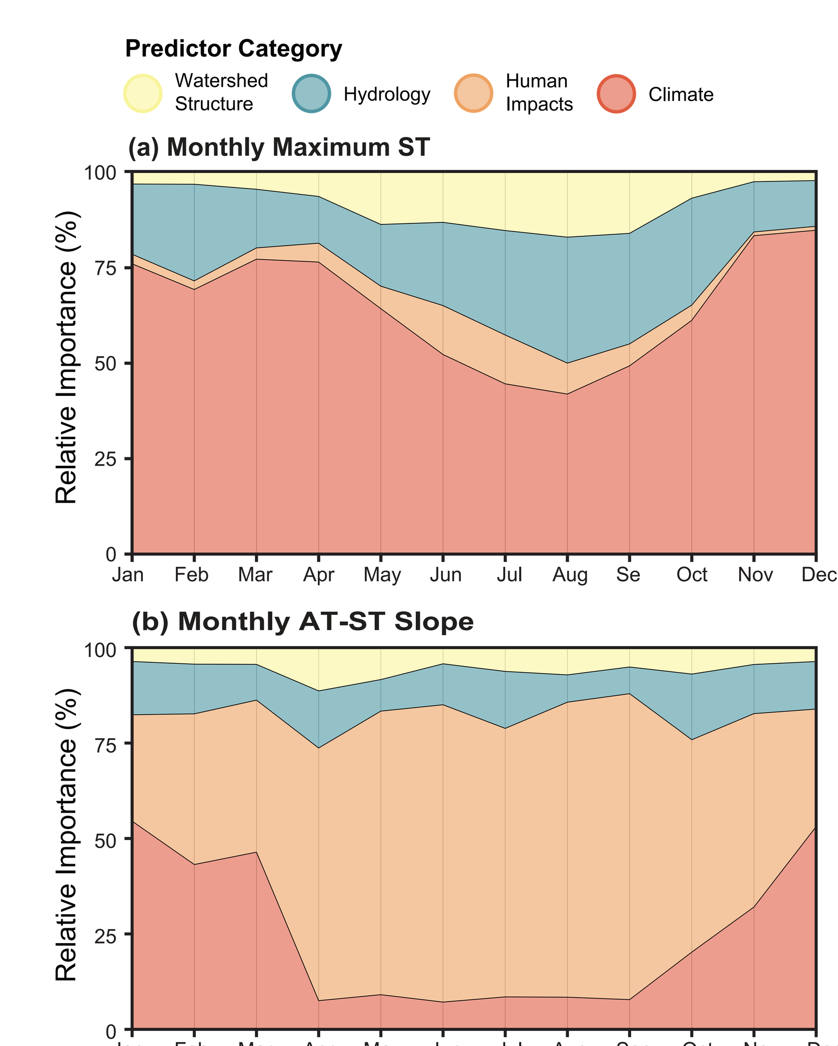
- Used to rank predictor variables based on their relative utility in explaining variability in model response
- Conditional variable importance accounts for predictor multicollinearity

Continental-scale models display seasonal shifts in predictor importance of ST metrics

- **Climate** a clear control on Max ST, but mediated by hydrologic and watershed characteristics in summer
- **Dam storage** a surprisingly important predictor of AT-ST Slope at a continental-scale
  - Dams tend to cause insensitive downstream water temperatures, representing a potential buffer against atmospheric effects

Continental Random Forest Models





Aggregate importance across variable categories to infer general seasonal patterns.

Geographic subset models emphasize diversity in regionally important predictors of ST magnitude

- While **climate** is an expected control in all regional models, each geographic subset has **unique secondary predictors** of Max ST
  - **New England/Mid-Atlantic** - Dam Storage, Snow Percentage, Elevation
  - **South Atlantic-Gulf** - Wetlands, Water Table Depth, Elevation
  - **Pacific Northwest** - Agricultural Land Cover, Snow Percentage, Riparian Forest Land Cover

Regional Random Forest Models (Monthly Max ST)