Projecting surface water change in the Southeastern U.S. under different scenarios

Mollie D. Gaines[a], Mirela G. Tulbure[a,b], Vinicius Perin[a]

[a] Center for Geospatial Analytics, North Carolina State University | [b] Department of Forestry and Environmental Resources, North Carolina State University
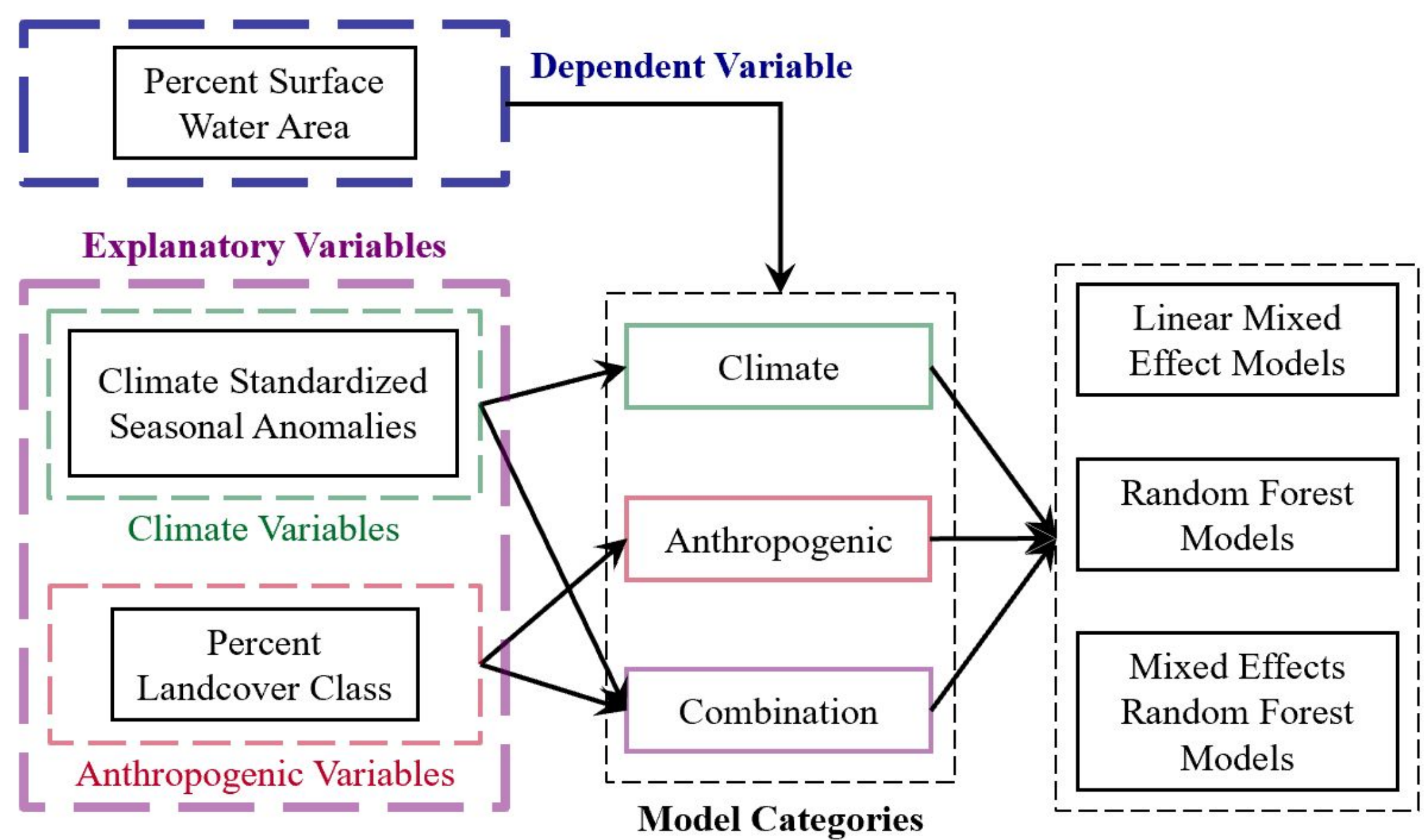
Why the Southeastern U.S.?

- Most land cover and land use change of any region in the U.S.
- Highest population growth rate in the U.S. leading to increased water demand
- Sprawling urban growth pattern increases impervious surface area
- Variable precipitation patterns that are project to become more seasonal

Objectives

- Determine what drivers - human, climate, or both - have the greatest impact on estimating surface water area based on satellite observation and train a machine learning model using these drivers.
- Use the machine learning model to predict surface water area through 2100 based on different climate and land cover/land use projections

Methods

- Aggregated Landsat-derived Dynamic Surface Water Extent product to seasonal surface water area per 8-digit HUC
- Aggregated climate and land cover data: maximum temperature and precipitation, and natural, intensive, and agricultural
- Train and test three data-driven statistical models with three different sets of explanatory variables
- Generated future climate and land cover data from annual trends
- Run the most accurate statistical model with simulated future explanatory variables
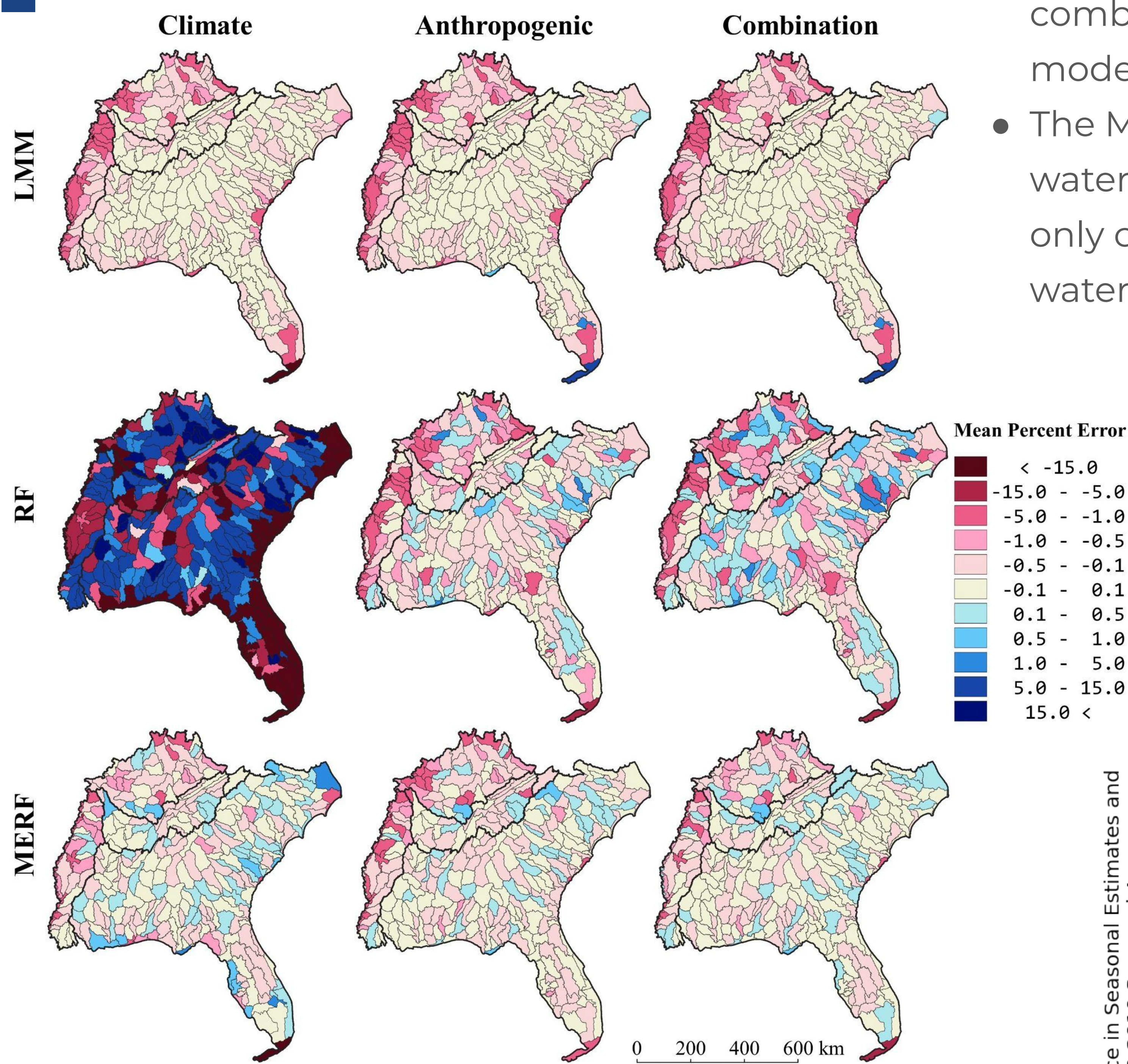
Main findings

Figure 1 Overall Mean Percent Error at the HUC level for all 9 models. Pink HUCs indicate an underestimation of percent surface water area with the magnitude of the underestimation increasing with the hue. Blue HUCs indicate an overestimation of percent surface water area with the magnitude of the overestimation increasing with the hue..

- For each statistical model, the combination of climate and human drivers produced the smallest absolute median watershed-level mean percent error
- Errors are not uniformly spatially distributed – consistent underestimation in the Lower Mississippi Water Resource Region
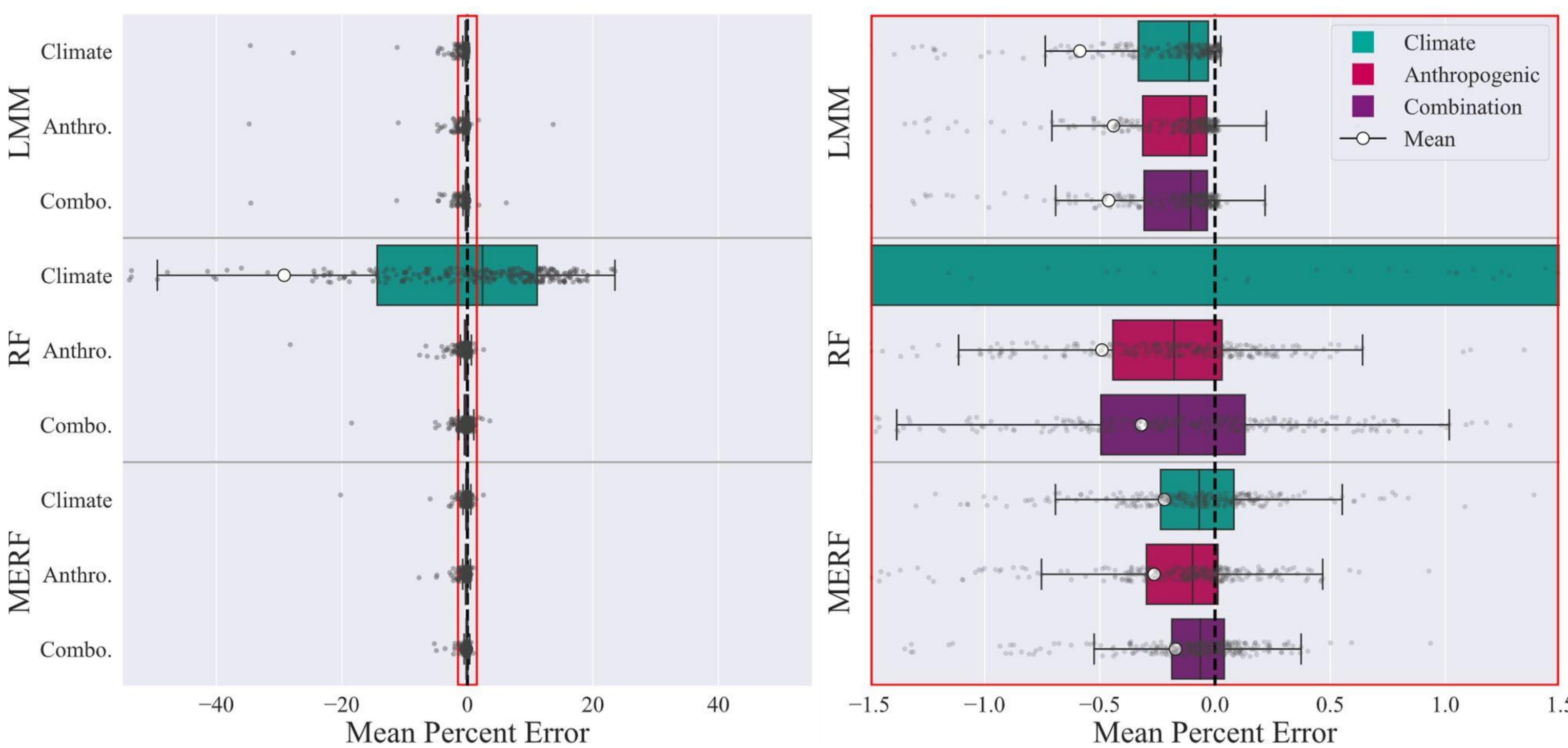

Figure 2 Boxplots of the mean percent error at the HUC level for all 9 models. The full range of MPEs, excluding those in climate RF <-55%, are shown on the left. The distribution of MPEs between -1.5 and 1.5 (red bounding box on both left and right) are shown on the right. The individual HUC MPEs (transparent grey dots) are shown on both left and right.

- The Mixed Effect Random Forest (MERF) model with a combination of climate and human drivers was the best model for estimating percent surface water area
- The MERF model had the smallest absolute average watershed-MPE (-0.17%) while the random forest model with only climate data had the highest absolute average watershed-MPE (~ -27%)
  - The MERF model is able to predict future percent surface water area (using simulated climate and land cover data)
  - Simulated predictions show a consistently higher percent surface water area at the seasonal, watershed scale
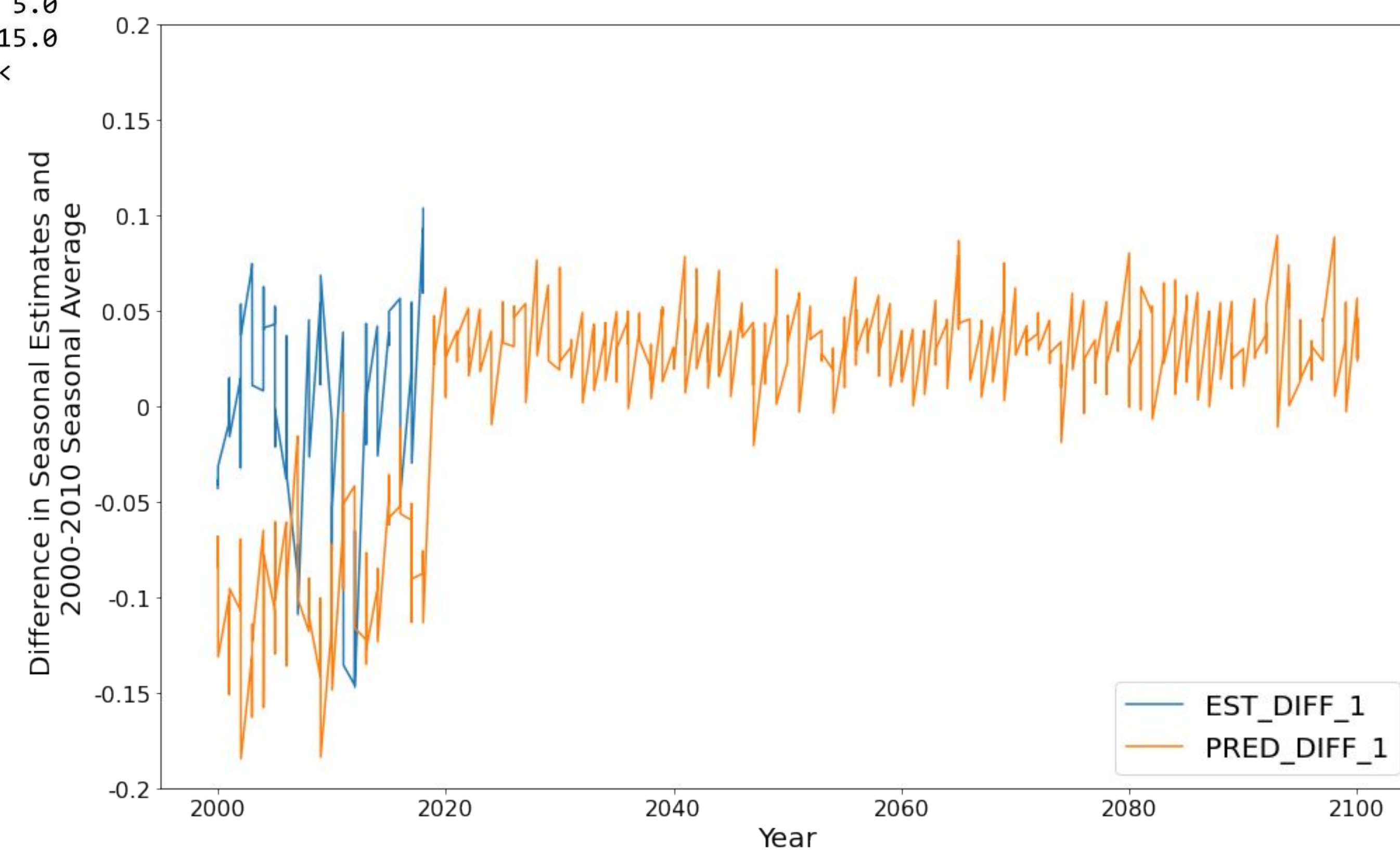

Figure 3. Comparison between estimated percent surface water area (blue) and projected surface water area (orange) differences from the 2000-2010 seasonal average, where the projected surface water area was calculated based on a set of simulated climate and land cover variables.

Next Steps

- Run trained MERF model on land use/land cover data projected under SSP2-RCP4.5 and climate data projected under RCP4.5
- Run trained MERF model on land use/land cover data projected under SSP5-RCP8.5 and climate data projected under RCP8.5
- Compare MERF projections to the Water Supply Stress Index surface water supply

Acknowledgement