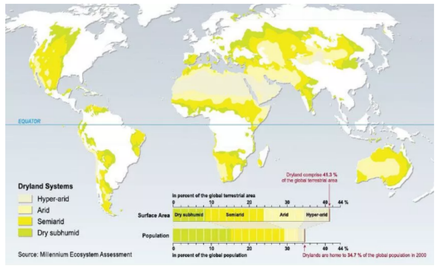
# Leveraging high resolution classifications for hindcasting decades of water resources dynamics in the Landsat time series with machine learning regression

Nick Kolarik, Nawaraj Shrestha, Jodi Brandt
*Boise State University*

1. **Machine learning regressions are a better choice for the mixed pixel problem than a traditional SMA.**
2. **A model built with recent training data can accurately estimate proportions in past and future images, thus "extending" higher resolution products temporally.**
3. **Longer time series can be analyzed statistically to identify major shifts in mesic ecosystems due to management and disturbance.**

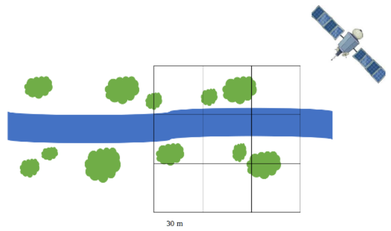
## Motivation

Dryland systems comprise >40% of the Earth's surface. Their ecological functions are at risk due to threats to the water resources (surface water, mesic vegetation) that sustain them.

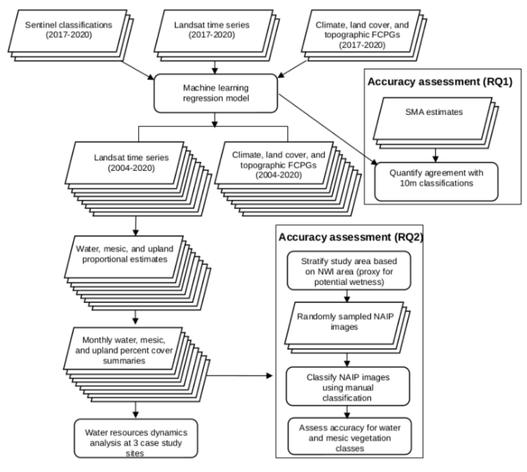



## Problem

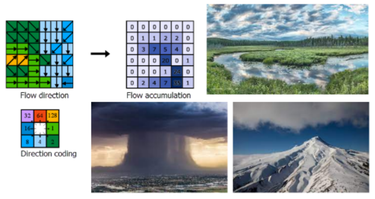
Earth observation datasets meant to monitor water resources are generally ineffective in dryland systems due to the relatively small size of these sustaining features.




## METHODS




We use 10 m Sentinel classifications and ancillary covariates to train a random forest regression to estimate water resources proportions (WRP) in monthly Landsat composites.

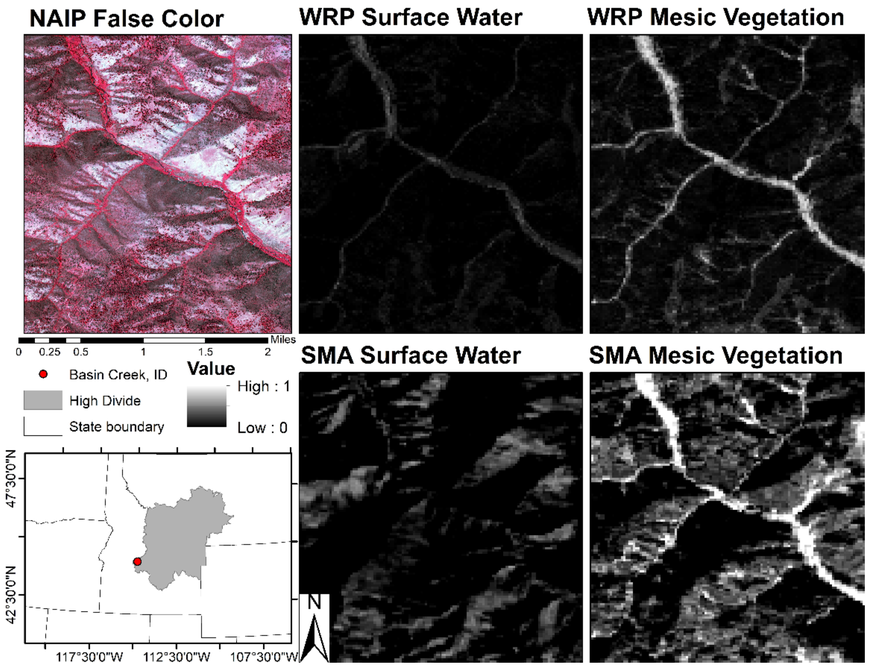



For climate and land cover variables, we used a flow conditioned parameter grids (Barnhart et al., 2021) approach that describes each pixel by the conditions in the contributing area.

### RESEARCH QUESTIONS

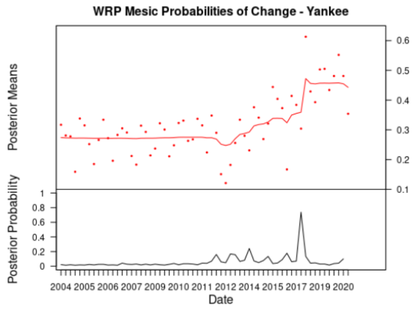
1. Do water resources proportions (WRP) estimated in 30-m Landsat pixels from random forest regression accurately capture water resources compared to 10-m Sentinel classifications?
2. Do random forest regressions capture water resource fractional coverage more accurately than spectral mixture analysis?
3. How accurate are fractional coverage estimates compared to "actual" water resource conditions as measured from high resolution aerial photography?

## RESULTS




A comparison of proportional estimates of surface water and mesic vegetation near Basin Creek, Idaho. Ancillary covariates help to reduce erroneous estimates of water resource areas in the WRP where these resources are unlikely.

Yankee Fork of the Salmon River restoration project led by Trout Unlimited. Beginning in 2012, this project sought to increase channel complexity, riparian vegetation, and fish habitat. Beavers established in 2015 and continued the work.




Mean time series of mesic vegetation extent at the case study site and the probabilities of a change point.

## CONCLUSIONS

- Using a machine learning approach, we can retrieve better estimates of surface water and mesic vegetation fractions in a mountainous semiarid system than by using SMA.
- We build on existing work for quantifying these relatively small resource areas that sustain semiarid systems and their land functions
- We consider this approach to be useful for extending the temporal record high resolution classifications backward, by using RFR and the stalwart Landsat time series to estimate fractional coverage of ecologically important land covers
- The WRP approach considering topographic features, climatic and land cover conditions, along with spectra is useful for estimating fractions of surface water and mesic vegetation in a given study area
- Resulting maps can be useful for quantifying surface water and mesic vegetation changes as a result of restoration activities, disturbance events, or climate change.

FUNDING and ACKNOWLEDGEMENTS:
This work is supported by a NASA FINESST Graduate student fellowship award (20-EARTH20-340).

Contact
nicholaskolarik@u.boisestate.edu