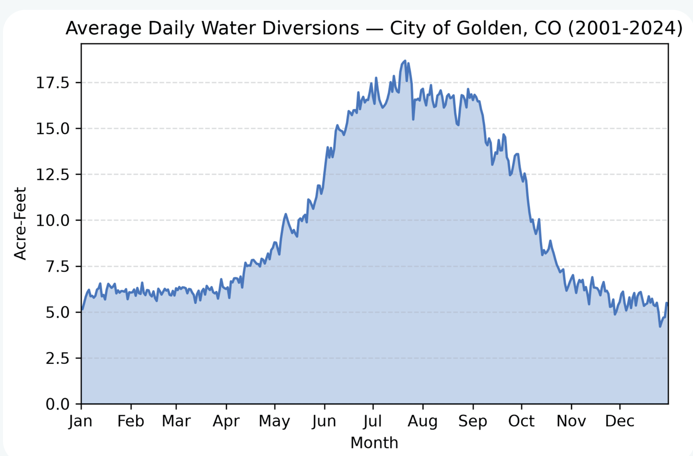
How Do You Prove Water Conservation Actually Works?

OTIS HALLEY GOTWAY[1,2,3,4], NELS BJARKE[1,2,3]

Western Water Assessment[1], University Colorado Boulder[2], Cooperative Institute for Research In Environmental Sciences[3], Research Experience for Community College Students[4]

INTRODUCTION



The west is in a crisis. Record low snowpack, depleted reservoirs, and a statewide drought emergency have made water conservation an absolute necessity. Every drop saved matters.

The City of Golden implemented water conservation measures, but faces a fundamental problem, due to **natural year to year weather variability makes it difficult to evaluate whether concervation efforts are actually saving water.**

To solve this, we build statistical models that predict what water demand would have been given weather conditions. If the observed demand post-conservation measures consistently falls below predictions, conservation is validated.

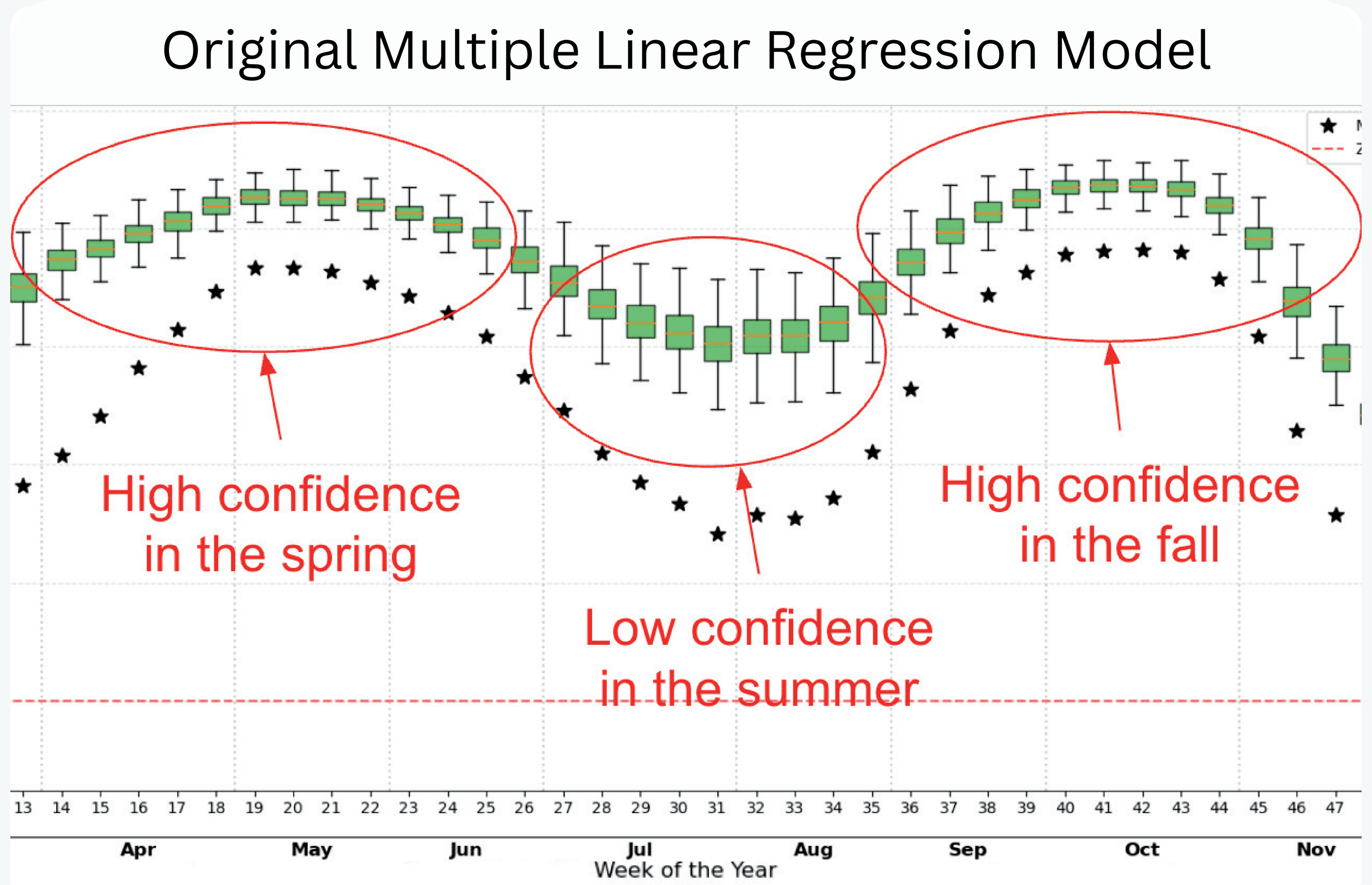
Previous work using MLR models showed **strong predictive skill in spring and fall, but a consistent drop in summer** exactly when demand is highest and accurate prediction matters most.

We investigate whether alternative model architectures can address the summertime drop in skill.

MOTIVATING INFORMATION

Prior modeling work on Golden's water system revealed a persistent pattern. Predictive skill drops sharply during summer, precisely when demand is highest and accurate prediction matters most. Both Multiple Linear Regression and Random Forest models showed this same dip despite being fundamentally different in how they work. **The model performs poorly in the summer when there is the most potental water to save.**

This raised a question: **is this a modeling problem or an observational problem?** This project systematically tests alternative model architectures to find out whether any of them improve summer skill, and if not, to confirm that the path forward runs through better data rather than better methodology.



METHODS

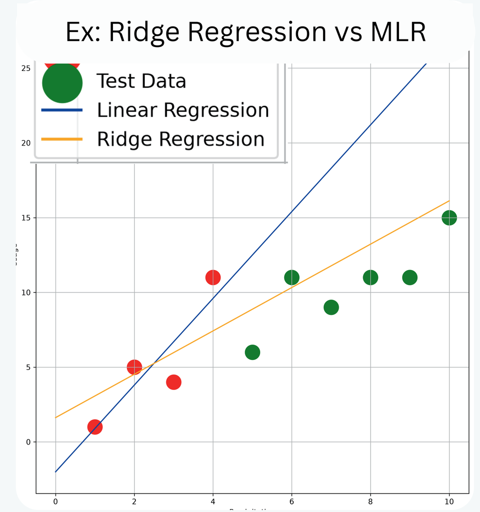
Dataset

1. **Water Diversions**: Weekly volume of water pulled into Golden's system
2. **Precipitation** (PR): Weekly rainfall and snowfall
3. **Maximum Temperature** (TMMX): Weekly peak air temperature
4. **Vapor Pressure Deficit** (VPD): How much more moisture the air can hold
5. **Potential Evapotranspiration** (PET): How much water would evaporate and transpire
6. **Antecedent Lags** (4 weeks): Prior week values of each weather variable. Capture delayed effects of past conditions on current demand

Program

A modular Python pipeline built from scratch handles everything from raw data ingestion to final model output. It cleans daily diversion data, averages gridded weather observations, generates lag features, and removes invalid rows before modeling begins. The bootstrap resampling loop runs 1000 iterations per model, randomly splitting available years into training and testing sets each time. Any statistical model can be dropped in and evaluated using the exact same framework.

Different Model Types



MLR

MLR (Multiple linear regression) is a type of statistical algorithm that can be used to find a line of best fit. This model was chosen as it shows all data being **equally weighted**.

Ridge

Ridge regression works similar to MLR. It still makes a line of best fit, but it also tries to minimize the slope of the line. Ridge regression will provide us with a model that weights smilar data less strongly and **prevents extreme outliers from having an outsized impact**.

Lasso

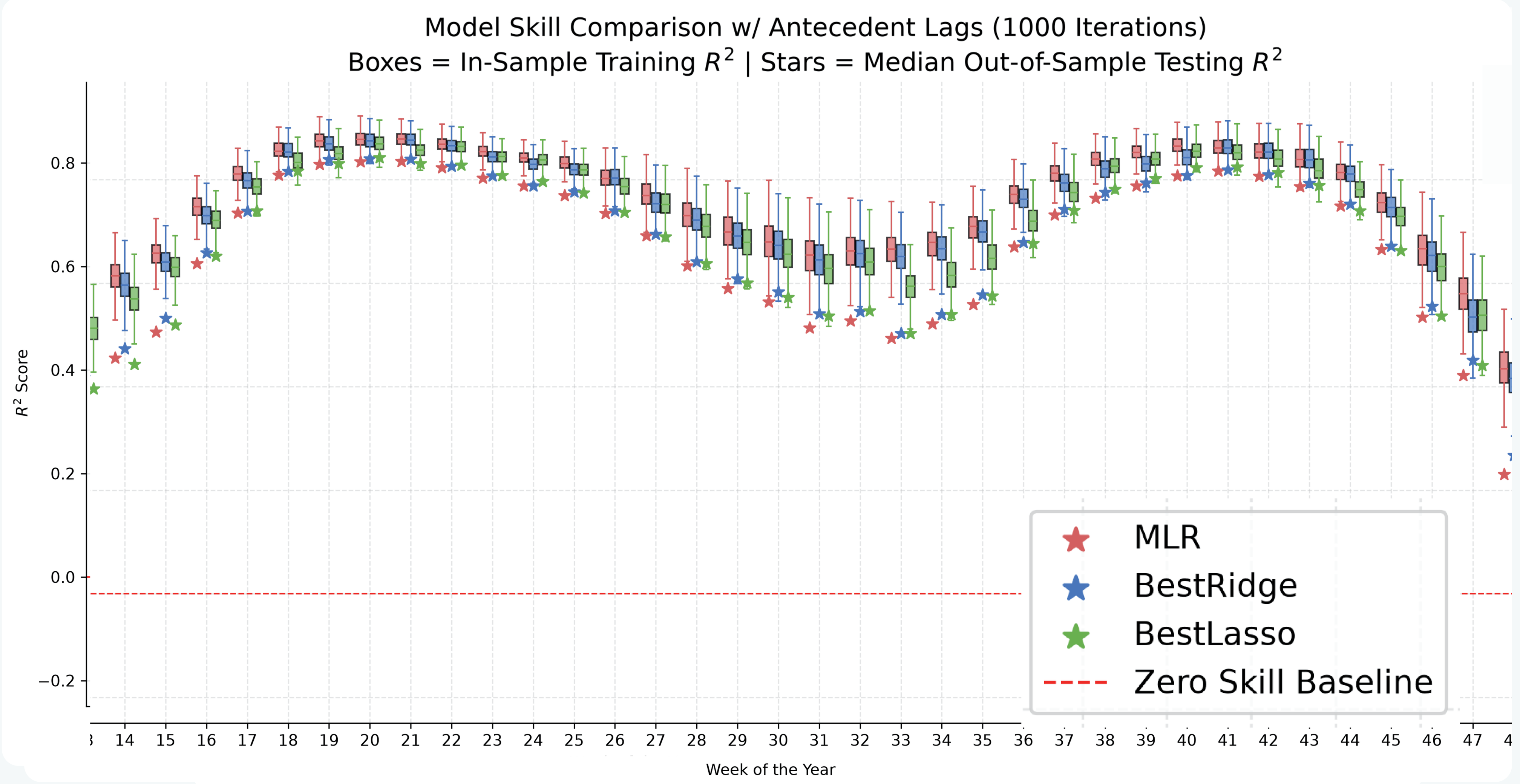
Lasso regression works much like ridge regression with one key difference. Lasso regression allows the model to completely get rid of datatypes that provide it no skill. This allows for **a "simpler" model than ridge regression**

Research question: How do these models compare?

RESULTS

Which model has the most skill week to week

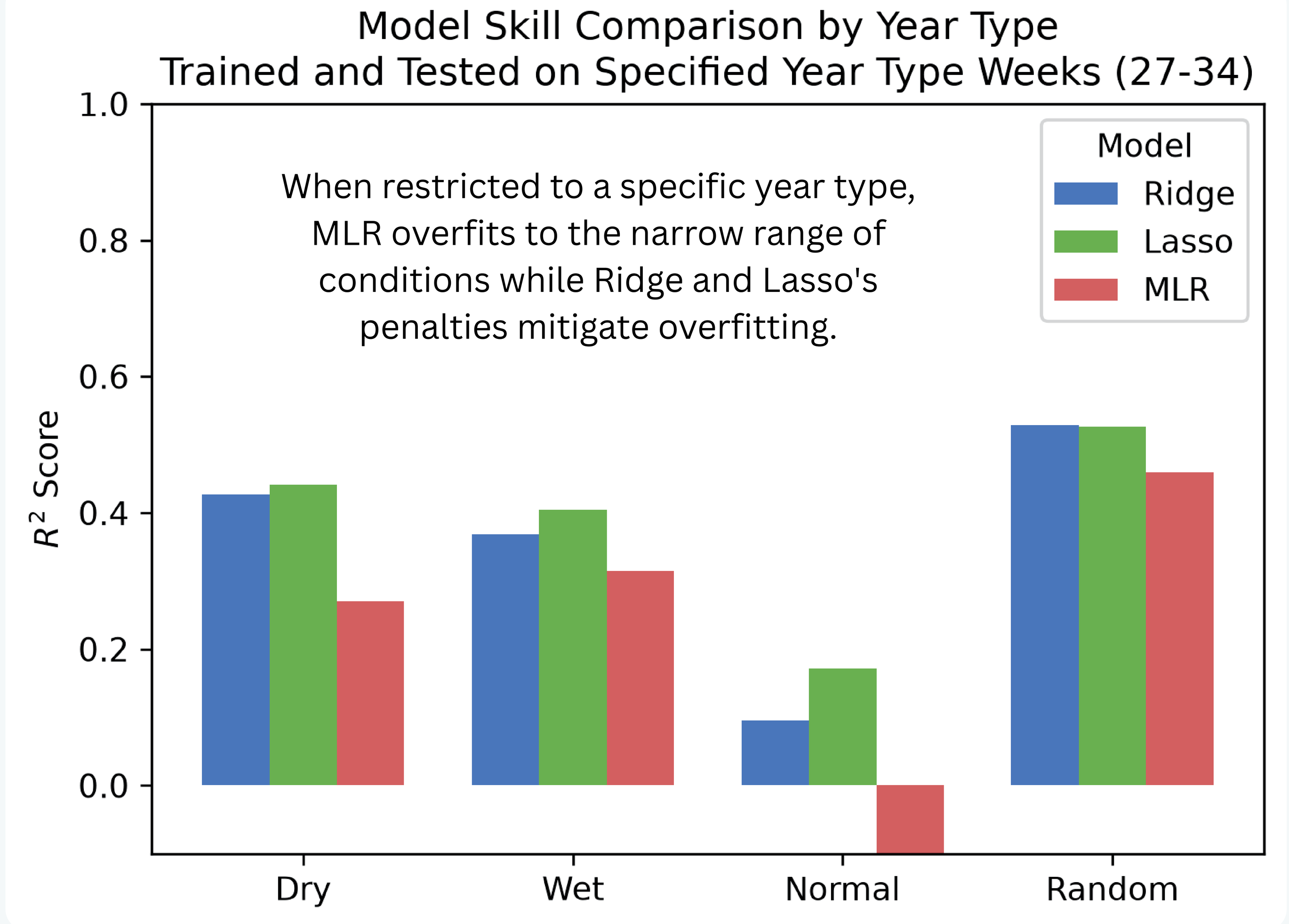
We have found that week to week, in the most water intensive months, **a change in statistical model does not signifigently improve model skill**, with a maximum added skill of 0.05 r^2. This may tell us that there is not much to be gleaned from changing model structure and that any major improvements should come in the form of more data. This is likely because the **summer predictability gap is driven by individual behavioral decisions around outdoor water use that weather data alone cannot capture**, regardless of model type. However, ridge regression comes out as the model type with the most skill



MORE RESULTS

Another question: Do the models perform differently during wet vs dry years?

We trained and tested models on years of different wetness, **normalized models outperform MLR across the board**. This process illuminates the challenges of different sample sizes.

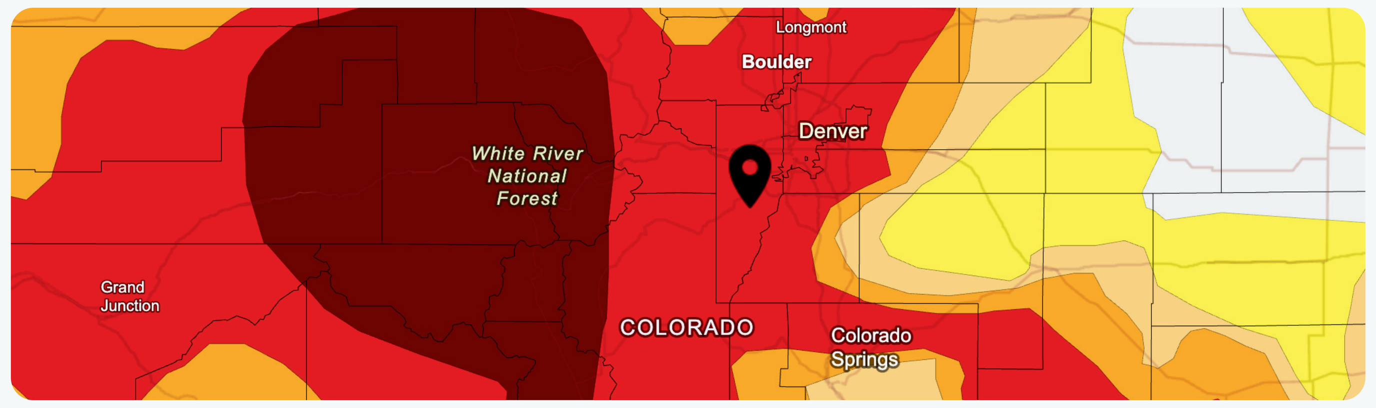


CONCLUSION

Weather driven statistical models can reliably predict Golden's water demand in spring and fall, providing a sound framework for validating conservation measures during shoulder seasons. A summer predictability gap persists across all model types, suggesting the problem lies in the data rather than the methodology, **summer demand is likely driven more by individual behavioral decisions than weather alone** or we are just not fully capturing what is driving differences in demand. When evaluating wet vurs dry years, normalization proves not just helpful but necessary. **Ridge Regression is the strongest overall performer and is recommended as the operational model going forward**. Future work should explore additional data sources to improve summer skill and further strengthen the conservation validation framework.

ACKNOWLEDGEMENTS

Thank you to my mentor Nels Bjarke for being a guiding hand along this journey.
Thank you to my cohort for being there and supporting me peer to peer.
Thank you to the RECCS leaders for teaching me new skills all along the way.
Thank you to my family for pushing me to seize this opportunity in the first place.



Credit: Drought.gov, NDMC, NOAA, USDA

University of Colorado Boulder

RECCS Research Experiences for Community College Students

WESTERN WATER ASSESSMENT
A NOAA RISA TEAM AT CIRES
UNIVERSITY OF COLORADO BOULDER

CIRES

NSF

NSF Award: EAR 1757930

References

GitHub code