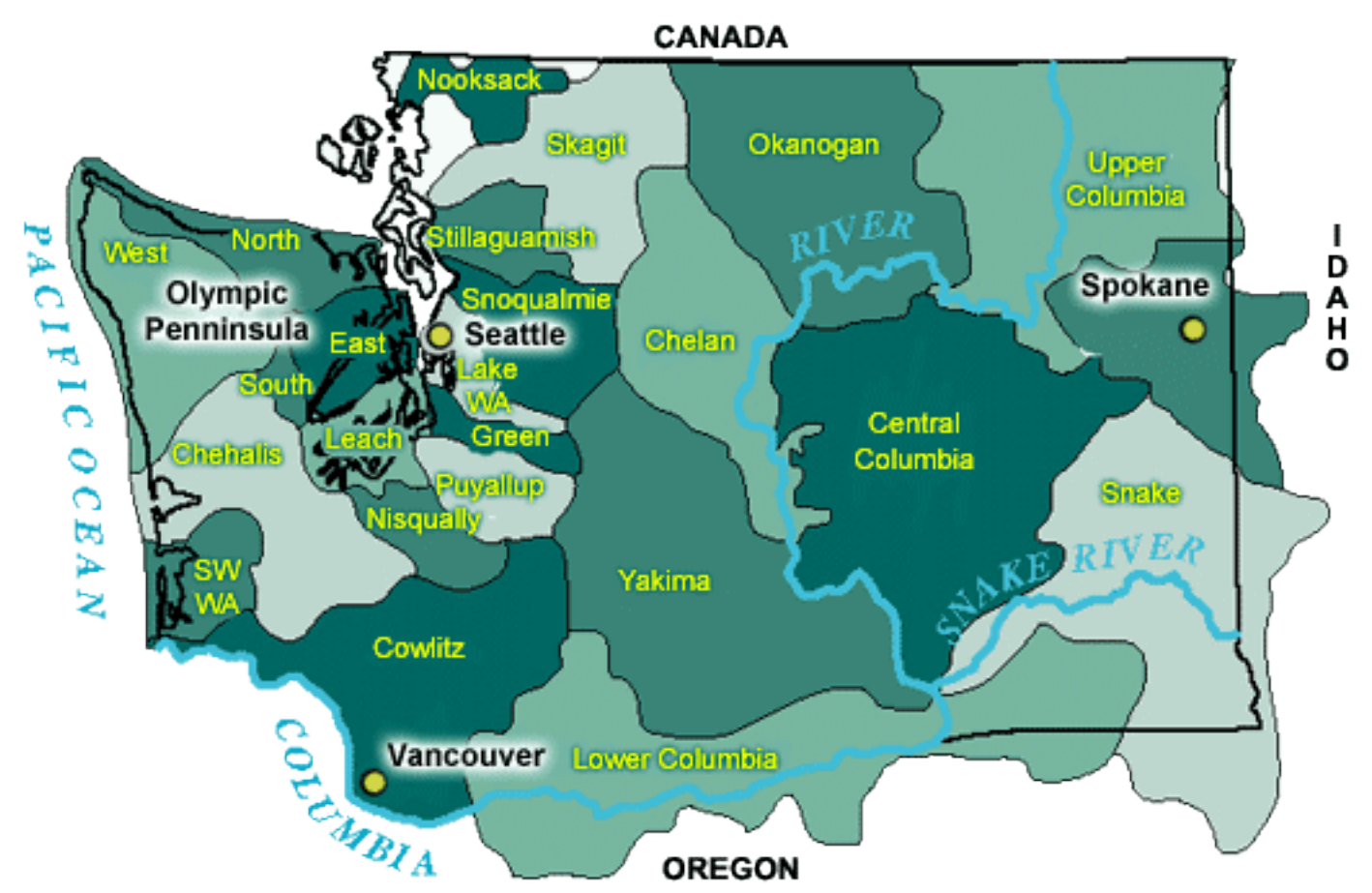


DEEP SNOW, DEEP LEARNING



Capstone Sponsor: Dr. Nicoleta Cristea
MSDS Team Members: Sue Boyd, Sarah Kilpatrick, and Derek Tropf

HOW MUCH WATER IS IN YOUR SNOWPACK?



Watersheds in Washington State. Watersheds, also known as hydrologic units, “HUCs”, are regions draining to a common portion of a body of water.

In the water cycle, snowpack acts like a natural reservoir, storing water in winter and gradually releasing it in spring and summer. Snow Water Equivalent (SWE) determines how much water the snowpack contains.

GLOSSARY

- **SWE – Snow Water Equivalent**
- **LSTM – Long Short-Term Memory, a type of Neural Network well-designed for time series data.**
- **HUC - Hydrologic Unit Code. The lower the code, the larger the area. Huc8 > Huc10 > Huc12.**
- **KGE – Klinge-Gupta Efficiency, a goodness-of-fit measure commonly used in Hydrology.**

ACKNOWLEDGEMENTS

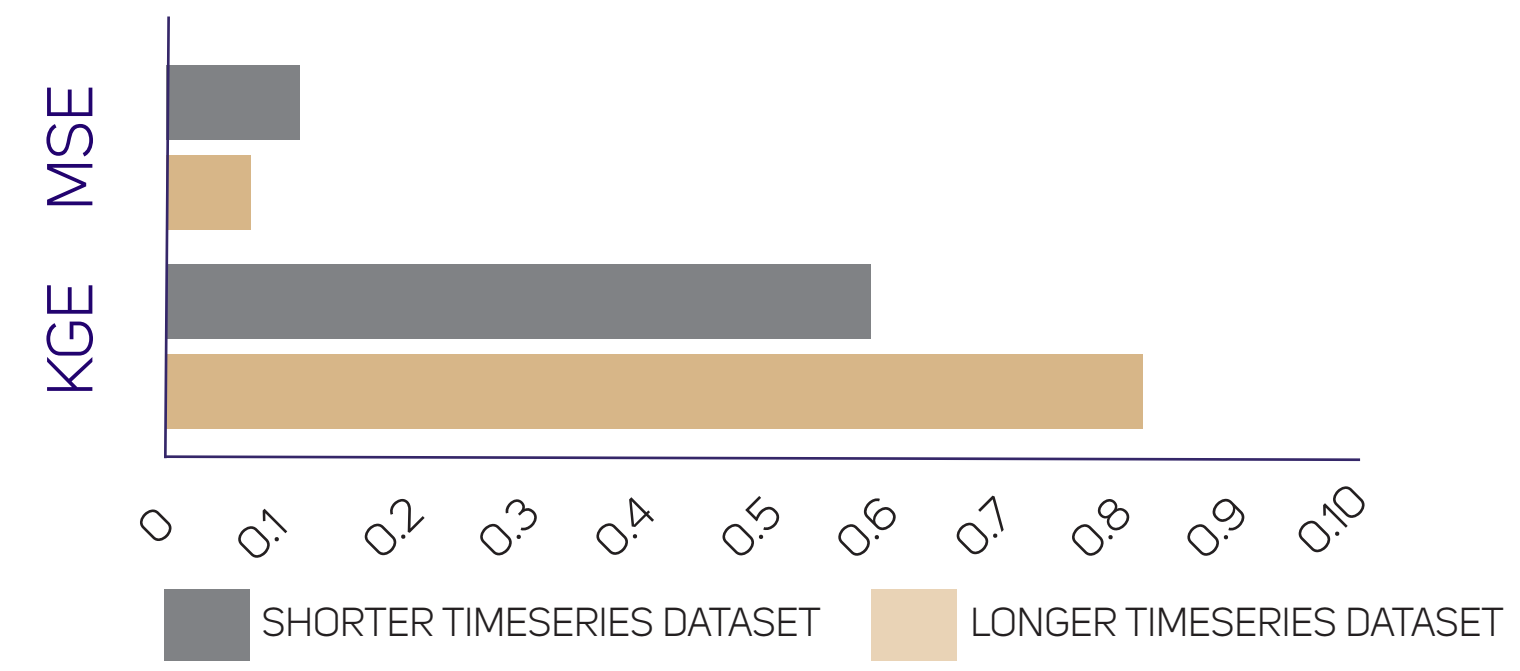
THANK YOU TO OUR SPONSOR DR. NICOLETA CRISTEA AND CAPSTONE DIRECTOR DR. MEGAN HAZEN FOR YOUR LEADERSHIP & EXPERTISE. YOUR TIME AND EFFORT INVESTED IN SUPPORTING OUR WORK HAS BEEN IMMESURABLY VALUABLE. THANK YOU TO THE ESCIENCE INSTITUTE AND AWS FOR THEIR COMPUTING SUPPORT.

LOCALLY-TRAINED MODELS

Can LSTM models represent physics-based model estimations across different watershed conditions?

MORE DATA = BETTER RESULTS

We trained a single-layer LSTM model with a linear layer for each of seven watersheds in the Skagit River Basin, achieving a superior model fit over previous attempts that relied on shorter time-series data.

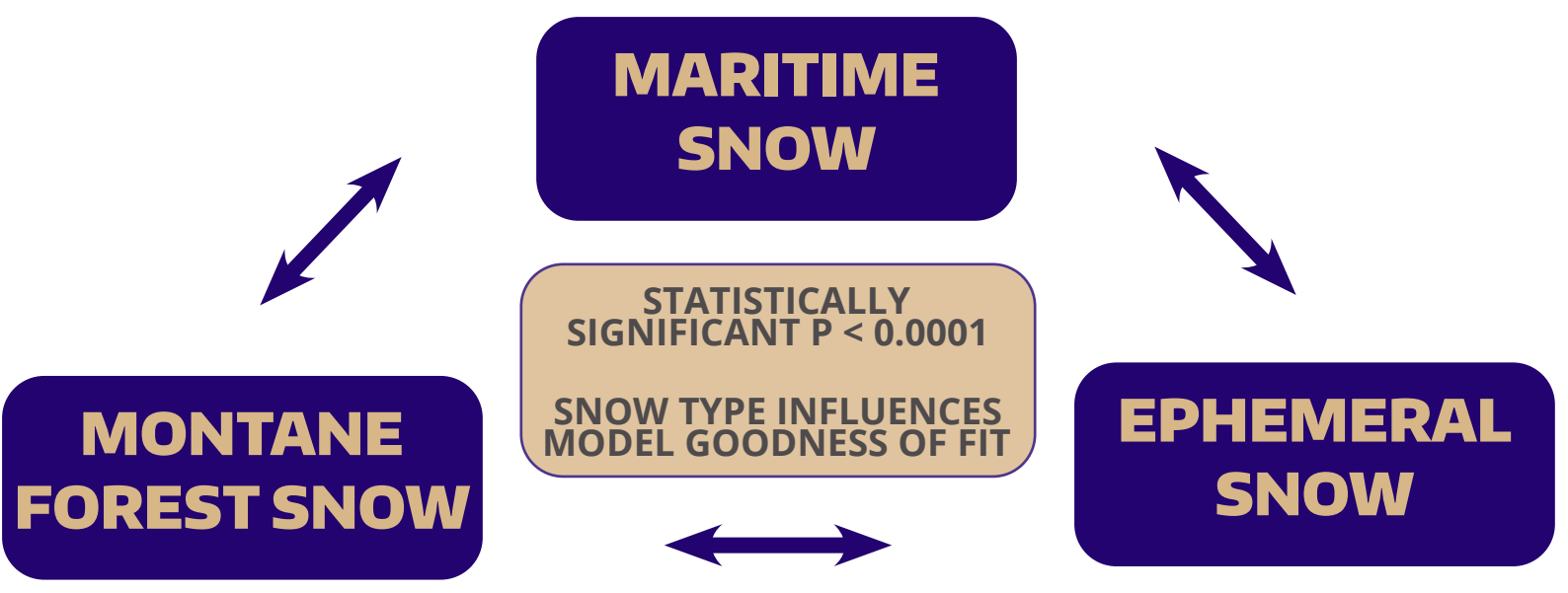


$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}$$

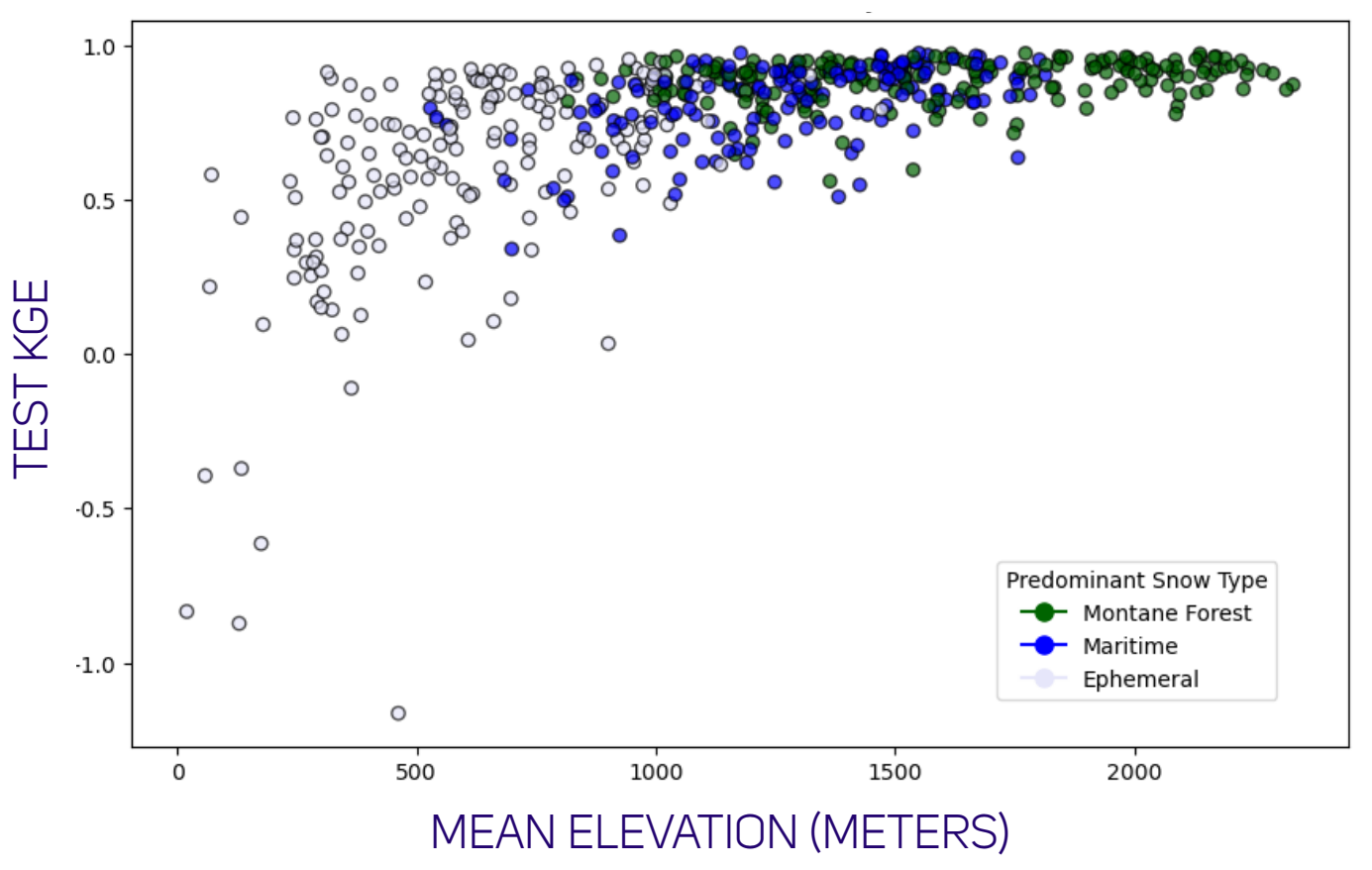
Klinge-Gupta Efficiency is a hydrologist-created metric designed to compare simulations to observations.

- r is the Pearson correlation coefficient
- α is a term representing the variability of prediction errors
- β is a bias term

PERFORMANCE VARIES BY SNOW TYPE AND ELEVATION, 500+ WATERSHEDS

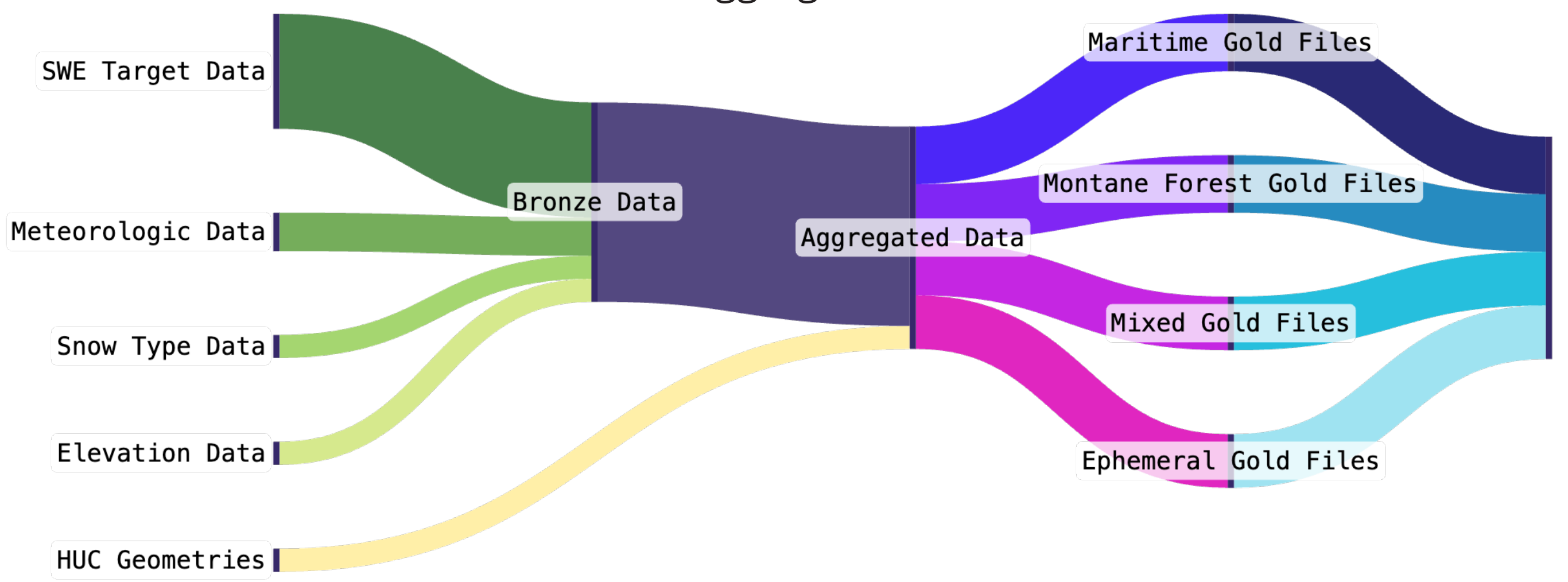


TEST KGE IN PACIFIC NW WATERSHEDS, BY ELEVATION AND SNOW TYPE



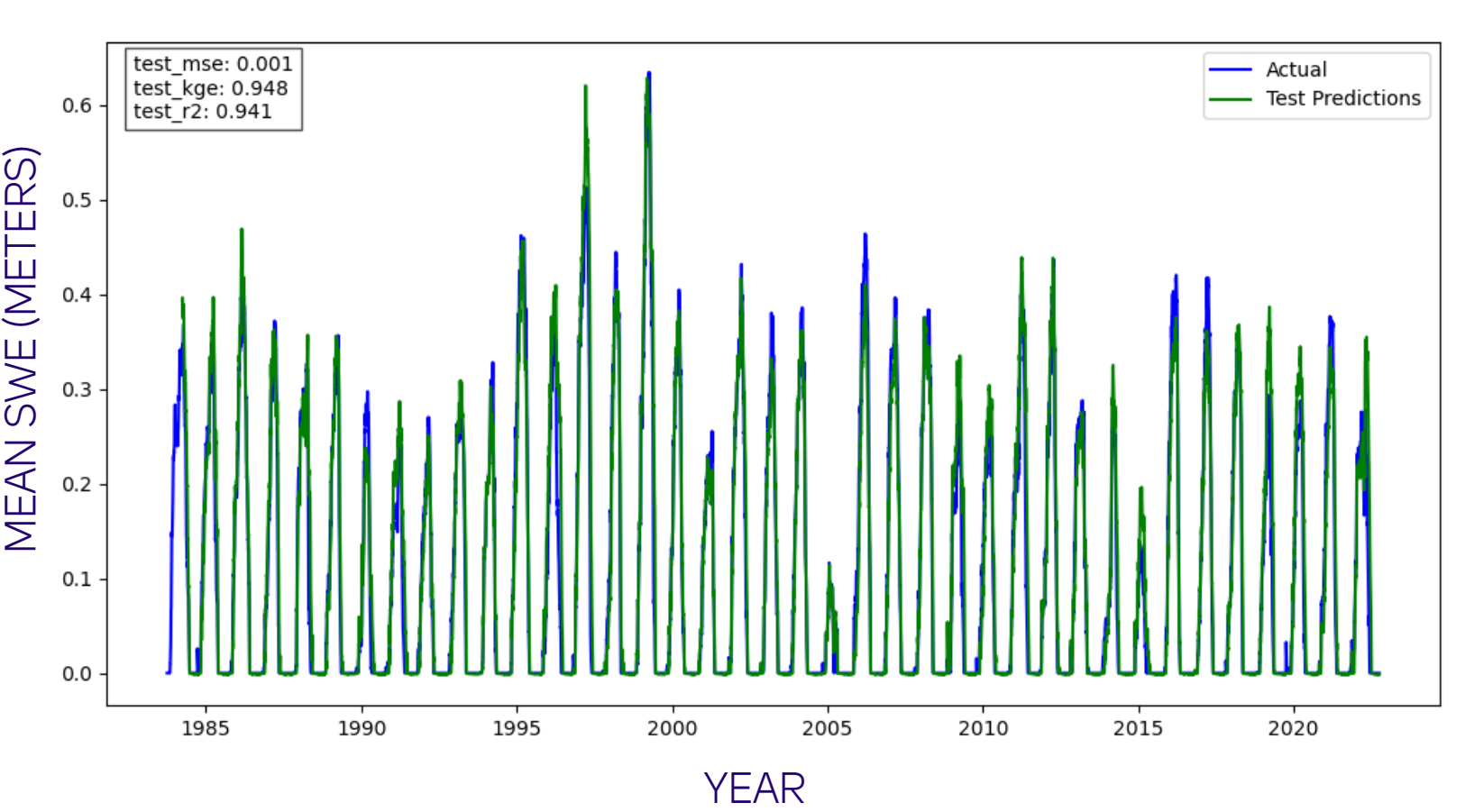
SCALABLE DATA PIPELINE

Daily measurements for water years 1983-2022. Medallion-inspired architecture. All variables aggregated to watershed scale.



SEPARATE REGIONAL DATA? NO PROBLEM!

SWE PREDICTIONS USING THE MULTI-HUC MODEL



ABOVE: SWE predictions for water years 1983-2022 in High Creek/ Naneum Creek Sub-Watershed, near Yakima valley. Our multi-watershed trained LSTM model predicts previously-unseen basins with accuracy comparable to the locally-trained models.

BELOW: Test results for SWE predictions in 48 HUC-12 sub-watersheds within the Naches and Upper Yakima sub-basins, important regions for Washington Agriculture:

	TEST MSE	TEST KGE	TEST R ²
MEAN	0.007	0.789	0.888
MEDIAN	0.004	0.821	0.919
INTERQUARTILE RANGE	(0.002 - 0.009)	(0.707 - 0.883)	(0.864 - 0.945)

USE OUR TOOLS

Our tools (Scalable Data Pipeline, Proof of Concept for Watershed Level Training, and the Interactive Dashboard) advance ML techniques in hydrology and support future research for the benefit of water resource managers.



WE'RE ON THE ROAD TO PUBLICATION!

Interested in our work? Check out our experiments and findings at the Frosty Dawgs Github!

INTERACTIVE DASHBOARD

Our interactive dashboard gives water resource managers real-time data visualizations.

