Oregon State University

River Level Forecasting

Orion Junkins & Patrick Donnelly

soundbendor lab

College of Engineering | Computer Science

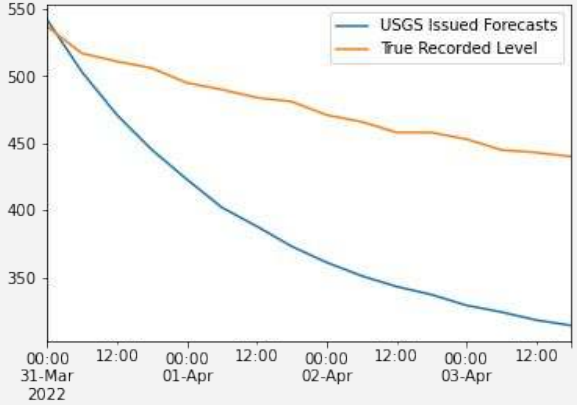
Introduction to Problem

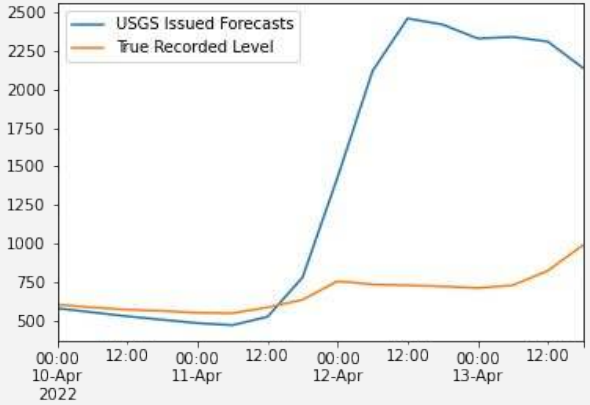
- Changes in river flow rate affect many stakeholders
  - Agricultural users
  - River adjacent landowners
  - Recreational users
- Flow rate is directly related to weather
  - Complex relationship
  - Challenging to model with traditional statistical methods

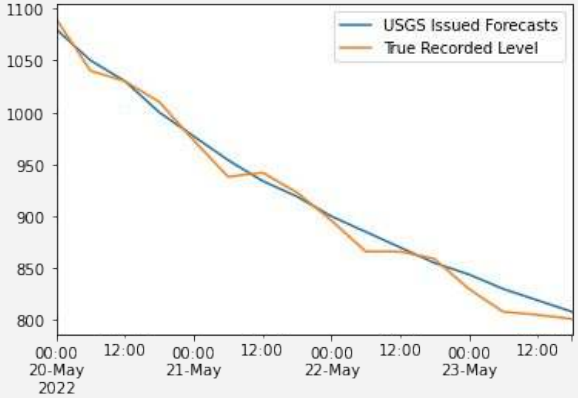
Existing Production Solutions

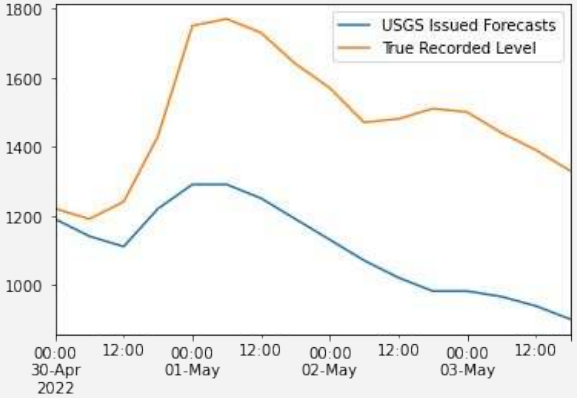
- National Oceanic and Atmospheric Administration (NOAA)
  - provides forecasts for a subset of rivers in the United States
- Leverage precipitation, temperature and snowpack data
  - proprietary statistical algorithm
  - lacks transparency
  - lacks interpretability
  - lacks accuracy

Examination of NOAA Accuracy









Error Metrics

MAE = 192.06 (ft^3/sec)

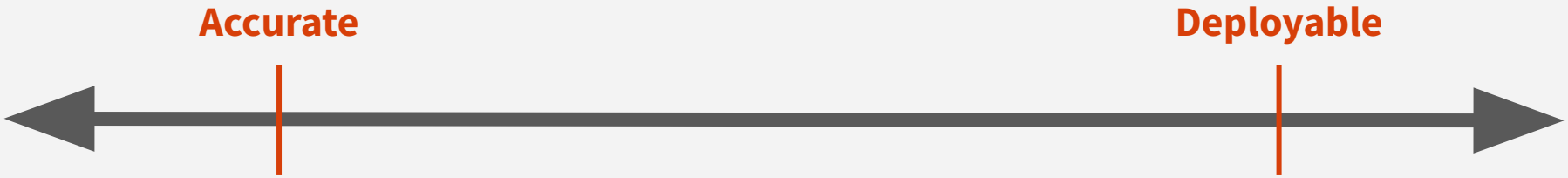
MAPE = 16.46%

Metrics computed for 24 hour advance forecasts across 110 samples

Existing Research Solutions

Oregon State University

- Neural networks are better suited to this task
- Prior solutions faced a trade off between **accuracy** and **deployability**



- large networks
- all available data
- extensive compute resources
- long training times
- slower inference
- **hard to deploy at scale**

- lightweight networks
- subset of available data
- hardware flexibility
- short training times
- rapid inference
- **less accurate**

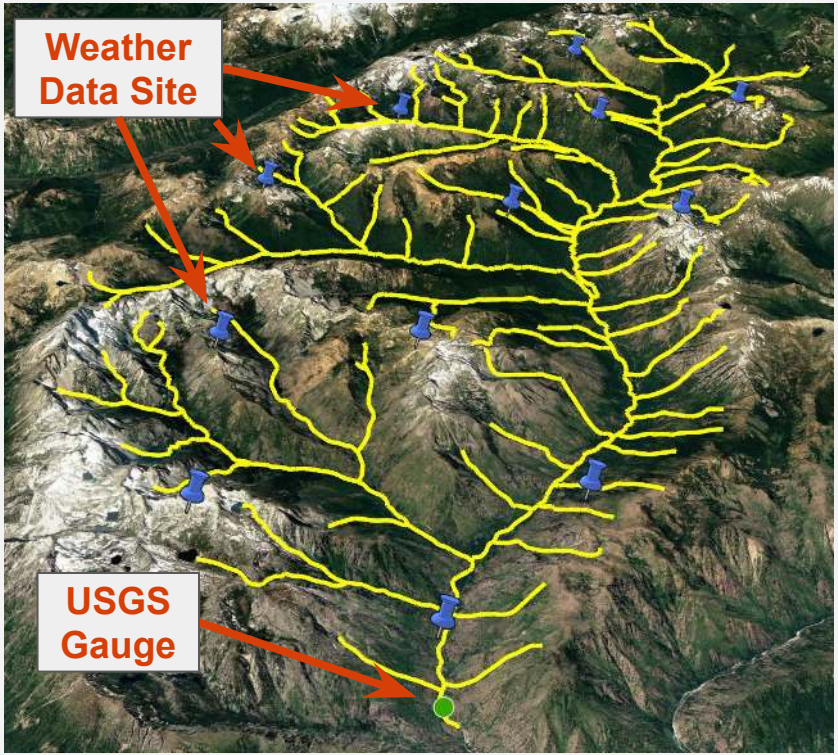
Research Goals

Oregon State University

1. Produce a forecasting model that achieves **high accuracy**, without sacrificing **real-world deployability**

2. Match or beat NOAA forecast accuracy with more **transparency** and **interpretability**

Data Collection



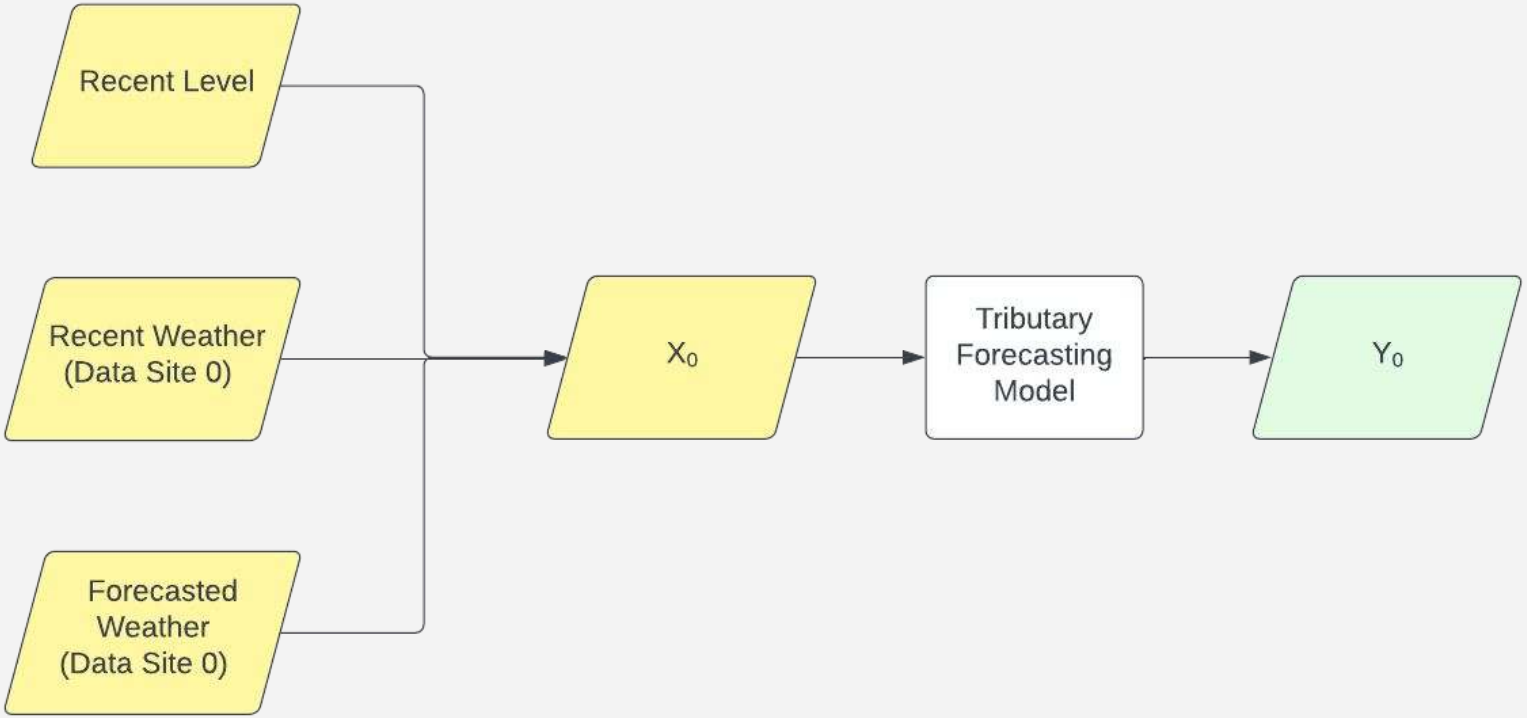
Icicle Creek Data Sources

- River Level Data
  - USGS a historical and current flow rate
  - Specific gauge sites
- Weather Data
  - OpenWeatherMap provides historical, current and forecasted
  - Any desired latitude & longitude

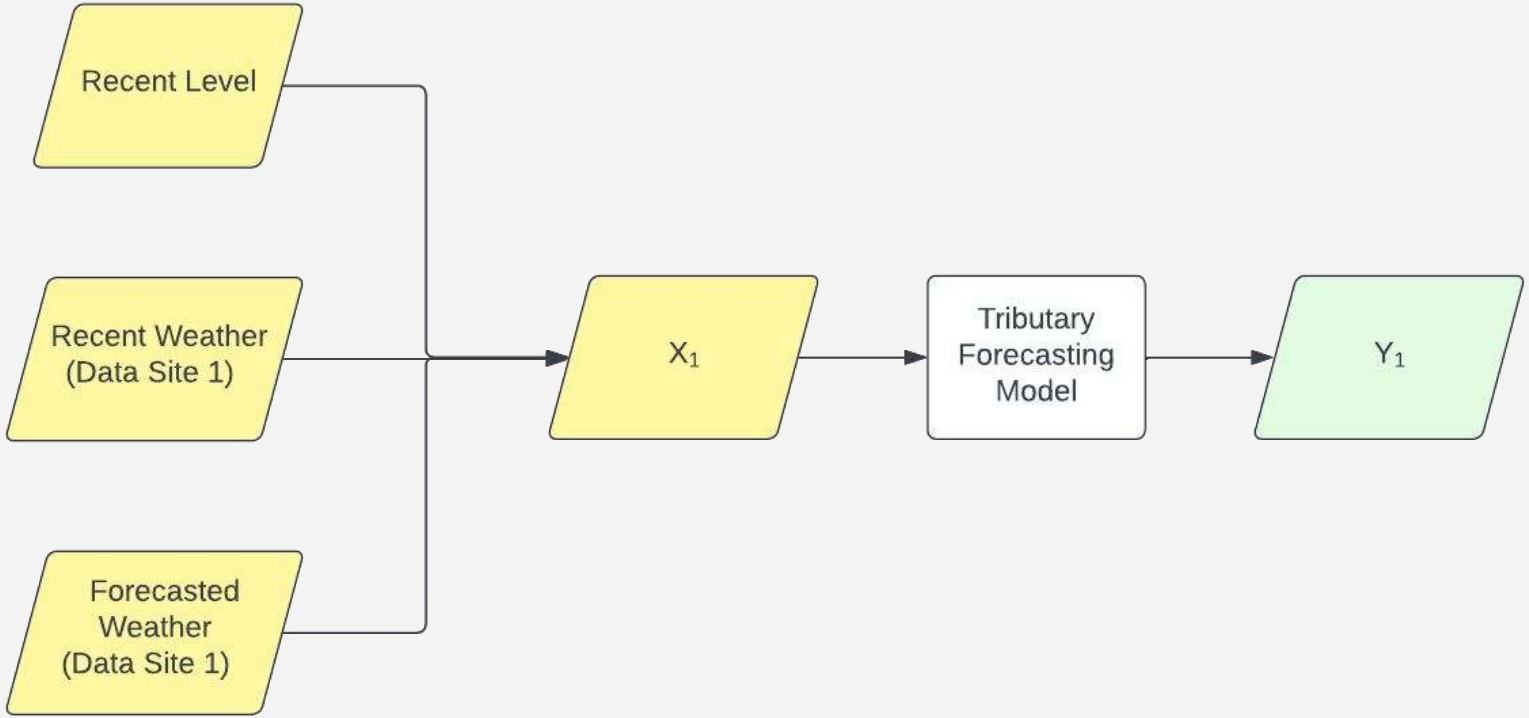
Dimensionality Reduction

- Multiple data sources results in high dimensionality
  - high order relationships are lost amidst noise
  - necessitates extensive compute resources/training time
- Reducing featurespace hinders accuracy
- Adopted ensemble approach
  - mitigates issues raised by high dimensionality
  - retains full featurespace to maximize accuracy

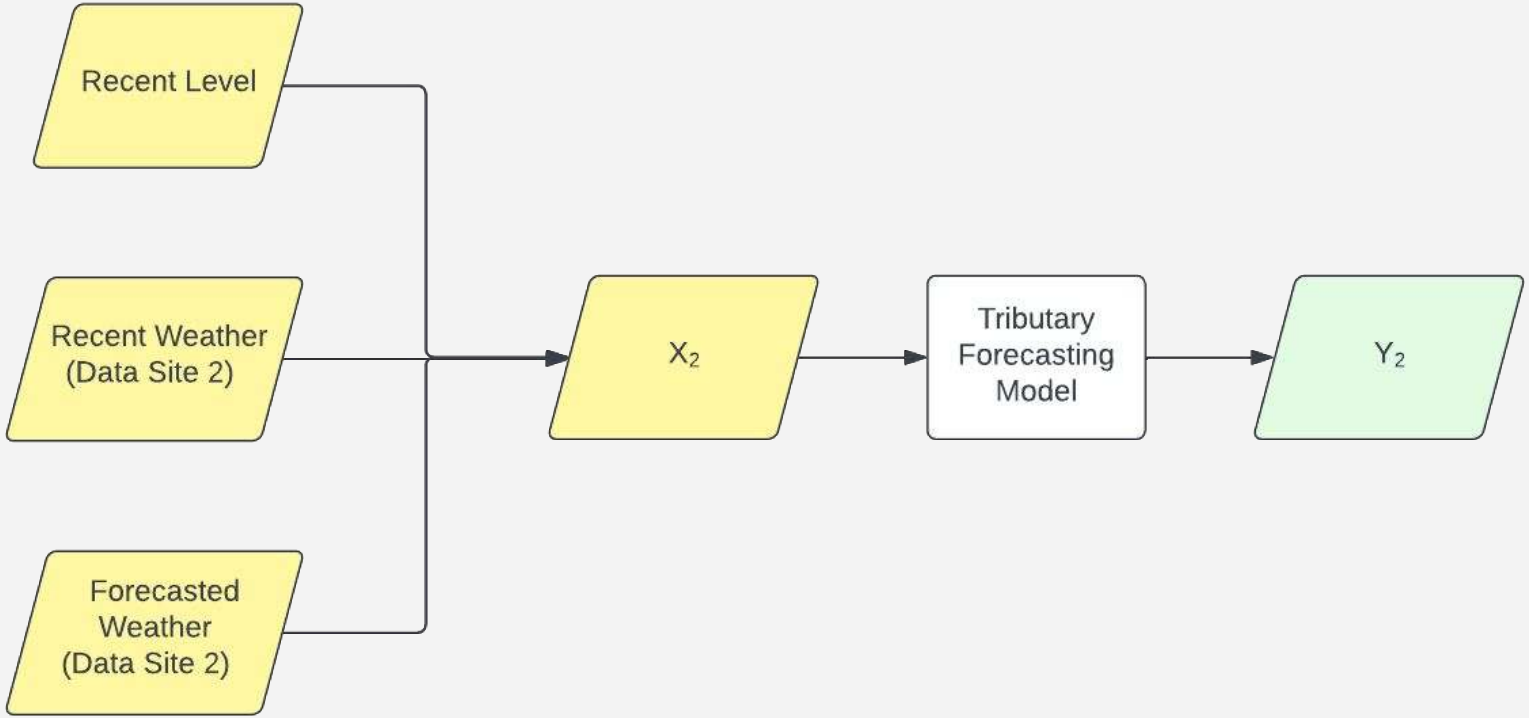
Tributary Forecasting Model



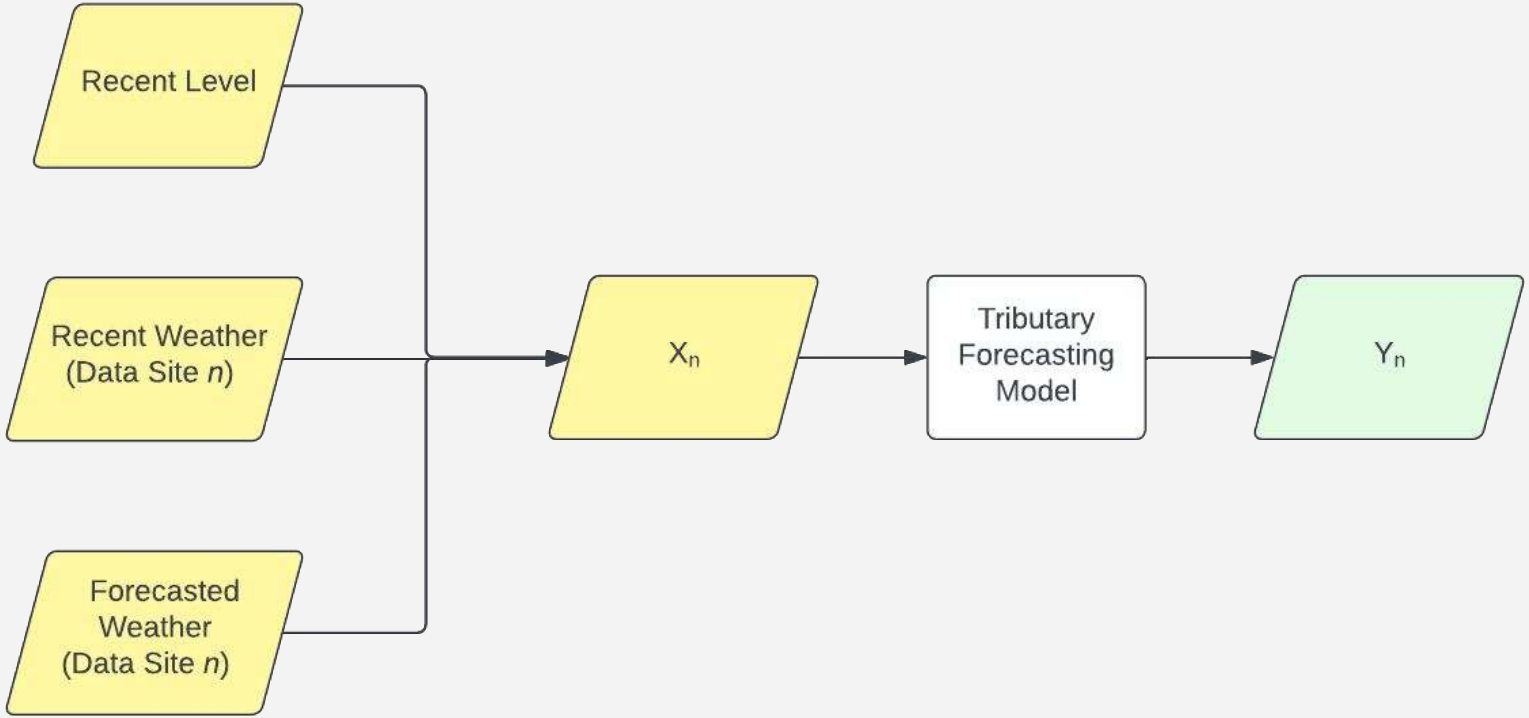
Tributary Forecasting Model



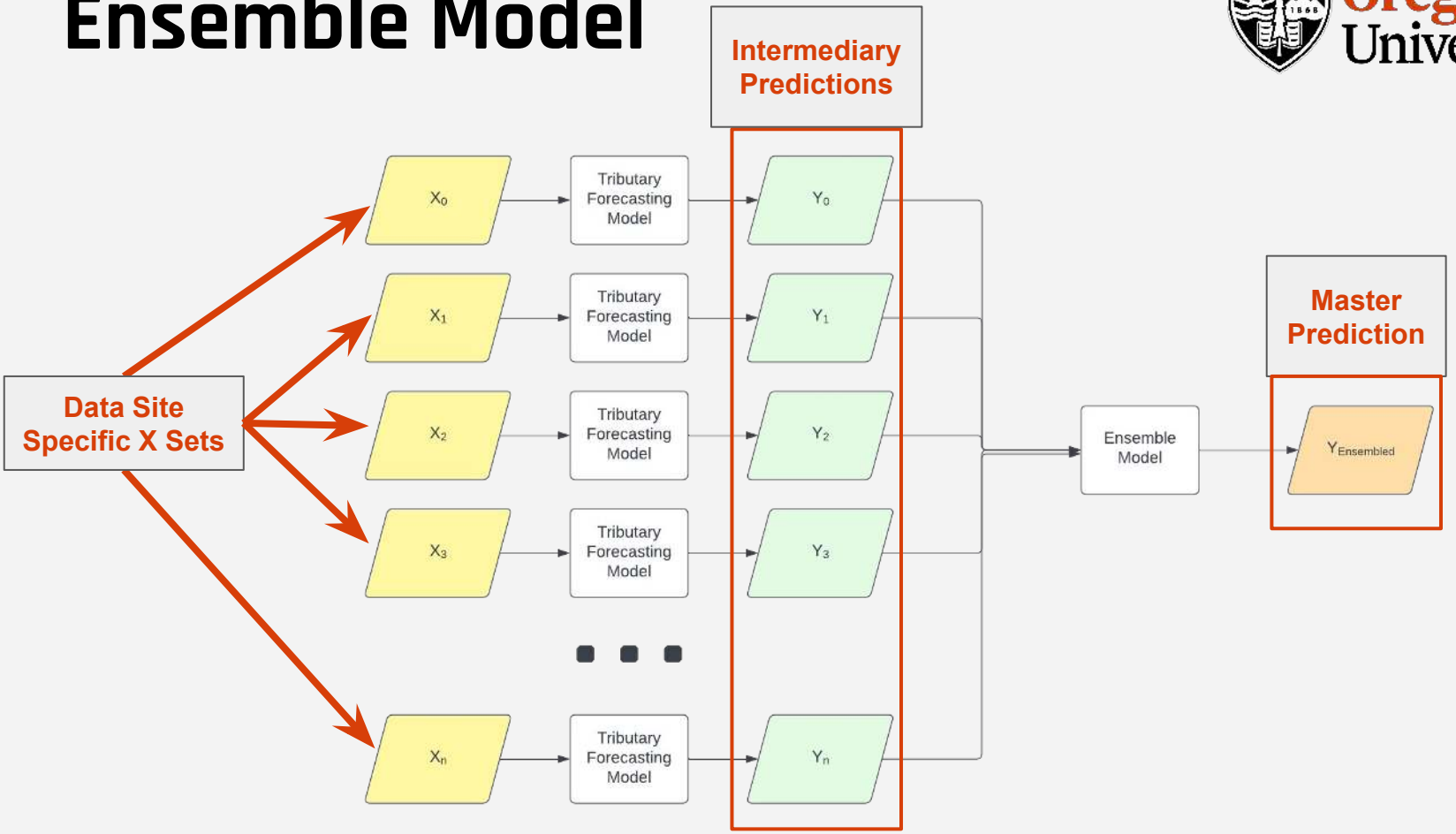
Tributary Forecasting Model



Tributary Forecasting Model



Ensemble Model



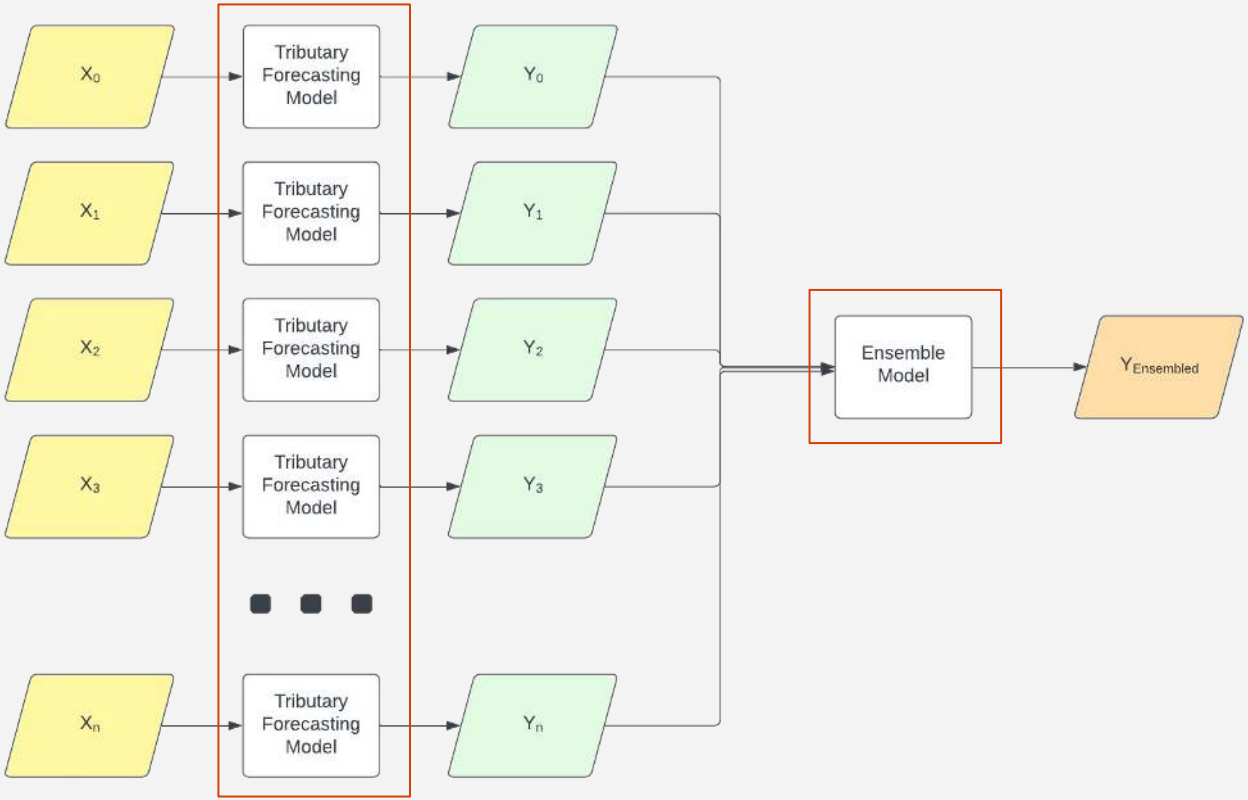
Advantages of Approach

Oregon State University

- Leverage all available data
- Minimize training time/compute resources needed
  - each individual model is lightweight
  - training is highly parallelizable
- Interpretability
- Modularity

Ensemble Model

Oregon State University



Results - Simulated Historical

Oregon State University

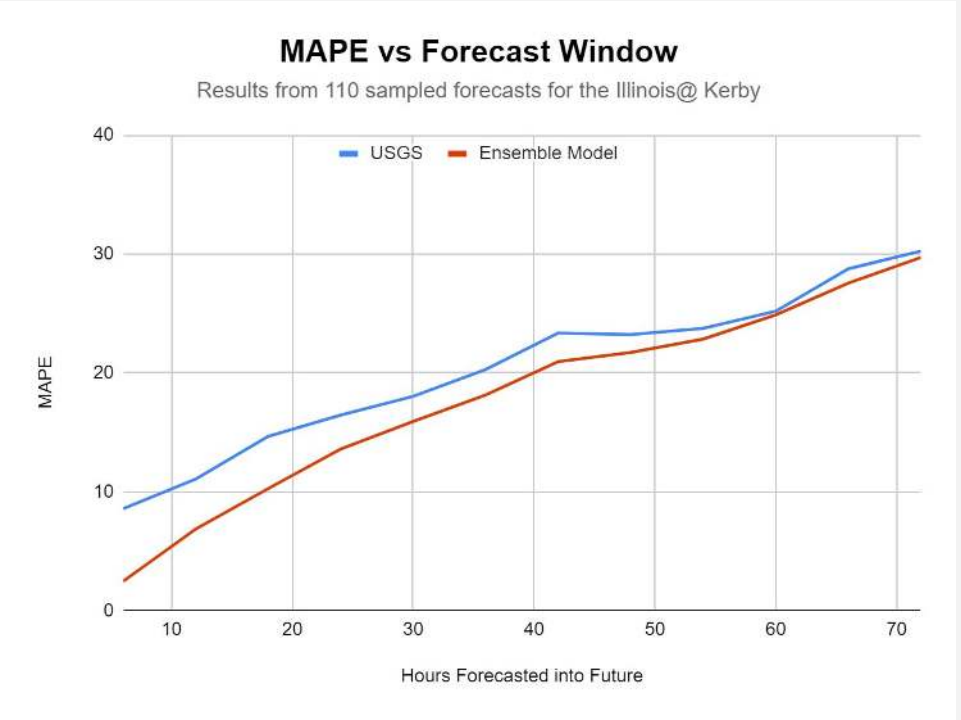
- 24 hour advance forecasts of the Illinois at Kerby
  - **MAPE of 8.57%**
  - 10.02% better than the best performing tributary model
  - demonstrates the efficacy of ensemble approach
- <12 hour train time, <30 second inference time (Nvidia 3070)
  - demonstrates viability of training/inference without HPC cluster
  - parallelization across a cluster can improve this substantially

Results - Real World Evaluation

Oregon State University

24 hour advance forecasts of Illinois @ Kerby over 110 days

- MAPE of **13.62%**
- **2.85%** better than forecasts issued by NOAA



Future Work

Oregon State University

- Experiment with more advanced model architectures
- Evaluate the efficacy of this approach on a more diverse set of rivers
  - larger river systems in particular
- Extrapolate a model from one river to another with transfer learning
- Explore domain generalizability of this approach
  - potentially effective for other weather related forecasting tasks
  - forest fire prediction and avalanche risk forecasting

References

[1] A. Mosavi, P. Ozturk, and K. Chau, “Flood prediction using machine learning models: Literature review,” Water, vol. 10, no. 11, p. 1536, 2018.

[2] H. R. Maier, A. Jain, G. C. Dandy, and K. P. Sudheer, “Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions,” Environmental modelling & software, vol. 25, no. 8, pp. 891–909, 2010.

[3] W. Li, Q. Duan, C. Miao, A. Ye, W. Gong, and Z. Di, “A review on statistical postprocessing methods for hydrometeorological ensemble forecasting,” Wiley Interdisciplinary Reviews: Water, vol. 4, no. 6, p. e1246, 2017.

[4] D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi, and A. Jamil, “A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin,” Earth Science Informatics, vol. 13, no. 3, pp. 915–927, 2020.

References

[5] W. Xu, Y. Jiang, X. Zhang, Y. Li, R. Zhang, and G. Fu, “Using long short-term memory networks for river flow prediction,” Hydrology Research, vol. 51, no. 6, pp. 1358–1376, 2020.

[6] A. M. Ahmed et al., “New double decomposition deep learning methods for river water level forecasting,” Science of The Total Environment, vol. 831, p. 154722, 2022.

[7] V. Nourani, “An emotional ANN (EANN) approach to modeling rainfall-runoff process,” Journal of Hydrology, vol. 544, pp. 267–277, 2017.

[8] T. C. Pagano et al., “Challenges of operational river forecasting,” Journal of Hydrometeorology, vol. 15, no. 4, pp. 1692–1707, 2014.