# Predicting Daily Snow Water Equivalent with Machine Learning Models

Shiheng Duan[1] Paul Ullrich[1]

[1]Department of Land, Air and Water Resources, University of California, Davis

## Model Design

### Motivation, Data and Area of Interested

As a vital component of the global hydroclimate system, precise snowpack prediction is of considerable value for science and society. Snow water equivalent (SWE) prediction can be viewed as a time series problem. In this study, we selected several machine learning (ML) models to predict daily SWE and extrapolated our model to estimate gridded SWE over Rocky Mountain area. The SWE data used for training is from SNOTEL stations and forcings are from gridMET dataset with 4km resolution.

### Model Inputs and Output

The dynamic input variables to our model are precipitation, minimum and maximum temperature, specific humidity, minimum and maximum relative humidity, solar radiation, vapor deficit and wind speed. The static features are latitude, longitude, elevation, diurnal anisotropic heating (DAH) index and topographic radiative aspect (TRASP) index. The time window is set to 180 days.

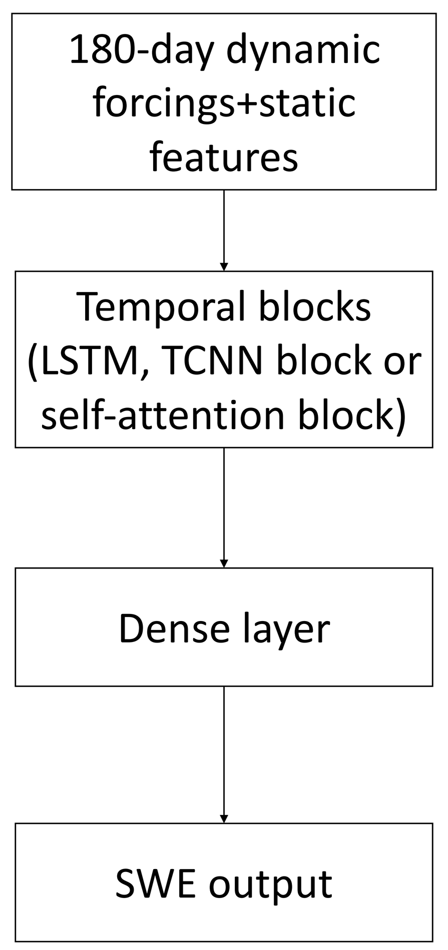


**Figure:** The data flow chart through the ML models.

### ML model used

We used long-short term memory (LSTM), temporal convolutional neural network (TCNN) and self-attention-based model to predict daily SWE.

### Model Training Setup

The model is trained with all the stations in western CONUS. The training period is from 1980-10-01 to 1999-09-30, validation from 1999-10-01 to 2008-09-30 and testing from 2008-10-01 to 2018-09-30. The loss function is set to mean squared error.

## Prediction

### SNOTEL Prediction Results

Since the neural networks use gradient based method to optimize the loss function, the initial weights will affect the model performance. In order to obtain a general idea for how our model performs for a random initial state, we run the model 10 times to get a distribution of NSE values.

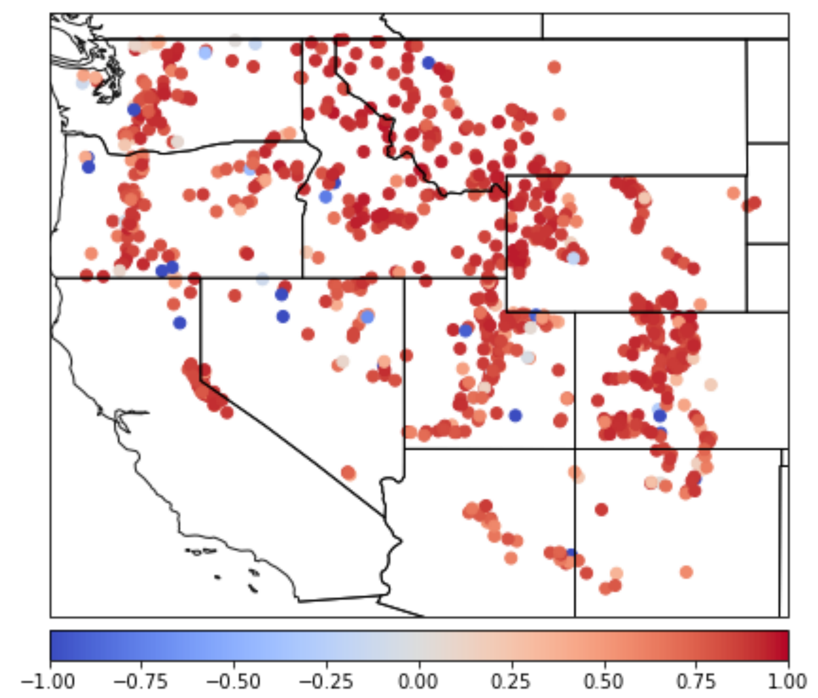

**Figure:** Ensemble NSE with TCNN model.

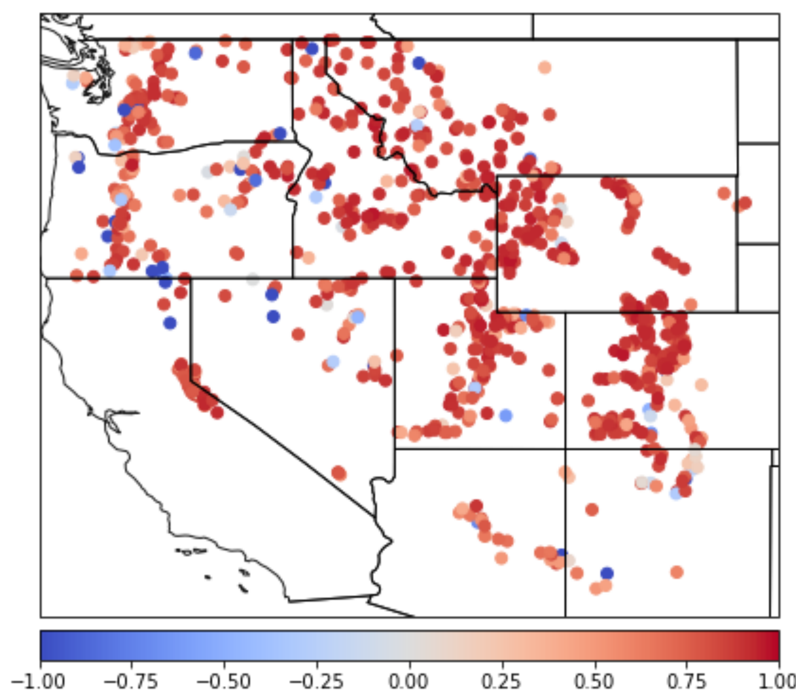

**Figure:** Ensemble NSE with LSTM model.

The machine learning models tend to have a good performance in the most stations. LSTM can achieve a median NSE value as 0.84 and 0.86 for TCNN. The stations with low performance are consistent for LSTM and TCNN.

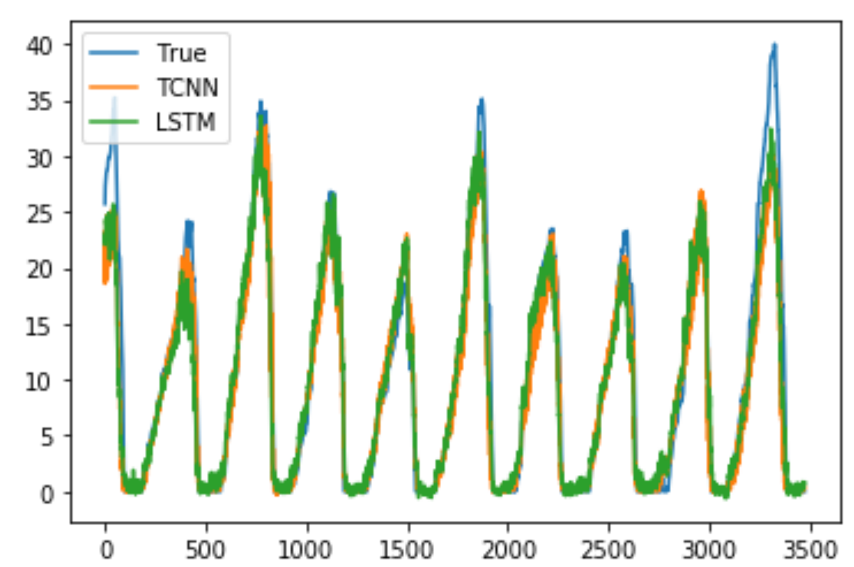


**Figure:** An ensemble prediction from TCNN and LSTM for one station. NSE=0.911 for TCNN and 0.914 for LSTM.

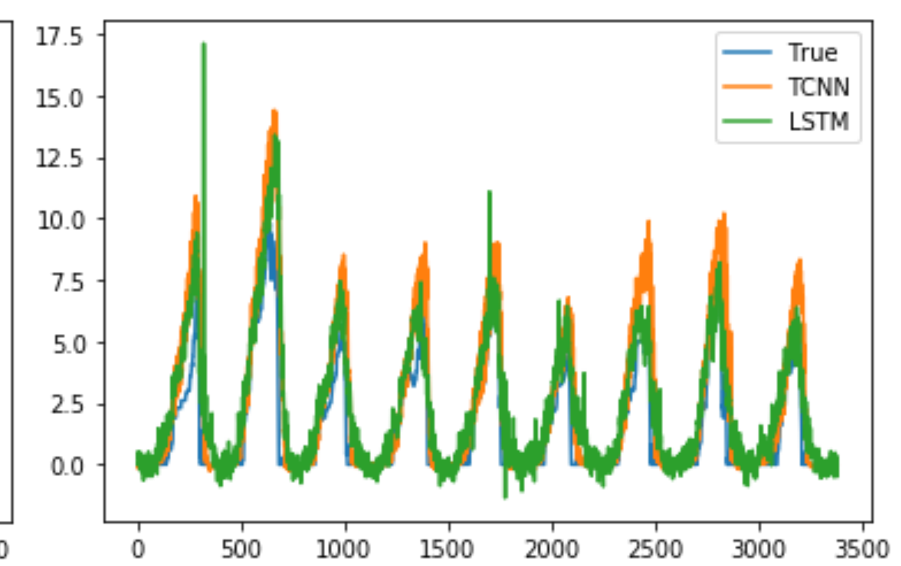


**Figure:** An ensemble prediction from TCNN and LSTM for one station. NSE=0.392 for TCNN and 0.597 for LSTM.

## Extrapolation

### Extrapolation Data

We used PRISM 800m gridded forcing data (relative short time period) to train a model and generate gridded daily SWE estimation in Rocky Mountain area.

The prediction is compared with University of Arizona daily gridded SWE dataset with 4km horizontal resolution.

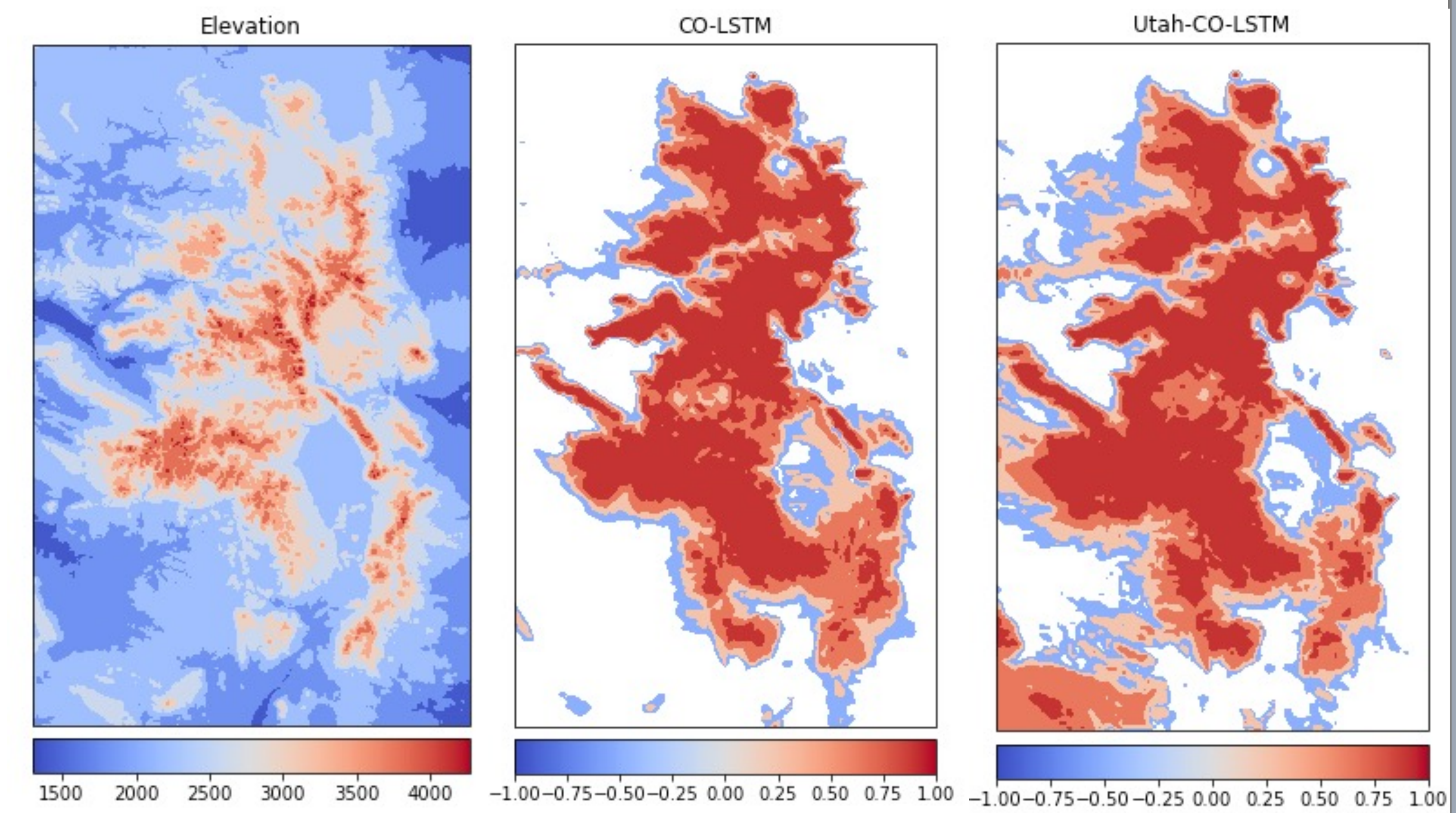


**Figure:** Elevation and r2 score for Rocky Mountain. R2 score less than -1 is masked.

Generally, for high-elevation area, the ML model agrees with the physics-based model. However, the ML model always overestimate SWE for low-elevation area. The reason could be that we trained the model with SNOTEL stations, which are in high-elevation area and always have high SWE.

### Future work

1. Use the model trained with 4km PRISM forcing for extrapolation.
2. Fine tune the model for western US and use it for extrapolation.
3. Use gridded satellite snow cover dataset which covers the low-elevation area to train a classification model to determine whether it is snow season.
4. Transfer the snow cover model to SWE prediction to see if we can better estimation at low elevation area.

### References

- Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems. 2017.
- Bai, Shaojie, J. Zico Kolter, and Vladlen Koltun. "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling." arXiv preprint arXiv:1803.01271 (2018).

### Acknowledgements

Support comes from the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research program under Award Number DE-SC0016605 "A framework for improving analysis and modeling of Earth system and intersectoral dynamics at regional scales."

### Contact

shiduan@ucdavis.edu