A Comprehensive Investigation of Machine Learning Models for Estimating Daily Snow Water Equivalent (SWE) over the Western U.S.

Shiheng Duan and Paul Ullrich
University of California, Davis
2021 AGU Fall Meeting

UCDAVIS
ATMOSPHERIC SCIENCE

U.S. DEPARTMENT OF ENERGY
Office of Science

U.S. Department of Energy | Office of Science
HYPERFACETS
A Framework for Improving Analysis and Modeling of Earth System and Intersectoral Dynamics at Regional Scales

Previous work on SWE estimation

- **From reanalysis datasets:**
  ANN model or random forest
- **From precipitation**: cGAN
- **From precipitation and snow-related variables:** LSTM
- Not all models can be used for projections.

- **Idea:** Use machine learning (deep learning) models for both SWE prediction and projection.
- The models should be able to handle time dependency.
- The models should mainly use atmospheric forcings for the projection purpose.

Effectively this task can be expressed as:

$$\mathrm{SWE}_t = f(P_t, P_{t-1}, P_{t-2}, \dots, P_{t-N+1}, T_t, T_{t-1}, T_{t-2} \dots, T_{t-N+1})$$

General Architecture

- **Dynamic input variables:** precipitation, temperature (min and max), solar radiation, specific humidity (min and max), relative humidity, vapor deficit and wind speed;
- **Static input variables:** latitude, longitude, elevation, diurnal anisotropic heat index (DAH) and solar radiation aspect index (TRASP).
- **Output variable:** SWE
- **Input window size:** 180 days.
- **Models:** Long-Short Term Memory (LSTM), Temporal Convolution Neural Network (TCNN), and Self-Attention model (Attention).

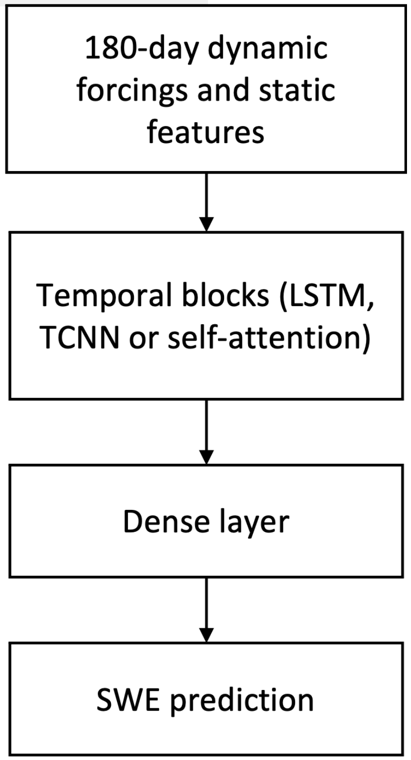


Figure: A flow chart depiction of our models.

Deep Learning Models for Time Series

Long-Short Term Memory (LSTM)

Temporal Convolutional Neural Network (TCNN)

Self-Attention

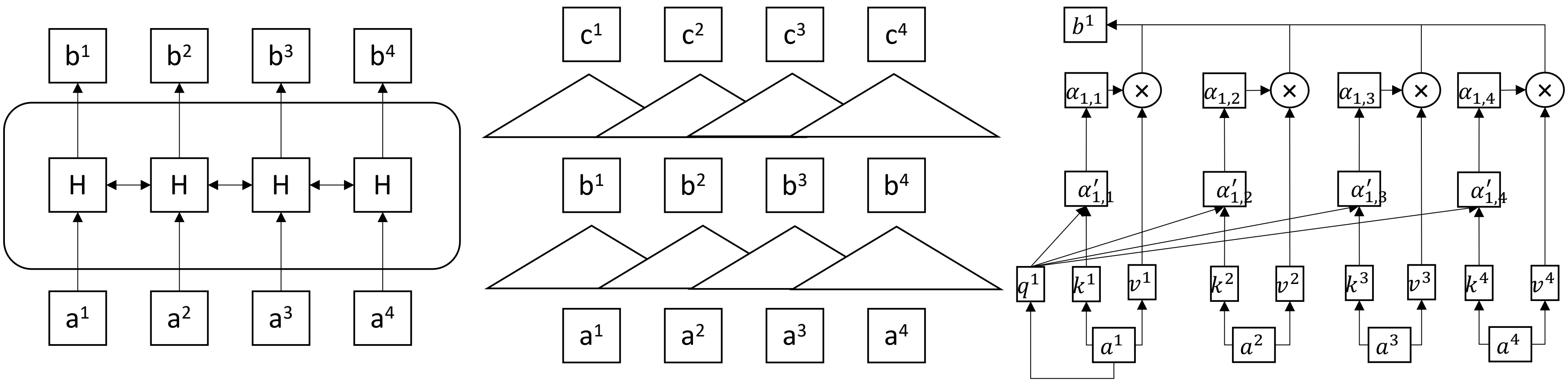


Figure: Our machine learning model architectures.

Training, Testing and Validation

- 581 SNOTEL stations are used to train the model. The variables are normalized with the mean and standard deviation from all the stations.
- Hyperparameters are determined with the validation data.
- Train each model 10 times and get the ensemble mean prediction.
- Training time with 1 RTX2080Ti GPU:
  - LSTM 5 hours
  - TCNN 10 hours
  - Attention 26 hours

Experiment Settings	
Loss function	Mean squared error
Training	1980-10-01 to 1999-09-30
Validation	1999-10-01 to 2008-09-30
Testing	2008-10-01 to 2018-09-30

SNOTEL Prediction Results

- Quantify the performance by **Nash-Sutcliffe model efficiency coefficient (NSE)** or **R square** score.
- Across all Western U.S. SNOTEL stations, the median NSE values for LSTM, TCNN and Attention are 0.909, 0.878 and 0.874, respectively.
- For comparison, the **NSIDC-UA** dataset has a median NSE value as 0.861.

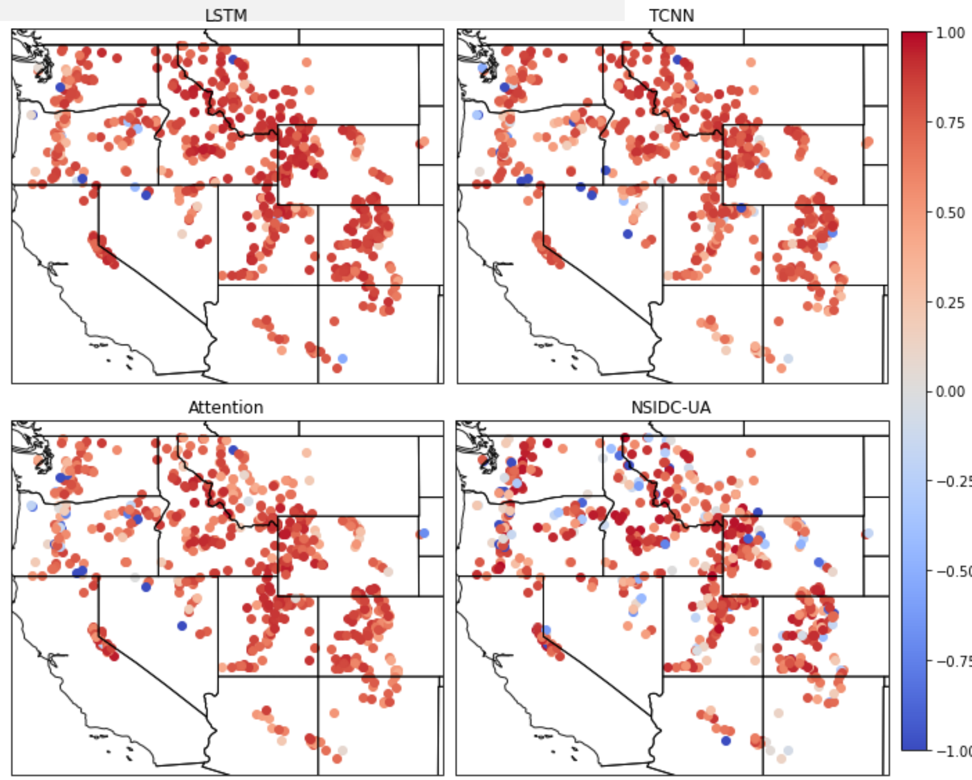


Figure: Prediction result from deep learning models and NSIDC UA dataset.

Prediction Results

- LSTM performance is the best: Highest median NSE value and more concentrated distributions over high NSE value regimes.
- TCNN and Attention are similar, although Attention is better at high NSE value ranges.
- NSIDC-UA dataset has more stations in low NSE regions compared with the deep learning models.
- There is a strong correlation among NSE values from different deep learning models. Pearson correlation is 0.945 between LSTM and TCNN and 0.818 between LSTM and Attention.

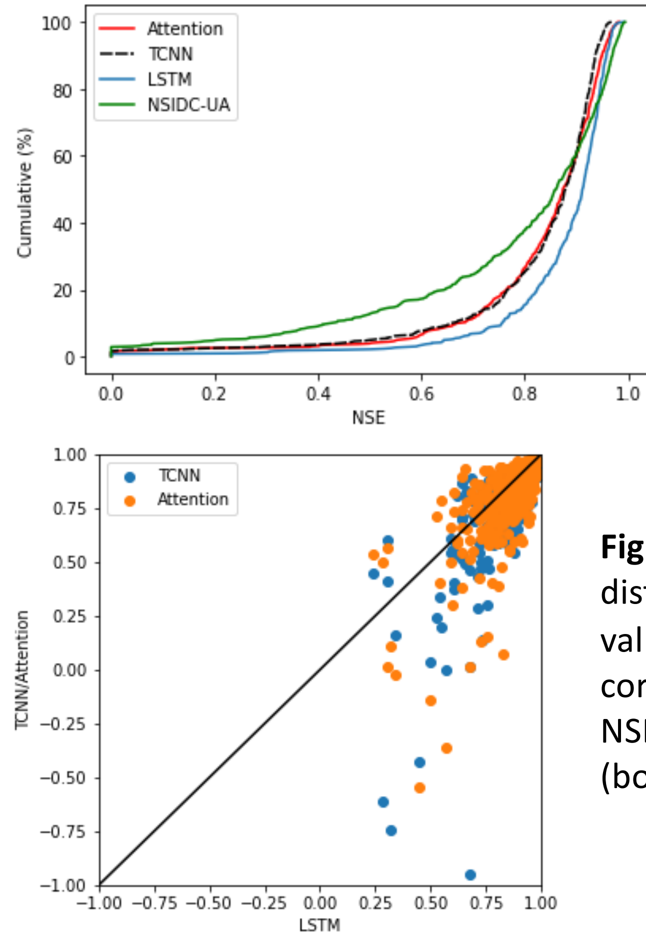


Figure: Probability distribution of NSE values (top) and correlation between NSE values (bottom).

Extrapolation

- Use the model trained on SNOTEL observations to generate a gridded SWE estimate.
- The statistical features of both input and output variables will be different. Models are in extrapolation regime.
- To deal with extrapolation, we focus on the seasonality of SWE instead of the actual SWE amount. Seasonality is calculated as the fraction of SWE with respect to the historical maximum SWE.

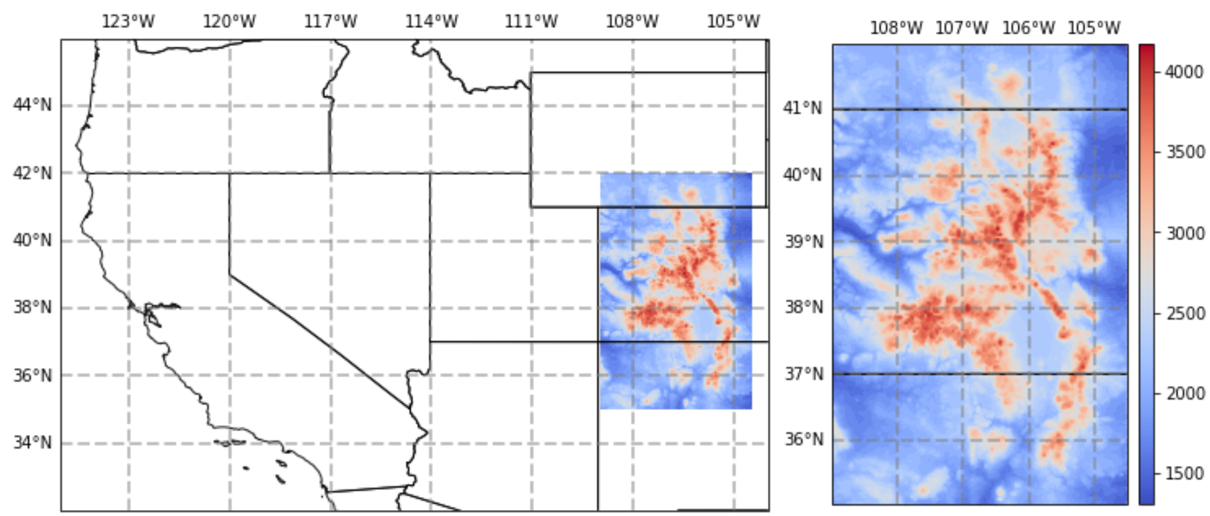


Figure: Rocky Mountain Domain (left) and elevation (right).

Extrapolation

- The seasonality itself improves the generalization.
- By training another set of models, the generalization performance is much better.
- We lose the information of the actual SWE but gain the information over a wider spatial domain. No free lunch.

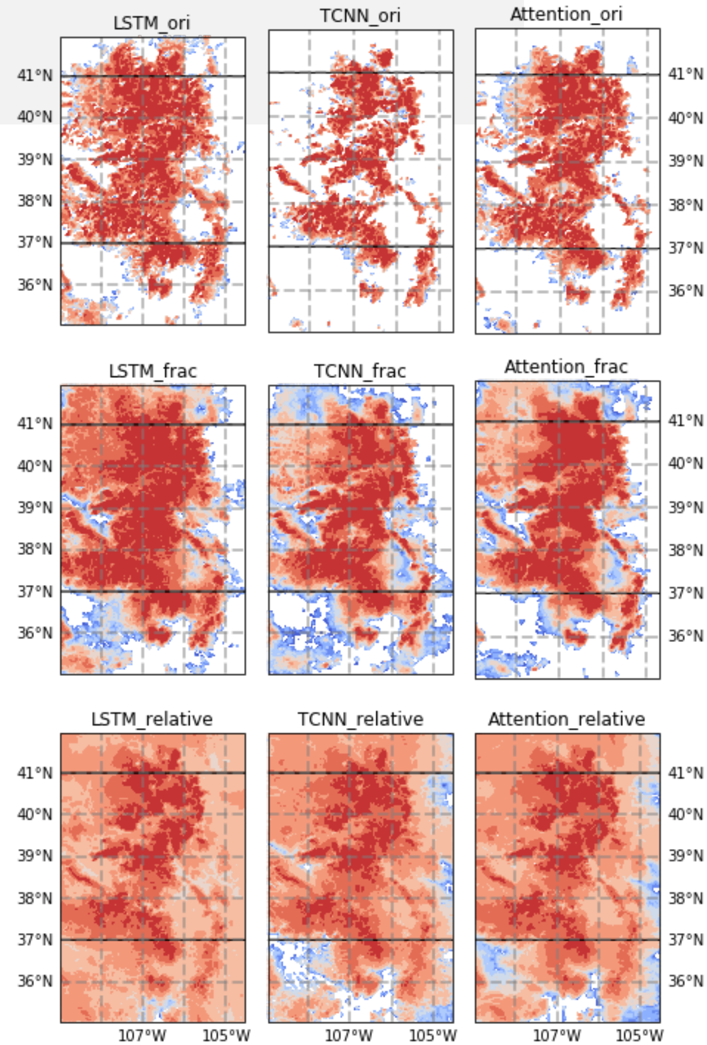


Figure: Extrapolation results.

Projection

- Continue with the SWE percentage and analyze the response of SWE to climate change.
- Use LOCA dataset as forcings. Select CESM-CAM5, CNRM-CM5, EC-EARTH, GFDL-ESM2M, HadGEM2-ES, and MIROC5.
- From the SWE seasonality, we used following metrics to assess the snowpack change.

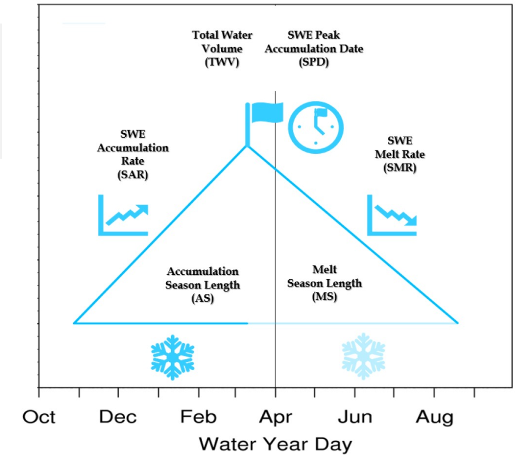


Figure: Our SWE Triangle Multi-Metric Framework (Rhoades et al., 2018)

Metric	Units	Assessment thresholds
Snowpack accumulation start date (SAD)	Days since Oct 1st	Day when SWE > 10% of maximum SWE
Snowpack peak accumulation date (SPD)	Days since Oct 1st	Day of maximum SWE
Complete melt date (CMD)	Days since Oct 1st	Day when SWE < 10% of maximum SWE
Snow season length	Number of days	Sum of days from SAD to CMD

Projection Results

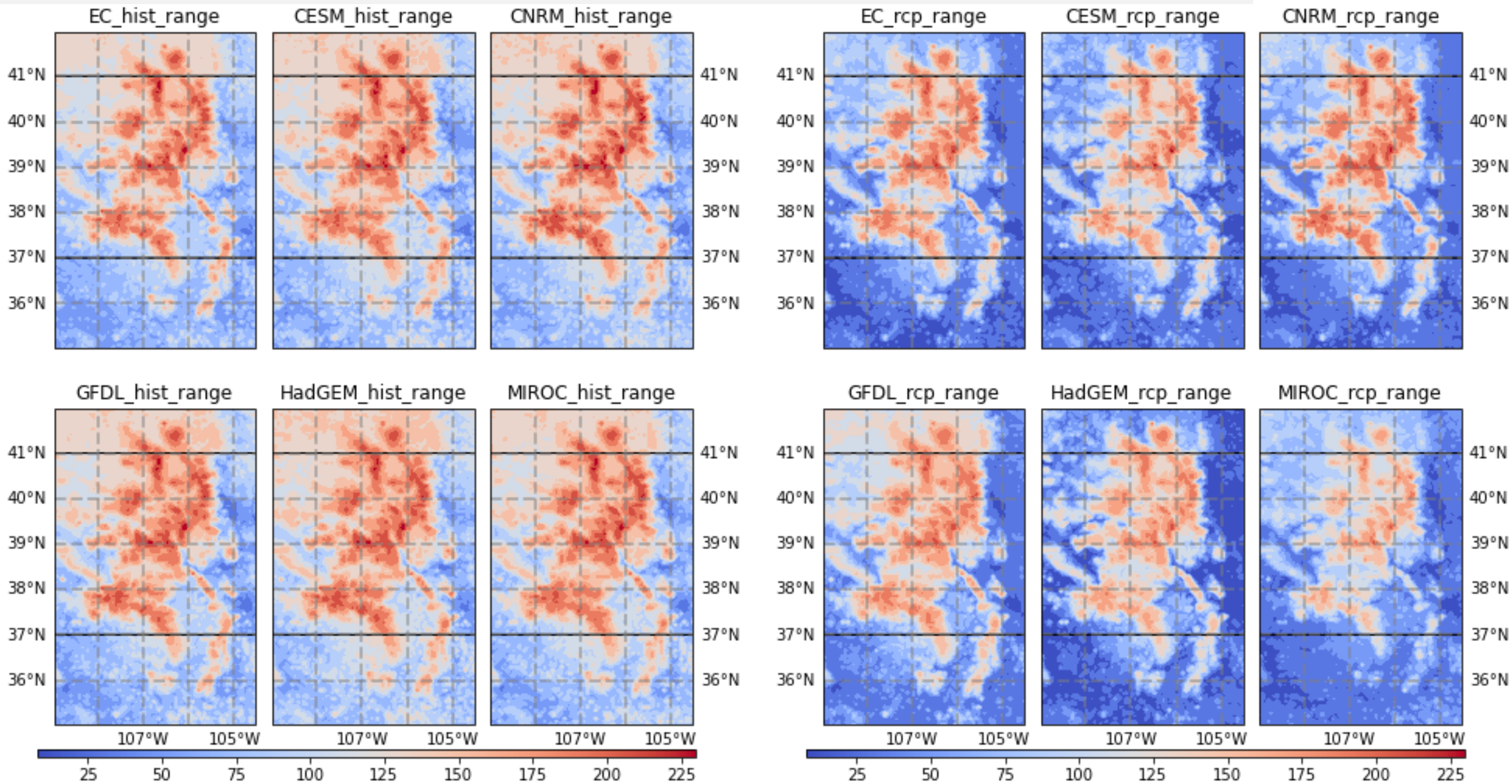
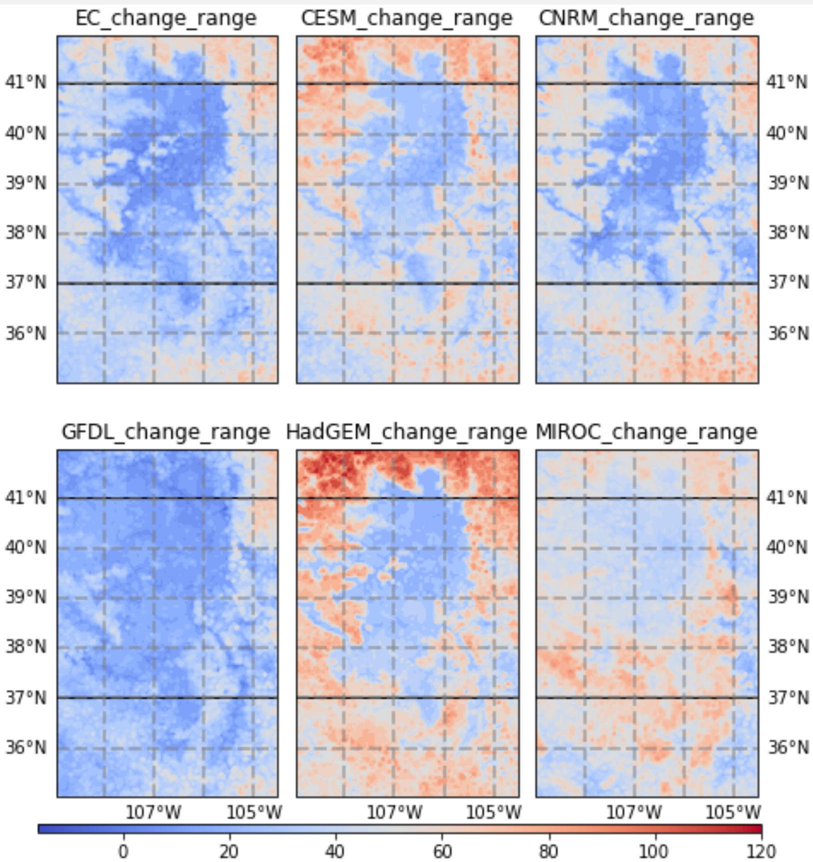


Figure: Historical (left) and RCP8.5 (right) projections of snow season length (days).

Projection Results



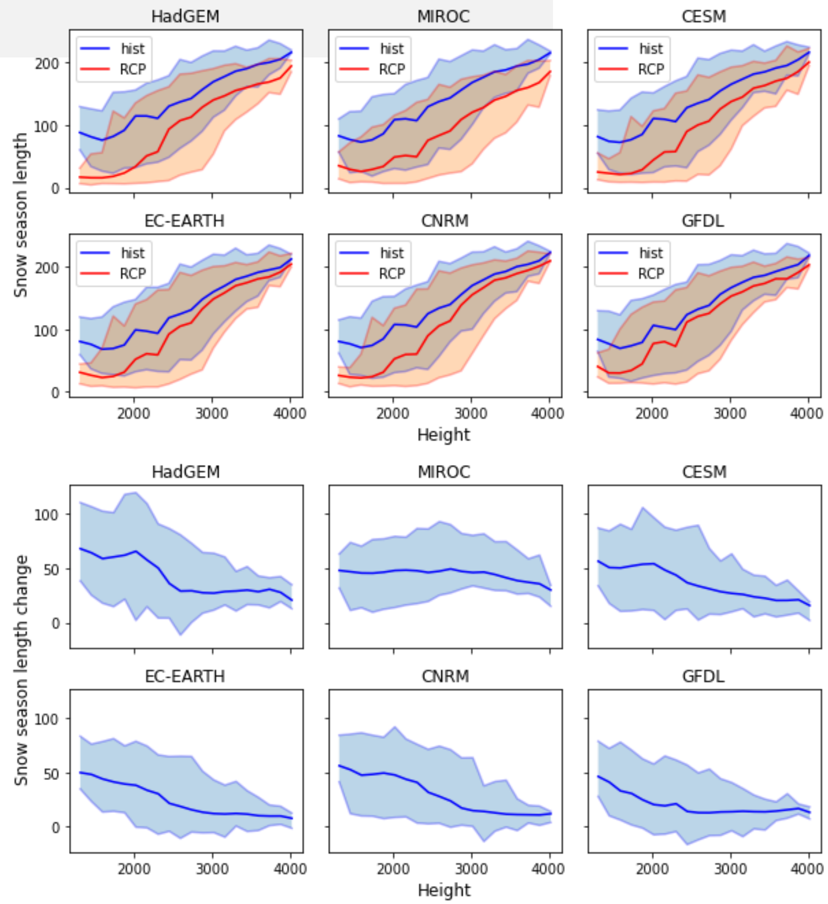


Figure: Snow season length changes in the future (left) and the height dependency (right).

Future Work

- Couple with a physical-based model to address extrapolation problems;
- Data assimilation from satellite-based products for low-elevation area;
- Explainable AI method to analyze the physical impactors;
- Generalize to a wider area;
- Projections with CMIP6 models when LOCA data is available.

Thanks

- We would like to acknowledge the helpful discussion with Chaopeng Shen, Wen-Ping Tsai from Pennsylvania State University, David John Gagne from NCAR, Mark Risser, Alan Rhoades and Chris Paciorek from Lawrence Berkeley National Laboratory.

- The computational platform used for this study is the Tempest GPU cluster at UC Davis. We also thank the Computational and Information System Lab for access to the Casper cluster through the Advanced Study Program at NCAR.

- Any further questions or suggestions, please contact at shiduan@ucdavis.edu