Exploring Predictability of Winter California Hydroclimate with Deep Learning

Anthony Chiado, Kristian Olsson, Luke Rohlwing, Michael Vaden, Antonios Mamalakis

UNIVERSITY OF VIRGINIA DATA SCIENCE INSTITUTE

Abstract

Predicting winter precipitation in California is crucial for policy decisions, ecosystem health, and inhabitants' well-being. However, high variability and difficulty in prediction pose significant challenges. This study explores the potential of a Convolutional Neural Network (CNN) trained on global summer sea surface temperatures (July-October) to forecast winter precipitation (November-March) across northern, central, and southern California. We leverage the Community Earth System Model 2 (CESM2) climate simulations for training and fine-tuning the CNN for real-world predictions. Our approach aims to build upon existing methods of prediction, establishing CNNs as a promising prospect for future research and forecasting accuracy.

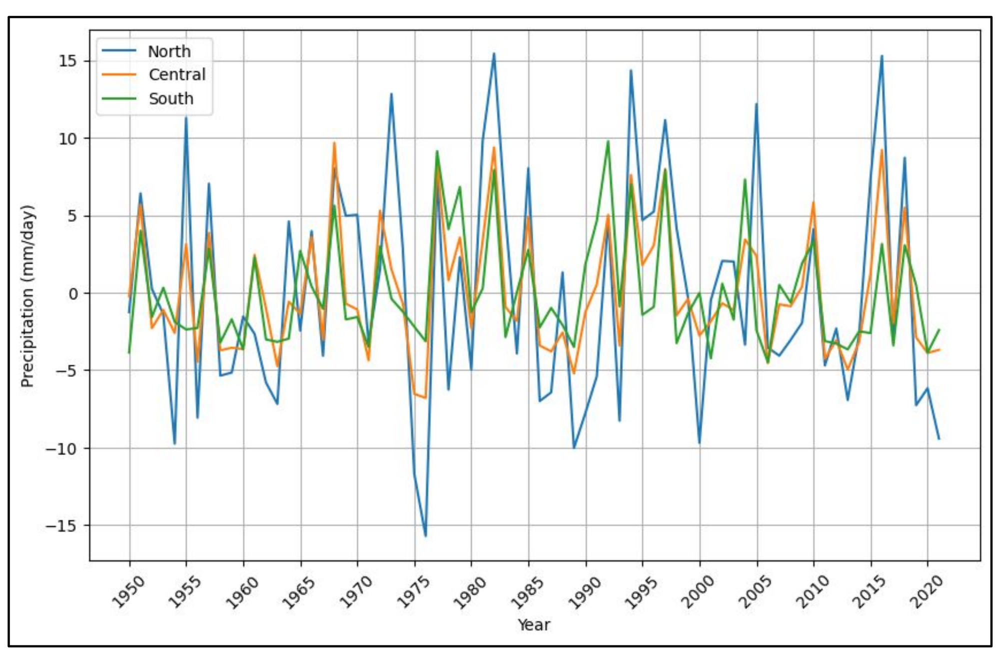

Fig 1: Time Series of Detrended Real-World Precipitation (mm/day) by Region

Methodology

Our methodology involves processing and preparing two key datasets before building our CNN: sea surface temperatures and precipitation data. Both datasets encompass monthly observations of the features across 100 climate simulations from the Community Earth System Model 2 (CESM2) spanning 1940-2013.

The proposed CNN architecture introduces a sophisticated framework for addressing the challenges inherent in predicting precipitation levels across California. Leveraging convolutional, pooling, and fully connected layers, the CNN undergoes iterative training loops to optimize its parameters, minimizing prediction errors and enhancing its predictive capabilities. Additionally, hyperparameter tuning of the learning rate, dropout probabilities, and activation functions further refine the model's performance.

We deploy three distinct models, each employing an identical CNN architecture which are individually trained with different hyperparameters optimized for their respective regions: northern, central, and southern California.

Subsequently, we tune the model weights from the final fully connected layer on real-world data from 1950-1989 and test this fine-tuned model on real-world data from 1990-2021 to test its predictive power.

Lastly, we employ saliency maps, which provide visualizations that highlight the most influential regions within the input data that contribute to the model's predictions. By analyzing these maps, we can discern which geographical features the model prioritizes when making predictions for each region. This interpretive tool provides valuable insights into the factors driving precipitation variability across different parts of California.

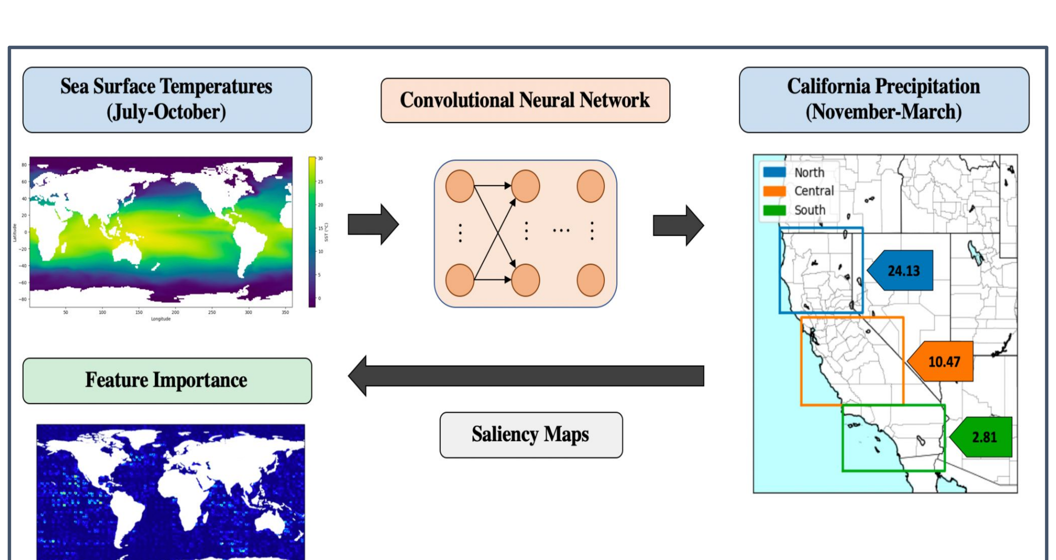

Fig 2: CNN Modeling Process with Saliency Maps

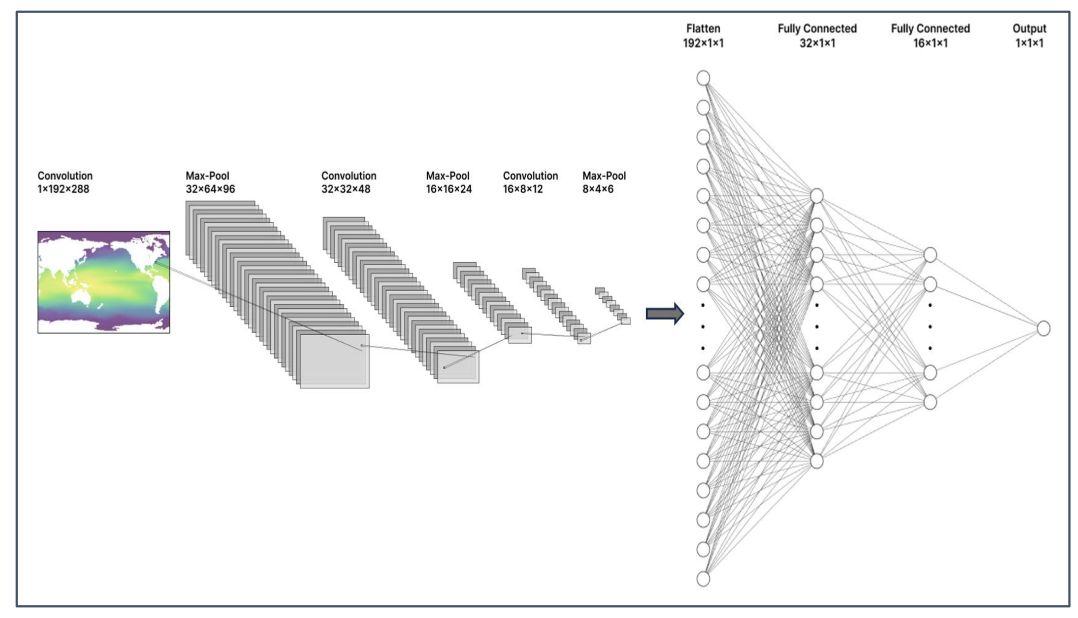

Fig 3: Final Model Architecture

Results

In general, our findings indicate that CNNs can be successfully applied to the problem at hand, and show promise for greatly improved prediction in the future.

Key Findings:

- CNN models outperform linear regression baseline on all test datasets for all regions
- Northern California is the most difficult to predict; Southern California is the easiest
- Fine tuning the model on real world data greatly improves predictive performance
- Saliency maps indicate predictions heavily influenced by sea surface temperatures in Indian Ocean, El Nino Southern Oscillation, Southern Ocean of the coast of Australia, and Northern Atlantic Ocean

R^2 Values

	CESM2 Data (Train: 80 sims Test: 10 sims)		*Real World Data (Test Years 1990-2021)*	
California Region	*El Niño Linear Model*	*CNN*	*El Niño Linear Model*	*CNN (Fine Tuned)*
Northern	0.000	0.037	0.005	0.089
Central	0.084	0.132	0.029	0.311
Southern	0.155	0.212	0.108	0.336

Fig 4: Predictive Performance on Model and Real-World Data

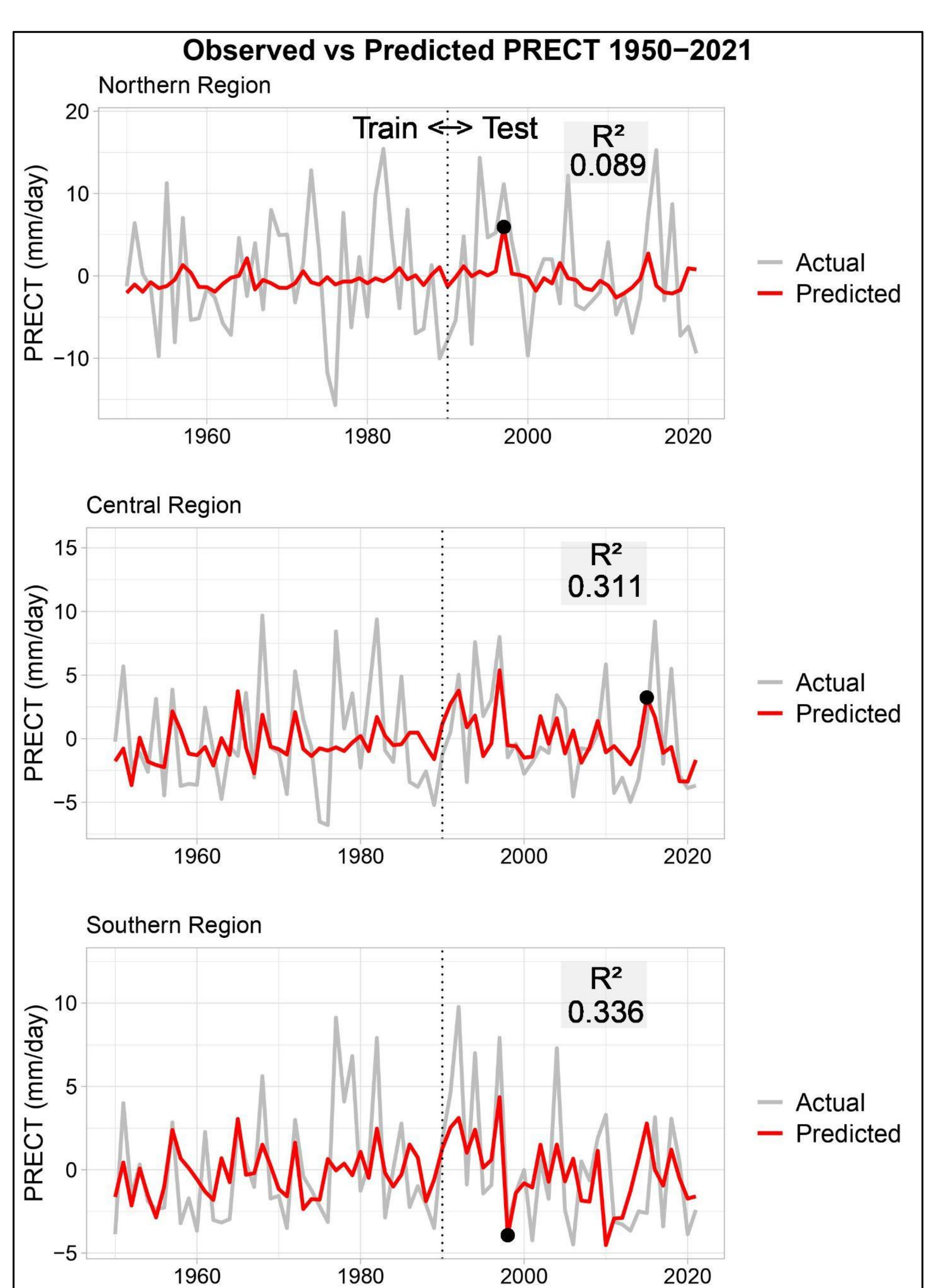

Fig 5: CNN Performance on Real World Data

Conclusion

Our Convolutional Neural Networks trained on maps of global summer sea surface temperatures were able to predict northern, central, and southern California precipitation to varying degrees of success. Our paper establishes CNNs as a successful and promising solution to predicting California precipitation, and there are many future steps to consider for this approach.

Limitations:

- Lack of real-world data
- Model complexity limited by data size

Areas of Future Work:

- Use data from different simulations
- Introduce more time-lags
- Use additional predictors (pressure, etc.)
- Continue hyperparameter/architecture tuning

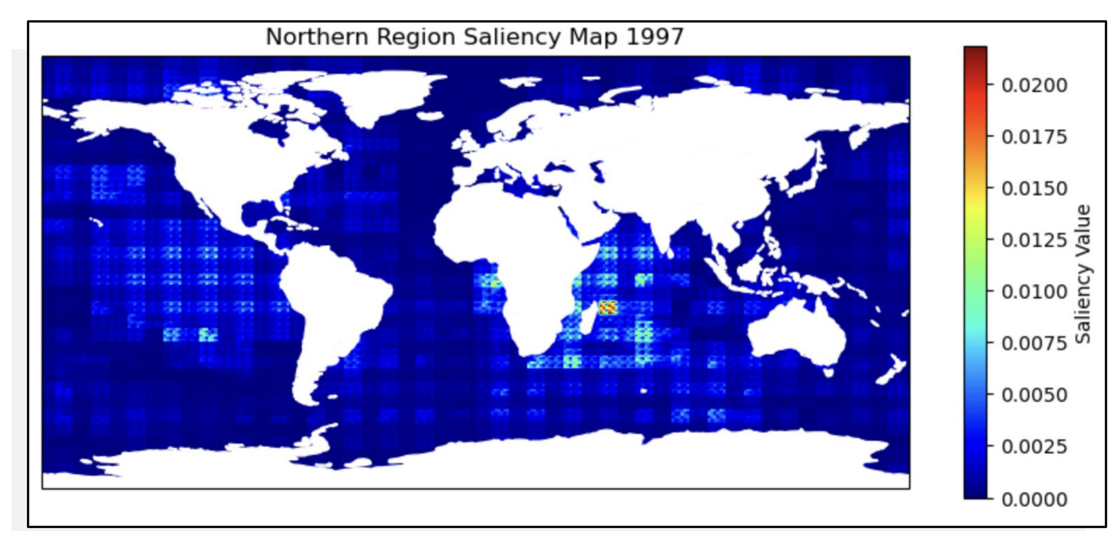

Fig 6: Saliency Map for Real-World Prediction of Northern Region in 1997 (See marked point in Fig 5 for prediction)

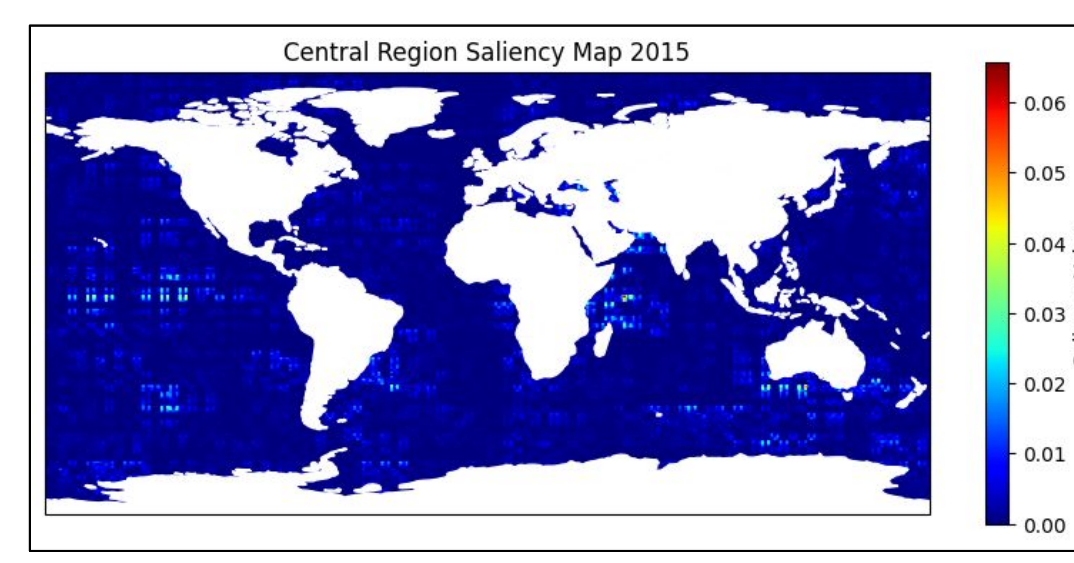

Fig 7: Saliency Map for Real-World Prediction of Central Region in 2015 (See marked point in Fig 5 for prediction)

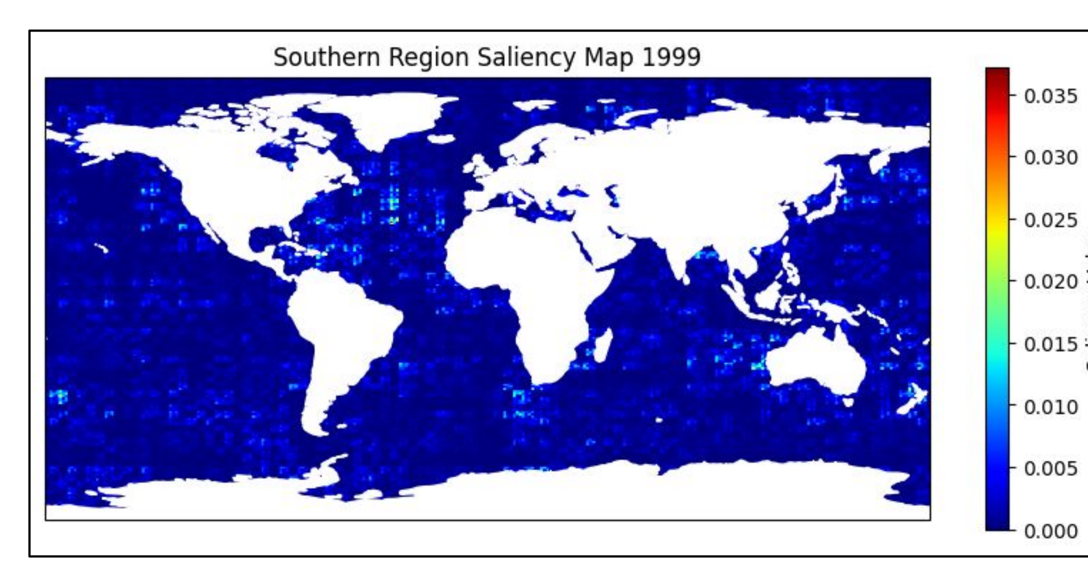

Fig 8: Saliency Map for Real-World Prediction of Southern Region in 1999 (See marked point in Fig 5 for prediction)

[1] Allen, R., Luptowitz, R. El Niño-like teleconnection increases California precipitation in response to warming. *Nat Commun* **8**, 16055 (2017). https://doi.org/10.1038/ncomms16055

[2] Gibson, P.B., Chapman, W.E., Altinok, A. et al. Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. *Commun Earth Environ* **2**, 159 (2021). https://doi.org/10.1038/s43247-021-00225-4

[3] Liu, T., Schmitt, R. W., & Li, L. (2018). Global search for autumn-lead sea surface salinity predictors of winter precipitation in southwestern United States. *Geophysical Research Letters*, **45**(16), 8445–8454. https://doi.org/10.1029/2018gl079293

[4] Mamalakis, A., Yu, JY., Randerson, J.T. et al. A new interhemispheric teleconnection increases predictability of winter precipitation in southwestern US. *Nat Commun* **9**, 2332 (2018). https://doi.org/10.1038/s41467-018-04722-7

[5] Stevens, A., Willett, R., Mamalakis, A., Foufoula-Georgiou, E., Tejedor, A., Randerson, J. T., Smyth, P., & Wright, S. (2020). Graph-Guided Regularized Regression of Pacific Ocean Climate Variables to Increase Predictive Skill of Southwestern U.S. Winter Precipitation. *Journal of climate*, **34**(2), 737–754. https://doi.org/10.1175/jcli-d-20-0079.1

UNIVERSITY of VIRGINIA | DATA SCIENCE