A Deep Learning Approach to Subseasonal Forecasting of Severe Weather over the United States

The UNIVERSITY of OKLAHOMA

Maria M. Madsen[1,3], Eleanor Salm[3,4], and Amy McGovern[1,2,3]

[1]School of Computer Science, University of Oklahoma, Norman, OK, USA; [2]School of Meteorology, University of Oklahoma, Norman, OK, USA

[3]NSF AI Institute for Research on Trustworthy AI In Weather, Climate and Coastal Oceanography (AI2ES), Norman, OK, USA; [4]University of Wisconsin-Madison, Madison, WI, USA

Motivation

The deficiency in predictability at subseasonal timescales relative to that at conventional weather prediction timescales is significant. With respect to severe weather over the United States, previous work has shown promise for prediction at longer timescales through consideration of various teleconnections. For example, tropical modes of variability like the Madden Julian Oscillation have been shown to modulate severe weather frequency in the weeks following specific phases. Also tied to the occurrence of severe weather are synoptic features like the jet stream, deep upper-level troughs, and tropopause polar vortices. However, many of these processes and teleconnections have been exclusively investigated in the context of severe weather outbreaks over the United States without consideration of interference between them.

Objective:

Use machine learning to improve and better understand subseasonal predictability of severe weather over the U.S..

Data

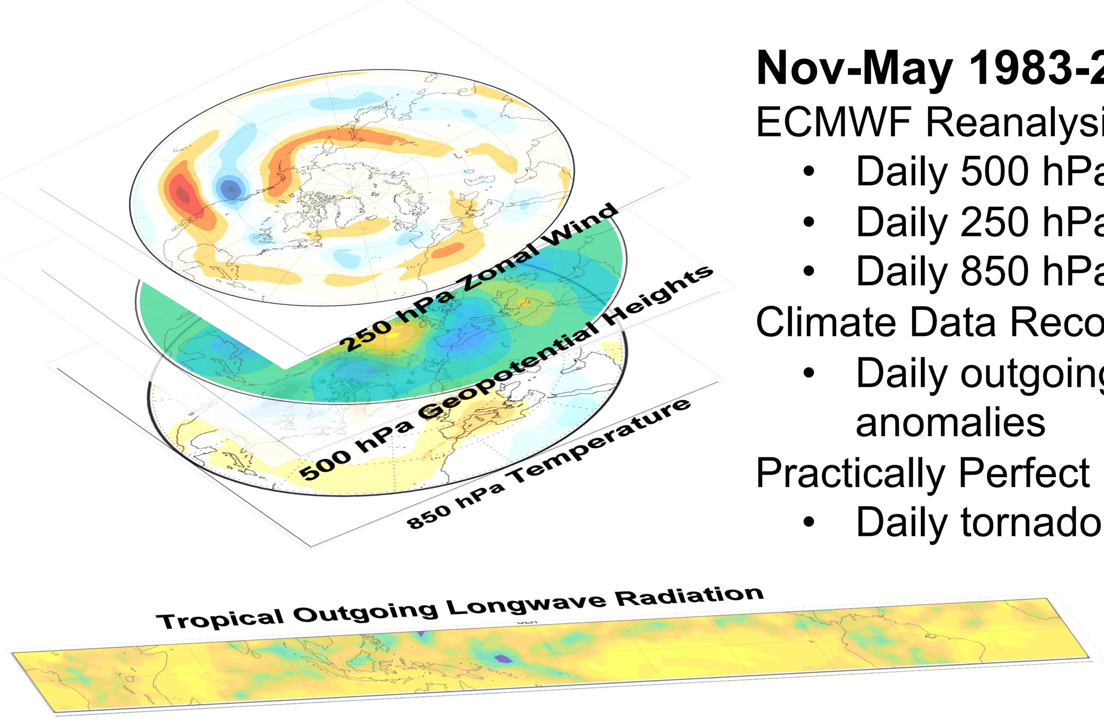

Fig. 1: Example of normalized daily input data for the ML model.

Nov-May 1983-2019

ECMWF Reanalysis v5 (ERA5)
- Daily 500 hPa geopotential height anomalies
- Daily 250 hPa zonal wind anomalies
- Daily 850 hPa temperature anomalies

Climate Data Record (CDR)
- Daily outgoing longwave radiation (OLR) anomalies

Practically Perfect Hindcasts (PPH)
- Daily tornado, hail, and wind probabilities

Practically Perfect Hindcasts

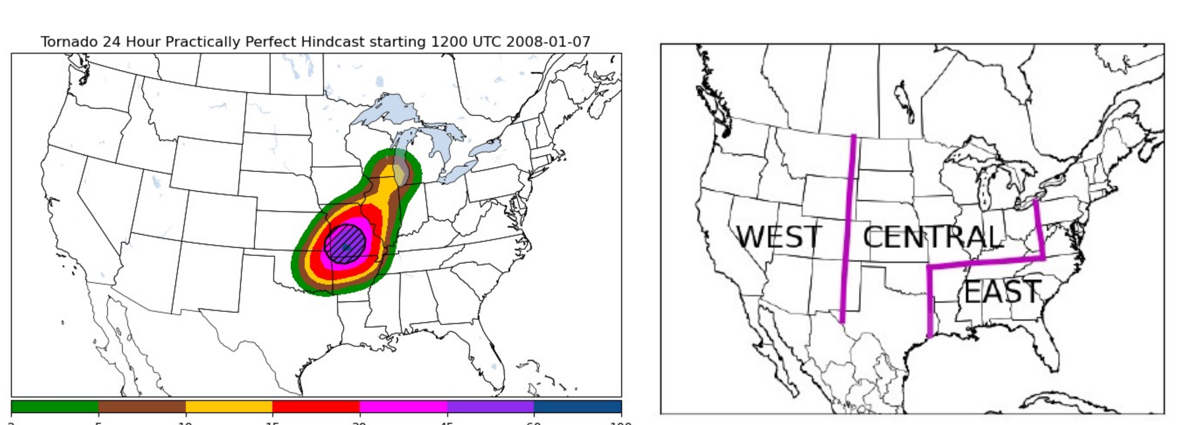

Fig. 2: (left) example of tornado PPH data. (right) Fig. 4 from Hill et al. (2023) showing the division of the three U.S .regions for the PPH data

PPH Details:
- PPH data (tornado, hail, wind probabilities) divided into 3 U.S. regions based on Hill et al. (2023)
- Averaged over 3 regions

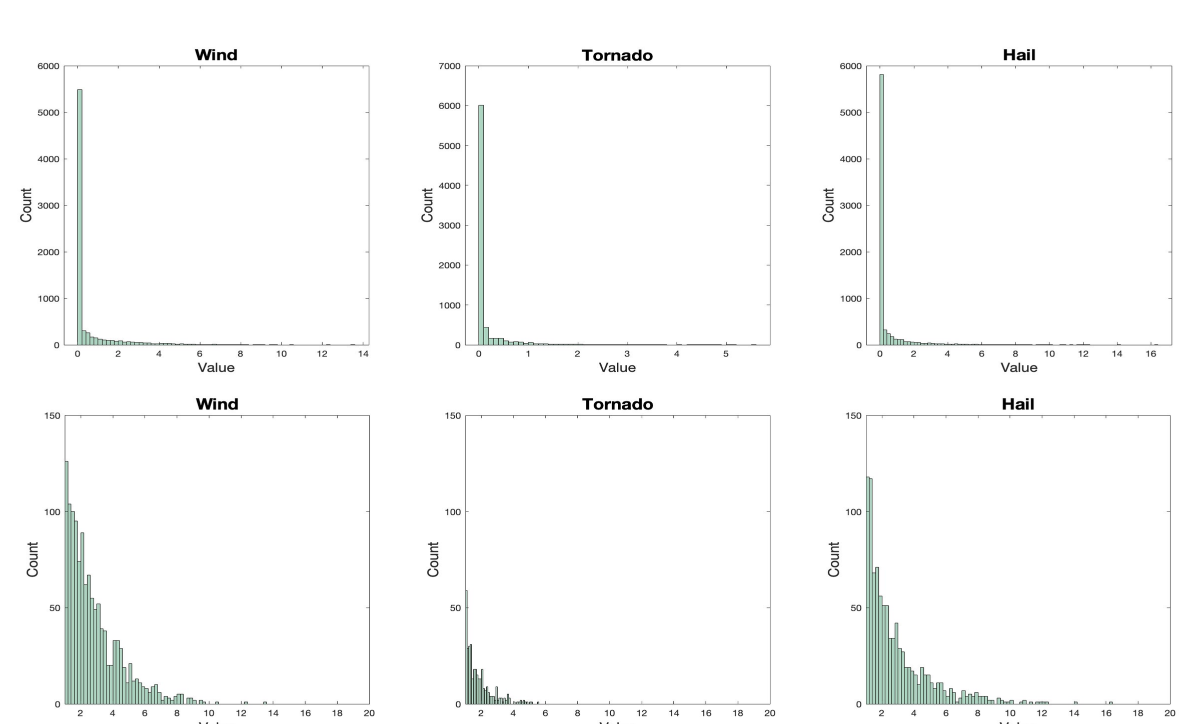

Fig. 3: Distribution of severe weather hazards over the central U.S. region for (top) all samples and (bottom) all nonzero samples.

CNN Model Architecture

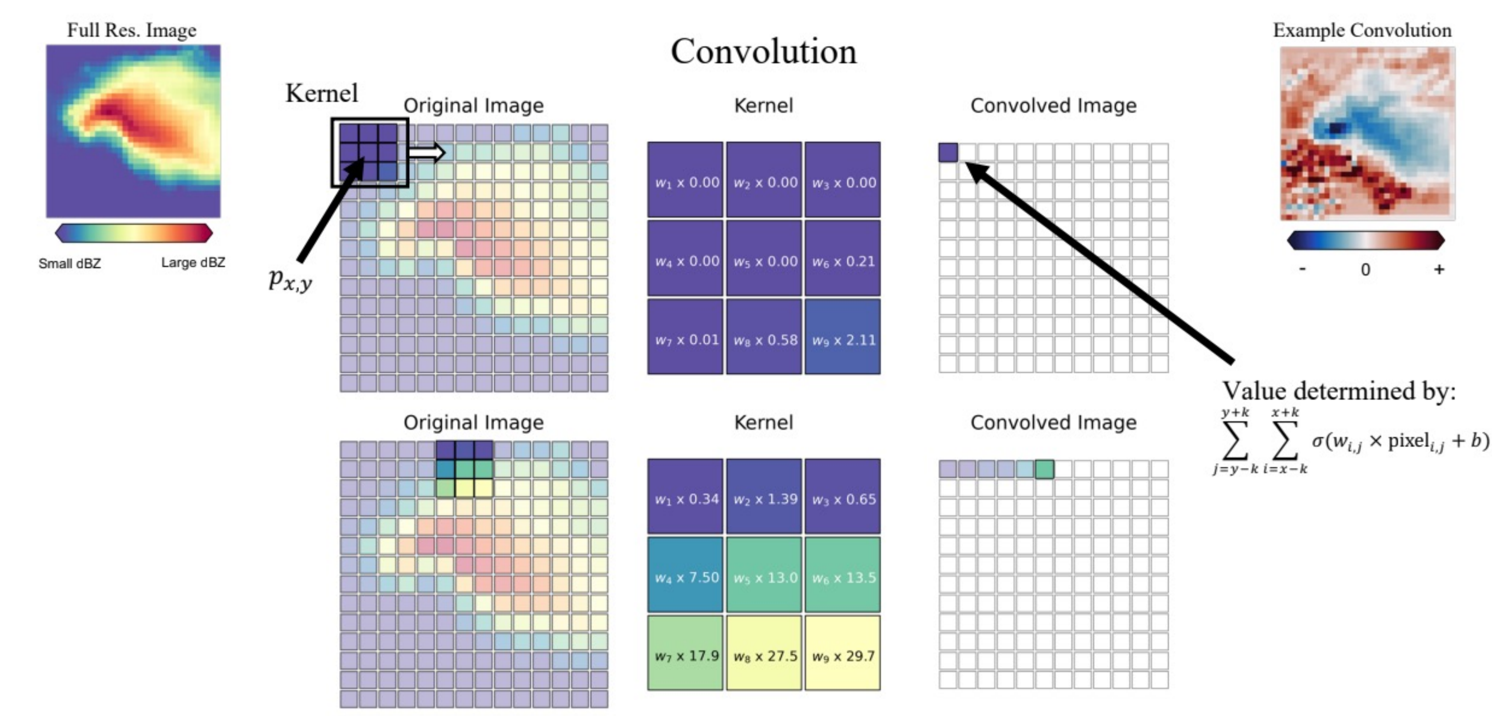

Fig. 4: From Fig. 3. of Chase et al. (2023). Schematic of a convolutional neural network on radar reflectivity at various steps. Original hook echo radar reflectivity is depicted in top left before the (top) step 0 convolution and (bottom) step 6 convolution. Final hook echo radar reflectivity is depicted in top right after full convolution.

- We utilize a convolutional neural network
- Hyperparameter searching is run with hundreds of combinations of varying parameters such at depth, kernel size, learning rate, etc.
- Regression CNN performed better than classification CNN
- Loss function : Mean Squared Error
- Training/Validation/Loss is an 80/10/10 split

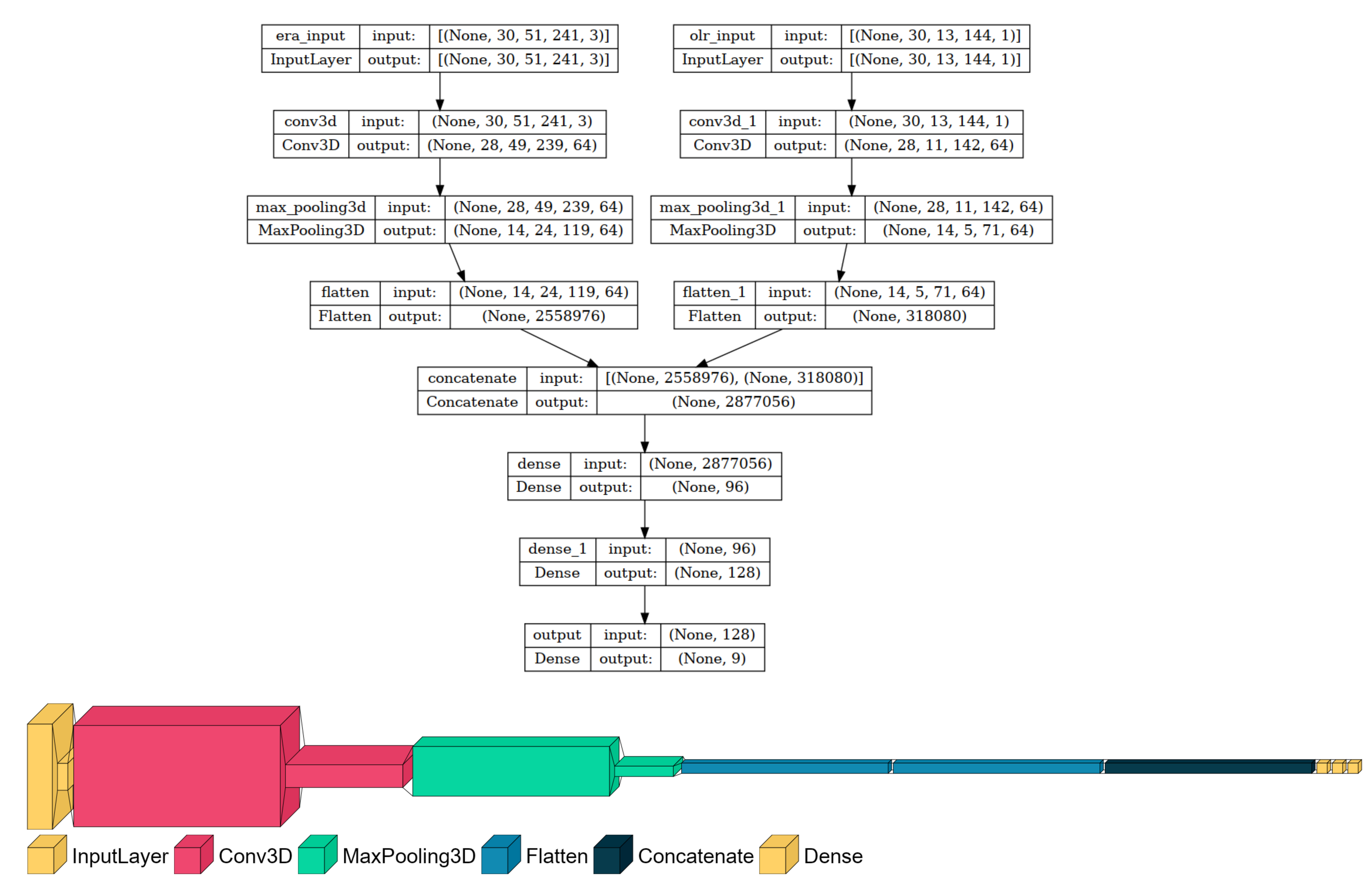

Fig. 5: CNN architecture with 2 input channels (OLR and ERA-5 data), and an output layer with 9 channels (tornado, wind, and hail for 3 U.S. regions)

Explainable AI (XAI)

- ML models are often seen as a 'black-box'.
    - XAI is a way to interpret the model with respect to the decisions made
- Utilize various XAI methods to investigate most important features in the inputs for best forecasts of medium range severe weather hazards
    - LRP_{COMP} : Uses layerwise backpropagation to determine the relevance of each input neuron to the output
    - Deep SHAP: Calculates Shapley values for small components of network and backpropagates to input layer

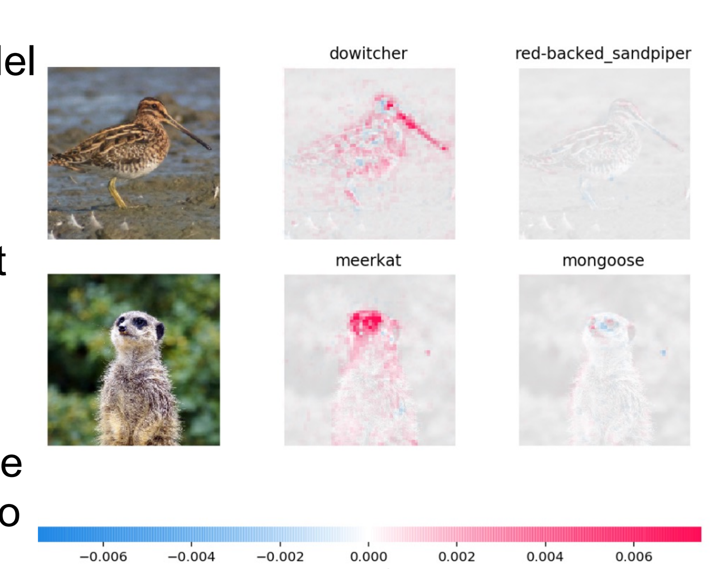

Fig. 8: GradientExplainer using SHAP from https://github.com/shap/shap

Model Performance

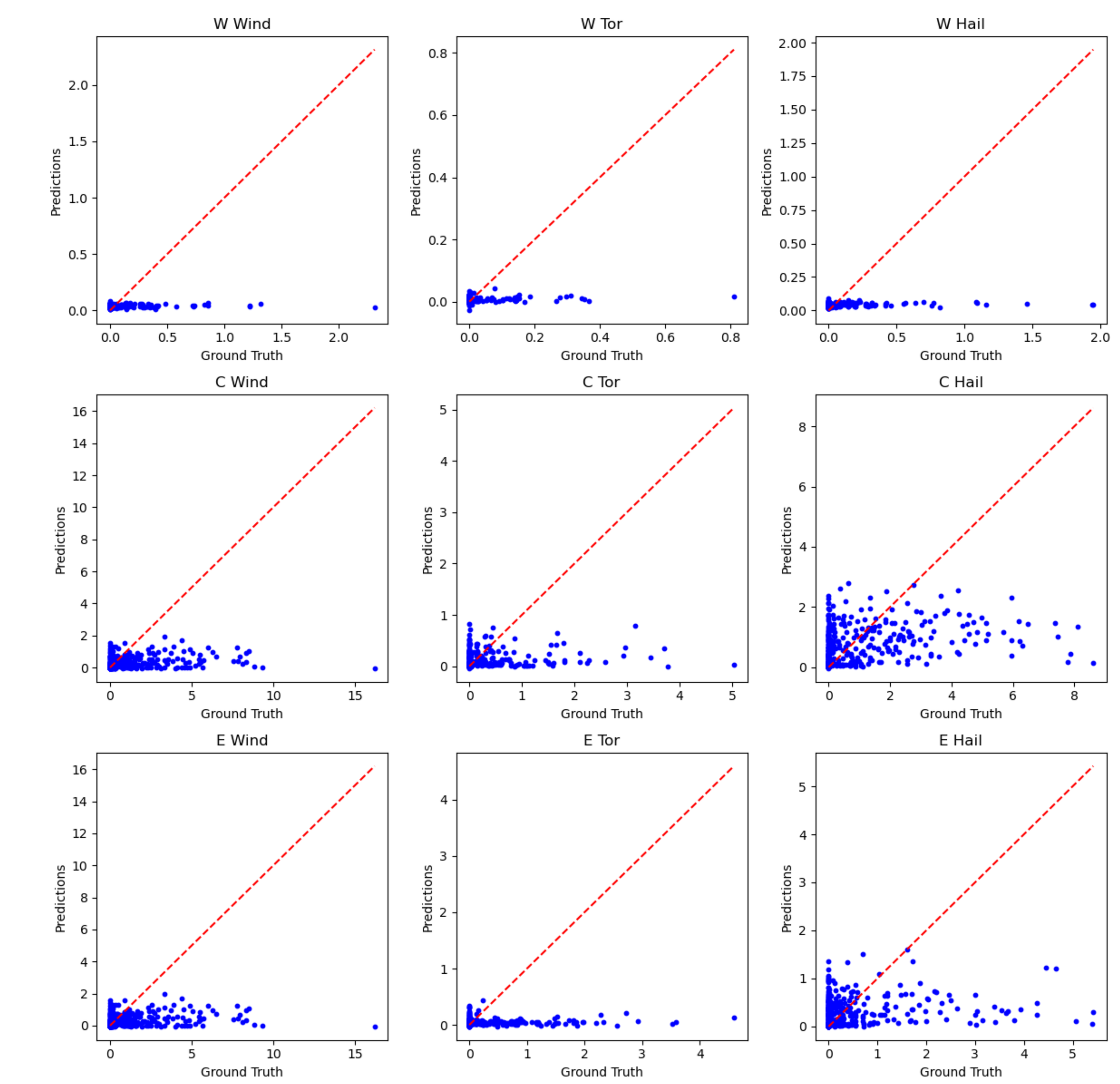

Fig. 6: Test set evaluation of CNN prediction. (Top) Western U.S. (middle) central U.S. and (bottom) eastern U.S.

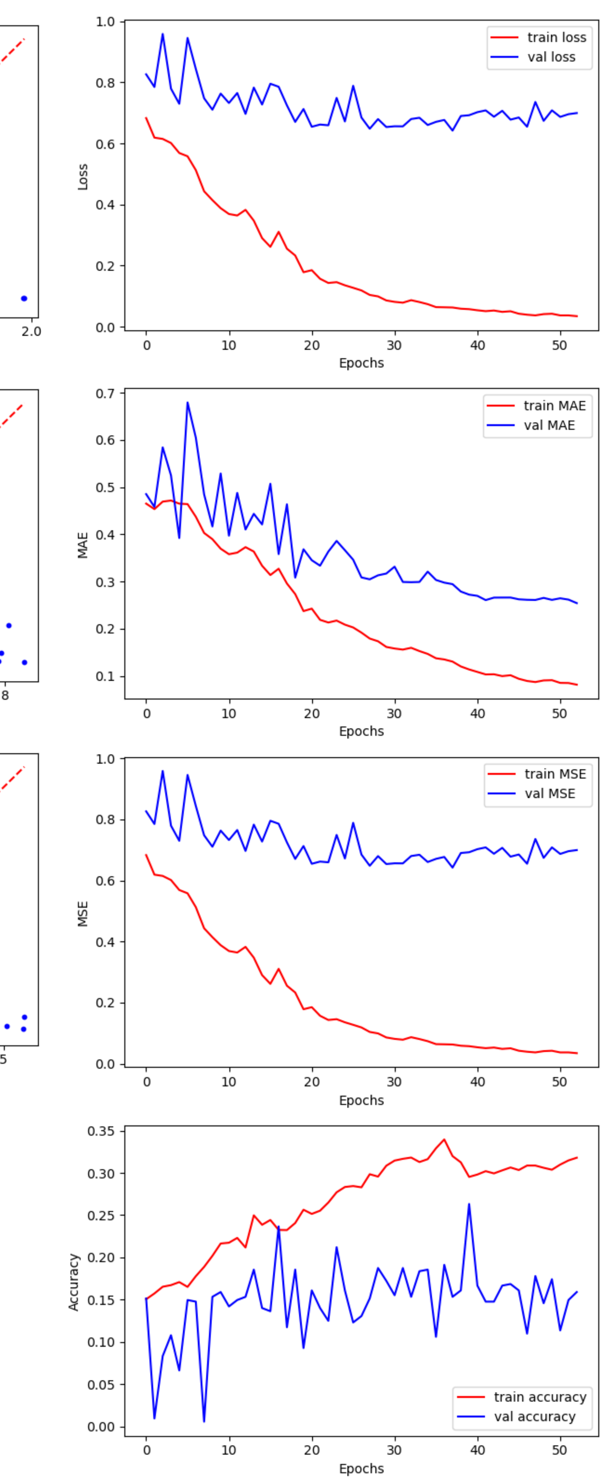

Fig. 7: Various metrics of CNN performance over epoch number.

- CNN needs improvement
- Training loss vs validation loss (Fig. 7) exhibit large gap suggesting an unrepresentative training dataset
- CNN performs better for hail than tornado or wind.
- CNN performs best over central U.S. and worst over western U.S.

Conclusions/Future Work

1. A CNN was trained on large-scale predictors including 250 hPa zonal wind, 500 hPa geopotential height, 850 hPa temperature, and tropical outgoing longwave radiation to predict probabilities for severe weather hazards over 3 different regions of the U.S.
2. Current CNN needs improvement
    1. CNN performs best for central and eastern U.S. hail probabilities.
    2. Worst performance noted over western U.S.
3. Ongoing work includes adding predictors
    1. Indices for Madden Julian Oscillation, El Niño–Southern Oscillation, Arctic Oscillation, etc.
    2. Tropopause polar vortices

References

Chase, R. J., D. R. Harrison, G. M. Lackmann, and A. McGovern, 2023: A Machine Learning Tutorial for Operational Meteorology, Part II: Neural Networks and Deep Learning. Wea. Forecasting, in press.

Gensini, V. A., A. M. Haberlie, and P. T. Marsh, 2020: Practically Perfect Hindcasts of Severe Convective Storms. Bull. Amer. Meteor. Soc., 101, E1259–E1278.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146 (730), 1999–2049.

Hill, A. J., R. S. Schumacher, and I. L. Jirak, 2023: A New Paradigm for Medium-Range Severe Weather Forecasts: Probabilistic Random Forest–Based Predictions. Wea. Forecasting, 38, 251–272.

Lundberg, S. M., and S. I. Lee, 2017: A Unified Approach to Interpreting Model Predictions. Proc. 31st Int. Conf. on Neural Information Processing Systems, Long Beach, CA, NIPS, 4768–4777.

Mamalakis, A., E. A. Barnes, and I. Ebert-Uphoff, 2022: Investigating the Fidelity of Explainable Artificial Intelligence Methods for Applications of Convolutional Neural Networks in Geoscience. Artif. Intell. Earth Syst., 1, e220012.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. ICER-2019758.