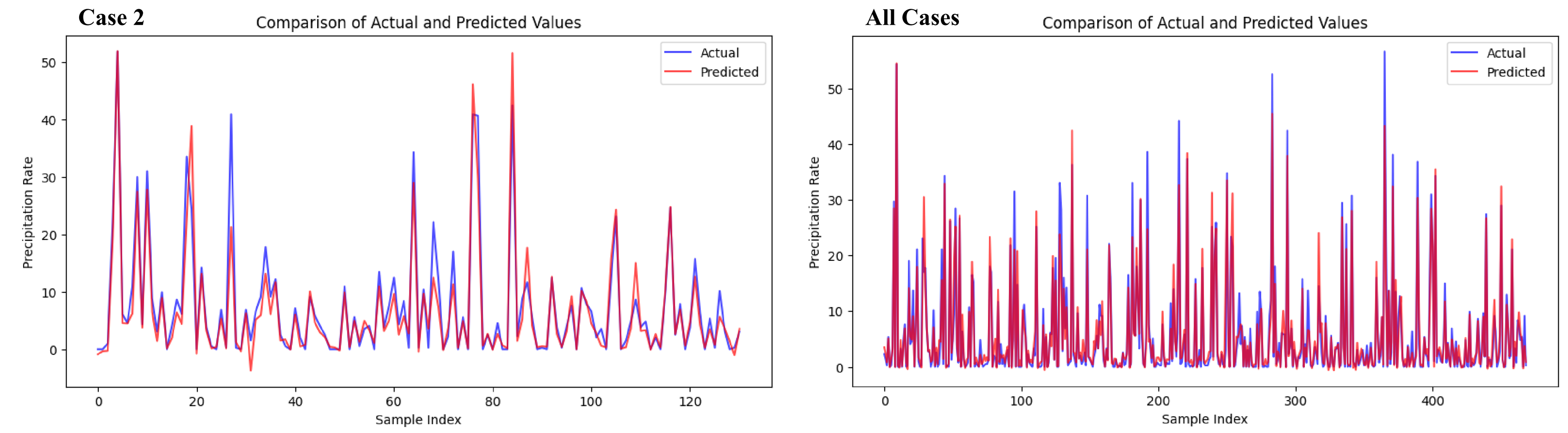
Polarimetric Radar QPE Using Neuro-Symbolic Learning (NSL)

Mostafa Cham[1], Haotong Jing[2], Olivia Zhang[2], Weikang Qian[2], Yixin Wen[2], Jianwu Wang[1]
[1] Department of Information Systems, University of Maryland, Baltimore County, Maryland, United States
[2] Department of Geography, University of Florida, Florida, United States

1. Introduction

- Precipitation is crucial for agriculture, water resource management, and disaster planning, influencing regional climate and weather extremes such as droughts and floods. **Understanding and predicting** precipitation patterns are essential for sustainable development and resilience against climate change.
- Though rain gauges are highly accurate in precipitation observation, their limited coverage, especially in remote areas, necessitates alternative methods. **Weather radar**, which provides extensive coverage and frequent updates, is one approach that offers significantly more coverage than rain gauges.
- To convert radar reflectivity (Z) to rain rate (R), traditional methods use **exponential Z-R relationships**. Polarimetric radar offers detailed data on raindrop shape, size, and distribution, enhancing **quantitative precipitation estimation (QPE)** accuracy by reducing reliance on traditional Z-R relationships, which can vary under different conditions.
- Most machine learning methods do well in improving the QPE accuracy, however, at the cost of lack of interpretation and domain understanding. This work proposes **Neuro-Symbolic Learning (NSL)** methods, combining **neural network** and **symbolic regression** algorithms, for QPE regime, to improve traditional QPE with both **higher accuracy and clearer interpretation insights**.



Metrics	1. Case 2	2. All Cases
MSE	11.333	7.105
RMSE	3.366	2.665
MAE	1.905	1.491
R^2	0.89	0.92

2. Data & Methods

- **Data**

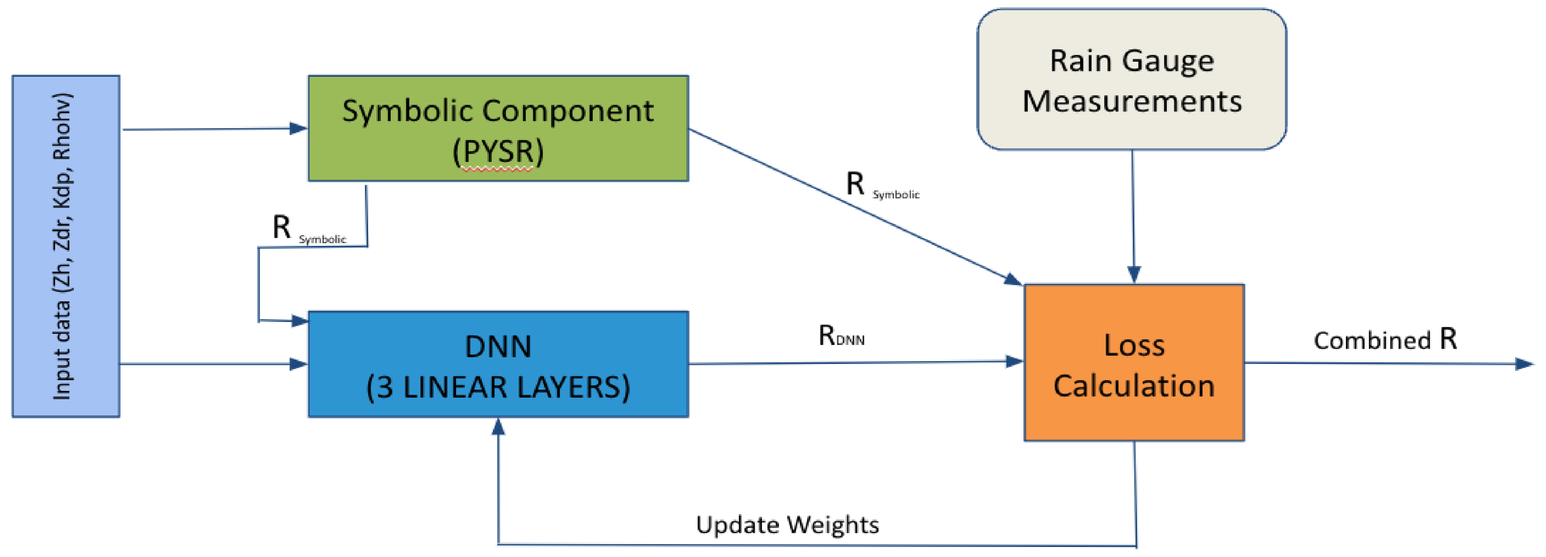
  3 cases from Florida and Oklahoma including convective warm tropical rain and continental stratiform rain events. Radar data are from WSR-88D radars KAMX (Miami) and KTLX (Oklahoma City). Precipitation data are from South Florida Water Management District (SFWMD) gauges and Oklahoma Mesonet gauges.

  Case 1: April 12, 2023; Miami; Convective and warm tropical rain

  Case 2: July 9, 2023; Oklahoma City; Continental stratiform rain

  Case 3: June 8, 2022; Oklahoma City; Continental stratiform rain

- **Method** (Neuro-Symbolic Learning)



$$W_s = \frac{1}{L_s}, \quad W_d = \frac{1}{L_d}$$

$$W_s = \frac{W_s}{W_s + W_d}, \quad W_d = \frac{W_d}{W_s + W_d}$$

$$R_{\text{combined}} = W_s R_{\text{symbolic}} + W_d R_{\text{DNN}}$$

3. Results

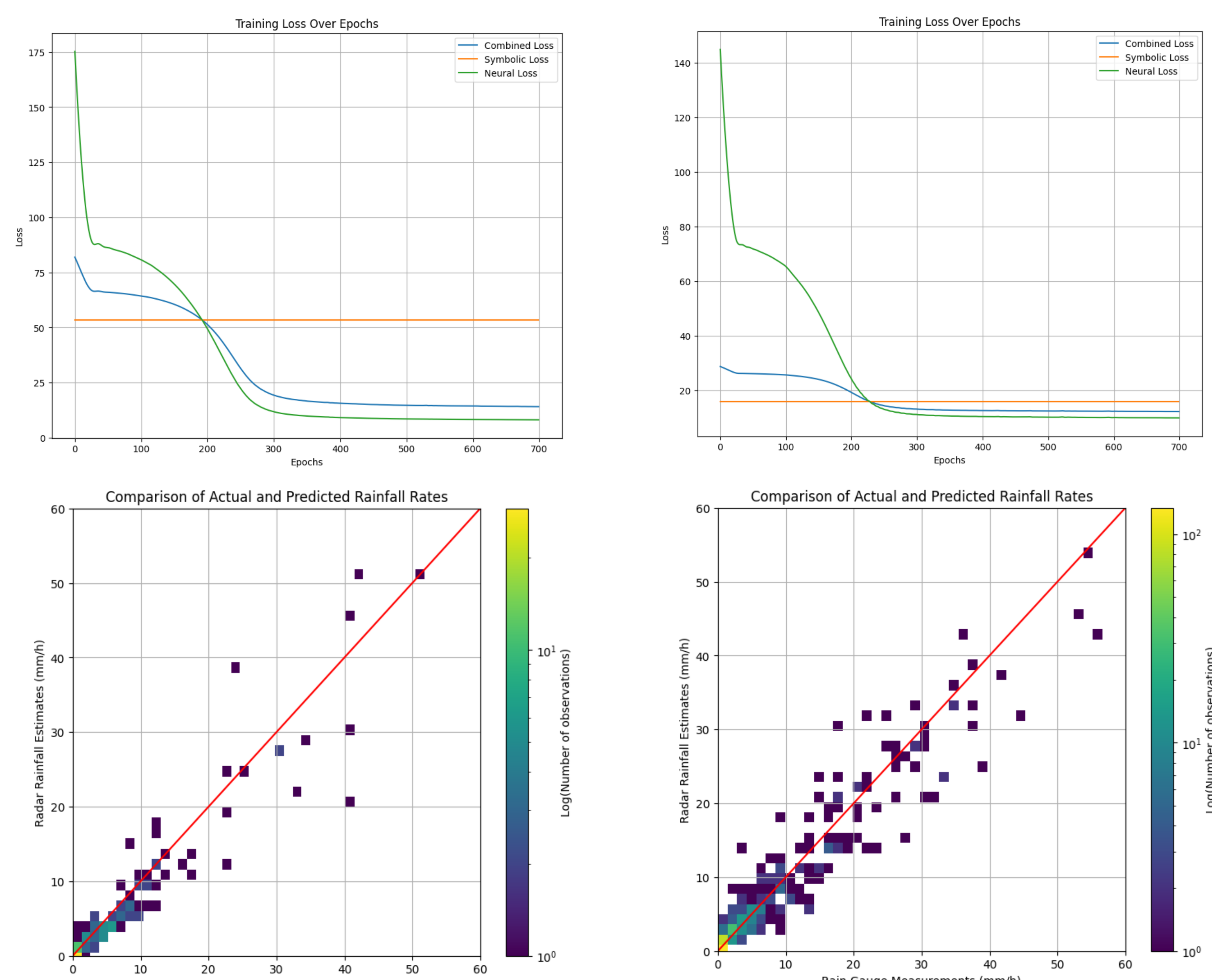
1. Empirical equations used for **Case 2** (July 9, 2023; Oklahoma City) [4, 5]: $\text{Heavy Rain Condition} = (Z_{dr} > 0.3) \wedge (K_{dp} > 0.4)$

$$R_{heavy} = 40.6 \times K_{dp}^{0.85} \qquad R_{light} = \left(\frac{\text{Reflectivity}}{200}\right)^{\frac{1}{1.6}}$$

2. Best equation found using symbolic regression for **all cases** [6]:

$$R = (K_{dp} + 0.07898629) \times (\text{Reflectivity} - Z_{dr})$$

Training and Testing Results (Left: 1, Right: 2):



4. Conclusion

- Initially, we tried the empirical equations as part of the symbolic component to estimate the precipitation rate for Case 2.
- Next, we employed the symbolic regression algorithm (PySR) [6] to extract the most fitted equation on the data (all 3 cases) and used that equation for precipitation rate estimation, which resulted in more accurate estimations.
- We combined the estimation from the symbolic component and deep neural network (DNN) component, utilizing the weighted average loss of each.
- The objective of the proposed model is to leverage the interpretability of Symbolic AI together with the high accuracy of DNNs in capturing spatial patterns in data.

5. Discussion & Future Work

Neuro-Symbolic AI for Science

- This work demonstrates the ability of NSL to discover and improve the understating of domain knowledge under the current data-driven regime with the help of explicit expressions.
- It is important to validate and evaluate the consistency of the best expressions learned from different powerful symbolic regression algorithms based on the same datasets in the future study.

Algorithm Development

- Human-machine interaction: Leverage human feedback in the optimization loop to improve symbolic learning for human knowledge-aided estimation.
- Efficient computation: Massive GPU parallel implementation and faster optimization for formula estimation.

[1] Wang, L., & Chen, H. (2023, January). Machine learning for polarimetric radar quantitative precipitation estimation. In 2023 United States National Committee of URSI National Radio Science Meeting (USNC-URSI NRSM) (pp. 298-299). IEEE.
[2] Osborne, A. P., Zhang, J., Simpson, M. J., Howard, K. W., & Cocks, S. B. (2023). Application of Machine Learning Techniques to Improve Multi-Radar Multi-Sensor (MRMS) Precipitation Estimates in the Western United States. Artificial Intelligence for the Earth Systems, 2(2), 220053. https://doi.org/10.1175/AIES-D-22-0053.1
[3] King, F., Duffy, G., Milani, L., Fletcher, C. G., Pettersen, C., and Ebell, K.: DeepPrecip: a deep neural network for precipitation retrievals, Atmos. Meas. Tech., 15, 6035–6050, https://doi.org/10.5194/amt-15-6035-2022, 2022.
[4] Ryzhkov, A., & Zrnić, D. (1996). Assessment of rainfall measurement that uses specific differential phase. Journal of Applied Meteorology and Climatology, 35(11), 2080-2090.
[5] Bringi, V. N., & Chandrasekar, V. (2001). Polarimetric Doppler weather radar: principles and applications. Cambridge university press.
[6] Cranmer, M. (2023). Interpretable machine learning for science with PySR and SymbolicRegression. jl. arXiv preprint arXiv:2305.01582.