Peeking Inside the Black Box: Explainable AI Methods for a Precipitation-type Model

Belen Saavedra[1,2], David John Gagne II[1], John Schreck[1], Charlie Becker[1], Gabrielle Gantos[1]
[1]National Center for Atmospheric Research (NCAR), Boulder, CO, USA, [2]Department of Computer Science at Berea College, Berea, KY, USA

1. MOTIVATION

Background Significance

Although machine learning is playing an increasingly important role in decision-making, understanding how ML models make their predictions is often a challenge for the model developers, forecasters and other end users [1]

XAI Pipeline:

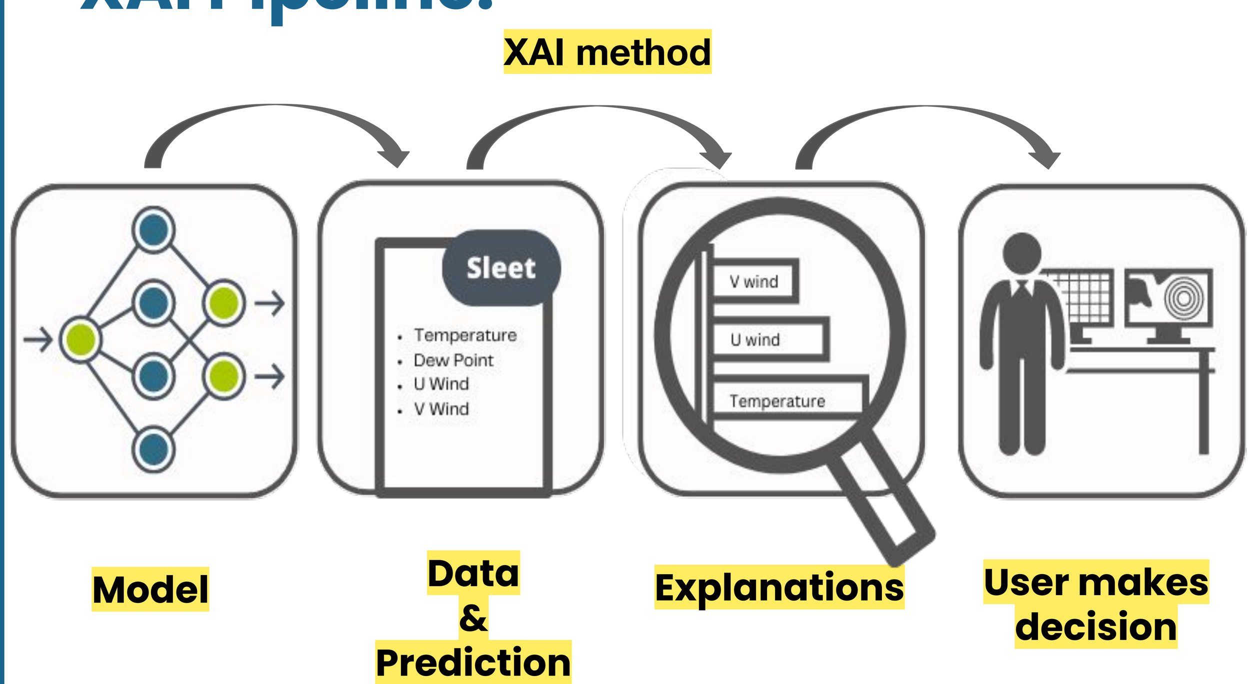


Fig 1. XAI pipeline

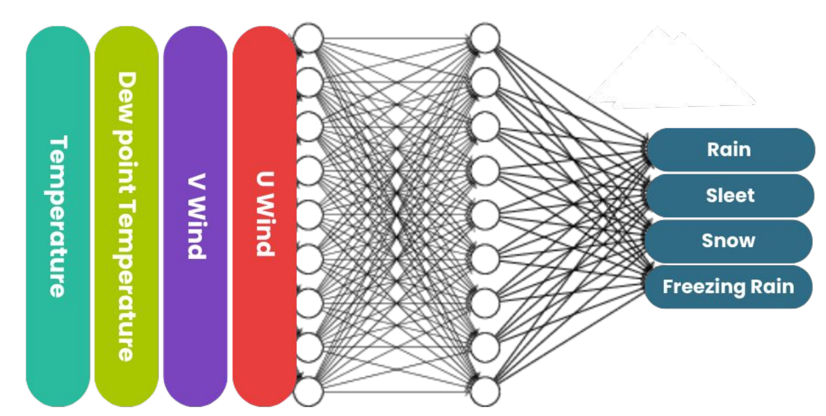
2. OBJECTIVES

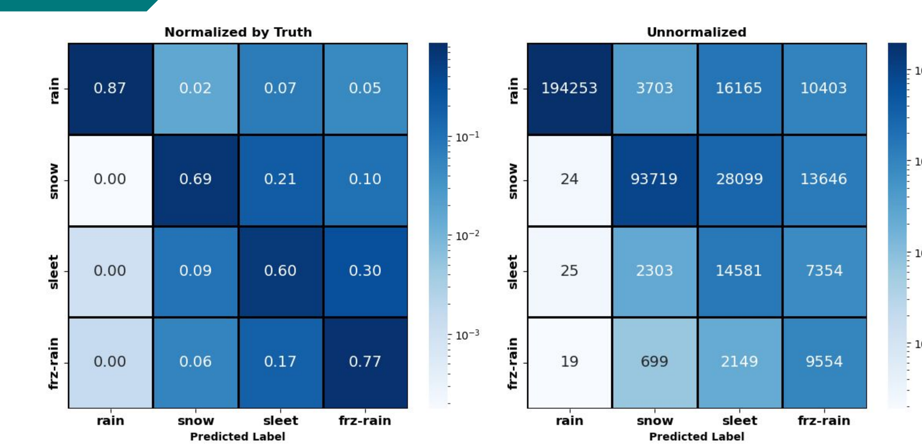
- ❖ **Evaluate the ML model performance** against existing numerical weather models
- ❖ Use different post-modelling XAI methods to **verify that the model is focusing on the most physically relevant features**

3. DATA

- ❖ Input data: The Rapid Refresh(RAP)
- ❖ True Labels: mPING (underreports freezing rain)

4. PRECIPITATION MODEL





Performance:

- ❖ Overprediction of rain
- ❖ Low accuracy predicting sleet and freezing rain

5. EXPLAINABLE AI (XAI)

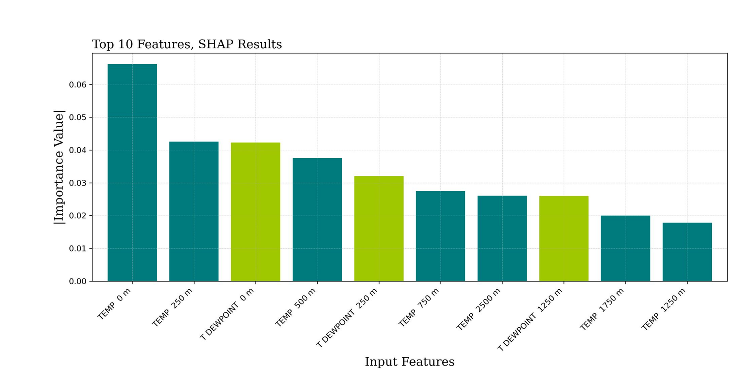


Fig 2. SHAP results

➔ Shapley Additive Explanations

- SHAP calculates the average contribution of each feature, representing how much each feature influences the model's prediction

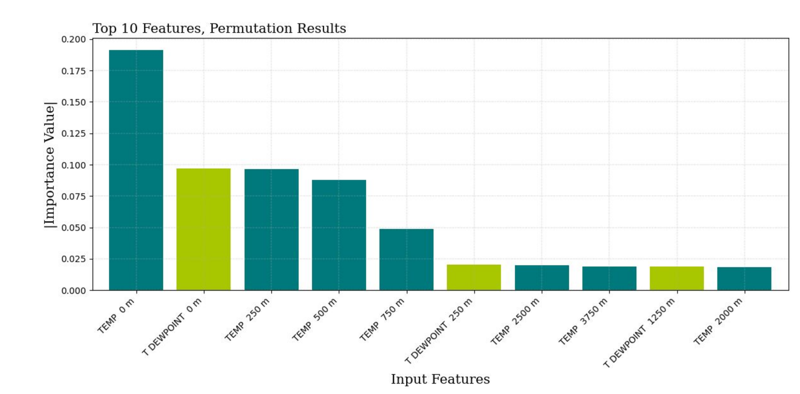


Fig 3. Permutation results

➔ Permutation Feature Importance

- Permutation feature importance works by randomly shuffling the values of a single feature and measuring the resulting change in the model's performance

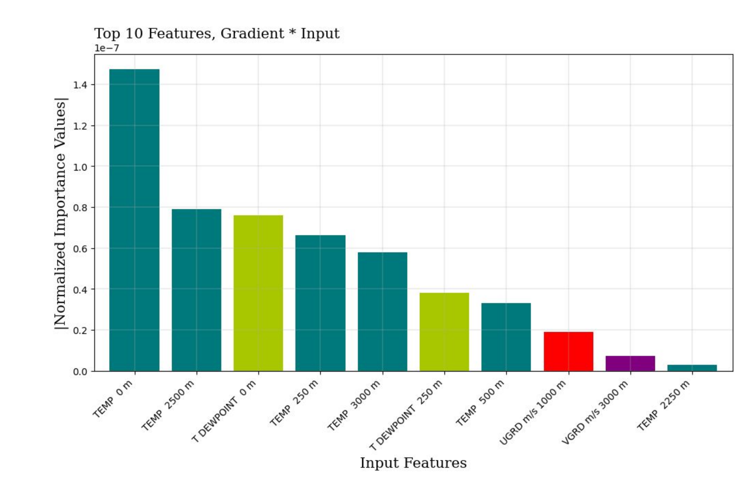


Fig 4. Gradient * Input results

➔ Gradient * Input

- This XAI method works by multiplying the gradient of the model's output with the input features

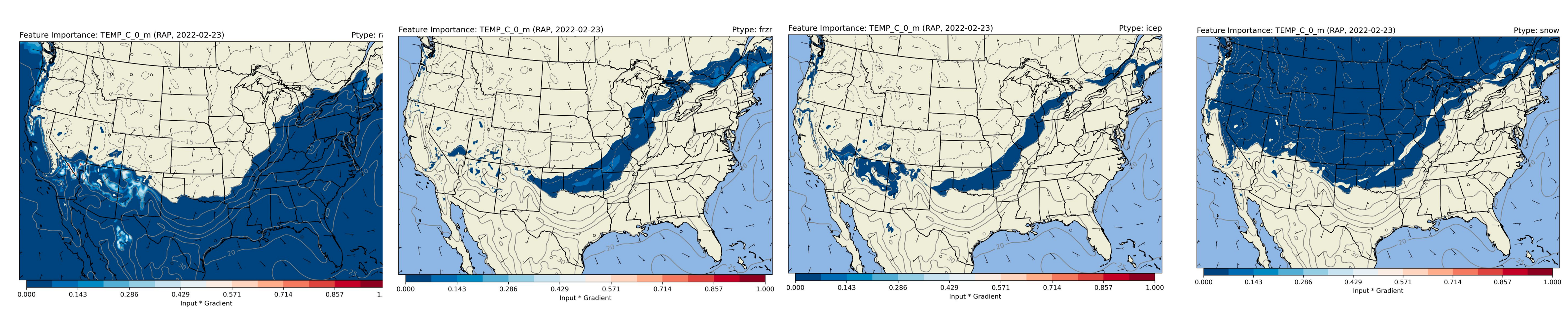


Fig 5. Absolute Gradient * Input values

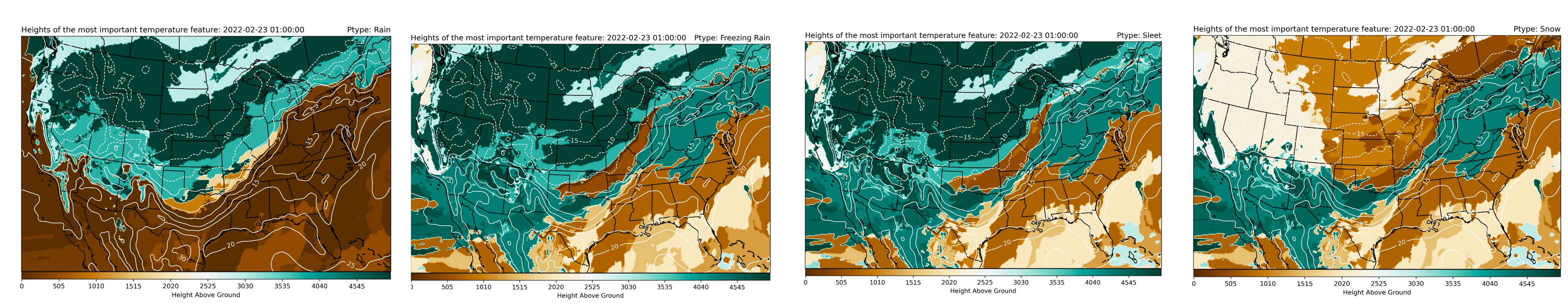


Fig 6. Gradient * Input by height per precipitation type

6. CONCLUSIONS

- ❖ Features that have a higher height are more important near the freezing line and fronts
- ❖ Each XAI method provides slightly different results
- ❖ Temperature at 0m is the top feature for each of the methods
- ❖ The Input features that are near the surface tend to be the most important

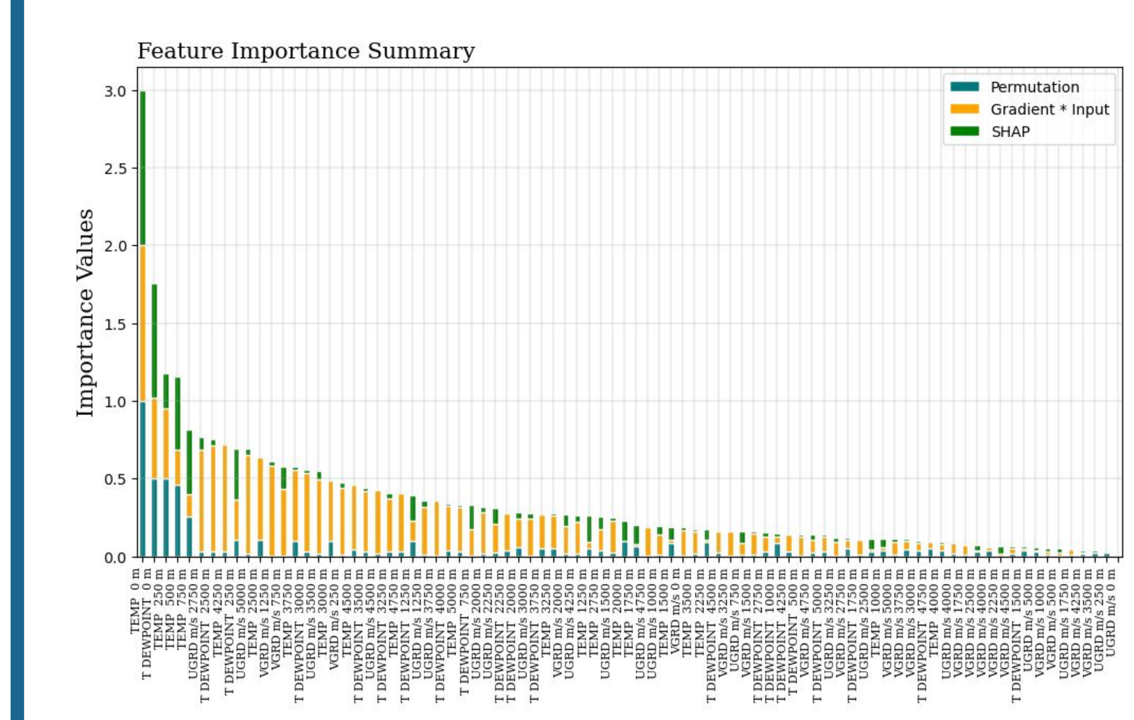


Fig 7. XAI methods summary

7. REFERENCES

[1] McGovern, A., Lagerquist, R., Gagne, D. J., Jergensen, G. E., Elmore, K. L., Homeyer, C. R., & Smith, T. (2019). Making the black box more transparent: Understanding the physical implications of machine learning

Fig 1. Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier. arXiv preprint arXiv:1602.04938, 2016

8. ACKNOWLEDGMENTS

- ❖ This project was supported by the National Science Foundation under Grant No. ICER-2019758
- ❖ SIParCS team: Virginia Do and Julius Owusu Afriyie