# Improving Tropical Cyclone Forecasts Using Machine Learning on Different Ensemble Combinations

*Abhay Bhaskar*[1,2]

[1]Edison Academy Magnet School, Edison, NJ 08837; [2]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794

## Introduction

- In 2023, twenty instances of tropical cyclones resulted in over $90 billion in damages and hundreds of casualties
- While there has been a decrease in the average track error of tropical cyclone forecasts (Fig. 1), it is still significant
- Current models can have particularly high errors for some tropical cyclones, such as Hurricane Joaquin (Fig. 2)

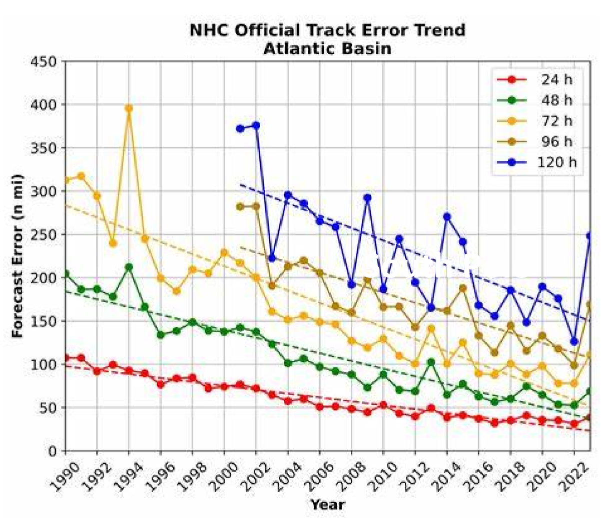



**Figure 1.** National Hurricane Center Average Annual Track Error (km) from 1970 to 1923 in the Atlantic Basin

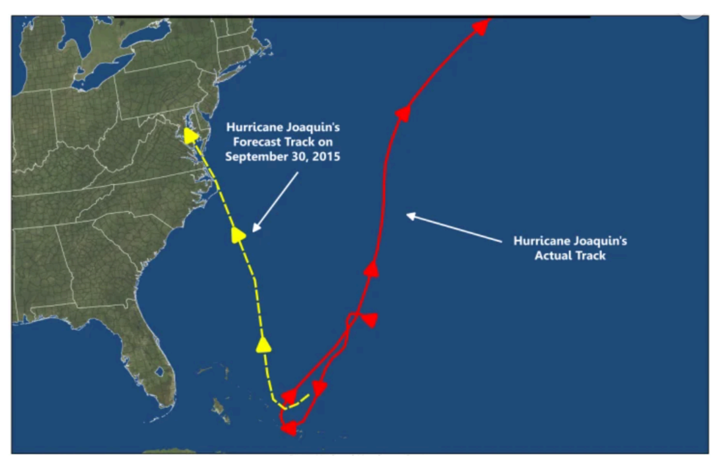



**Figure 2.** Error in National Hurricane Center Forecast of Hurricane Joaquin (2015), which was predicted to hit the United States but instead continued out to sea

## GEFS and ECMWF Ensemble Models

- The Global Ensemble Forecast System (GEFS) and European Center for Medium-Range Weather Forecasts (ECMWF) are widely used physics models for predicting tropical cyclone tracks
- The 20 GEFS and 50 ECMWF Ensemble Members are created by using slightly different starting conditions to account for the inherent uncertainties of the atmosphere
- GEFS and ECMWF generate forecasts by taking an average of their ensemble members, but experience a systematic slow bias (Fig. 3, Leonardo and Colle 2017)

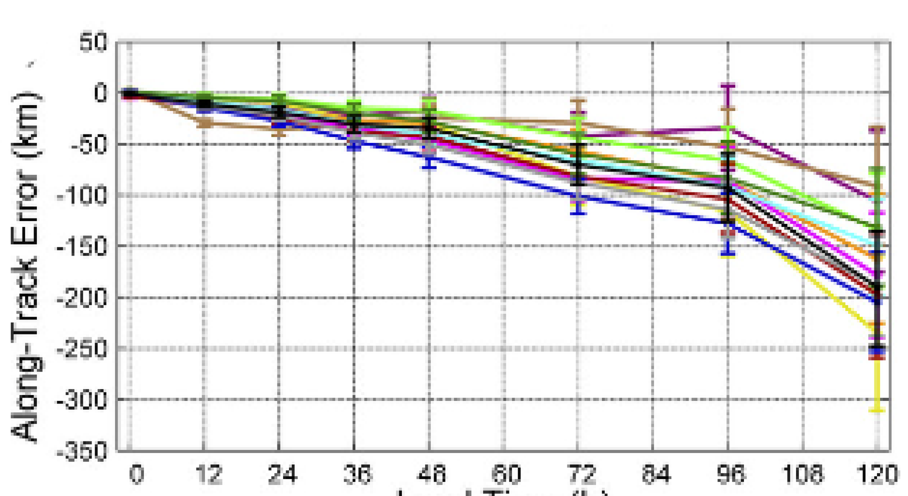



**Figure 3.** Along-Track Error for GEFS, ECMWF, and other Forecasting Models at various lead times (tau); this error makes up the vast majority of the total tropical cyclone track error due to the slow bias plaguing current models

## Data Collection

- GEFS Ensemble Members and Best Tracks were gathered from the National Hurricane Center
    - 1954 forecasts of tau values (lead times) from 0 to 120 hours (increments of six)
- ECMWF Ensemble Members and Best Tracks were accessed through the THORPEX Interactive Grand Global Ensemble (TIGGE)
    - 1866 forecasts of tau values from 0 to 120 hours (increments of twelve)

**Figure 4.** Sample Data View (XML File) Available Online from TIGGE

## Methodology

Machine Learning was applied to three different combinations of ensemble members: 20 GEFS Members, 50 ECMWF Members, and the first 20 ECMWF Members

Five-Day Ensemble Member Forecasts → CNN Model and K-fold Cross-Validation → Five-Day Forecast: Predicted (lat, long) up until day five → Calculate Average Track Error at Each Tau-Value and Plot Results

## Machine Learning Approach and Evaluation Process

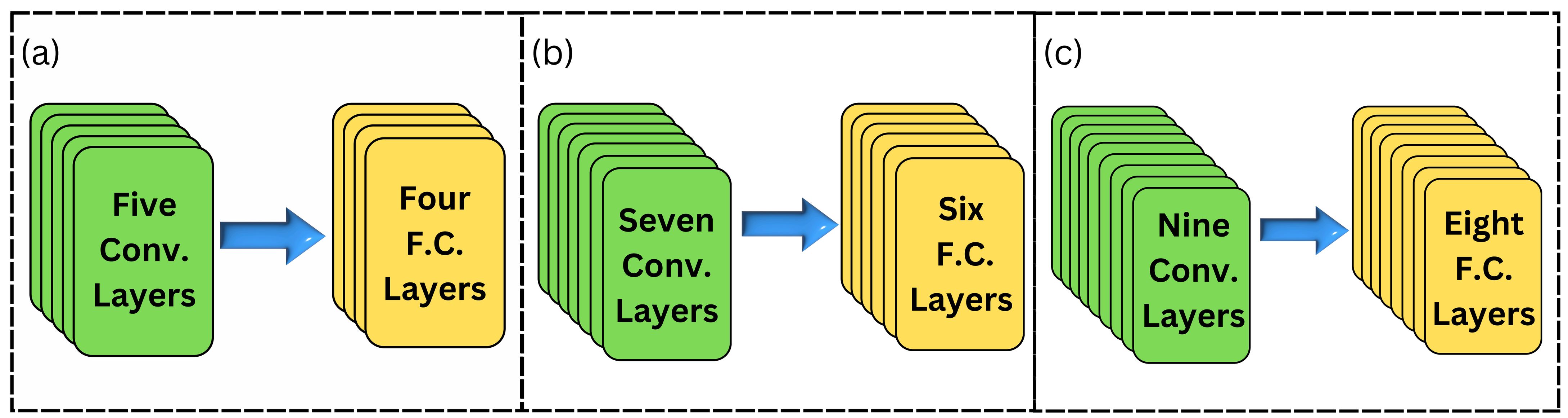



**Figure 5.** CNN Architectures for (a) 20 GEFS Ensemble Members, (b) first 20 ECMWF Ensemble Members, (c) 50 ECMWF Ensemble Members; Conv. denotes Convolutional layers and F.C. refers to Fully Connected layers

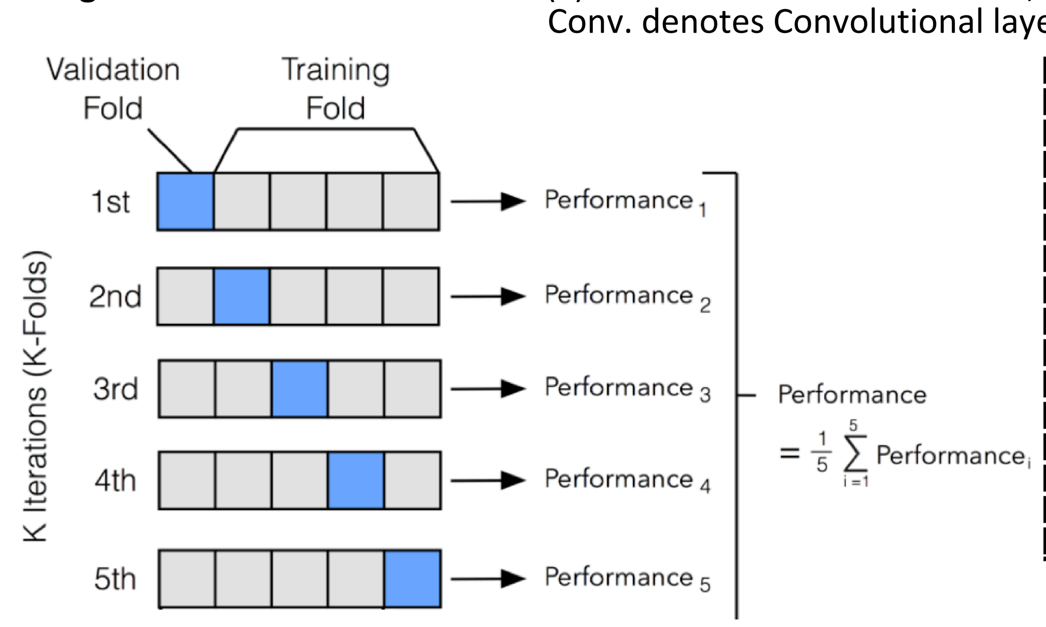
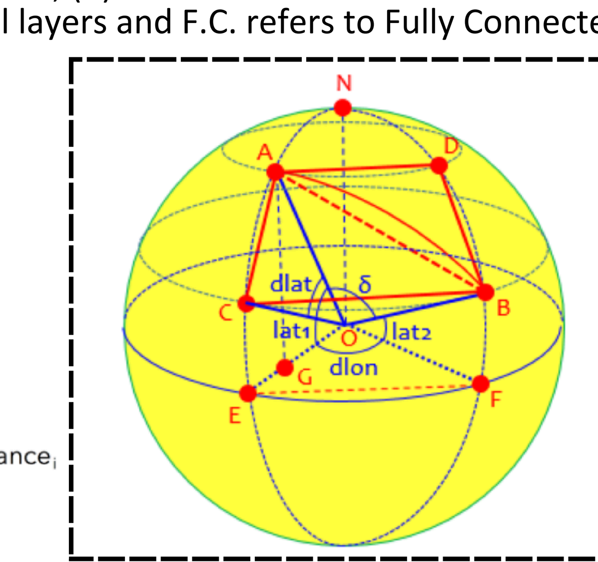



**Figure 6.** K-fold cross-validation was run with k=5, indicating that the data was split into five folds, with four used for training and one for validation




$$AB^2 = \left[2\sin\left(\frac{\Delta\mathrm{lat}}{2}\right)\right]^2 + \left[2\sin\left(\frac{\Delta\mathrm{lon}}{2}\right)\cos(\mathrm{lat}_1)\right] \times \left[2\sin\left(\frac{\Delta\mathrm{lon}}{2}\right)\cos(\mathrm{lat}_2)\right] = 4\sin^2\left(\frac{\Delta\mathrm{lat}}{2}\right) + 4\sin^2\left(\frac{\Delta\mathrm{lon}}{2}\right)\cos(\mathrm{lat}_1)\cos(\mathrm{lat}_2)$$

**Figure 7.** The ML model's forecasts and best tracks are a series of points (latitude, longitude), so the distance between them was calculated with the haversine formula

## Example Tracks of CNN Model and Ensemble Means

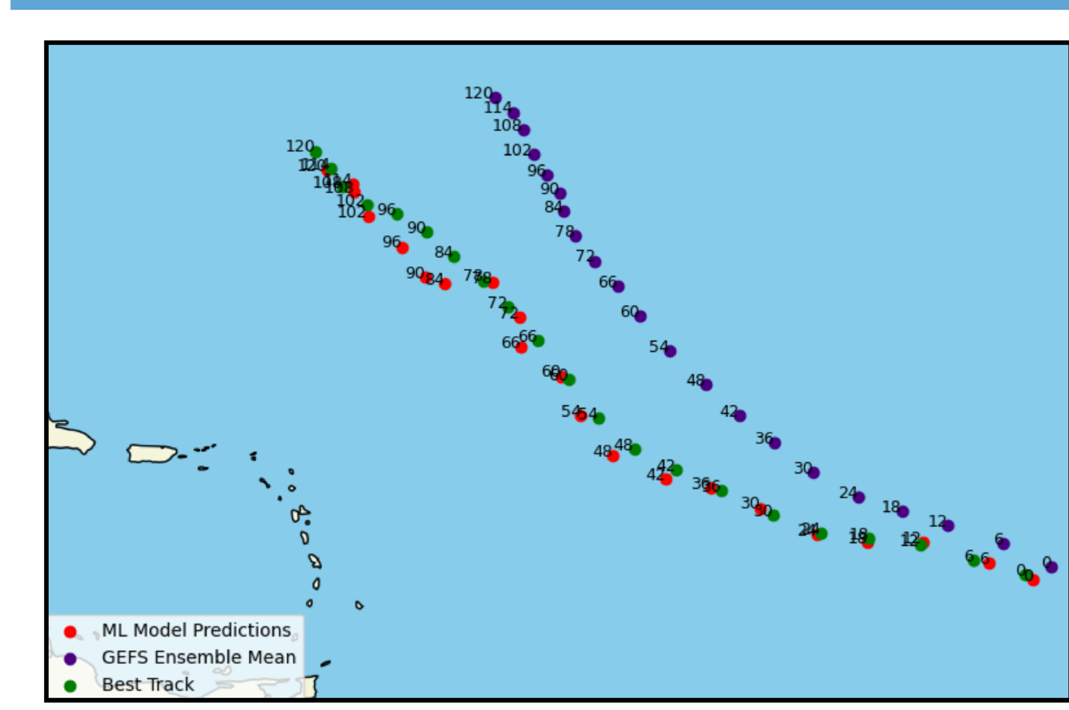



**Figure 8.** CNN Model 5-day Forecast (red), GEFS Ensemble Mean 5-day forecast (purple), and Best Track (green) for Hurricane Danielle

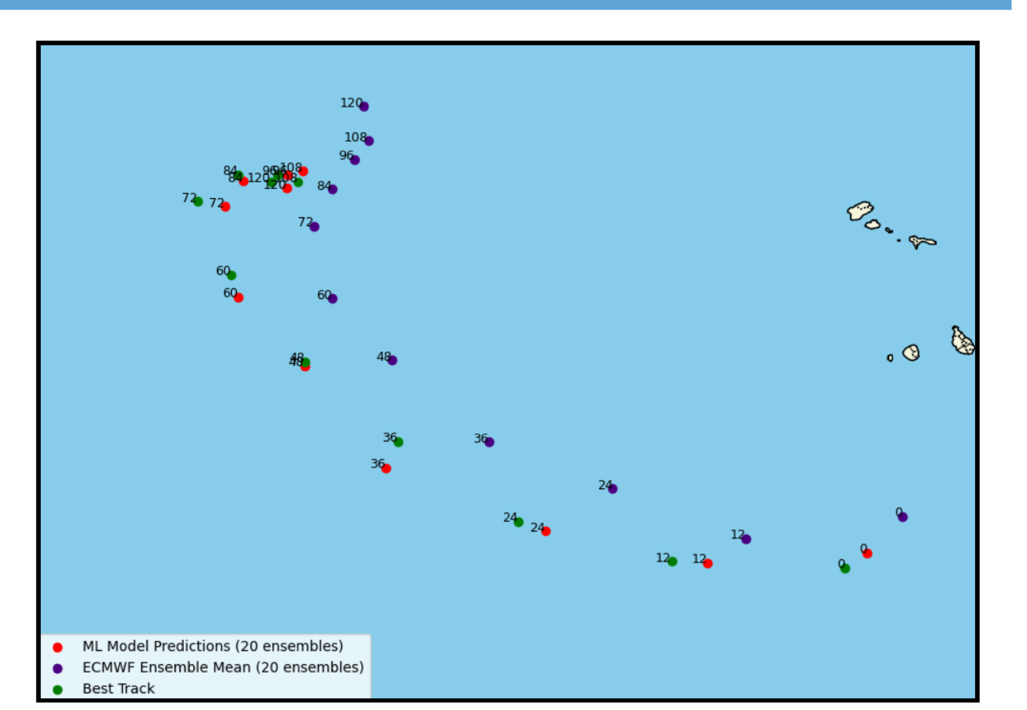



**Figure 9.** CNN Model 5-day Forecast (red), ECMWF Mean 5-day forecast (purple), and Best Track (green) for Hurricane Enrique

## Analysis

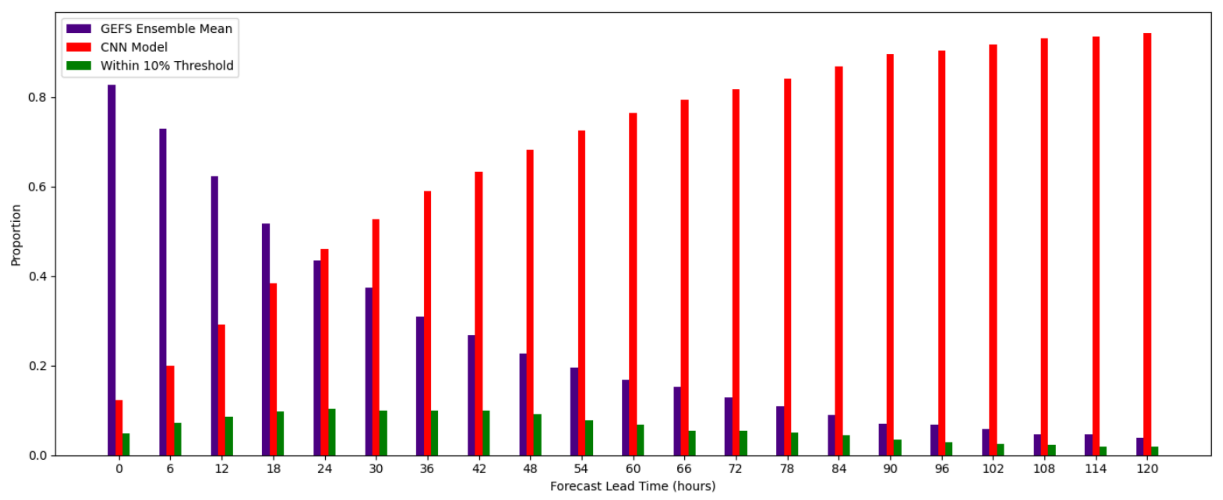



**Figure 10:** Proportion of forecasts where either the CNN model (red) or the GEFS Ensemble Mean (purple) outperformed the other.

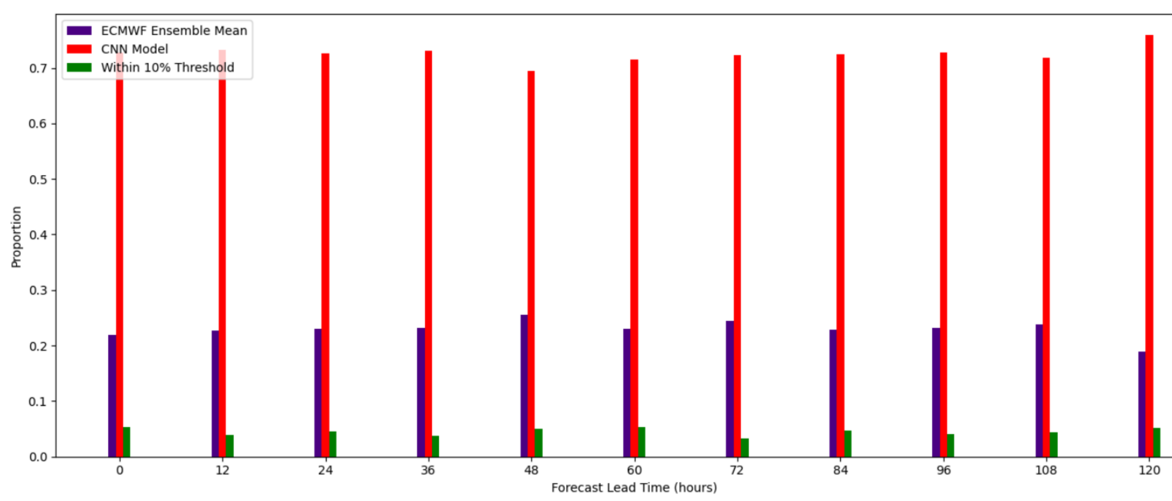



**Figure 11:** Proportion of forecasts where either the CNN model (red) or the ECMWF Ensemble Mean with 50 Members (purple) outperformed the other.

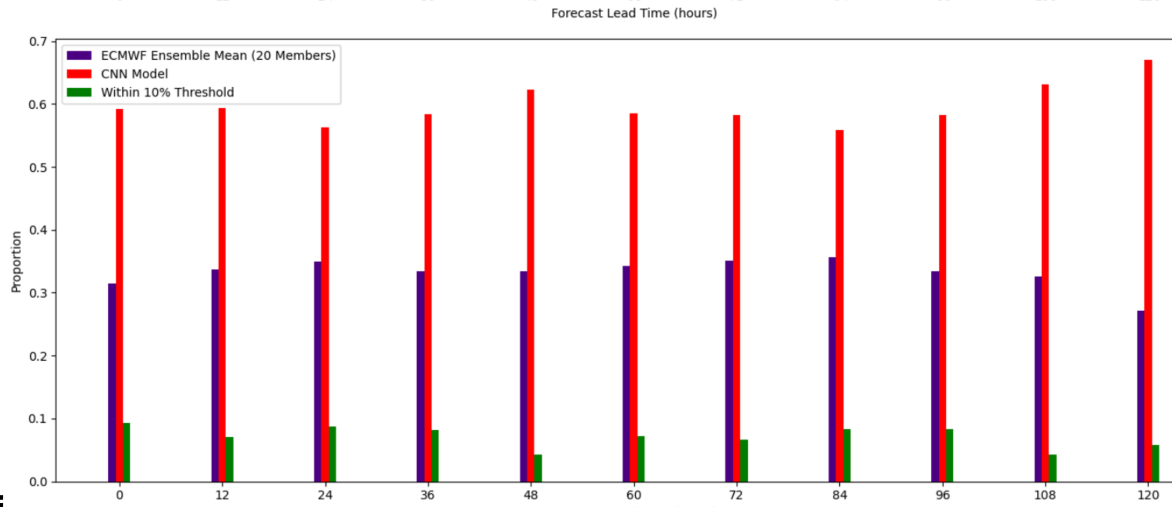



**Figure 12:** Proportion of forecasts where either the CNN model (red) or the ECMWF Ensemble Mean with 20 Members (purple) outperformed the other.

## Conclusions and Future Work

- The Machine Learning Model saw great improvements over the GEFS mean in later forecast lead times (tau>18 hours)
- Machine Learning led to more accurate tracks in comparison to the ECMWF mean (20 and 50 ensembles) for all values of tau tested
- Future research will use reanalysis data from ERA5, incorporating twenty atmospheric, land, and oceanic climate variables
- I also aim to test other machine-learning models, including a Long Short-Term Memory Network

## Acknowledgements

I would like to thank Professor Brian Colle and Nikita Agrawal for their guidance on this project. I would also like to thank Post Doctoral Associates Nicholas Leonardo and Joshua Carter for their support. Finally, I am grateful to the Simons Summer Research Program and Stony Brook University for providing me with this opportunity.

## References

NCAR. (2024, July 6). THORPEX Interactive Grand Global Ensemble. https://rda.ucar.edu/datasets/ds330.3/dataaccess/

NHC. (2024, June 7). National Hurricane Center Forecast Verification. https://www.nhc.noaa.gov/verification/verify5.shtml.

Leonardo, N. M., & Colle, B. A. (2017). Verification of Multimodel Ensemble Forecasts of North Atlantic Tropical Cyclones. *Weather and Forecasting*, *32*(6), 2083-2101, https://doi.org/10.1175/WAF-D-17-0058.1

Leonardo, N. M., and B. A. Colle, 2021: An Investigation of Large Cross-Track Errors in North Atlantic Tropical Cyclones in the GEFS and ECMWF Ensembles. Mon. Wea. Rev., 149, 395–417, https://doi.org/10.1175/MWR-D-20-0035.1.