Accelerating the WRF-CHEM Model Using a Machine Learning Emulation

Phuong Nguyen, Z. Yang, S. U. Patil, S. Shivadekar, D. Chapman, and M. Halem NASA: Carlos Cruz from Science system and Application

UMBC

Abstract

- Incorporating aerosol with microphysics and chemistry parameterizations in climate simulation and numerical weather prediction can significantly impact sub-seasonal weather forecasts. However, performing such simulations and predictions at high resolution is severely limited by computer resources.
- WRF-Chem is the Weather Research and Forecasting (WRF) model coupled with Chemistry. The model simulates the emission, transport, mixing, and chemical transformation of trace gases and aerosols simultaneously with the meteorology.
- We propose to accelerate an implementation of the WRF-CHEM Microphysics parameterization with a Machine Learning (ML) emulation using the Auto-Keras Neural Net Architecture Search (NAS) algorithm.
- We have trained and tested NAS-ML emulations using WRF-CHEM output with 30 levels including Chemistry and Microphysics over US nested domains at a coarse 2.50 and at a high resolution of 0.090 to assess potential speed ups over conventional computational approaches. Our results indicate that the NAS ML emulation of the Thompson cloud microphysics scheme, correlated well with WRF-CHEM model output when predicting precipitation (RAINNC) and the fraction of frozen precipitation (SR) with low Root Mean Square Errors.

Background Work

- Kasim et al (Oxford Univ) performed evaluations of the ECHAM6.3-HAM2.0 model using NAS ML to speed up performance. Jan 2020. billion time speed up.
- Gorman et al (MIT) used Random Forest to learn a parameterization, sub-grid processes from output of System for Atmospheric Modeling (SAM) at different resolution gained speed up 120x .Jan 2020.
- Gentile (Columbia Univ) used customized loss functions to emulate physic parametrizations at high resolution.

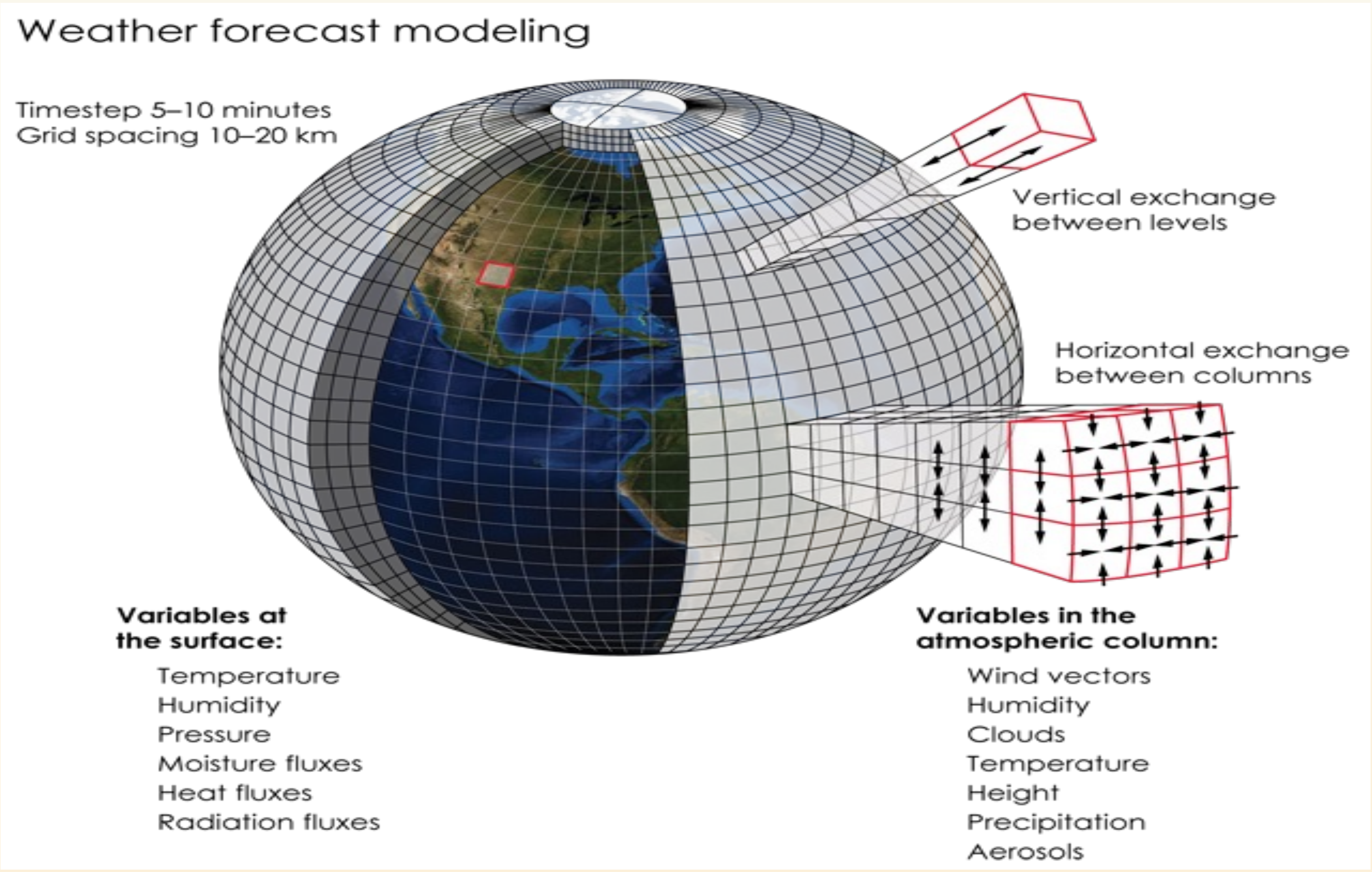
Dataset

- WRF-CHEM experiment using the YSU scheme started from 00:00UTC Jan 2018 to May 31,2019 over the North America at 2.50x2.50 degrees with 29 vertical levels 5N-70N and 160W-32W domain, hourly output.
- Emulating grid scale precipitation (RAINNC) using 2018 January month, first 20 days as training and 10 days for testing.
- Input as potential temperature, water vapor mixing ratio (QVAPOR), inverse density (alt), pressure (PB)
- Output: Predict accumulated total grid scale precipitation (RAINNC) variable.

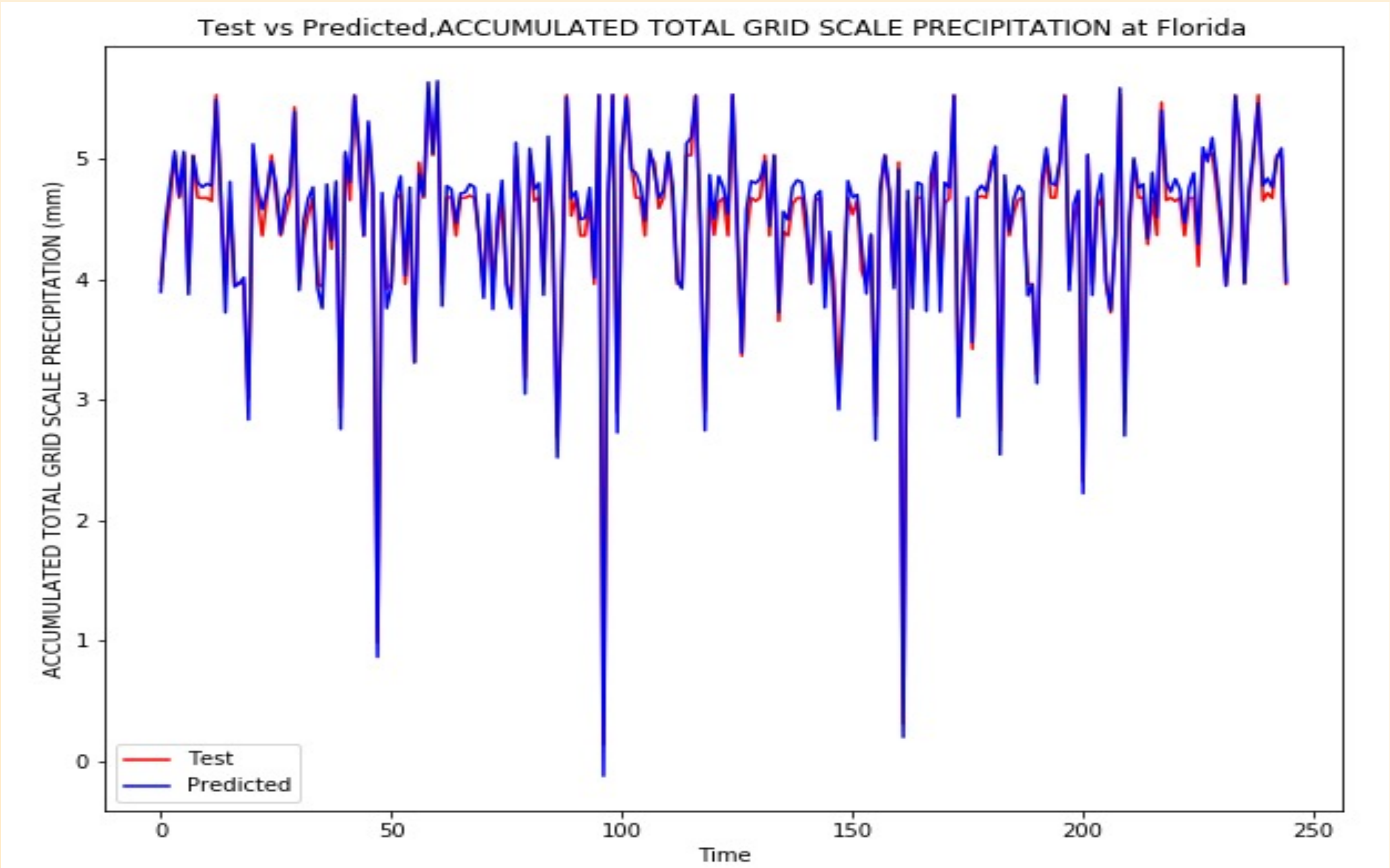
Emulation of Microphysics

- The current resolution of the NASA Unified-Weather Research Forecasting model (NU-WRF) runs at 4km and 30 levels and includes physical parameterizations of Aerosols, Chemistry and Microphysics. NU-WRF takes 3.5 months to simulate a month on the NASA/GSFC computer using 800 processor cores.
- TO accelerate the computation, we are using Neural Architecture Search Machine Learning model from Auto-Keras.
- Auto-Keras NAS use the Bayesian optimization to guide the network morphism for searching for an optimum neural architecture during the training process which laugh multiple ML models. After training, the NAS search algorithm gives the best performance ML model.

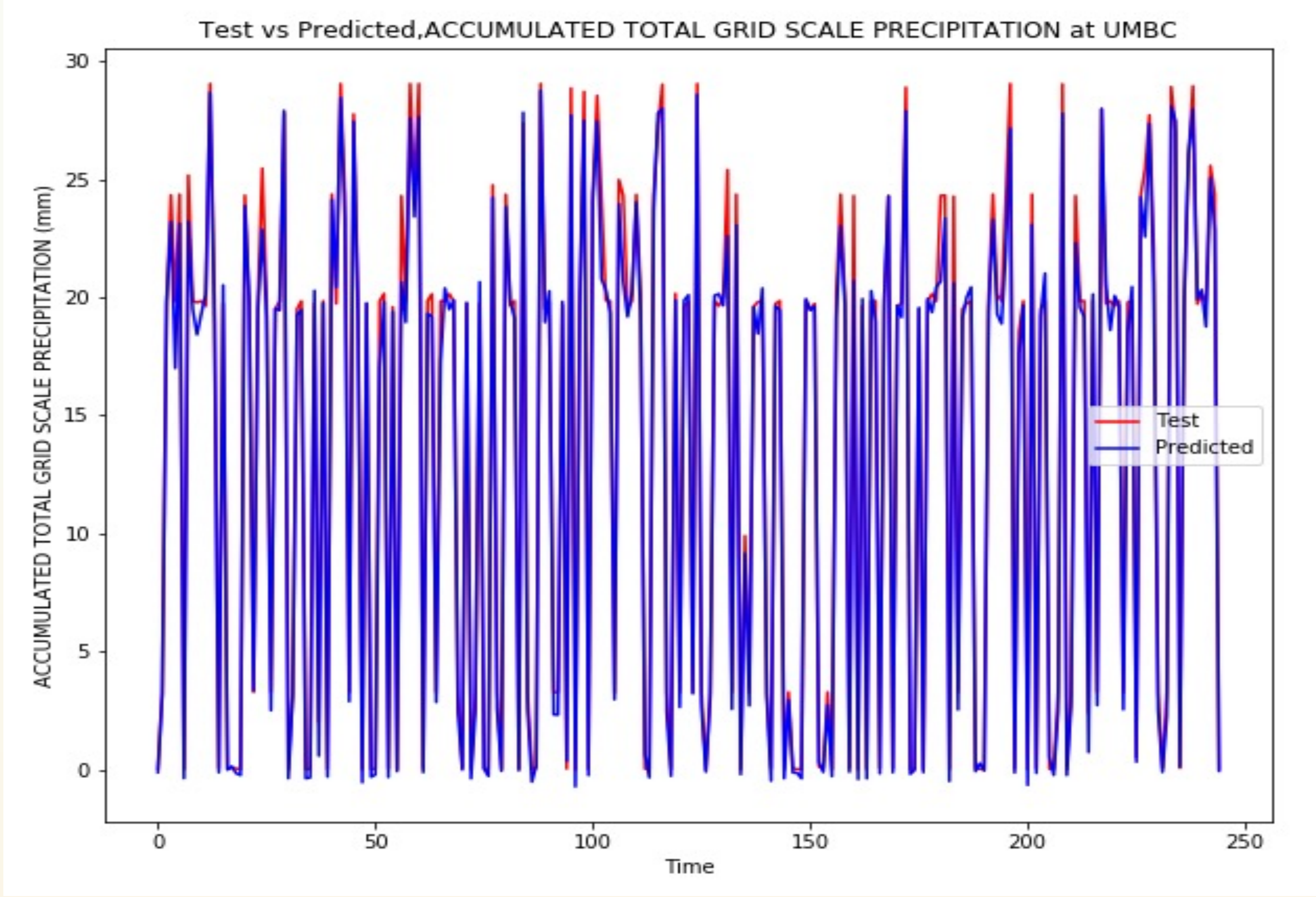
Weather Research Forecast Model



Florida : R square = 0.989 RMSE: 0.013

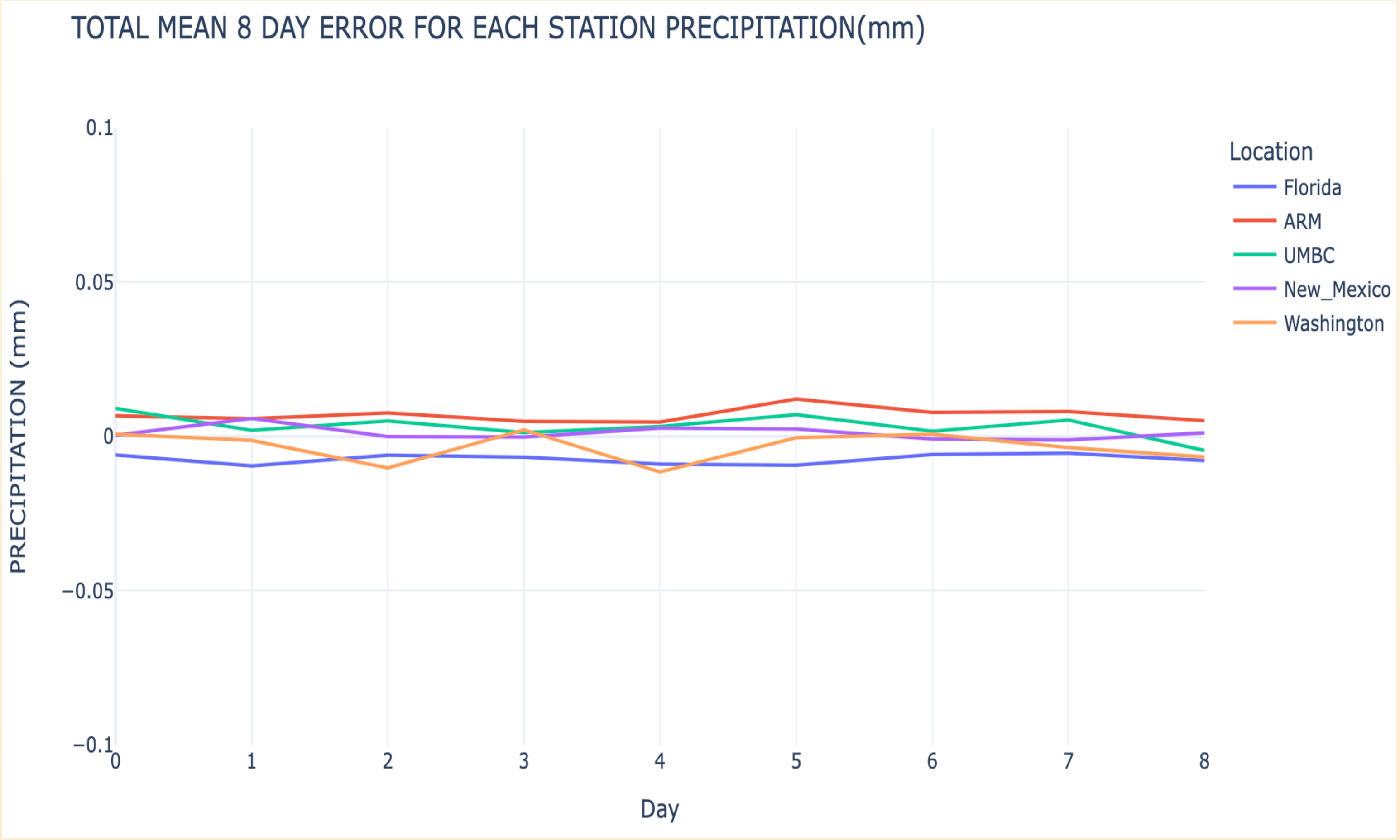


Maryland : R square = 0.99 and RMSE= 0.015



- To train on 20 days data and predict 10 days of accumulated precipitation for single location NSA model takes ~7 minutes.
- Estimated time to predict the data for all the location is ~48 hours.

Test Error comparison with WRF model output 24 hours avg



Future Work

- Predict the total grid scale precipitation(RAINNC) and frozen precipitation (SR) for entire grid cell.
- Run the model on high resolution ~9 km. Currently model is running on ~250 km resolution data.
- Embed this NSA Machine Learning model into WRF-CHEM model