

INTRODUCTION and MOTIVATION

- Three-dimensional Eulerian chemical transport models such as CMAQ often report a significant model-measurement error due to uncertainties in the treatment of physical processes and also require higher run-time.
- Machine models are more computationally efficient and are currently used widely for forecasting purposes. Deep Neural network (DNN) techniques comprise a popular class of machine learning methods.
- Predicting hourly air quality, especially ozone, is challenging due to its highly varying and complex behavior in the atmosphere. Here, we used modeled meteorological parameters (by MCIP) along with selected modeled gaseous species (by CMAQ) as our inputs for predicting future ozone concentrations.

MATERIALS and METHODS

Deep Learning Algorithm

- A deep convolutional neural networks (CNN) was implemented for predicting hourly ozone concentration. Inspired by biological processes, CNN is a class of deep, feed-forward artificial neural networks.
- CNN uses relatively little pre-processing compared to other machine learning techniques. This means that the network learns the features that in traditional algorithms were hand-engineered.
- In CNN (Figure 1), the convolutional layer applies a convolution operation to the input, passing the result to the next layer. The fully connected layer connects every neuron in the last convolutional layer to every neuron in the output layer, similar to the traditional multi-layer perceptron neural network (MLP).

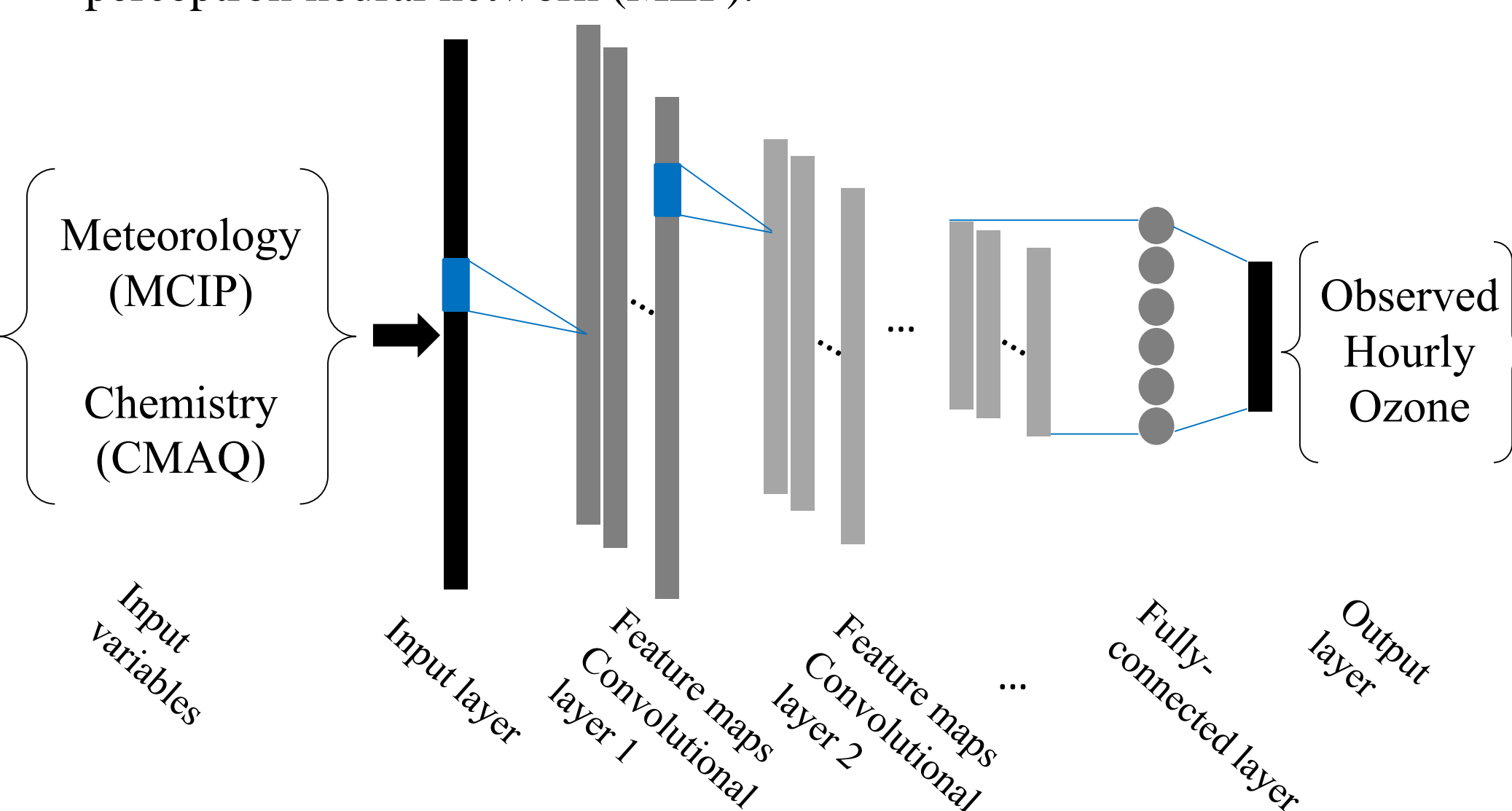


Figure 1: Schematic for the regressive 1D convolutional neural networks (ConvNet-1D).

Table 1: Meteorological variables (left) and chemical species (right) used as input in CMAQ-CNN model.

Abb.	Variable Name	units	Abb.	Variable Name	units
PRSEC	Surface Pressure	Pascal	NO2	Nitrogen dioxide	ppbv
USTAR	cell averaged friction velocity	M/S	NO	Nitrous oxide	ppbv
WSTAR	convective velocity scale	M/S	O3	Ozone	ppbv
PBL	PBL height	M	NO3	Nitrates	ppbv
MOLI	inverse of Monin-Obukhov length	1/M	OH	Hydroxide	ppbv
HEX	sensible heat flux	WATTS/M ²	HO2	Hydroperoxyl	ppbv
QFX	latent heat flux	WATTS/M ²	N2O5	Dinitrogen pentoxide	ppbv
RADYNI	inverse of aerodynamic resistance	M/S	HNO3	Nitric acid	ppbv
RSTOMI	inverse of bulk stomatal resistance	M/S	FORM	Formaldehyde	ppbv
TEMPG	skin temperature at ground	k	ALD2	Aldehyde	ppbv
TEMP2	temperature at 2 m	k	ISOP	Isoprene	ppbv
Q2	mixing ratio at 2 m	kg/kg	XYL	Xylene	ppbv
WSPD10	wind speed at 10 m	M/S	TOL	Toluene	ppbv
WDIR10	wind direction at 10 m	deg			
GLW	longwave radiation at ground	WATTS/M ²			
GSW	solar radiation absorbed at ground	WATTS/M ²			
RGRND	solar rad reaching sfc	WATTS/M ²			
RN	nonconv. pcpn per met TSTEP	cm			
RC	convective pcpn per met TSTEP	cm			
CFRAC	total cloud fraction	-			
CLDT	cloud top layer height (m)	M			
CLDB	cloud bottom layer height (m)	M			
WBAR	avg. liquid water content of cloud	G/M ³			
SNOCOV	snow cover	decimal			
VEG	vegetation coverage (decimal)	decimal			
LAI	leaf-area index	-			
SEACE	sea ice (fraction)	-			
WR	canopy moisture content	m			
SOIM1	volumetric soil moisture in top cm	M ³ /M ³			
SOIM2	volumetric soil moisture in top m	M ³ /M ³			
SOIT1	soil temperature in top cm	k			
SOIT2	soil temperature in top m	k			
SETYP	soil texture type by USDA category	-			

Since the CMAQ is used in this model, some limitations of the CNN model can be mitigated, especially those with interpreting the physical variables and their relationship with ozone concentrations, and unavailability of important meteorological predictors such as PBL height and chemical predictors such as OH.

RESULTS

ConvNet modeling time period:

- Training data: 2011 – 2013
- Real-time prediction: 2014
- We trained/tested the model only on ozone season (April-October).
- The focus is to improve CMAQ output concentration, particularly in predicting ozone peaks.

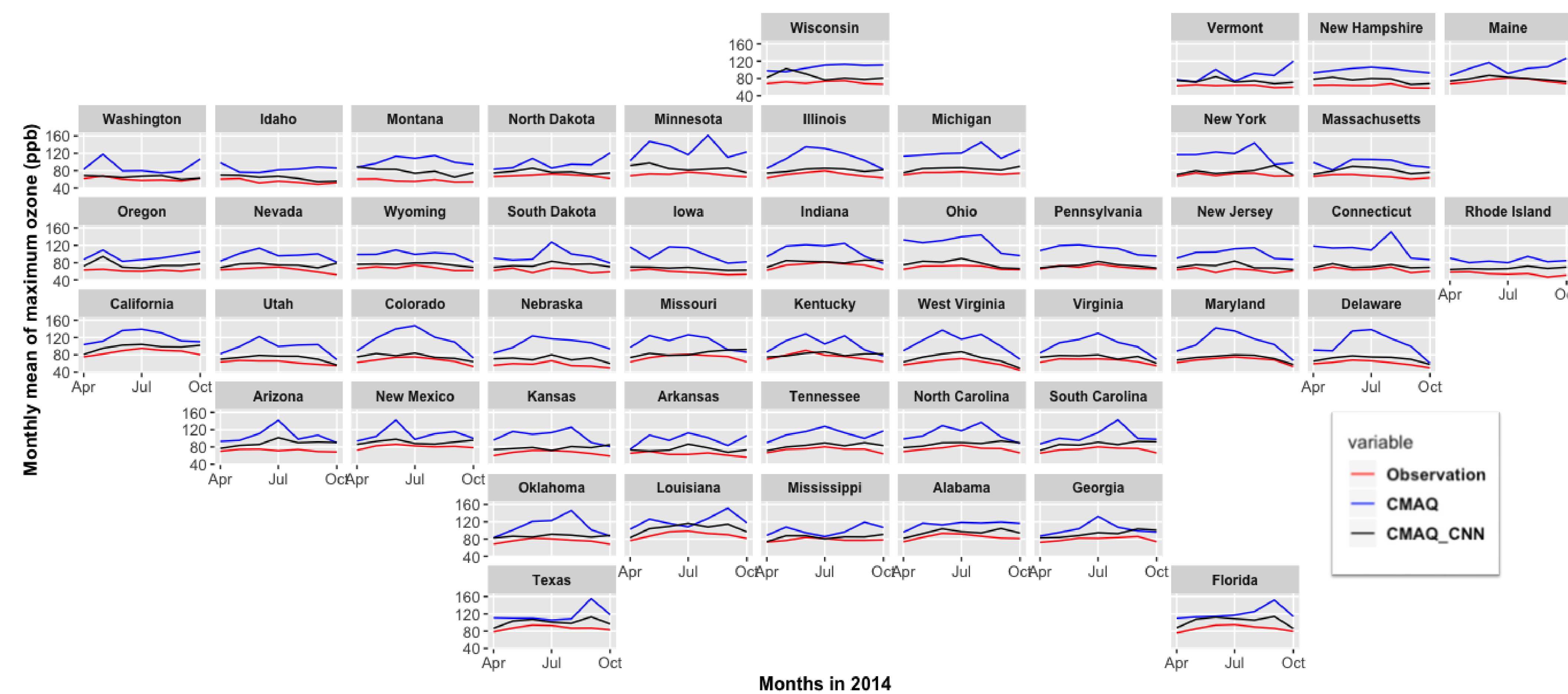


Figure 2. Monthly mean of maximum ozone for 7 months of ozone season (April-October) in 2014.

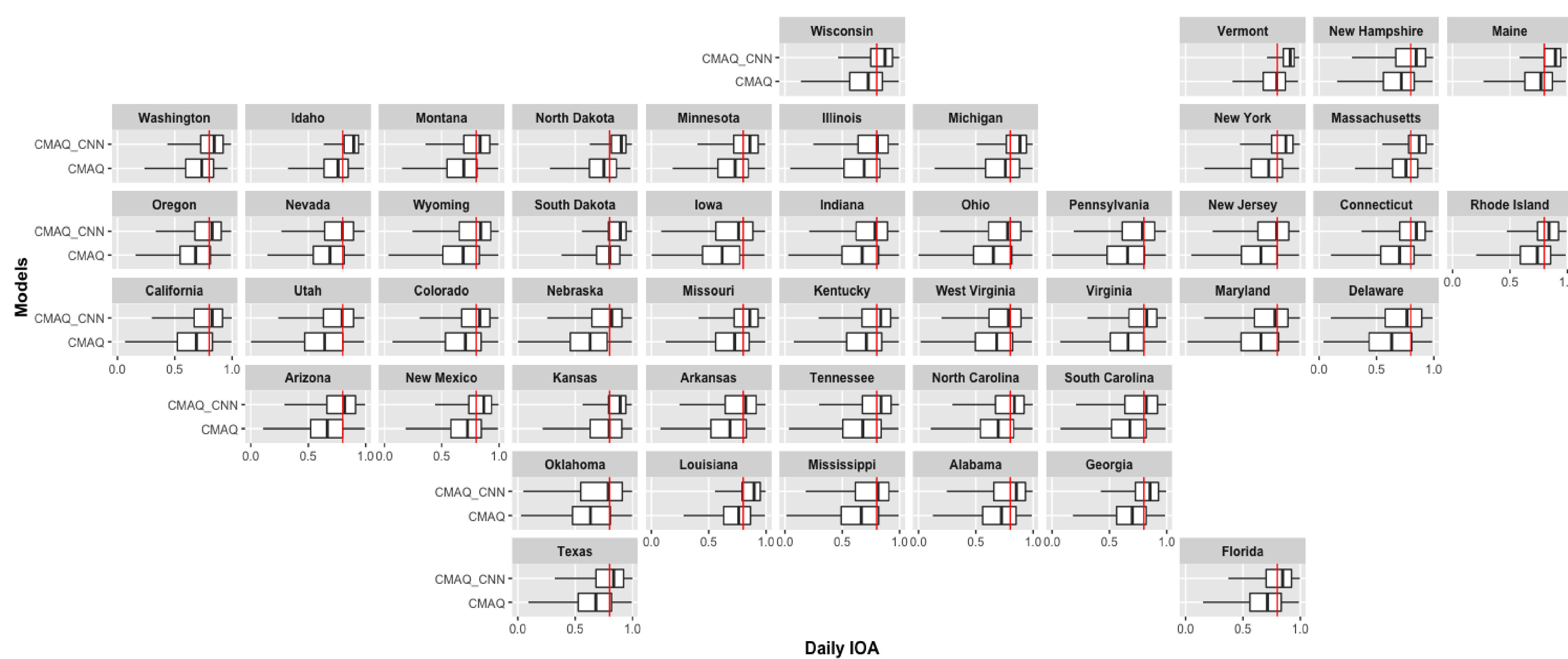


Figure 3. Daily index of Agreement (IOA) of CMAQ, and CMAQ-CNN models for 2014 ozone season (April-October) in 2014. The vertical red lines indicate IOA=0.8 as a reference.

Key message:

- The CMAQ model generally overpredicted ozone peaks during the ozone season (Figure 2). This overprediction was more pronounced in central and southeast United States.
- The CMAQ-CNN model recovered the CMAQ model's overprediction (Figure 2). However, the CMAQ-CNN generally underpredicted the high ozone peaks.
- The CMAQ-CNN model predicts next day's hourly ozone concentration with promising improvement in accuracy (IOA) (Figure 3). The monthly median of the daily IOAs was mostly above 0.8 threshold indicating reasonable prediction level.
- The number of low accuracy days was also decreased significantly for the CMAQ-CNN model.

RESULTS (cont.)

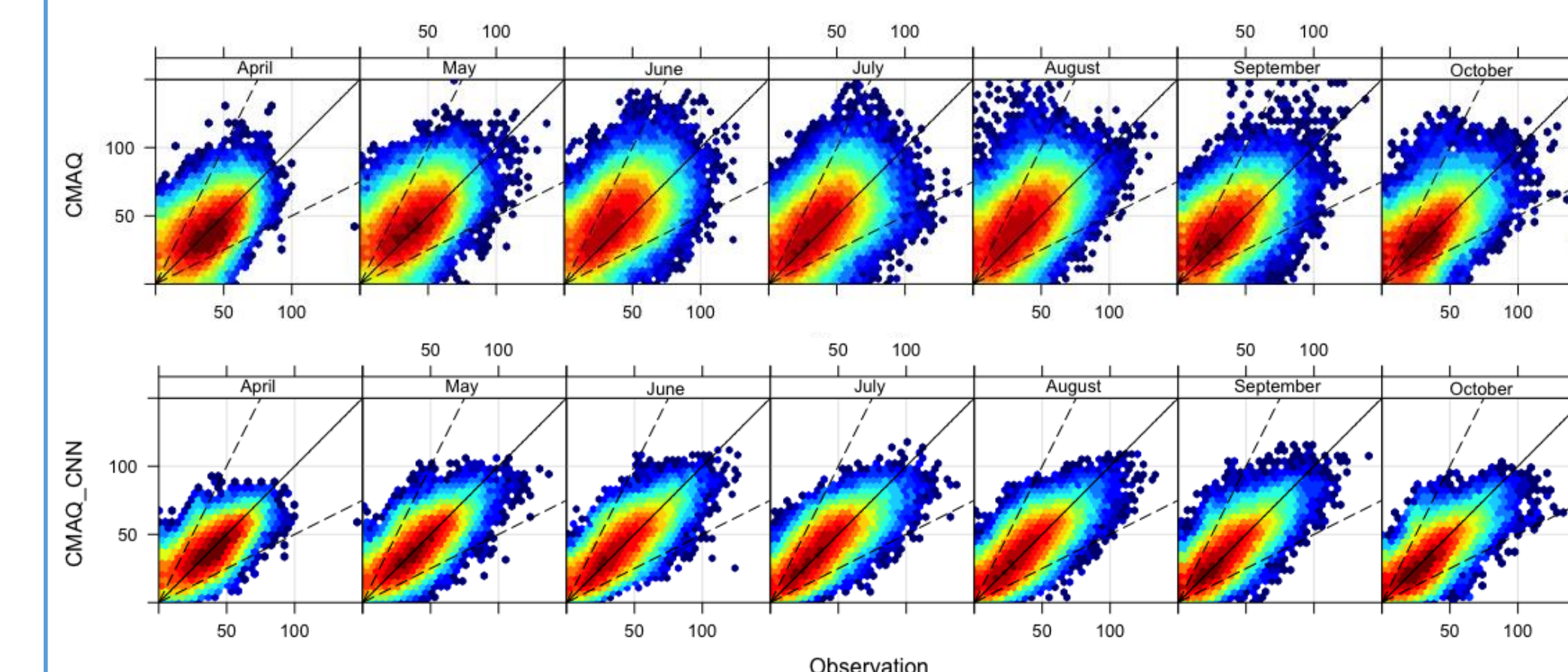


Figure 5: Scatter plots of CMAQ and CMAQ-CNN models in different months in 2014.

We developed two different CMAQ-CNN models:

- CNN_generalized: we trained/tested the model by combining all input samples from the entire domain (United States).
- CNN_Standalone: we trained/tested the model for each station individually.

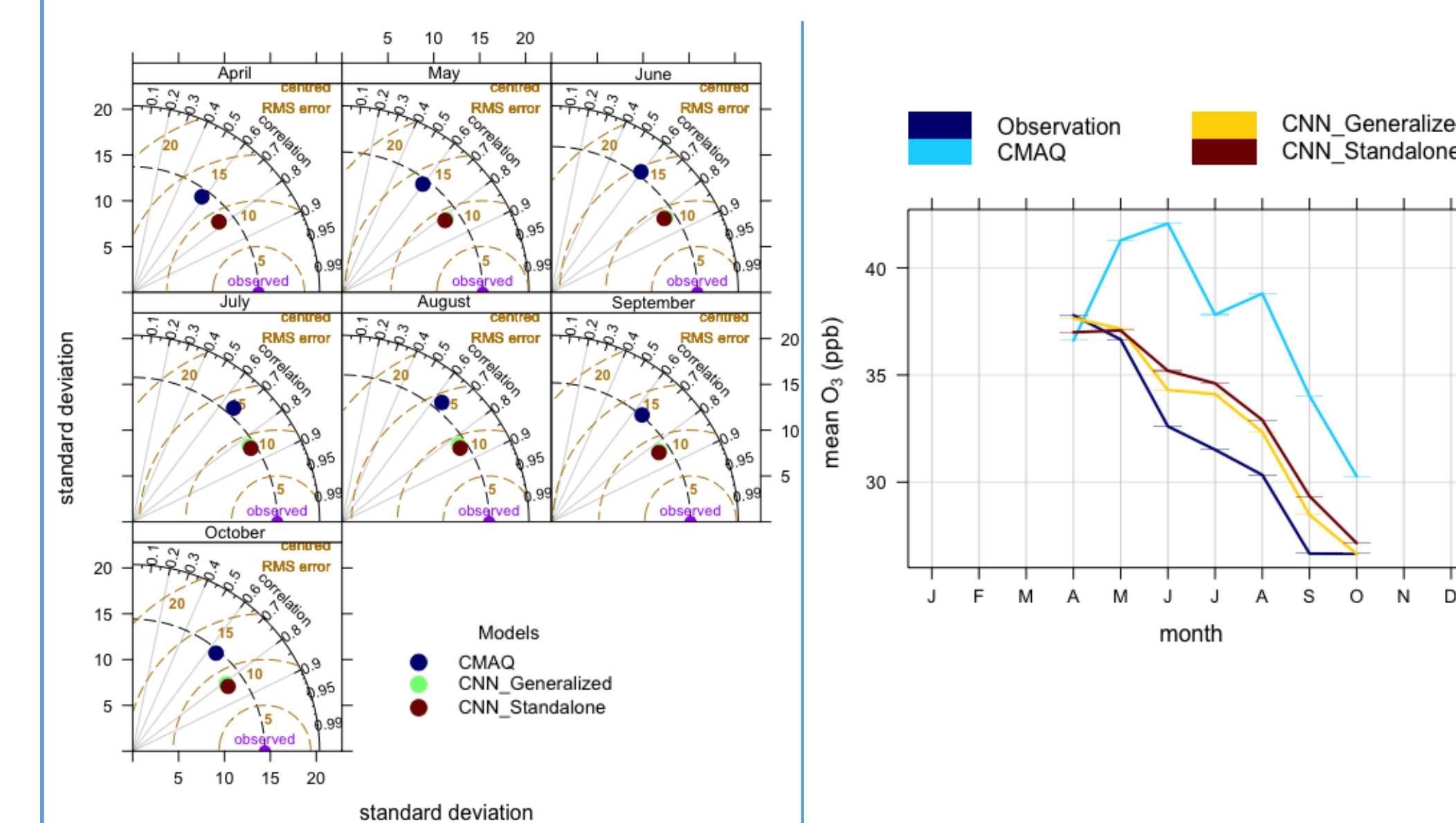


Figure 4: Taylor diagram showing the performance of all models in different months of 2014 averaged over 1048 AQS stations in the continental US.

Figure 5: Monthly mean ozone concentration of all models 1048 AQS stations in the continental US.

CONCLUSIONS

- A timely-efficient 1D deep convolutional neural network (ConvNet 1D), called CMAQ-CNN, was implemented and trained on using CMAQ outputs as inputs to predict hourly ozone concentration in real-time across the continental US (1081 AQS stations in 48 states).
- The CMAQ-CNN model significantly improved the performance of the CMAQ model in term of both accuracy (IOA) and bias (maximum daily ozone).
- IOA improved around 0.06 in average and up to 0.3 across the United States by using CMAQ-CNN model.
- The CMAQ-CNN model shows mediocre performance on capturing very high ozone peaks (over 90 ppb).

ACKNOWLEDGEMENTS

This study was partially supported by funding from the Department of Earth and Atmospheric Sciences (EAS Research Grant) of the University of Houston.

* corresponding author: Dr. Yunsoo Choi, PI of UH EAS ICAS air quality forecasting/modeling group, email: ychoi6@uh.edu