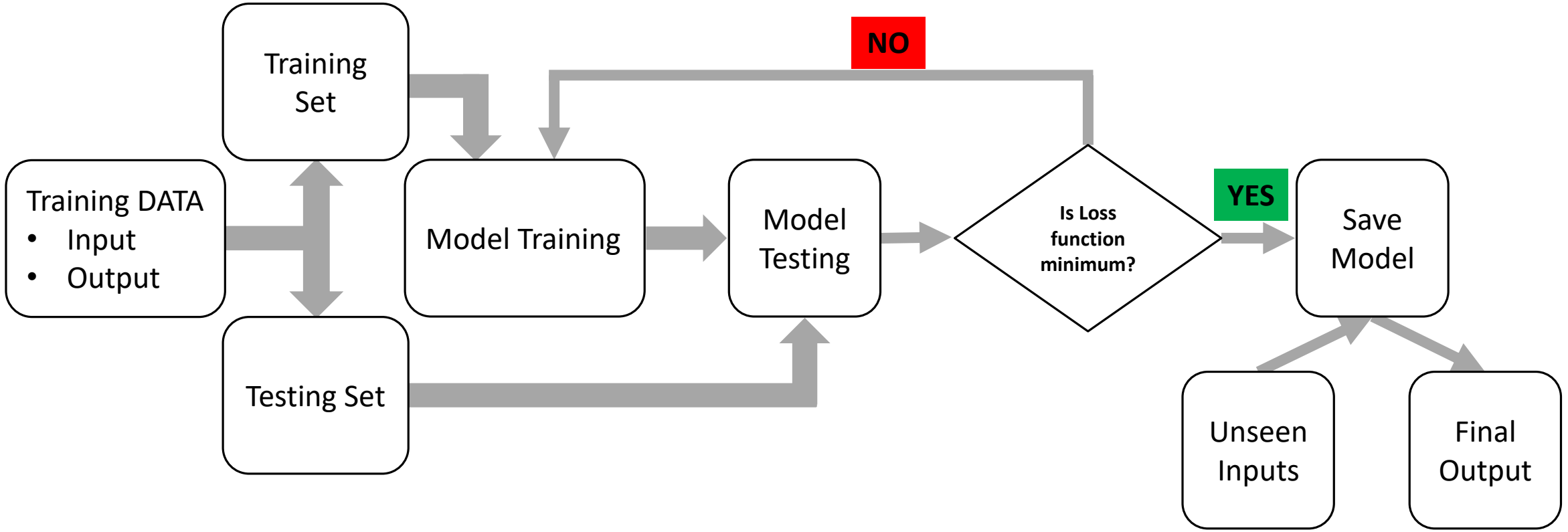
A Novel CMAQ-CNN Hybrid Model to Forecast Hourly Surface-Ozone Concentrations Fourteen Days in Advance

Alqamah Sayeed1, Yunsoo Choi1*, Ebrahim Eslami1, Jia Jung1, Yannic Lops1,
Ahmed Khan Salman1, Min-Hyeok Choi2, Hyun-Ju Park2, and Jae-Bum Lee2

Presenter: Dr. Yunsoo Choi (ychoi6@uh.edu)
1 Department of Earth and Atmospheric Sciences, University of Houston
2 National Institute of Environmental Research, Air Qualify Forecasting Center

CMAS Conference, October 2020

Neural Network Model



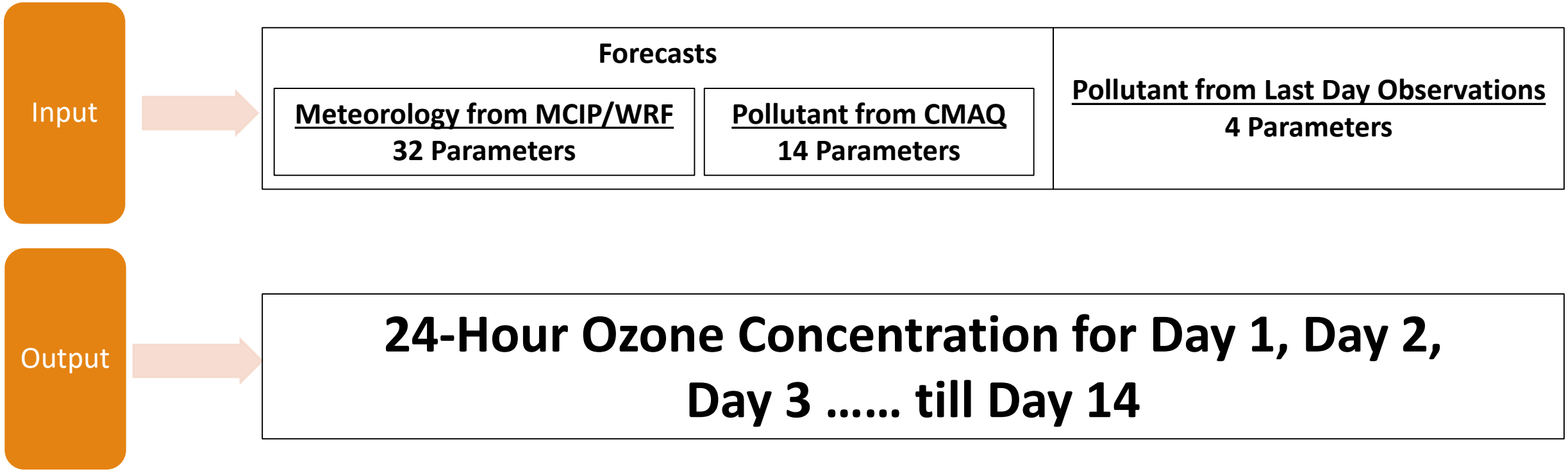
Our deep learning application on short-term forecasting

- Prediction of Air pollutants
  - Ozone
  - Pollen
  - PM
- Prediction of various weather phenomenon
  - Wind Speed
  - Temperature
  - Relative Humidity
  - Precipitation
- Bias correction of numerical modeling outputs (e.g., from CMAQ and WRF)

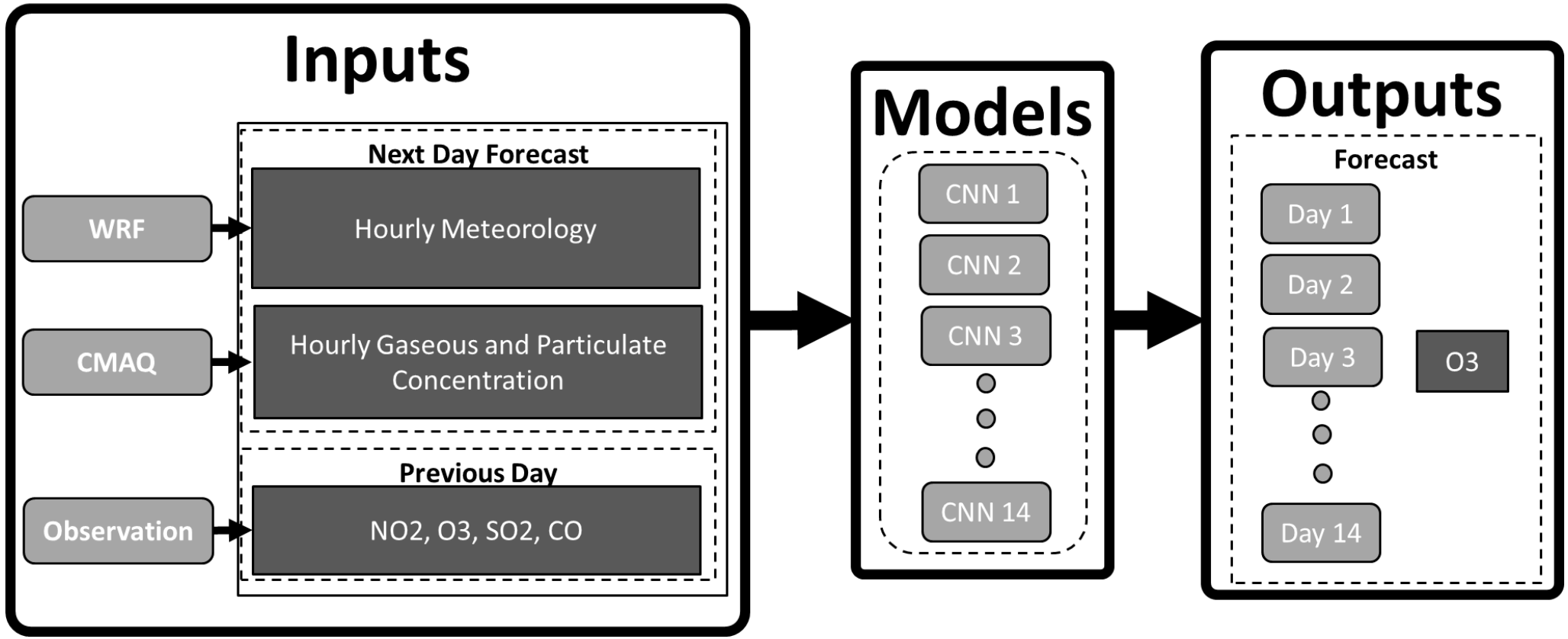
Motivation of this study

- Is there a hybrid model to combine numerical and deep learning modeling?
- Can this hybrid model be used to forecast long-term (e.g., one week or two weeks) concentrations of air pollutants?

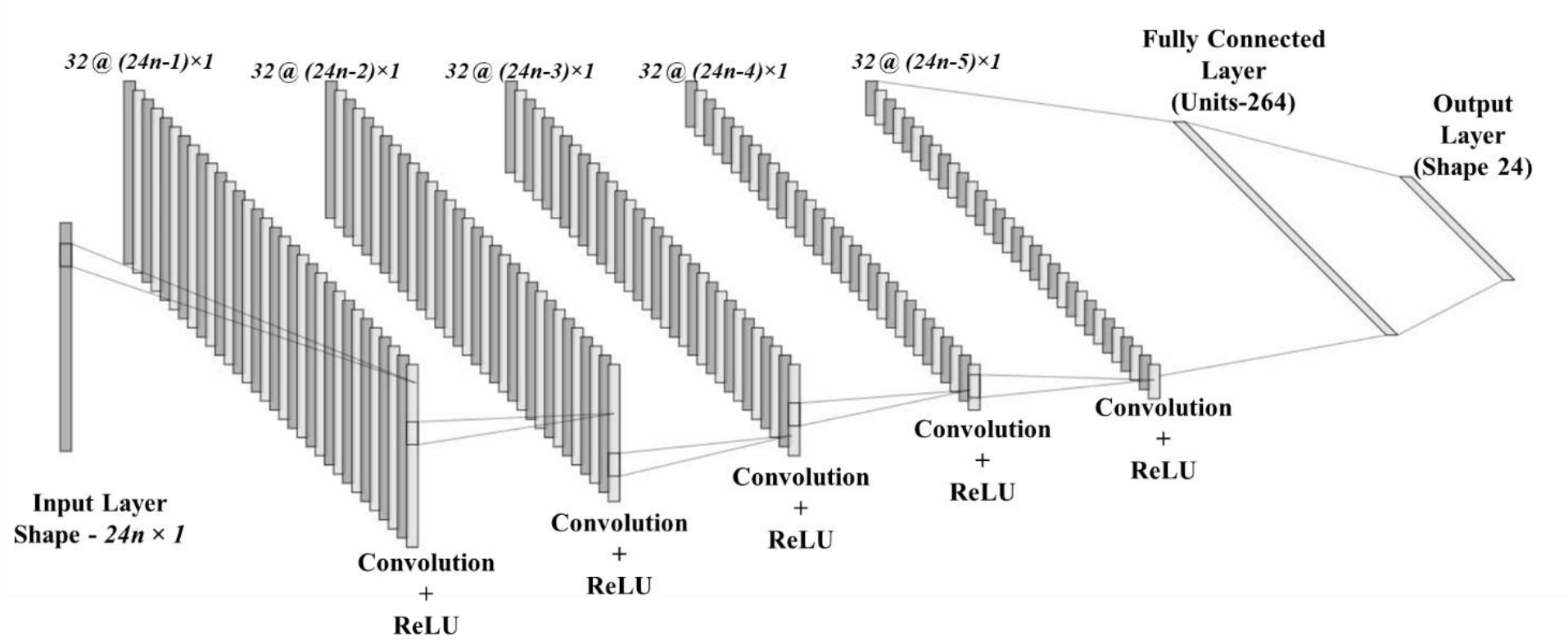
Data Set-up


Input
Forecasts
Meteorology from MCIP/WRF
32 Parameters
Pollutant from CMAQ
14 Parameters
Pollutant from Last Day Observations
4 Parameters
Output
24-Hour Ozone Concentration for Day 1, Day 2,
Day 3 …… till Day 14

Data Set-up


Inputs
Next Day Forecast
WRF
Hourly Meteorology
CMAQ
Hourly Gaseous and Particulate Concentration
Previous Day
Observation
NO2, O3, SO2, CO
Models
CNN 1
CNN 2
CNN 3
CNN 14
Outputs
Forecast
Day 1
Day 2
Day 3
Day 14
O3

Model Architecture



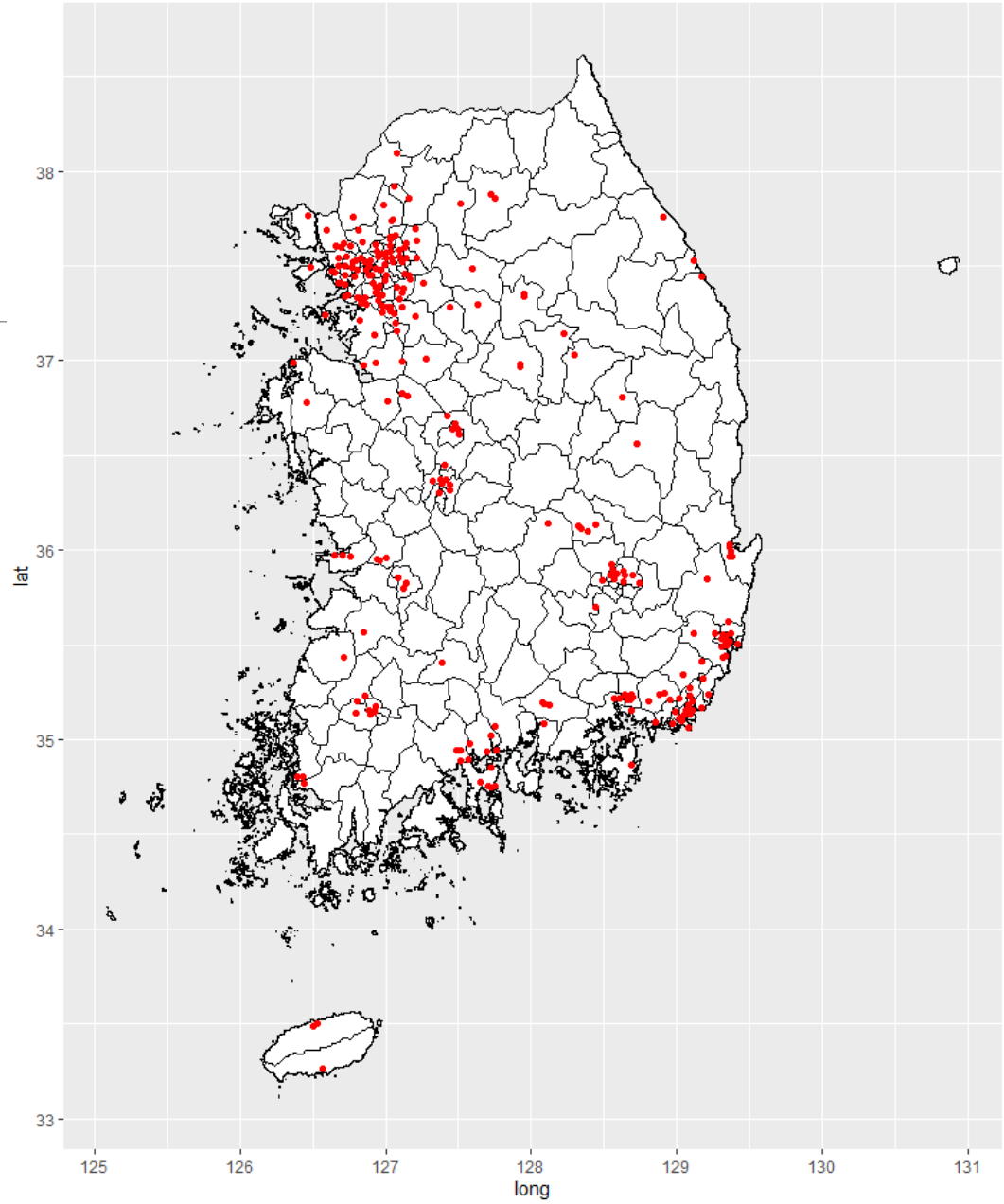
Loss Function

- Method to Evaluate the Performance of Model on a Sample Dataset.
- Common Loss Functions:
  - Mean Squared Error (Method 1)
  - Mean Absolute Error
  - Mean Absolute Percentage Error etc.
- Loss Function
  - Index of Agreement (Method 2)

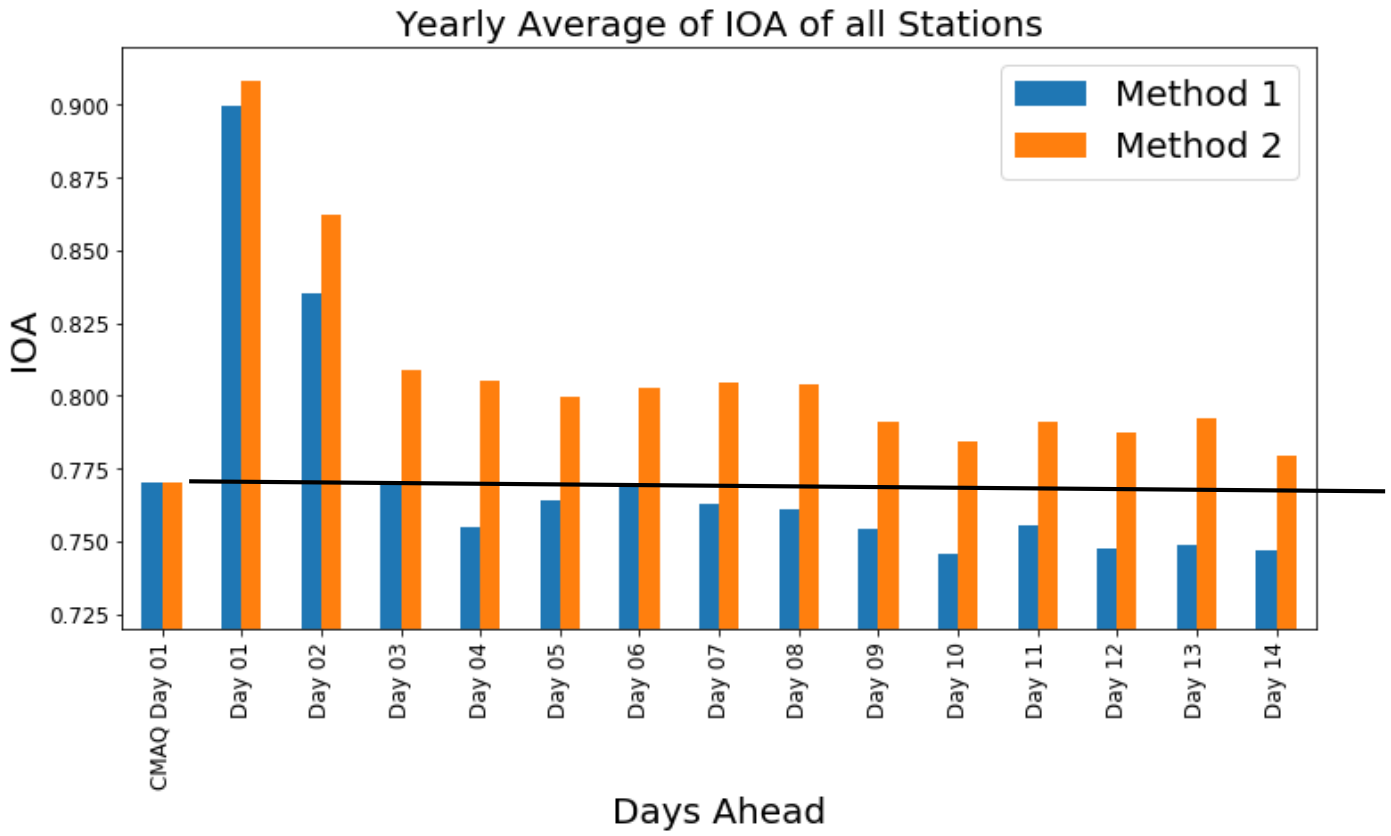
$$\mathrm{IOA} = 1 - \frac{\sum(\mathrm{O_i} - \mathrm{P_i})^2}{\sum(\mathrm{abs}(\mathrm{O_i} - \overline{\mathrm{O}}) + \mathrm{abs}(\mathrm{P_i} - \overline{\mathrm{O}}))^2}$$

Study Area and Data

- 255 Air Quality Monitoring Station of South Korea
- WRF Surface Meteorology Processed by MCIP (Meteorology-Chemistry Interface Processor).
- CMAQ predicted Surface Chemistry (Air Pollutant)
- Training: 2015-2017 Forecasting:2018

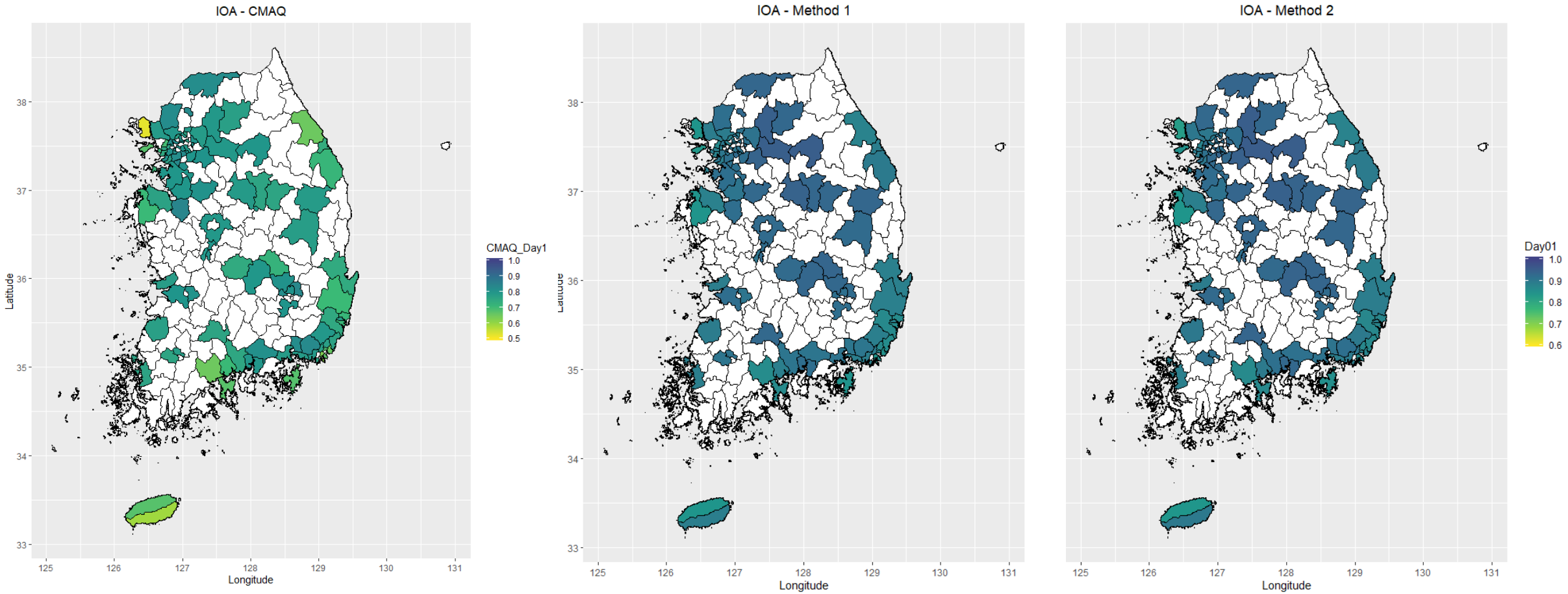


Results : Based on Loss Function

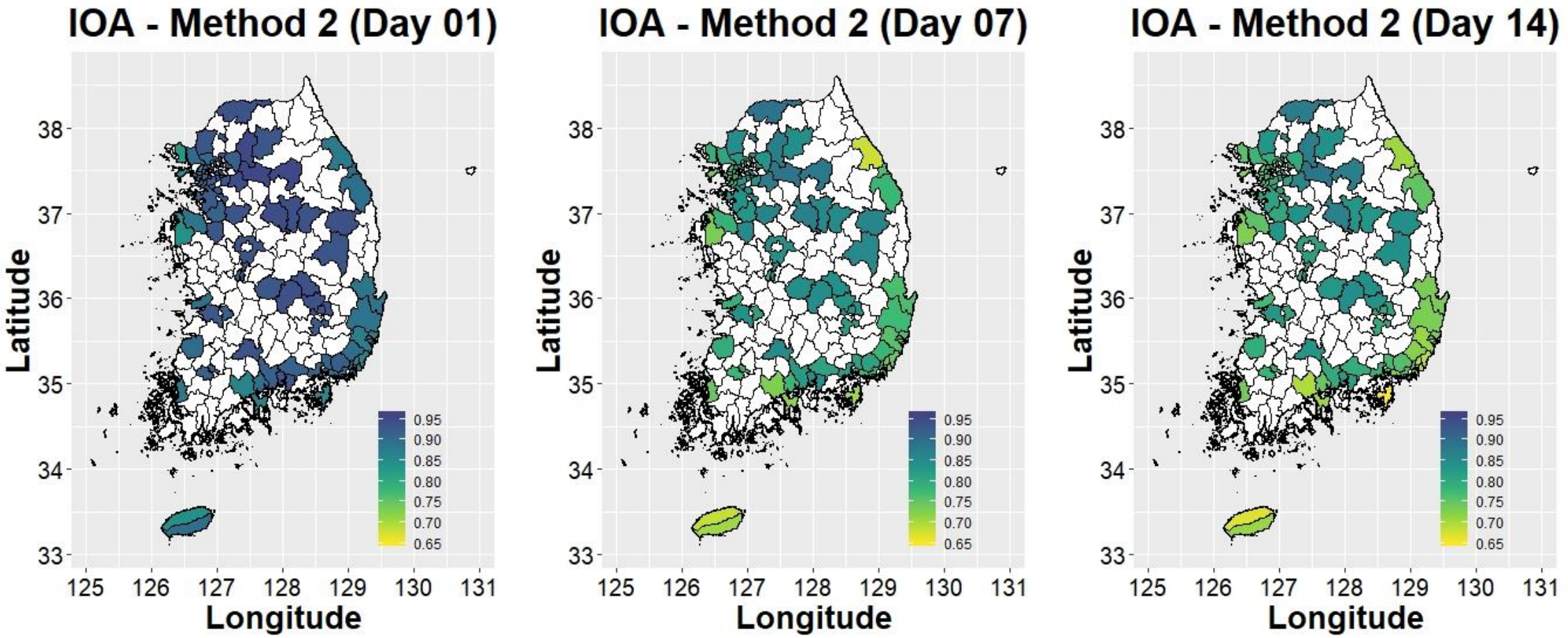


MSE – Method 1
IOA – Method 2

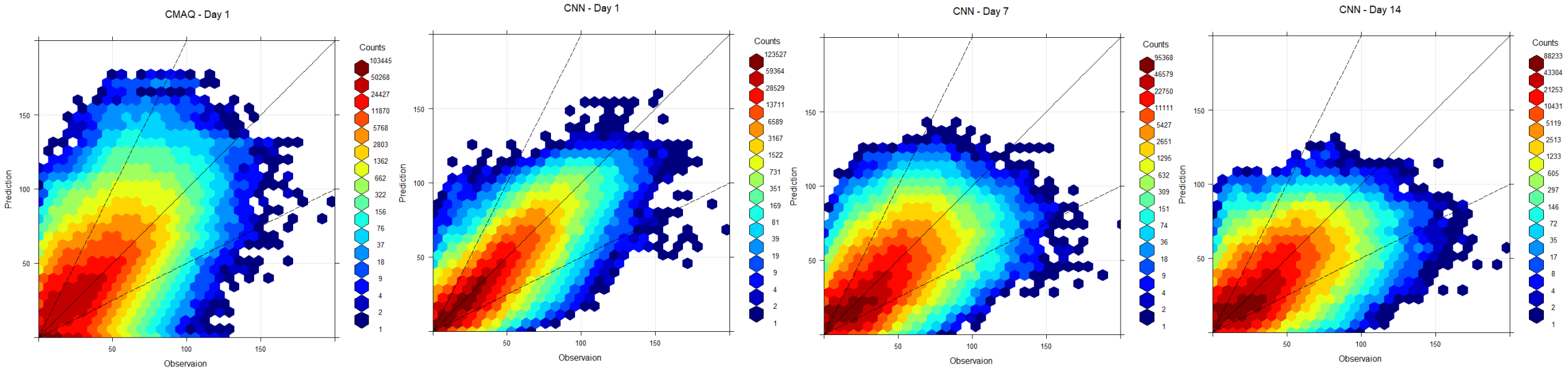
Results : Different Method Comparison



Results : Performance Comparison (Days)



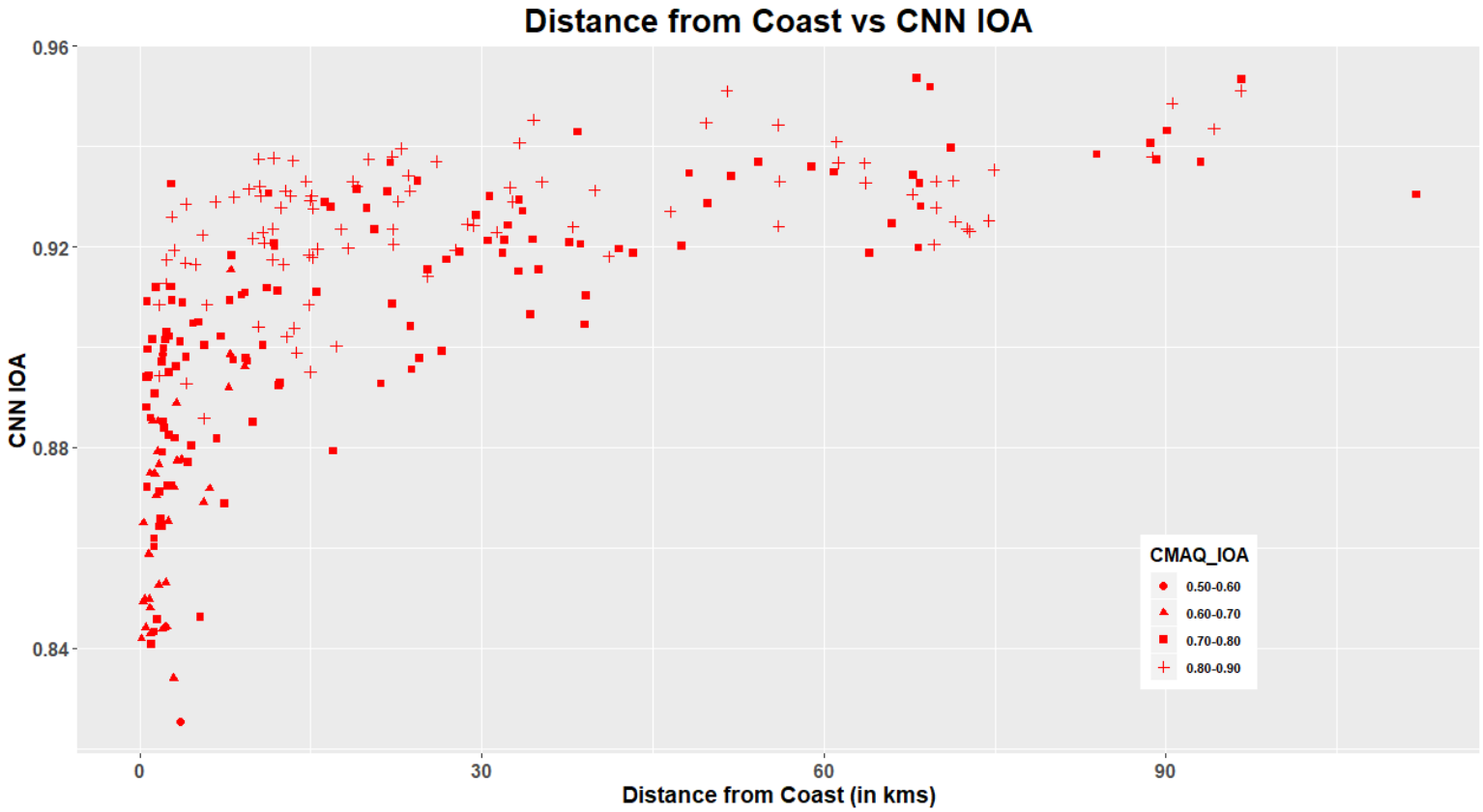
Results


CMAQ - Day 1
Prediction
Observaion
Counts
103445
50268
24427
11870
5768
2803
1362
662
322
156
76
37
18
9
4
2
1
CNN - Day 1
Prediction
Observaion
Counts
123527
59364
28529
13711
6589
3167
1522
731
351
169
81
39
19
9
4
2
1
CNN - Day 7
Prediction
Observaion
Counts
95368
46579
22750
11111
5427
2651
1295
632
309
151
74
36
18
9
4
2
1
CNN - Day 14
Prediction
Observaion
Counts
88233
43304
21253
10431
5119
2513
1233
605
297
146
72
35
17
8
4
2
1

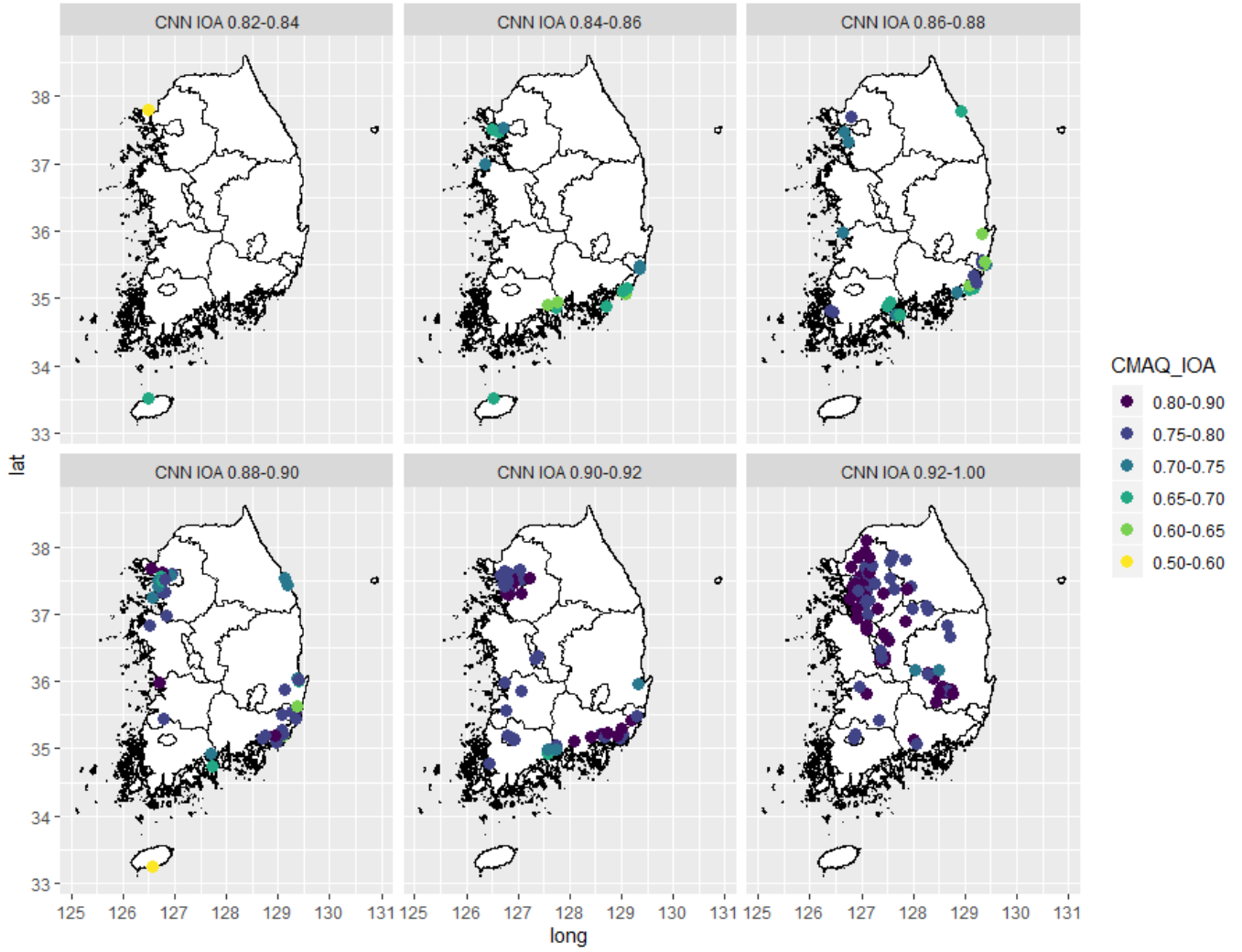
Factors Affecting Model's Performance

- Distance from Coast
- Base Model Performance (CMAQ)
- Ozone diurnal variation
- Urbanization

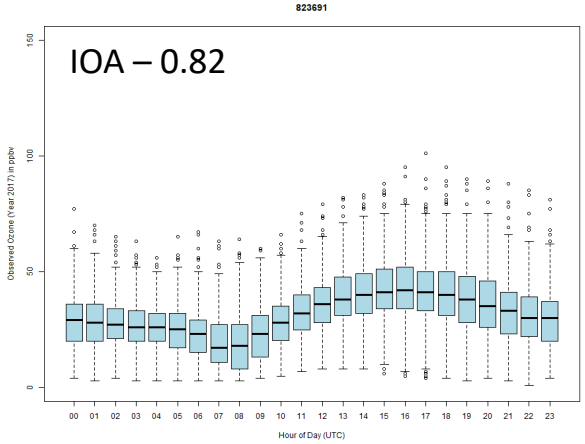
Distance from Coast

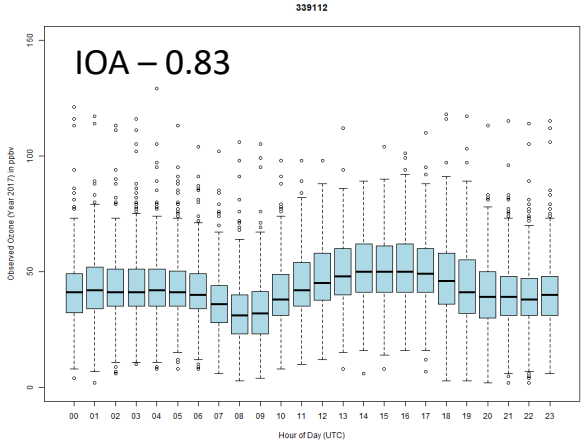


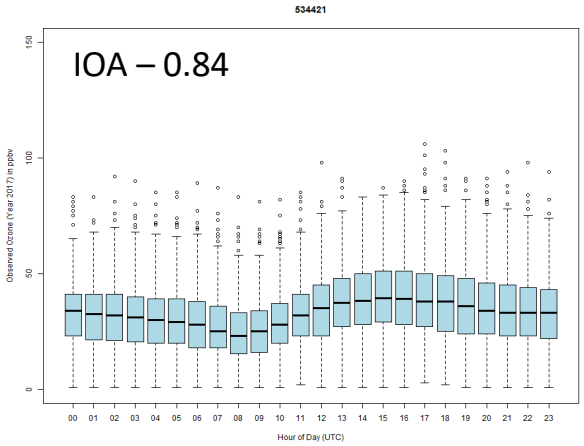
Dependence on CMAQ

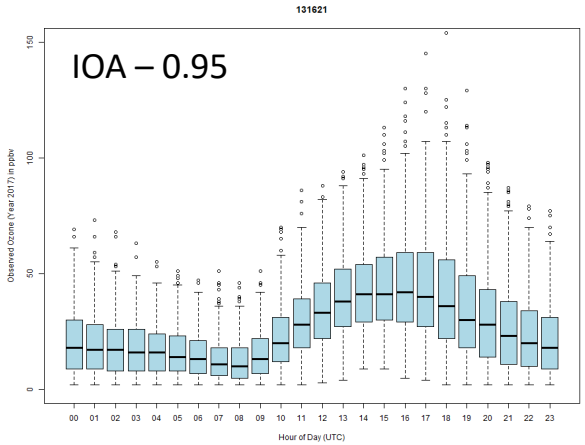


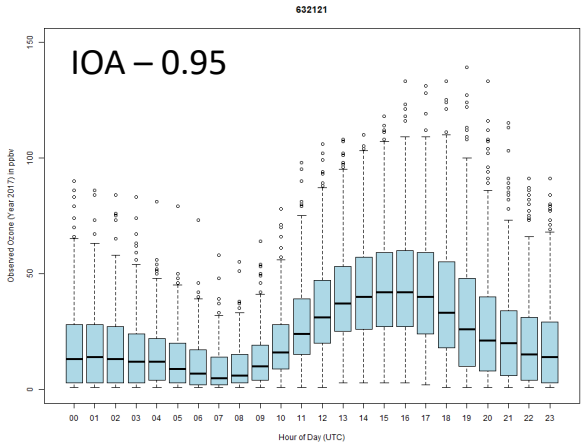
Ozone diurnal variation

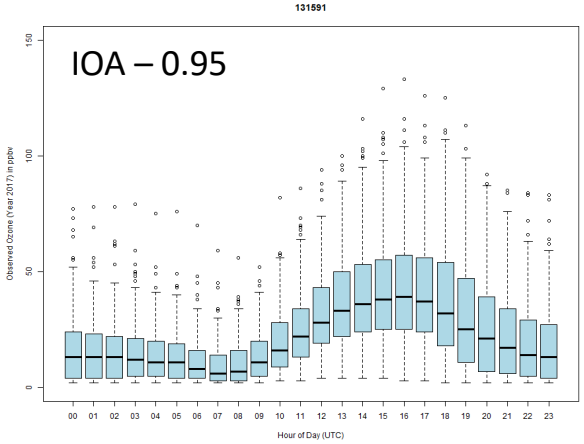




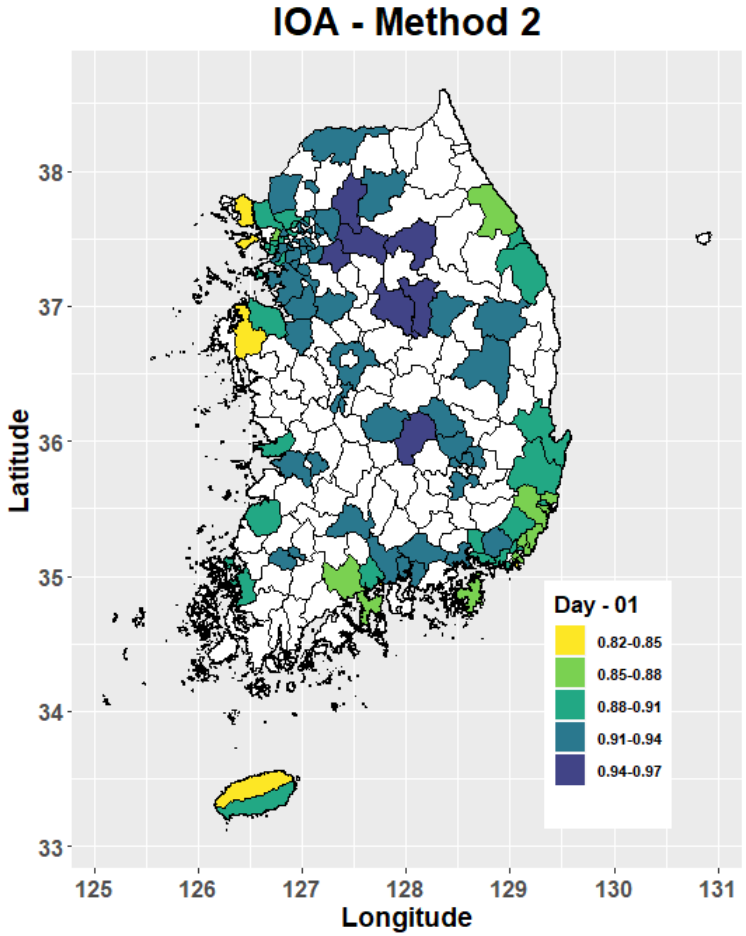


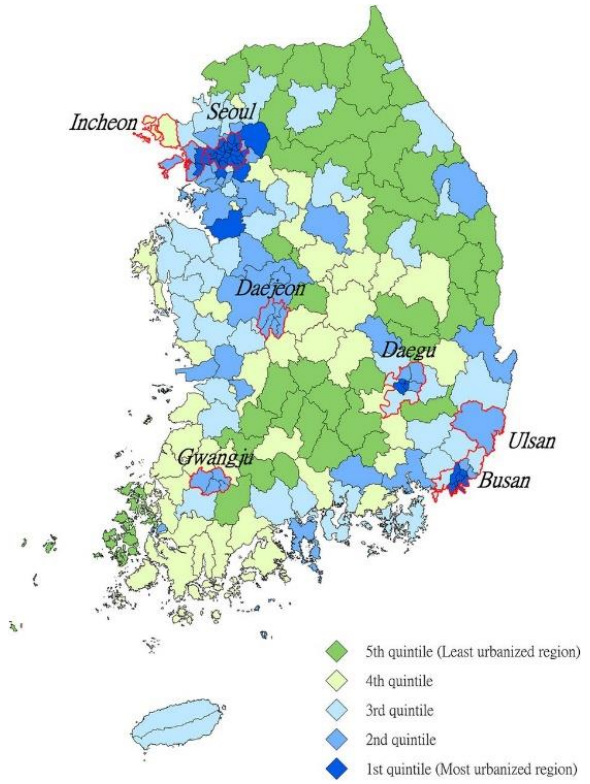






Urbanization





Source: Chan, C. H., Caine, E. D., You, S. & Yip, P. S. F. Changes in South Korean urbanicity and suicide rates, 1992 to 2012. *BMJ Open* **5**, e009451 (2015).

Summary & Funding

- Developed a Hybrid CMAQ-CNN model to predict 2-week Hourly Ozone Concentration
- Model significantly improved performance from the CMAQ Day 1.
- Average IOA for 1st Day was 0.91 that gradually reduces to 0.78 for 14th day.
- Computational time was significantly reduced.
- This approach can be used for other pollutants and weather varaibles.

- Funding: The National Institute of Environment Research (NIER), NIER-2018-04-02-056