



Supplement of

Using wavelet transform and dynamic time warping to identify the limitations of the CNN model as an air quality forecasting system

Ebrahim Eslami et al.

Correspondence to: Yunsoo Choi (ychoi6@uh.edu)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.



Figure S1. The number of machine learning-related publications on different topics from the Web of Science database.



Figure S2. Schematics comparing wavelet transform to Fourier transform (FT).



Figure S3. Time series and wavelet transform analysis of ozone concentrations from the NIER stations #111121 in Seoul, South Korea. Index at the x-axis as well as the period is in hours.



Figure S4. Comparison of time steps based on the DTW (black) and Euclidean distance (red) analysis.



Figure S5. Distance plot calculated by DTW comparing similarities (red line) between observations and two prediction models (CNN in the top right and CMAQ in the top left panels).



Daily IOA of CNN prediction: average over 25 stations

Figure S6. Box plot of daily IOAs of the CNN model in different weeks of 2017, averaged over 25 stations in Seoul, South Korea (from Eslami et al., 2019a).



Figure S7. Time series showing daily CNN model predictions (in red) and observed ozone concentrations (in black) during the daytime and the nighttime, averaged over 25 stations in Seoul, South Korea (from Eslami et al., 2019a).



Figure S8. Polar frequency (with respect to wind fields) of errors in CNN ozone predictions, averaged over all stations in Seoul, South Korea, for different months.



Observation vs. CNN by levels of RH

Figure S9. Scatter plots comparing CNN predictions with observed values in different seasons during the daytime and the nighttime.



Figure S10. Hourly time series of a high ozone episode in Seoul, South Korea, with respect to (a) levels of CNN prediction bias and (b) levels of RH%.



Figure S11. Variation in daylight and nighttime hours throughout the year in Seoul, South Korea.



Figure S12. Scatter plots comparing CNN predictions and observations with respect to levels of (a) wind speed and (b) pressure.



Figure S13. Taylor diagram comparing CMAQ and CMAQ-CNN models in different months of the 2014 ozone season, averaged over all stations across the continental United States.

		40 60 80 100 120		40 60 80 100 120		40 60 80 100 120		40 60 80 100 120
	Alabama	Arizona	Arkansas	California	Colorado	Connecticut	Delaware	Florida
CMAQ_CNN		-						
CMAQ	·		 	·•	·•			
Observation				·		i -	i - i	
	Georgia	Idaho	Illinois	Indiana	lowa	Kansas	Kentucky	Louisiana
CMAQ_CNN		 	 		-	-		
CMAQ	-		- •	·	·		·i	·
Observation		∳ ;	i ● i				<u> </u>	·•
	Maine	Maryland	Massachusetts	Michigan	Minnesota	Mississippi	Missouri	Montana
CMAQ_CNN	÷- ● + -}	 -	÷- ● +	÷-●	 -	-	 	- --
CMAQ		·		÷			••	
Observation	[•]		-		·	 	-	1 8 -1
	Nebraska	Nevada	New Hampshire	New Jersey	New Mexico	New York	North Carolina	North Dakota
CMAQ_CNN			÷-∎-÷	●				i ● i
CMAQ	·····	֥	÷		·i	÷	·	
Observation		 				֥	·	
	Ohio	Oklahoma	Oregon	Pennsylvania	Rhode Island	South Carolina	South Dakota	Tennessee
CMAQ_CNN	i ● i	÷- ●	÷- ● ;	i ● i			; - -;	
CMAQ	·	· · · · · • • · · · · · · · · · · · · ·		·		·	·	·····
Observation			- -			-		
	Texas	Utah	Vermont	Virginia	Washington	West Virginia	Wisconsin	Wyoming
CMAQ_CNN		i - -	֥	-	i +		i ● i	÷- ● ;
CMAQ		••		·····	i ● i		·•	
Observation	·		ii	¦ 	÷- ●	⊧ - • ∔		¦ [● }
	40 60 80 100 120		40 60 80 100 120		40 60 80 100 120		40 60 80 100 120	
Maximum daily ozone (ppb)								

Figure S14. Box plots comparing daily maximum ozone of CMAQ and CMAQ-CNN models with observed values in 48 states in the continental United States. Red vertical lines represent the yearly mean of ozone peaks as a reference.