



Supplement of

Development of a deep neural network for predicting 6 h average PM_{2.5} concentrations up to 2 subsequent days using various training data

Jeong-Beom Lee et al.

Correspondence to: Jae-Bum Lee (gercljb@korea.kr)

The copyright of individual parts of the supplement might differ from the article licence.

Table S1. Statistical performance results according to the number of layers.

Model	Day	MSE ((μgm^{-3}) ²)	RMSE (μgm^{-3})	R	IOA
2-layer	D+0	59.3	7.7	0.91	0.94
	D+1	92.1	9.6	0.86	0.89
	D+2	156.3	12.5	0.75	0.80
4-layer	D+0	54.7	7.4	0.91	0.95
	D+1	88.3	9.4	0.86	0.90
	D+2	134.5	11.6	0.77	0.84
5-layer (DNN-ALL)	D+0	53.3	7.3	0.91	0.95
	D+1	81.0	9.0	0.85	0.90
	D+2	112.4	10.6	0.79	0.86
6-layer	D+0	174.2	13.2	0.81	0.66
	D+1	292.4	17.1	0	0.17
	D+2	292.4	17.1	0	0.17
8-layer	D+0	302.7	17.4	0	0.15
	D+1	292.4	17.1	0	0.17
	D+2	292.4	17.1	0	0.17

Table S2. AQI performance results according to the number of layers.

Model	Day	ACC (%)		POD (%)		FAR (%)		F1-score (%)
2-layer	D+0	70.0	63/90	81.8	18/22	28.0	7/25	77
	D+1	55.6	50/90	81.0	17/21	39.3	11/28	69
	D+2	51.1	46/90	81.0	17/21	50.0	17/34	61
4-layer	D+0	71.1	64/90	81.8	18/22	28.0	7/25	76
	D+1	60.0	54/90	85.7	18/21	35.7	10/28	73
	D+2	60.0	54/90	81.0	17/21	45.2	14/31	65
5-layer (DNN-ALL)	D+0	77.8	70/90	72.7	16/22	11.1	2/18	80
	D+1	64.4	58/90	71.4	15/21	31.8	7/22	70
	D+2	61.1	55/90	76.2	16/21	40.7	11/27	67
6-layer	D+0	55.6	50/90	50	11/22	8.3	1/12	64
	D+1	47.8	43/90	0	0/21	0	0/0	0
	D+2	47.8	43/90	0	0/21	0	0/0	0
8-layer	D+0	45.6	41/90	0	0/22	0	0/0	0
	D+1	47.8	43/90	0	0/21	0	0/0	0
	D+2	47.8	43/90	0	0/21	0	0/0	0

5 Recall about "good" category = $\frac{a1}{a1+ b1+c1+d1}$, (S1)

Recall about "moderate" category = $\frac{b2}{a2+ b2+c2+d2}$, (S2)

Recall about "bad" category = $\frac{c3}{a3 b3+c3+d3}$, (S3)

Recall about "very bad" category = $\frac{d4}{a4+ b4+c4+d4}$, (S4)

Precision about "good" category = $\frac{a1}{a1+ a2+a3+a}$, (S5)

10 Precision about "moderate" category = $\frac{b2}{b1+ b2+b3+b4}$, (S6)

Precision about "bad" category = $\frac{c3}{c1+ c2+c3+c4}$, (S7)

Precision about "very bad" category = $\frac{d4}{d1 d2+d3+d4}$, (S8)

Table S3. Statistical evaluation results of CMAQ, DNN-OBS, DNN-OPM, and DNN-ALL models in the training set (2016 to 2018).

Model	Day	MSE (($\mu\text{g}\text{m}^{-3}$) ²)	RMSE ($\mu\text{g}\text{m}^{-3}$)	R	IOA
CMAQ	D+0	136.9	11.7	0.76	0.86
	D+1	146.4	12.1	0.74	0.84
	D+2	185.0	13.6	0.67	0.80
DNN-OBS	D+0	79.2	8.9	0.78	0.87
	D+1	139.2	11.8	0.54	0.65
	D+2	158.8	12.6	0.43	0.54
DNN-OPM	D+0	53.3	7.3	0.86	0.92
	D+1	88.4	9.4	0.75	0.83
	D+2	108.2	10.4	0.68	0.77
DNN-ALL	D+0	39.7	6.3	0.90	0.94
	D+1	57.8	7.6	0.84	0.90
	D+2	72.3	8.5	0.80	0.87

15

Table S4. Statistical evaluation results of CMAQ, DNN-OBS, DNN-OPM, and DNN-ALL models in the validation set (2019).

Model	Day	MSE (($\mu\text{g}\text{m}^{-3}$) ²)	RMSE ($\mu\text{g}\text{m}^{-3}$)	R	IOA
CMAQ	D+0	123.2	11.1	0.82	0.90
	D+1	132.3	11.5	0.80	0.89
	D+2	156.3	12.5	0.75	0.86
DNN-OBS	D+0	92.2	9.6	0.84	0.88
	D+1	182.3	13.5	0.63	0.65
	D+2	216.1	14.7	0.52	0.52
DNN-OPM	D+0	65.6	8.1	0.89	0.92
	D+1	123.2	11.1	0.78	0.81
	D+2	166.4	12.9	0.66	0.72
DNN-ALL	D+0	42.3	6.5	0.93	0.95
	D+1	77.4	8.8	0.88	0.90
	D+2	108.2	10.4	0.81	0.84

Table S5. Statistical performance of the DNN-ALL and Random Forest models.

Model	Day	MSE ((μgm^{-3}) ²)	RMSE (μgm^{-3})	R	IOA
DNN-ALL	D+0	53.3	7.3	0.91	0.95
	D+1	81.0	9.0	0.85	0.90
	D+2	112.4	10.6	0.79	0.86
Random Forest	D+0	62.4	7.9	0.90	0.93
	D+1	106.1	10.3	0.83	0.85
	D+2	156.3	12.5	0.73	0.76

20 Table S6. Precision and recall of DNN-ALL and CMAQ models by four categories: "good" ($PM_{2.5} \leq 15 \mu g m^{-3}$), "moderate" ($16 \mu g m^{-3} \leq PM_{2.5} \leq 35 \mu g m^{-3}$), "bad" ($36 \mu g m^{-3} \leq PM_{2.5} \leq 75 \mu g m^{-3}$), and "very bad" ($76 \mu g m^{-3} \leq PM_{2.5}$).

Model	Day	Precision					Recall				
		Good	Moderate	Bad	Very bad	Total	Good	Moderate	Bad	Very bad	Total
DNN-ALL	D+0	0.83	0.71	0.88	1.0	0.86	0.70	0.85	0.71	1.0	0.82
	D+1	0.83	0.61	0.64	0.0	0.52	0.38	0.79	0.70	0.0	0.47
	D+2	0.79	0.59	0.56	0.0	0.49	0.42	0.67	0.75	0.0	0.46
CMAQ	D+0	0.74	0.67	0.64	0.0	0.51	0.64	0.68	0.67	0.0	0.50
	D+1	0.85	0.69	0.54	0.0	0.52	0.65	0.67	0.70	0.0	0.51
	D+2	0.82	0.69	0.42	0.0	0.48	0.69	0.63	0.55	0.0	0.47

Table S7. AQI performance of the DNN-ALL and Random Forest model.

Model	Day	ACC (%)		POD (%)		FAR (%)		F1-score (%)
DNN-ALL	D+0	77.8	70/90	72.7	16/22	11.1	2/18	80
	D+1	64.4	58/90	71.4	15/21	31.8	7/22	70
	D+2	61.1	55/90	76.2	16/21	40.7	11/27	67
Random Forest	D+0	75.6	68/90	77.3	17/22	19.0	4/21	79
	D+1	61.1	55/90	76.2	16/21	33.3	8/24	71
	D+2	48.9	44/90	71.4	15/21	50.0	15/30	58