

# **Supporting Information for**

## **Satellite-based PM<sub>2.5</sub> Estimation Directly from Reflectance at the Top of the Atmosphere Using a Machine Learning Algorithm**

Jianjun Liu<sup>1</sup>, Fuzhong Weng<sup>2\*</sup> and Zhanqing Li<sup>3</sup>

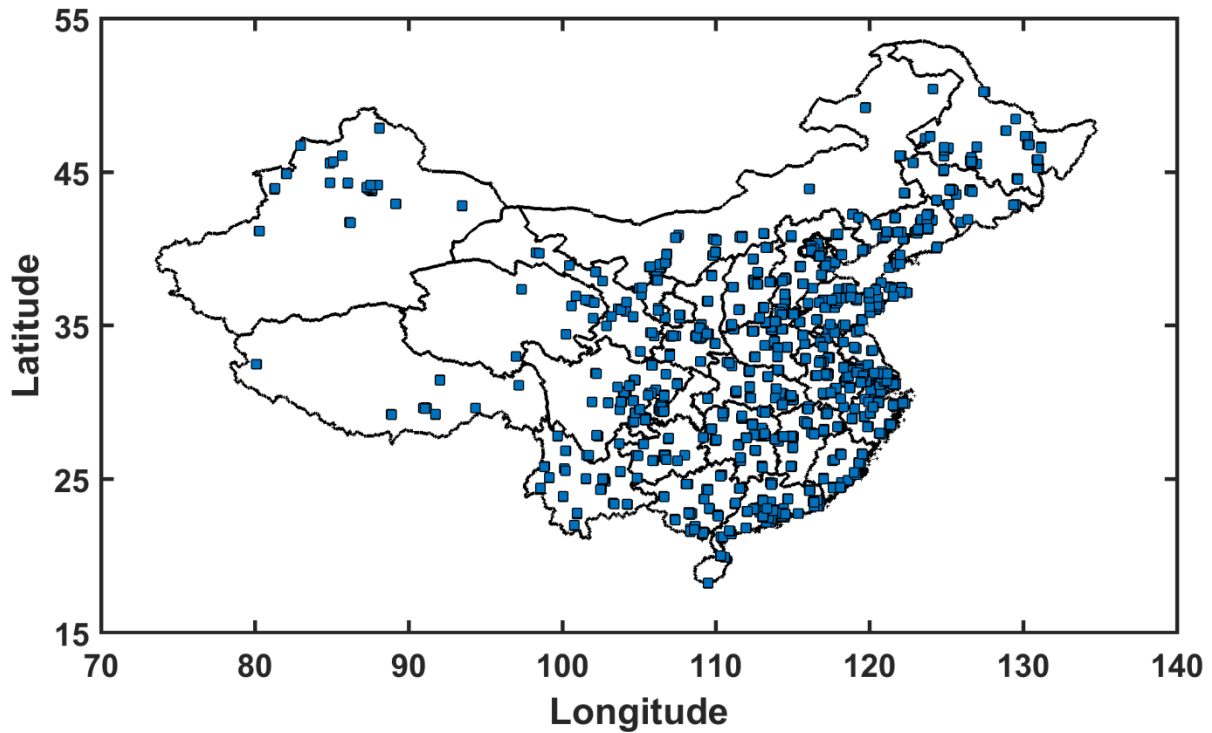
<sup>1</sup>Laboratory of Environmental Model and Data Optima (EMDO), Laurel, MD, USA

<sup>2</sup>Laboratory of Severe Weather, Beijing, China

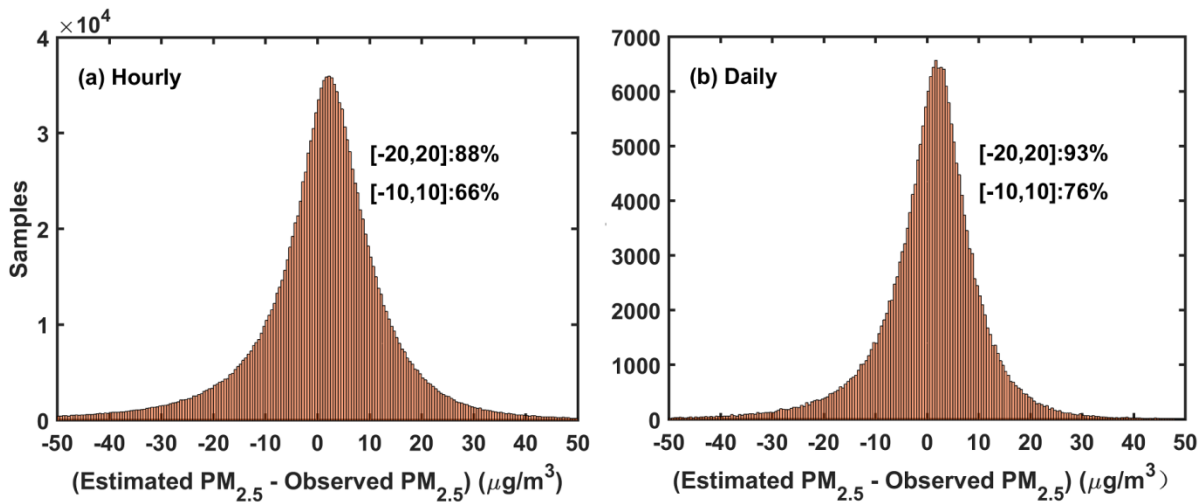
<sup>3</sup>Earth System Science Interdisciplinary Center and Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA

### **Contents of this file**

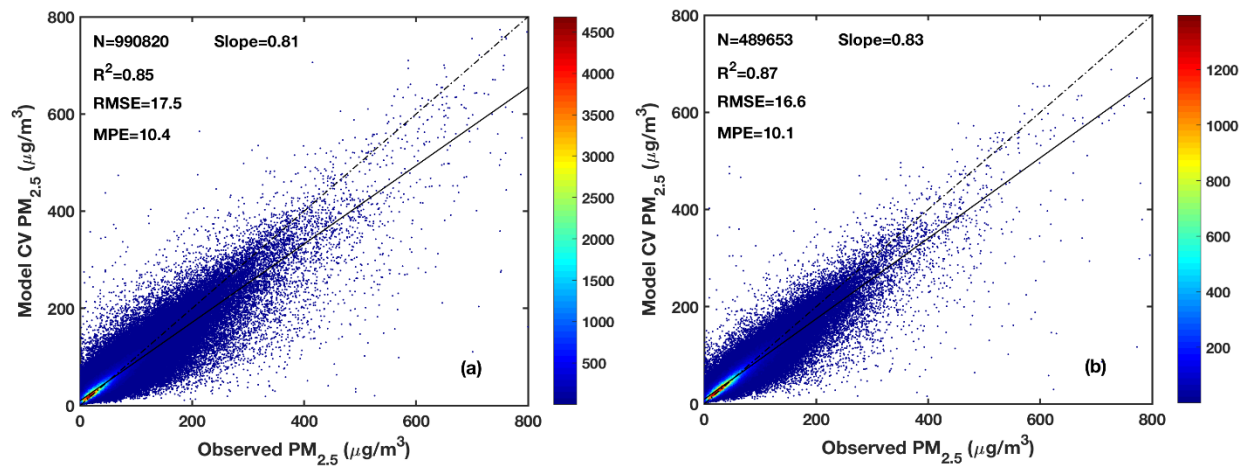
Figure S1, Figure S2, Figure S3 and Tables S1 and Table S2



**Figure S1.** The spatial distribution of PM<sub>2.5</sub> monitoring sites in mainland China in 2016 used in the study.



**Figure S2.** Histograms of the (a) hourly and (b) daily cross-validation biases of the Ref-PM<sub>2.5</sub> model. The percentage of samples in two ranges of bias values are shown in each panel.



**Figure S3.** Cross-validation (CV) results of the Ref- $\text{PM}_{2.5}$  model for samples (a) with TOA reflectance but without AOD retrievals available and (b) with both AOD and TOA reflectance available. The color bar shows the counts of data points. The dashed line is the 1:1 line. N: number of samples;  $R^2$ : coefficient of determination; RMSE: root-mean-square error of the predictions ( $\mu\text{g}/\text{m}^3$ ); MPE: mean prediction error ( $\mu\text{g}/\text{m}^3$ ).

**Table S1.** Descriptive statistics of the Ref-PM modeling variables in the training dataset.

	MAM				JJA				SON				DJF			Annual				
	Mean	Std	Max	Min	Mean	Std	Max	Min	Mean	Std	Max	Min	Mean	Std	Max	Min	Mean	Std	Max	Min
<b>PM<sub>2.5</sub></b>	48	39	569	1	31	21	996	1	49	41	1000	1	67	61	998	1	49	45	1000	1
<b>Ref1</b>	0.26	0.08	0.99	0.08	0.24	0.08	0.99	0.06	0.29	0.09	0.99	0.07	0.33	0.12	0.99	0.10	0.28	0.1	0.99	0.06
<b>Ref3</b>	0.20	0.07	0.88	0.04	0.17	0.06	0.95	0.03	0.21	0.08	0.90	0.03	0.24	0.10	0.92	0.04	0.2	0.08	0.95	0.03
<b>Ref6</b>	0.16	0.06	0.78	0.001	0.11	0.04	0.72	0.001	0.15	0.06	0.66	0.001	0.17	0.06	0.72	0.001	0.15	0.06	0.78	0.003
<b>P</b>	956	754	1038	553	951	675	1019	557	954	816	1043	555	964	819	1045	547	957	769	1045	547
<b>T</b>	288.2	8.9	310.3	251.4	300.2	5.8	313.3	269.6	286.3	10.0	307.9	245.1	275.3	10.8	303.1	230.9	287.3	12.8	313.3	230.9
<b>RH</b>	47.2	22.6	99.9	2.0	62.5	20.7	99.9	1.9	59.3	20.3	99.9	6.9	49.8	20.7	99.9	1.2	54.3	22.1	99.8	1.2
<b>U<sub>10</sub></b>	0.87	2.5	14.0	-10.2	0.2	2.0	11.2	-9.8	0.3	2.4	12.5	-15.6	1.0	2.2	13.6	-8.2	0.62	2.3	14	-15.6
<b>V<sub>10</sub></b>	0.49	3.0	11.1	-11.2	0.7	2.5	10.6	-13.0	-0.4	2.8	17.3	-15.3	-0.8	2.5	9.0	-15.3	-0.002	2.78	17.3	-15.3
<b>TCW</b>	12.8	9.9	69.9	0.4	34.8	14.9	75.4	1.6	18.8	15.3	79.4	0.5	6.2	5.4	49.6	0.32	17.7	16	79.4	0.3
<b>Ozone</b>	0.007	0.0009	0.01	0.005	0.006	0.0005	0.009	0.005	0.006	0.0007	0.01	0.005	0.007	0.001	0.01	0.005	0.007	0.001	0.011	0.005
<b>PBLH</b>	1025	819	4192	9	94	794	4270	10	800	577	3444	10	686	493	3297	8.0	866	702	4270	9
<b>NDVI</b>	0.28	0.13	0.94	0.02	0.43	0.12	0.98	0.01	0.34	0.13	0.96	0.01	0.26	0.13	0.96	0.01	0.33	0.14	0.98	0.01

MAM: March, April, May; JJA: June, July, August; SON: September, October, November; DJF: December, January, February

Std: Standard deviation; Max: maximum; Min: Minimum

Ref1, Ref3, and Ref6: Reflectance at 0.47, 0.64, and 2.3  $\mu\text{m}$

P: surface pressure (hPa); T: 2-m temperature (K); RH: surface relative humidity (%); U<sub>10</sub>: 10-m u-wind (m s<sup>-1</sup>);

V<sub>10</sub>: 10-m v-wind (m s<sup>-1</sup>); TCW: total column water (kg m<sup>-2</sup>); Ozone: total column ozone (kg m<sup>-2</sup>);

PBLH: planetary boundary layer height (m); NDVI: normalized difference vegetation index

**Table S2.** Pearson correlation coefficients for prediction and all predictor variables.

Variables	PM <sub>2.5</sub>	Ref1	Ref3	Ref6	P	T	RH	U10	V10	TCW	Ozone	PBLH	NDVI
PM <sub>2.5</sub>	1.0	0.24	0.21	0.17	0.12	-0.22	0.09	-0.02	0.07	-0.17	-0.01	-0.06	-0.17
Ref1	0.24	1.0	0.93	0.47	0.03	-0.53	0.19	0.03	-0.07	-0.23	0.23	-0.25	0.03
Ref3	0.21	0.93	1.0	0.63	-0.05	-0.51	0.1	0.06	-0.06	-0.26	0.26	-0.15	-0.11
Ref6	0.17	0.47	0.63	1.0	-0.24	-0.57	-0.09	0.15	-0.09	-0.45	0.29	0.16	-0.68
P	0.12	0.03	-0.05	-0.24	1.0	0.1	0.1	-0.11	-0.07	0.17	0.14	0.03	0.15
T	-0.22	-0.53	-0.51	-0.57	0.1	1.0	0.07	-0.22	0.24	0.74	-0.43	0.03	0.43
RH	0.09	0.19	0.1	-0.09	0.1	0.07	1.0	-0.23	0.12	0.45	-0.3	-0.06	0.34
U <sub>10</sub>	-0.02	0.03	0.06	0.15	-0.11	-0.22	-0.23	1.0	0.11	-0.28	0.35	0.11	-0.19
V <sub>10</sub>	0.07	-0.07	-0.06	-0.09	-0.07	0.24	0.12	0.11	1.0	0.16	-0.05	-0.06	0.08
TCW	-0.17	-0.23	-0.26	-0.45	0.17	0.74	0.45	-0.28	0.16	1.0	-0.43	-0.03	0.53
Ozone	-0.01	0.23	0.26	0.29	0.14	-0.43	-0.3	0.35	-0.05	-0.43	1.0	0.18	-0.26
PBLH	-0.06	-0.25	-0.15	0.16	0.03	0.03	-0.06	0.11	-0.06	-0.03	0.18	1.0	-0.2
NDVI	-0.17	0.03	-0.11	-0.68	0.15	0.43	0.34	-0.19	0.08	0.53	-0.26	-0.2	1.0

Ref1, Ref3, and Ref6: Reflectance at 0.47, 0.64, and 2.3  $\mu\text{m}$

P: surface pressure (hPa); T: 2-m temperature (K); RH: surface relative humidity (%); U<sub>10</sub>: 10-m u-wind ( $\text{m s}^{-1}$ );

V<sub>10</sub>: 10-m v-wind ( $\text{m s}^{-1}$ ); TCW: total column water ( $\text{kg m}^{-2}$ ); Ozone: total column ozone ( $\text{kg m}^{-2}$ );

PBLH: planetary boundary layer height (m); NDVI: normalized difference vegetation index