Supplementary Material

Evaluation of using satellite-derived aerosol optical depth in land use regression models for fine particulate matter and its elemental composition

Chun-Sheng Huang [1], Ho-Tang Liao [1], Tang-Huang Lin [2], Jung-Chi Chang [1], Chien-Lin Lee [1], Eric Cheuk-Wai Yip [1], Yee-Lin Wu [3], Chang-Fu Wu [1,4,*]

[1] Institute of Environmental and Occupational Health Sciences, National Taiwan University, Taipei, Taiwan
[2] Center for Space and Remote Sensing Research, National Central University, Taoyuan, Taiwan
[3] Department of Environmental Engineering, National Cheng Kung University, Tainan, Taiwan
[4] Department of Public Health, National Taiwan University, Taipei, Taiwan

* Corresponding author at: Institute of Environmental and Occupational Health Sciences, National Taiwan University, Room 717, No.17, Xu-Zhou Road, Taipei 100, Taiwan. Tel/Fax: +886 2 3366 8096; E-mail address: changfu@ntu.edu.tw (C.-F. Wu).

Table S1: Certified and measured values for PM measures using National Institute of Standards and Technology Standard Reference Material 2783 (n=3)

PM Measures	Certified Value (ng/cm^2)	Measured Value (ng/cm^2)	Percent Recovery [a] (%)
Al	2330.32 ± 53.21	2309.64 ± 39.75	99
Ca	1325.3 ± 170.68	1240.62 ± 7.86	94
Cr	13.55 ± 2.51	12.89 ± 1.21	95
Fe	2660.64 ± 160.64	2579.39 ± 7.2	97
K	530.12 ± 52.21	472.01 ± 1.78	89
Mn	32.13 ± 1.2	31.97 ± 1.65	100
S	105.42 ± 26.1	93.87 ± 1.73	89
Si	5883.53 ± 160.64	5458.59 ± 135.81	93
Ti	149.6 ± 24.1	140 ± 1.06	94
V	4.87 ± 0.6	5.18 ± 0.19	106
Zn	179.72 ± 13.05	188.46 ± 0.34	105

[a] Percent Recovery = (Measured Value/Certified Value) × 100

Table S2: The definitions of predictor variables in constructions of LUR models

Descriptions of Variables [a]	Predictors Name	Unit	Radius (m)	Effect
Land Use				
High density residential area	HDRES	km^2		+
Low density residential area	LDRES	km^2		+
Industrial area	INDUSTRY	km^2	100, 300, 500, 1000, 5000	+
Port area	PORT	km^2		+
Semi-natural and forested area	SEMINATURAL	km^2		–
Urban green area	URBANGREEN	km^2		–
Road Information and Related Variables				
Major road area	MAJORROADAREA	km^2		+
All road area	ALLROADAREA	km^2	25, 50, 100, 300, 500, 1000	+
Length of major road	MAJORROADLEN	km		+
Length of all road	ALLROADLEN	km		+
Inverse of distance to the nearest major road	DISTINVMR1	m^{-1}	-	+
Inverse of distance squared to the nearest major road	DISTINVMR2	m^{-2}	-	+
Inverse of distance to the nearest road	DISTINVAR1	m^{-1}	-	+
Inverse of distance squared to the nearest road	DISTINVAR2	m^{-2}	-	+
Elevation				
Elevation	ELV	m	-	–
Demographic Data				
Number of population	POPULATION	thousand	100, 300, 500, 1000, 5000	+
Number of household	HOUSEHOLD	thousand		+
Stationary Emission Sources				
Number of stationary emission sources	POINT_N	N	100, 300, 500, 1000, 5000	+
Temples				
Number of temples	TEMPLE	N	100, 300, 500, 1000, 5000	+

Satellite Data				
AOD	AOD	n.a.	-	+
AOD percentage	AOD_PER	n.a.	-	+/−

[a] The predictors were acquired from diverse sources and listed as follows: land use data were obtained from National Land Survey and Mapping Center (https://www.nlsc.gov.tw/); road information were obtained from Ministry of Transportation and Communications (https://www.motc.gov.tw/); demographic data were obtained from Ministry of the Interior (https://www.moi.gov.tw/); elevation data were obtained from Advanced Spaceborne Thermal Emission and Reflection Radiometer satellite, which is managed by National Aeronautics and Space Administration (NASA) of the US and Ministry of Economy, Trade, and Industry of Japan (https://asterweb.jpl.nasa.gov/); the coordinates of stationary emission sources were obtained from Taiwan Emission Data System, which is managed by Taiwan Environmental Protection Administration (https://teds.epa.gov.tw/); locations of temples were obtained from the cultural resources GIS, which is built by the Center for GIS of Academia Sinica (http://gis.rchss.sinica.edu.tw/); satellite data were obtained from Terra and Aqua satellites, which are managed by NASA (https://www.nasa.gov/).

Table S3. Descriptive statistics of $PM_{2.5}$ and elemental composition in annual value, HPS and LPS

	PM Measures	Median [a]	Mean	SD	Max	Min	Range/mean (%)
Annual value	**$PM_{2.5}$**	21.2	21.0	7.4	33.3	8.7	117%
	Al	183.3	176.8	44.8	247.2	79.3	95%
	Ca	113.0	114.7	20.8	148.8	68.1	70%
	Cr	13.5	13.5	1.9	15.9	9.4	49%
	Fe	210.8	191.9	65.3	265.8	63.6	105%
	K	470.8	448.5	145.2	691.8	190.9	112%
	Mn	16.1	16.2	8.9	40.3	4.2	222%
	S	3392.1	3164.4	602.9	3869.2	1944.3	61%
	Si	359.7	343.7	98.5	527.9	140.5	113%
	Ti	12.1	11.5	4.0	18.4	3.6	129%
	V	9.2	9.0	3.3	18.1	3.8	159%
	Zn	77.6	76.5	24.6	113.0	32.8	105%
High $PM_{2.5}$ season (HPS)	**$PM_{2.5}$**	26.7	27.0	11.7	48.5	9.2	145%
	Al	261.9	255.0	63.7	332.1	108.8	88%
	Ca	142.1	139.3	26.5	167.7	70.1	70%
	Cr	17.9	17.7	2.5	20.6	12.4	47%
	Fe	277.9	242.1	80.2	322.1	80.7	100%
	K	697.8	662.2	189.7	892.5	238.5	99%
	Mn	22.7	22.1	9.5	38.7	6.1	147%
	S	3791.0	3443.4	926.4	4664.6	1801.0	83%
	Si	475.4	456.0	129.8	620.5	179.1	97%
	Ti	14.5	14.3	4.9	21.1	3.9	120%
	V	8.9	9.0	3.3	13.7	3.6	112%
	Zn	92.5	92.3	37.0	149.6	31.9	128%
Low $PM_{2.5}$ season (LPS)	**$PM_{2.5}$**	16.4	16.6	4.6	24.4	8.3	97%
	Al	110.6	120.9	39.3	186.5	58.2	106%
	Ca	93.4	97.1	22.5	156.4	66.7	92%
	Cr	9.9	10.4	1.7	12.6	7.2	53%
	Fe	170.4	156.0	56.4	235.4	51.4	118%
	K	244.9	295.8	154.2	608.4	113.3	167%
	Mn	11.4	12.0	9.4	42.9	2.8	333%
	S	3121.3	2965.1	453.5	3736.5	2046.6	57%
	Si	271.7	263.5	80.8	461.7	112.9	132%

Ti	10.1	9.5	3.5	16.5	3.3	139%
V	8.5	8.9	4.4	24.0	3.9	225%
Zn	67.1	65.2	18.9	113.1	33.4	122%

[a] Units: $\mu g/m^3$ for $PM_{2.5}$ and ng/m^3 for elemental composition.

Table S4. Summary of land use regression models of $PM_{2.5}$ and elemental composition using annual averages (Method 1)

PM Measures	LUR model [a]	R^2 of model	Adjusted R^2	LOOCV R^2	LOOCV RMSE [b]	p-value of Moran's I
$PM_{2.5}$	9.96 + 2.31 × HDRES_5000 − 3.14 × URBANGREEN_5000	0.49	0.41	0.35	6.16	0.93
Al	213.95 − 41.24 × SEMINATURAL_1000	0.44	0.41	0.33	37.95	0.81
Ca	76.97 + 457.54 × MAJORROADAREA_500 + 96.48 × MAJORROADLEN_100	0.70	0.66	0.58	14.02	0.58
Cr	13.80 −145.95 × SEMINATURAL_100	0.31	0.27	0.24	1.72	0.11
Fe	63.34 + 1226.52 × INDUSTRY_300 + 14.41 × HDRES_5000 + 271.87 × MAJORROADAREA_1000	0.62	0.53	0.40	52.14	0.04
K	225.88 + 47.67 × HDRES_5000 − 68.57 × URBANGREEN_5000	0.54	0.48	0.40	116.56	0.67
Mn	10.16 + 50.08 × INDUSTRY_1000 + 1.29 × INDUSTRY_1000_5000	0.64	0.59	0.05	8.94	0.13
S	2241.55 + 189.34 × HDRES_5000 − 249.28 × URBANGREEN_5000	0.49	0.41	0.34	504.04	0.63
Si	196.09 + 25.00 × INDUSTRY_5000 + 9.28 × MAJORROADLEN_500	0.60	0.54	0.39	79.15	0.55
Ti	4.43 + 0.10 × MAJORROADLEN_1000 + 0.03 × TEMPLE_5000	0.57	0.51	0.41	3.14	0.19
V	7.25 + 1.62 × TEMPLE_300	0.35	0.31	0.07	3.28	0.80
Zn	44.30 + 5.93 × INDUSTRY_5000 + 4.47 × HDRES_5000 − 8.92 × URBANGREEN_5000	0.68	0.61	0.44	19.09	0.54

[a] The surface area (km^2) of major road (MAJORROADAREA_X), industry (INDUSTRY_X), high density residence (HDRES_X) , urban green area (URBANGREEN_X), semi-natural and forested area (SEMINATURAL_X), the total length (km) of major roads (MAJORROADLEN_X), the number (N) of temples (TEMPLE_X). The _X indicates the buffer size (in meters). INDUSTRY_1000_5000 indicates the surface area of industry between buffer size of 1000 and 5000 meters.

[b] The concentration of $PM_{2.5}$ and elemental composition are represented as $\mu g/m^3$ and ng/m^3, respectively.

Table S5. Summary of land use regression models of $PM_{2.5}$ and elemental composition using annual averages (Method 2)

PM Measures	LUR model [a]	R^2 of model	Adjusted R^2	LOOCV R^2	LOOCV RMSE [b]	p-value of Moran's I
$PM_{2.5}$	3.62 + 35.68 × AOD	0.56	0.53	0.44	5.74	0.47
Al	66.81 + 6.08 × MAJORROADLEN_100 + 115.62 × AOD	0.58	0.52	0.44	34.70	0.11
Ca	76.97 + 457.55 × MAJORROADAREA_500 + 96.48 × MAJORROADLEN_100	0.70	0.66	0.58	14.02	0.58
Cr	13.80 − 145.95 × SEMINATURAL_100	0.31	0.27	0.24	1.72	0.11
Fe	7.70 + 1181.99 × MAJORROADAREA_500 + 0.14 × POINT_N_5000 + 296.24 × AOD	0.77	0.71	0.63	41.23	0.62
K	150.35 + 615.23 × AOD	0.43	0.39	0.30	125.52	0.72
Mn	−0.54 + 38.38 × INDUSTRY_1000 + 30.21 × AOD	0.75	0.72	0.70	5.06	0.76
S	1816.45 + 2781.53 × AOD	0.51	0.48	0.40	481.95	0.50
Si	119.07 + 18.74 × INDUSTRY_5000 + 8.18 × MAJORROADLEN_500 + 212.89 × AOD	0.69	0.61	0.47	74.38	0.41
Ti	4.43 + 0.10 × MAJORROADLEN_1000 + 0.03 × TEMPLE_5000	0.57	0.51	0.41	3.14	0.19
V	7.25 + 1.62 × TEMPLE_300	0.35	0.31	0.07	3.28	0.80
Zn	17.80 + 68.33 × MAJORROADAREA_1000 + 4.60 × INDUSTRY_5000 + 85.95 × AOD	0.75	0.69	0.57	16.75	>0.99

[a] The surface area (km^2) of major road (MAJORROADAREA_X), industry (INDUSTRY_X), semi-natural and forested area (SEMINATURAL_X), the total length (km) of major roads (MAJORROADLEN_X), the number (N) of stationary emission sources (POINT_N_X), temples (TEMPLE_X). The _X indicates the buffer size (in meters). AOD denotes the extracted AOD value at the site.

[b] The concentration of $PM_{2.5}$ and elemental composition are represented as $\mu g/m^3$ and ng/m^3, respectively.

Table S6. Summary of land use regression models of $PM_{2.5}$ and elemental composition in high $PM_{2.5}$ season (HPS) (Method 3)

PM Measures	LUR model [a]	R^2 of model	Adjusted R^2	LOOCV R^2	LOOCV RMSE [b]	p-value of Moran's I
$PM_{2.5}$	−14.18 + 95.24 × ALLROADAREA_300 −729.50 × URBANGREEN_100 + 0.03 × POINT_N_5000 + 35.74 × AOD + 134.33 × AOD_PER	0.97	0.95	0.91	3.54	0.46
Al	74.19 − 7758.20 × URBANGREEN_100 + 5.76 × HOUSEHOLD_1000 + 20.58 × TEMPLE_300 + 990.00 × AOD_PER	0.86	0.81	0.70	36.10	0.37
Ca	56.30 + 4.00 × HDRES_5000 − 2215.07 × URBANGREEN_100 + 34.78 × MAJORROADLEN_100 + 5.36 × HOUSEHOLD_500 + 293.51 × AOD_PER	0.92	0.88	0.73	14.34	0.20
Cr	11.74 + 0.75 × INDUSTRY_5000 −447.00 × URBANGREEN_100 − 76.30 × URBANGREEN_100_500 + 8.72 × HOUSEHOLD_100 + 0.01 × HOUSEHOLD_100_5000 + 50.04 × AOD_PER	0.85	0.76	0.62	1.61	0.09
Fe	46.40 + 302.04 × MAJORROADAREA_1000 + 10.56 × HDRES_5000 + 850.42 × INDUSTRY_300 + 9.83 × INDUSTRY_300_5000 + 716.82 × AOD_PER	0.89	0.85	0.80	36.76	0.20
K	176.93 −15692.82 × URBANGREEN_100 + 7.82 × HOUSEHOLD_1000 + 1.05 × TEMPLE_5000 + 2959.48 × AOD_PER	0.83	0.78	0.67	112.94	>0.99
Mn	−0.75 − 978.74 × URBANGREEN_100 − 262.83 × URBANGREEN_100_500 + 1.87 × TEMPLE_300 + 0.08 × POINT_N_5000 + 187.59 × AOD_PER	0.95	0.92	0.90	3.12	0.63
S	1707.04 + 200.49 × INDUSTRY_5000 −40114.77 × URBANGREEN_100 + 13449.25 × AOD_PER	0.92	0.90	0.87	350.45	0.14
Si	167.56 + 1861.66 × ALLROADAREA_300 + 0.31 × HOUSEHOLD_5000 + 1672.75 × AOD_PER	0.71	0.64	0.49	96.13	0.58
Ti	0.89 −313.56 ×URBANGREEN_100 + 0.27 × HOUSEHOLD_1000 + 0.03 × TEMPLE_5000 + 64.79 × AOD_PER	0.92	0.90	0.86	1.89	0.92
V	4.88 + 42.93 × AOD_PER	0.64	0.61	0.53	2.32	0.65

Zn	3.44 + 491.45 × ALLAROADAREA_300 + 12.85 × INDUSTRY_5000 − 2886.23 × URBANGREEN_100 − 633.20 × URBANGREEN_100_500 + 577.89 × AOD_PER	0.97	0.96	0.91	11.47	0.44

[a] The surface area (km^2) of all road (ALLROADAREA_X), major road (MAJORROADAREA_X), high density residential area (HDRES_X), industry (INDUSTRY_X), urban green area (URBANGREEN_X), the total length (km) of major roads (MAJORROADLEN_X), the number (N) of household (HOUSEHOLD_X), temples (TEMPLE_X) and stationary emission sources (POINT_N_X). The _X indicates the buffer size (in meters). URBANGREEN_100_500 indicates the surface area of urban green between buffer size of 100 and 500 meters, and so on. AOD denotes the extracted AOD value at the site. AOD_PER denotes the AOD percentage at the site.

[b] The concentration of $PM_{2.5}$ and elemental composition are represented as $\mu g/m^3$ and ng/m^3, respectively.

Table S7. Summary of land use regression models of $PM_{2.5}$ and elemental composition in low $PM_{2.5}$ season (LPS) (Method 3)

PM Measures	LUR model [a]	R^2 of model	Adjusted R^2	LOOCV R^2	LOOCV RMSE [b]	p-value of Moran's I
$PM_{2.5}$	8.56 + 19.36 × AOD	0.51	0.48	0.40	3.69	0.68
Al	24.58 + 1.70 + MAJORROADLEN_1000 + 101.48 × AOD	0.64	0.59	0.55	27.25	0.94
Ca	55.94 + 364.52 × MAJORROADAREA_500 + 3809.75 × HDRES_100 −1529.21 × URBANGREEN_100 + MAJORROADLEN_100 + 1.01 × POINT_N_1000	0.82	0.74	0.59	14.81	0.84
Cr	7.43 + 0.57 × INDUSTRY_5000 − 248.97 × URBANGREEN_100 + 42.72 × AOD_PER	0.58	0.48	0.29	1.47	0.83
Fe	−2.53 + 9755.61 × ALLROADAREA_100 + 1064.55 × MAJORROADAREA_500 + 949.85 × INDUSTRY_300 + 226.23 × AOD	0.81	0.74	0.43	44.05	0.96
K	52.32 + 588.95 × AOD	0.42	0.38	0.33	130.18	0.15
Mn	0.30 + 58.54 × INDUSTRY_1000 + 20.55 × AOD	0.87	0.86	0.80	4.34	0.60
S	2093.35 + 6912.75 × MAJORROADAREA_500 + 1826.00 × AOD	0.54	0.47	0.36	373.76	0.45
Si	53.29 + 21.33 × INDUSTRY_5000 − 1551.79 × URBANGREEN_300 + 5.64 × MAJORROADLEN_500 + 0.51 × TEMPLE_5000 + 1334.43 × AOD_PER	0.88	0.82	0.76	41.07	0.19
Ti	1.01 + 0.69 × MAJORROADLEN_300 + 0.04 × TEMPLE_5000 + 43.26 × AOD_PER	0.66	0.58	0.49	2.62	0.97
V	6.31 + 2.49 × TEMPLE_300	0.46	0.43	0.07	4.36	0.46
Zn	4.92 + 219.17 × ALLROADAREA_300 + 40.18 × MAJORROADAREA_1000 + 8.15 × INDUSTRY_5000 − 2296.31 × URBANGREEN_100 + 2611.95 × DISTINVMR2 + 5.57 × TEMPLE_300 + 416.50 × AOD_PER	0.93	0.88	0.71	10.60	0.22

[a] The surface area (km^2) of all road (ALLROADAREA_X), major road (MAJORROADAREA_X), high residential area (HDRES_X), industry (INDUSTRY_X), urban green area (URBANGREEN_X), the total length (km) of major roads (MAJORROADLEN_X), the number (N) of stationary

emission sources (POINT_N_X) and temples (TEMPLE_X). The _X indicates the buffer size (in meters). DISTINVMR2 represents the inverse of distance squared to the nearest major road. AOD denotes the extracted AOD value at the site. AOD_PER denotes the AOD percentage at the site.

[b] The concentration of $PM_{2.5}$ and elemental composition are represented as $\mu g/m^3$ and ng/m^3, respectively.

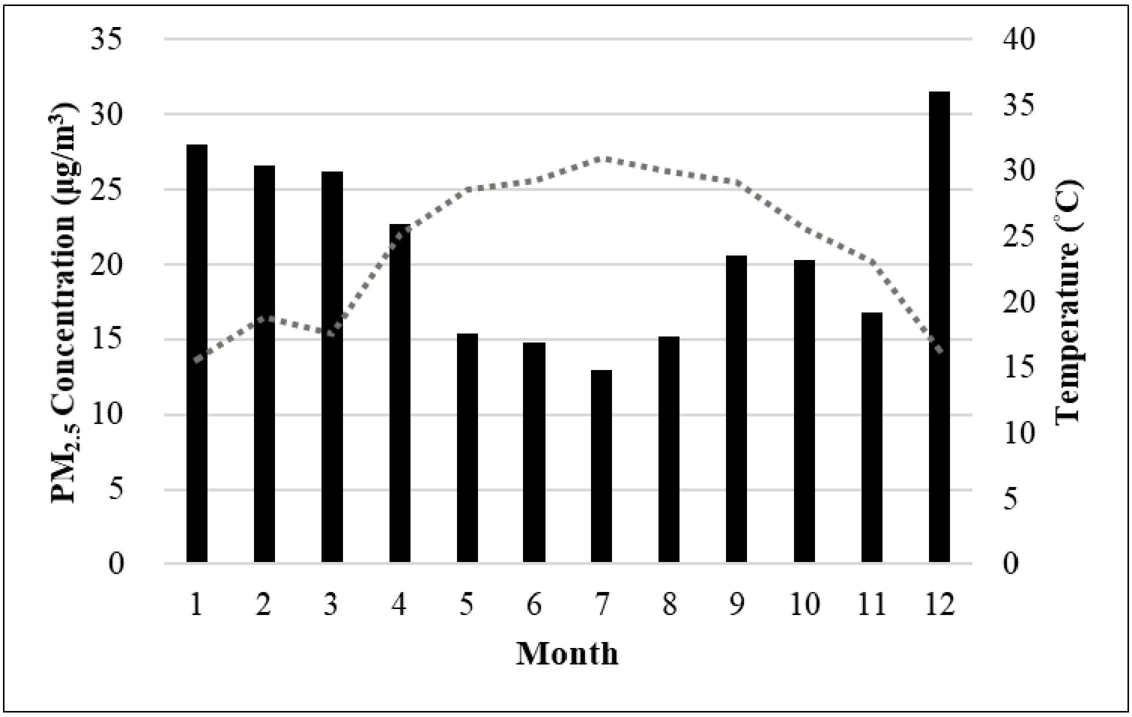

Figure S1: Comparison of monthly $PM_{2.5}$ concentration (black column) and ambient temperature (dotted line)