Supporting Information for

Satellite-derived 1-km-resolution PM_1 concentrations from 2014 to 2018 across China

Jing Wei[†, ‡], Zhanqing Li[‡*], Jianping Guo[§], Lin Sun[//], Wei Huang[⊥],

Wenhao Xue[†], Tianyi Fan[†], Maureen Cribb[‡]

[†] State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China

[‡] Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland 20742, United States

[§] State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China

[//] College of Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

[⊥] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100101, China

Contents of this file

Table S1. Summary of data sources used in this study.

Dataset	Variable	Content	Unit	Spatial resolution	Temporal resolution	Data source
PM_1	PM_1	PM_1	$\mu g/m^3$	-	5 min	CAWN
AOD	AOD	MAIAC AOD	-	1 km	Daily	MCD19A2
Meteorological	BLH	Boundary layer height	m	0.125°	3-hour	ERA-Interim reanalysis product
	PRE	Total precipitation	mm	0.125°	3-hour	
	ET	Evaporation	mm	0.125°	3-hour	
	RH	Relative humidity	%	0.125°	3-hour	
	TEM	2-m air temperature	K	0.125°	6-hour	
	SP	Surface pressure	hPa	0.125°	6-hour	
	WS	10-m wind speed	m/s	0.125°	6-hour	
	WD	10-m wind direction	m/s	0.125°	6-hour	
Pollution emission	IDT	Industry	Mg/grid	0.25°	Monthly	MEIC emission data
	POW	Power	Mg/grid	0.25°	Monthly	
	RST	Residential	Mg/grid	0.25°	Monthly	
	TST	Transportation	Mg/grid	0.25°	Monthly	
Land use	NDVI	NDVI	-	500 m	Monthly	MOD13A3
	LUC	Land use cover	-	500 m	Annually	MCD12Q1
Topographic	DEM	DEM	m	90 m	-	SRTM
	Relief	Surface relief	m	90 m	-	
	Aspect	Surface aspect	°	90 m	-	
	Slope	Surface slope	°	90 m	-	
Traffic	Road	Road	M	1 km	-	-
Population	NTL	Night light	$Wcm^{-2}\ sr^{-1}$	500 m	Monthly	VIIRS

Table S2. Correlations and effects of independent variables on PM_1 concentrations from all the monitoring stations in China from 2014 to 2018.

Positive variables				Negative variables			
No.	Variable	R	*p*-value	No.	Variable	R	*p*-value
1	AOD	0.378	$p < 0.01$	1	BLH	-0.351	$p < 0.01$
2	RSD	0.372	$p < 0.01$	2	DEM	-0.235	$p < 0.01$
3	SP	0.236	$p < 0.01$	3	TEM	-0.220	$p < 0.01$
4	ET	0.192	$p < 0.01$	4	WS	-0.188	$p < 0.01$
5	TSP	0.188	$p < 0.01$	5	NDVI	-0.155	$p < 0.01$
6	LUC	0.138	$p < 0.01$	6	PRE	-0.111	$p < 0.01$
7	NTL	0.135	$p < 0.01$	7	Relief	-0.076	$p < 0.01$
8	IDT	0.129	$p < 0.01$	8	Aspect	-0.039	$p < 0.01$
9	POW	0.068	$p < 0.01$	9	RH	-0.009	$p < 0.05$
10	WD	0.047	$p < 0.01$	10	Slope	-0.007	$p < 0.05$
11	Road	0.031	$p < 0.01$				

Table S3. Model performance for different extremely randomized trees models in China.

Year	R^2				RMSE				MAE			
	ERT	SET	TET	STET	ERT	SET	TET	STET	ERT	SET	TET	STET
2014	0.69	0.71	0.72	**0.74**	19.9	19.2	18.9	**18.2**	12.5	12.2	11.9	**11.5**
2015	0.73	0.74	0.74	**0.76**	14.7	14.7	14.4	**14.1**	9.5	9.4	9.2	**9.0**
2016	0.72	0.72	0.72	**0.74**	15.0	15.0	14.9	**14.5**	9.2	9.1	9.0	**8.7**
2017	0.73	0.74	0.76	**0.77**	13.5	13.3	12.8	**12.5**	8.5	8.4	8.1	**7.9**
2018	0.71	0.72	0.73	**0.76**	10.2	10.1	9.9	**9.5**	6.3	6.2	6.1	**5.9**
All	0.73	0.74	0.75	**0.77**	15.6	15.3	15.0	**14.5**	9.6	9.4	9.2	**8.9**

ERT: extremely randomized trees; SET: space extremely randomized trees; TET: time extremely randomized trees; STET: space-time extremely randomized trees

Table S4. Annual and seasonal mean spatial coverage and concentrations of 1-km-resolution PM_1 estimates in China.

Year	Spatial coverage	Annual mean PM_1	Season	Spatial coverage	Seasonal mean PM_1
2014	98.9 %	32.3±12.8 μg/m^3	Spring	97.1 %	22.4±7.5 μg/m^3
2015	99.0 %	24.5±8.7 μg/m^3	Summer	95.9 %	16.4±5.8 μg/m^3
2016	99.1 %	24.7±8.7 μg/m^3	Autumn	98.8 %	24.4±8.8 μg/m^3
2017	99.0 %	23.8±8.0 μg/m^3	Winter	85.8 %	36.3±14.5 μg/m^3
2018	98.8 %	16.8±7.3 μg/m^3			
Mean	99.5 %	24.4±8.7 μg/m^3			

Table S5. Comparison of the model performances of different regression models in China.

Model	Model validation			Spatial resolution	Satellite	Time period	Reference
	R^2	RMSE	MAE				
GAM	0.59	22.5	-	10 km	MODIS	2014	Chen et al. (2018)
MLR	0.23	32.8	22.5				
LME	0.41	28.7	19.4				
BPNN	0.52	26.1	17.8	5 km	Himawari-8	2015–2017	Zang et al. (2018)
GRNN	0.61	24.1	14.3				
PCA-GRNN	0.65	22.0	13.8				
PCA-GRNN	0.74	19.0	11.4	5 km	Himawari-8	2015–2017	Zang et al. (2019)
Two-stage	0.80	15.4	9.3	5 km	Himawari-8	2015–2017	Wang et al. (2019)
STET	**0.72**	**18.2**	**11.5**	**1 km**	MODIS	2014	Our study
STET	**0.76**	**13.7**	**8.5**	**1 km**	MODIS	2015–2017	Our study

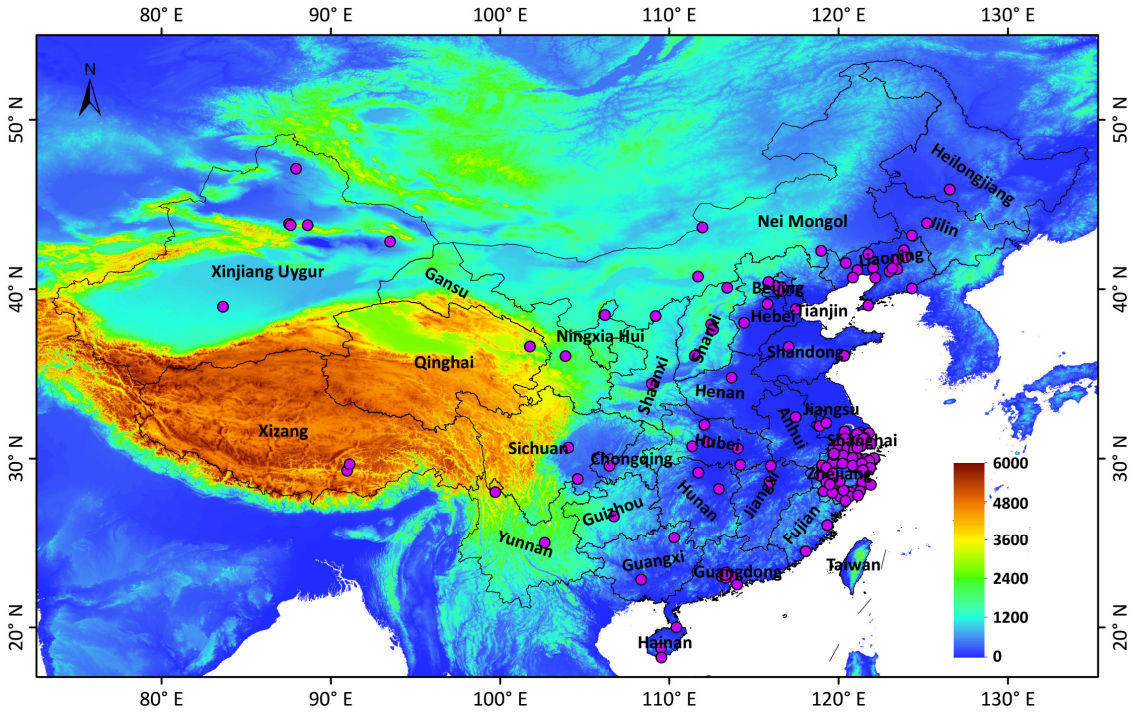

Figure S1. Spatial distribution of surface PM_1 monitoring stations in China (marked as purple dots). The background map shows the digital elevation model (DEM) data (unit = m).

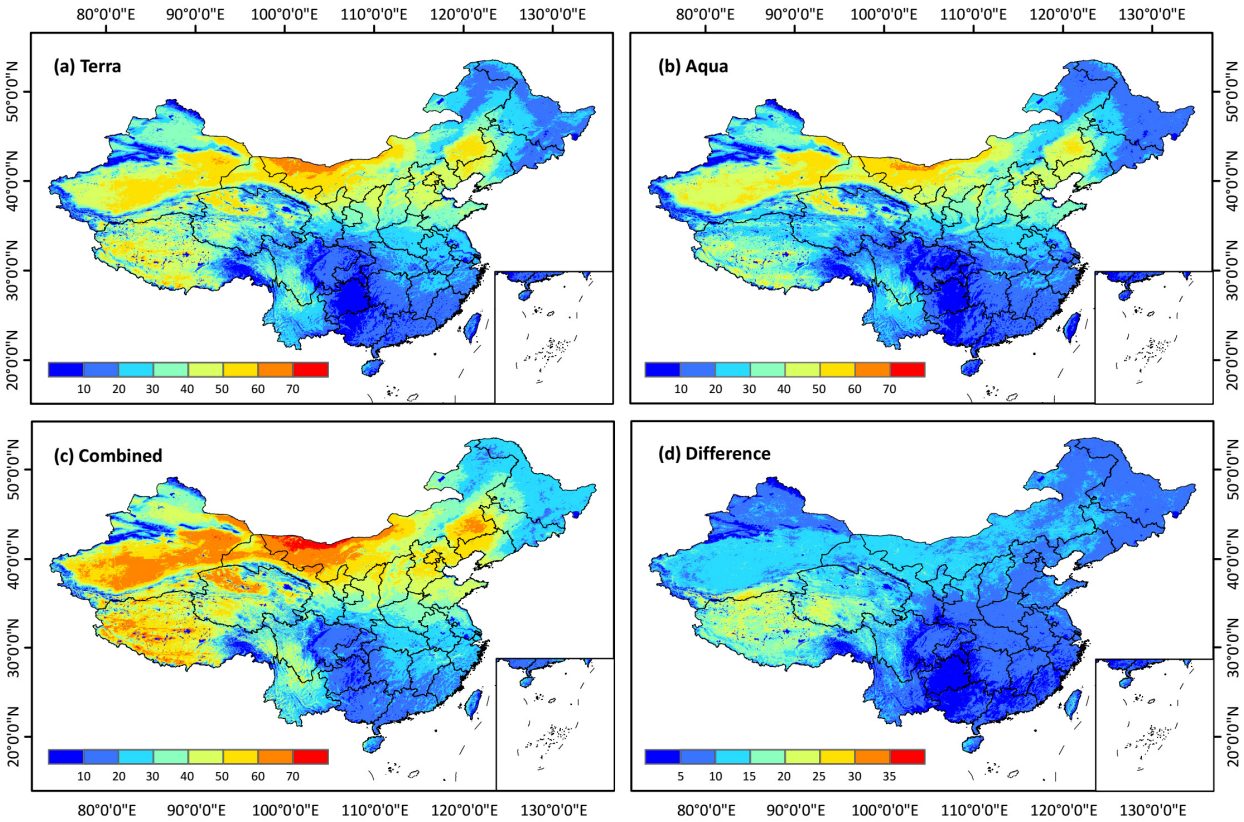

Figure S2. Spatial coverage (%) for (a) Terra, (b) Aqua, (c) Terra and Aqua combined MAIAC products, and (d) their difference (Combined - Terra) from 2014 to 2018 in China.

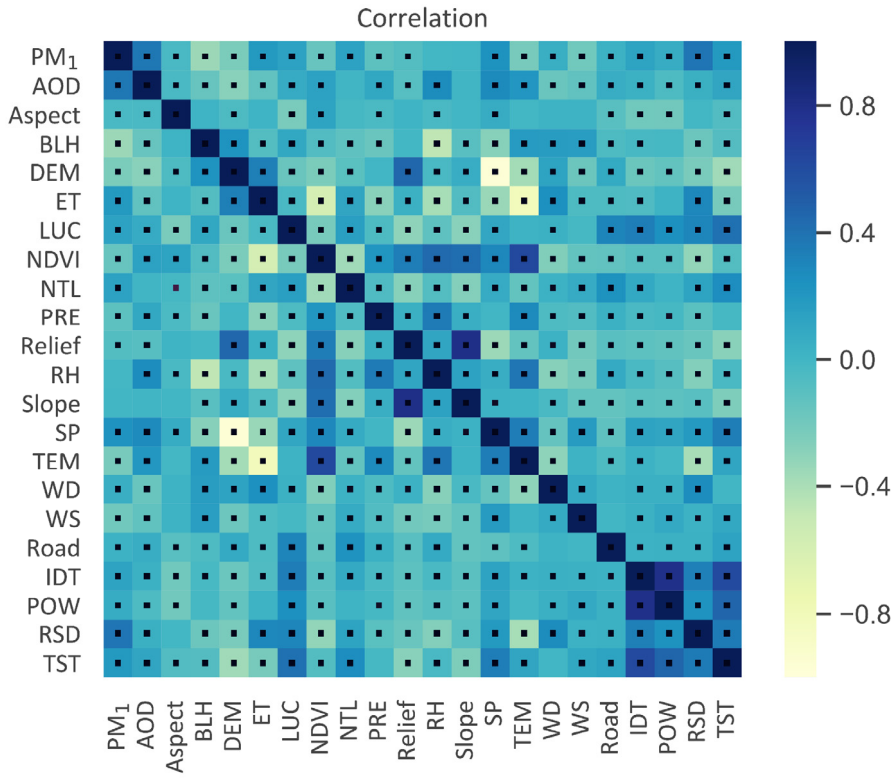

Figure S3. Correlation coefficient matrix between surface PM_1 measurements and all independent variables in China. Black dots indicate a significant trend at the 99% confidence level (p-value < 0.01).

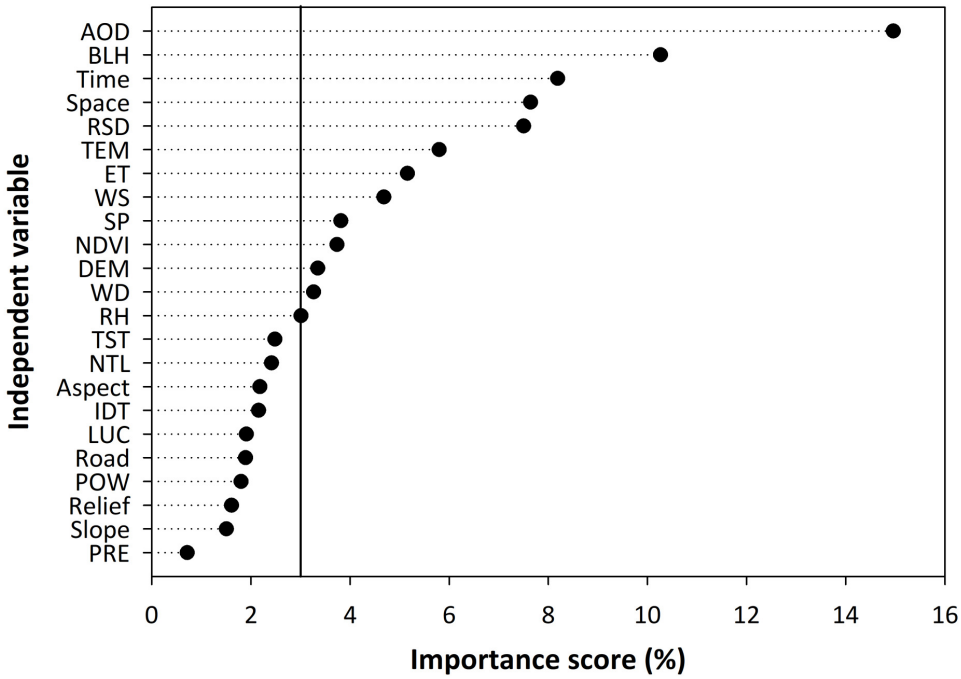

Figure S4. Importance scores of all independent variables for the space-time extremely randomized trees (STET) model.

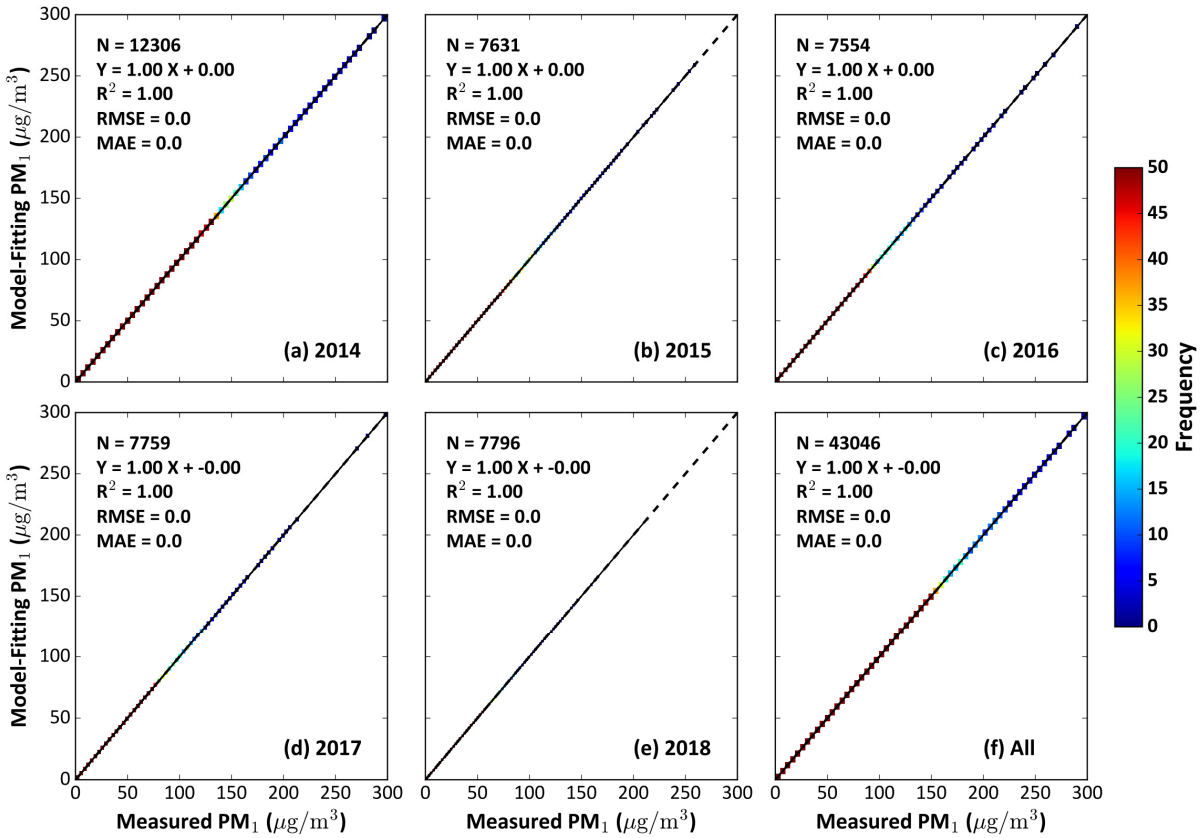

Figure S5. Density scatterplots of STET-model-estimated PM_1 as a function of measured PM_1 across China for the years (a)-(e) 2014 to 2018 and (f) all years. Statistical metrics are given in each panel: the number of samples (N), the coefficient of determination (R^2), the root-mean-square error (RMSE; μg/m^3), and the mean absolute error (MAE; μg/m^3). The linear regression relationship is also given in each panel. Dashed lines are the 1:1 lines.

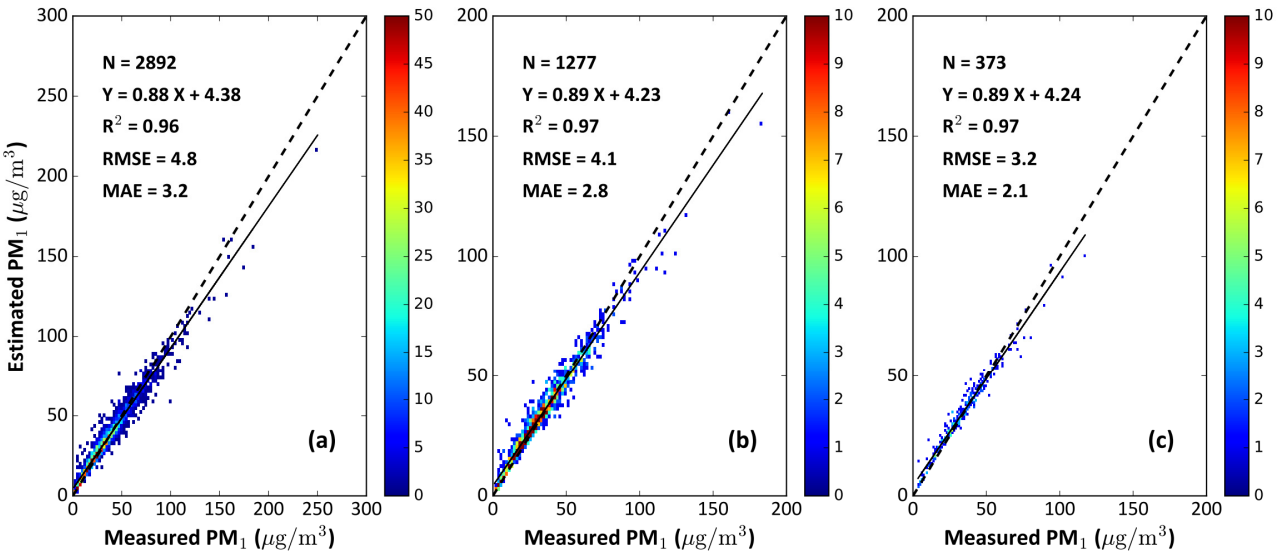

Figure S6. Validation of (a) monthly, (b) seasonal, and (c) annual mean PM_1 estimates from the STET model during 2014–2018 in China. Statistical metrics are given in each panel: the number of samples (N), the coefficient of determination (R^2), the root-mean-square error (RMSE; $\mu g/m^3$), and the mean absolute error (MAE; $\mu g/m^3$). The linear regression relationship is also given in each panel. Dashed lines are the 1:1 lines, and solid lines are the linear best-fit lines through the data points.

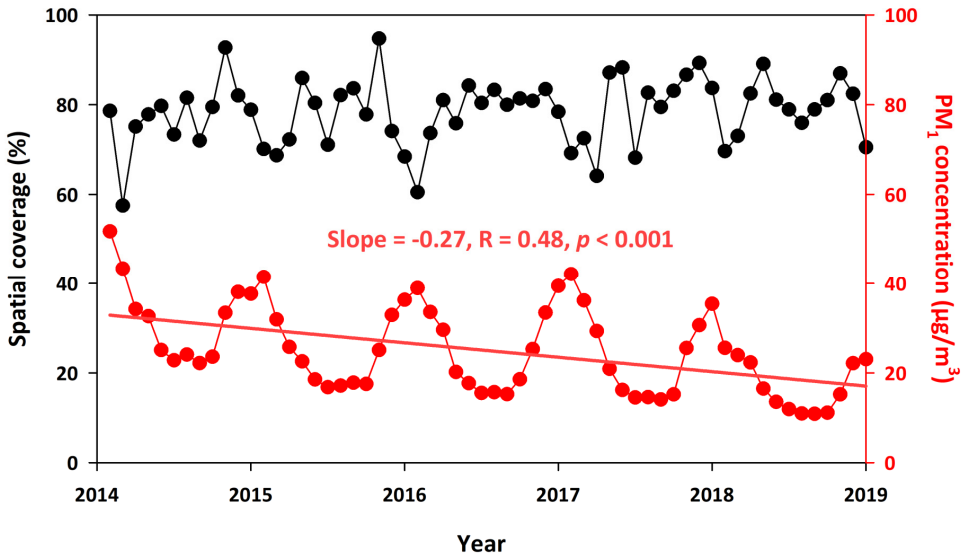

Figure S7. Time series of monthly spatial coverage (in black) and mean PM_1 concentration (in red) from 2014 to 2018 across China.

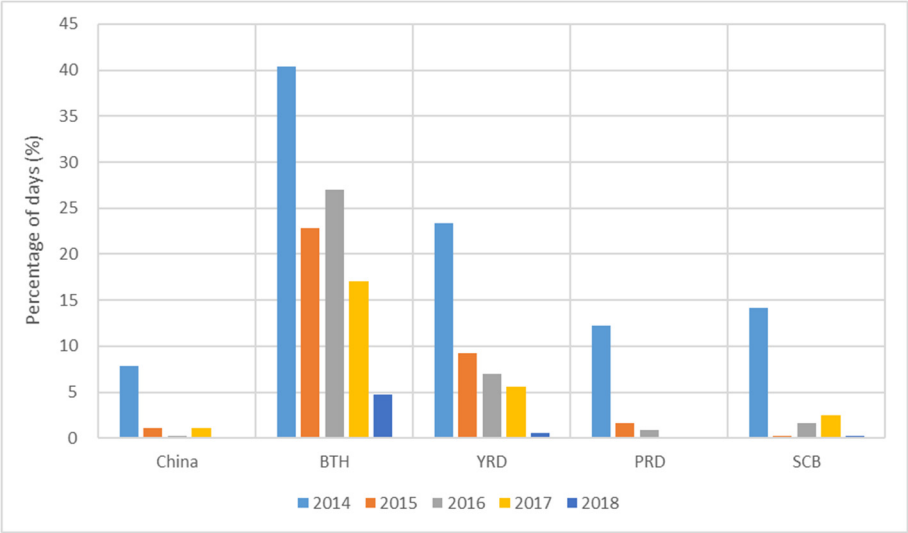

Figure S8. Percentage of high-PM_1 days, i.e., days when daily mean PM_1 concentrations exceed 50 μg/m^3, in each year from 2014 to 2018 across China, and in the Beijing-Tianjin-Hebei (BTH) region, the Yangtze River Delta (YRD), the Pearl River Delta (PRD), and the Sichuan Basin (SCB).