Supporting information --

Integrating augmented in-situ measurements and a spatiotemporal machine learning model to back extrapolate historical particulate matter pollution over the United Kingdom: 1980-2019

Riyang Liu[1,2,3], Zongwei Ma[1], Antonio Gasparrini[4], Arturo de la Cruz[4], Jun Bi[1], Kai Chen[2,3*]*

[1] State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing 210023, China

[2] Department of Environmental Health Sciences, Yale School of Public Health, New Haven, Connecticut 06520, United States

[3] Yale Center on Climate Change and Health, Yale School of Public Health, New Haven, Connecticut 06520, United States

[4] Environment & Health Modelling (EHM) Lab, Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, London WC1H 9SH, UK

*** Corresponding Authors**

Summary:

Number of pages:45
Number of figures:23
Number of tables: 11
Number of texts: 4

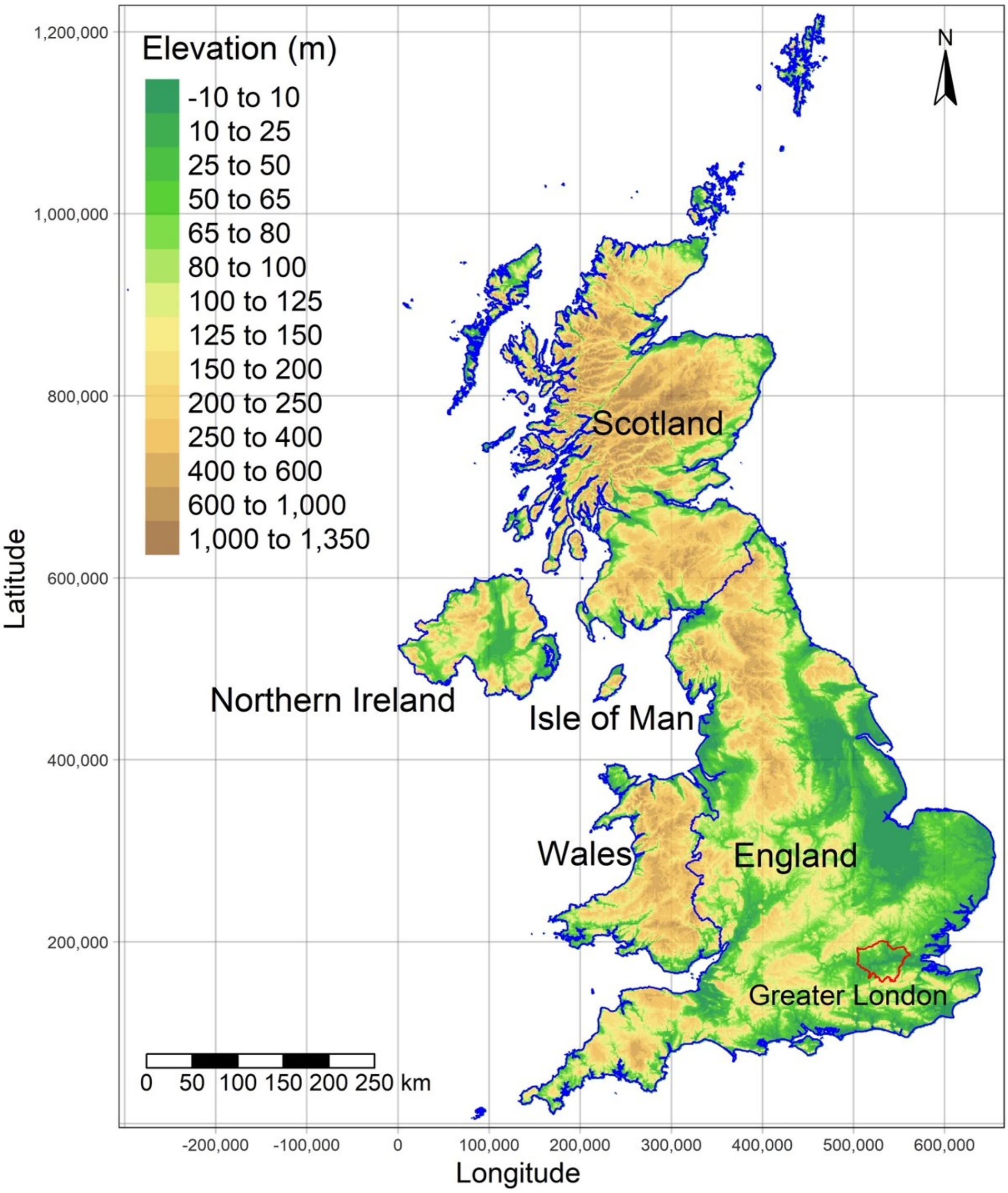


Figure S1. Subregions in the study area. The area within the red line is the Greater London area.

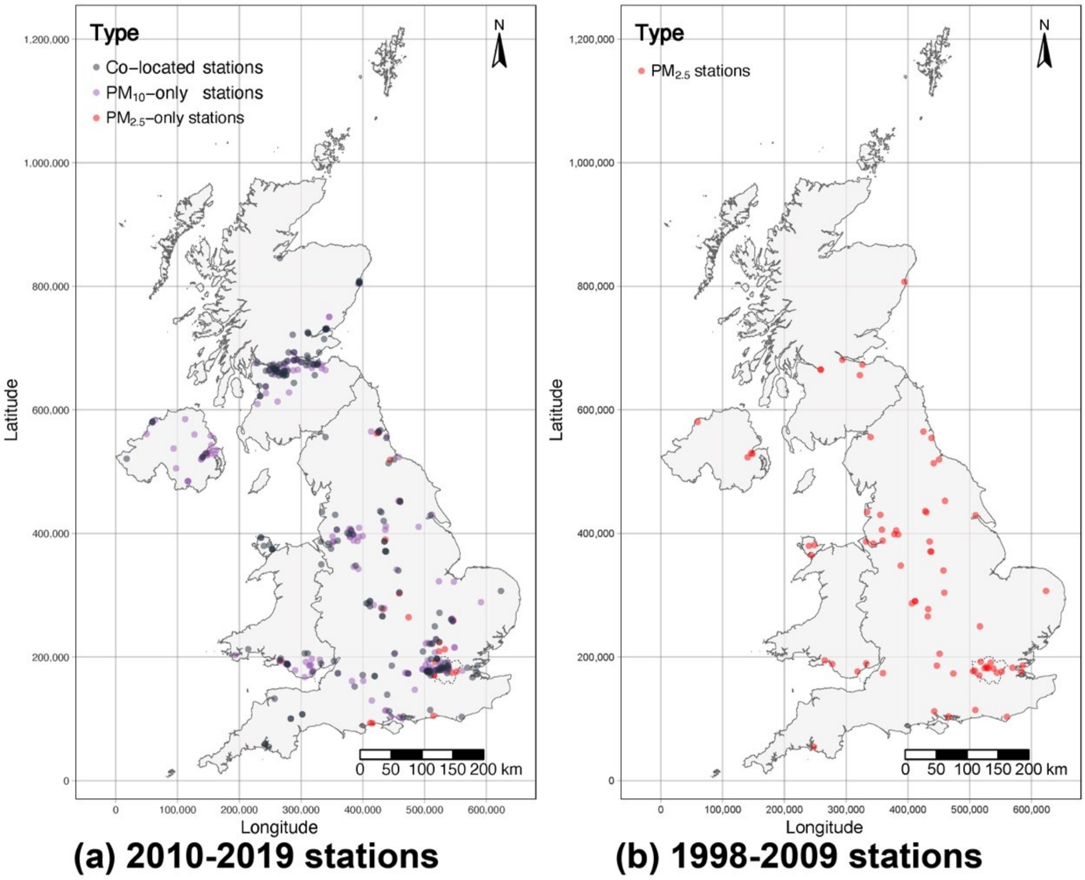


Figure S2. Spatial distribution of PM monitoring stations from national networks in the UK from (a) 2010 to 2019 and (b) 1998 to 2009. Note that some clustered stations are overlapped because of their proximity.

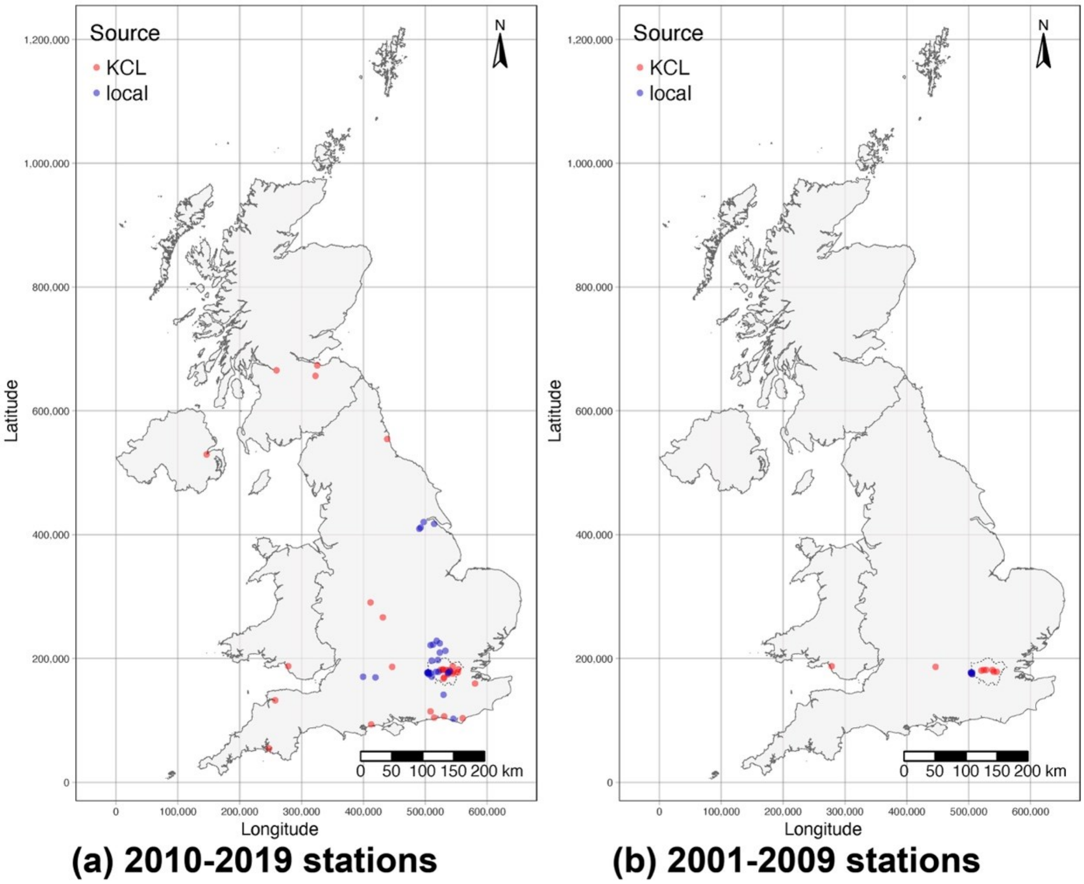


Figure S3. Spatial distribution of $PM_{2.5}$ monitoring stations from regional networks in the UK from 2010 to 2019 (a) and from 2001 to 2009 (b). Note that some clustered stations are overlapped because of their proximity.

Table S1. Summary of datasets used in this study

Category	Variable name	Description	Original Spatial Resolution	Unit	Temporal Resolution	Period	Data source
Ground-level monitoring data	$PM_{2.5}$	$PM_{2.5}$	-(stations)	μg/m^3	Hourly	1998-2019	AURN, AQE, WAQN, SAQN, NI, KCL, local
	PM_{10}	PM_{10}		μg/m^3	Hourly	2010-2019	
Meteorological factors	blh	Boundary layer height	0.25°×0.25°	m	Hourly	1980-2019	ERA5[1]
	lcc	Low cloud cover		(0-1)			
	tcc	Total cloud cover		(0-1)			
	v10	10m v-component of wind	0.1°×0.1°	m/s^1	Hourly	1980-2019	ERA5-land[2]
	u10	10m u-component of wind		m/s^1			
	strd	Surface thermal radiation downwards		J/m^2			
	ssrd	Surface solar radiation downwards		J/m^2			
	sp	Surface pressure		Pa			
	d2m	2m dewpoint temperature		K			
	tasmax	Daily maximum temperature	1 km	℃	Daily	1980-2019	HadUK-Grid[3, 4]
	tasmin	Daily minimum temperature		℃			
	rainfall	Precipitation		mm			

Category	Variable name	Description	Original Spatial Resolution	Unit	Temporal Resolution	Period	Data source
Aerosol reanalysis	BCSMASS	Black Carbon Surface Mass Concentration	0.5°× 0.625°	kg/m^3	Hourly	1980-2019	MERRA-2[5]
	OCSMASS	Organic Carbon Surface Mass Concentration		kg/m^3			
	SO4SMASS	SO4 Surface Mass Concentration		kg/m^3			
	DUSMASS25	Dust Surface Mass Concentration - $PM_{2.5}$		kg/m^3			
	SSSMASS25	Sea Salt Surface Mass Concentration - $PM_{2.5}$		kg/m^3			
Emission inventory	BC	Black carbon emission	0.1°×0.1°	$kg/m^2/s$	Daily	1980-2019	CEDS[6, 7]
	OC	Organic carbon emission		$kg/m^2/s$			
	SO2	SO_2 emission		$kg/m^2/s$			
	NOx	Nitrogen oxides emission		$kg/m^2/s$			
	NMVOC	Non-methane volatile organic compounds emission		$kg/m^2/s$			
	NH3	NH_3 emission		$kg/m^2/s$			

Category	Variable name	Description	Original Spatial Resolution	Unit	Temporal Resolution	Period	Data source
Land-cover	Settlement	The area proportion of settlement in each grid cell	300 m	/	Annual	1992-2019	Land cover classification gridded maps[8]
	wetland	The area proportion of wetland in each grid cell					
	grassland	The area proportion of grassland in each grid cell					
	forest	The area proportion of forest in each grid cell					
	agricultural	The area proportion of agricultural in each grid cell					
Road network	Tertiary_density	The length of tertiary road in each grid cell	-(vector)	/	The latest	The latest	OpenStreetMap[9]
	secondary_density	The length of secondary road in each grid cell					
	primary_density	The length of primary road in each grid cell					
	trunk_density	The length of trunk road in each grid cell					
	motorway_density	The length of motorway in each grid cell					

Category	Variable name	Description	Original Spatial Resolution	Unit	Temporal Resolution	Period	Data source
	Road_density	The length of all 5 types of road in each grid cell					
Terrain data	Altitude	DEM	1 arc second	m	-	2000	NASADEM[10]
	slope	Slope derived from merged height	1 arc second	degree	-		NASADEM[11]
Anthropogenic activities	pop	The number of people per cell	1 km	/	Every 5 years	1980-2020	GHSL[12]
	SMOD	The Degree of Urbanization		/			GHSL[13]
	Nighttime_light	Nighttime light	30 arc second	/	Annual	1992-2019	Harmonization of DMSP and VIIRS nighttime light data, version 5[14]

Notes. AURN: Automatic Urban and Rural Network; AQE: Air Quality England network; WAQN: Air Quality Wales network; SAQN: Air Quality Scotland network; NI: Northern Ireland network; KCL: King's College London network; local: locally managed AQ networks in England; link: https://uk-air.defra.gov.uk/data/ (accessed 2022-02-20). ERA5: the fifth generation of European ReAnalysis, link: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form (accessed 2022-04-27). ERA5-Land: the land component of ERA5, link: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=form (accessed 2022-04-27). HadUK-Grid link: https://data.ceda.ac.uk/badc/ukmo-hadobs/data/insitu/MOHC/HadOBS/HadUK-Grid/v1.0.3.0/1km (accessed 2022-06-08). MERRA-2: The Modern-Era Retrospective Analysis for Research and Applications, Version 2. CEDS: the Community Emissions Data System, link: https://data.pnnl.gov/dataset/CEDS-4-21-21 (accessed 2022-08-09). OpenStreetMap link: http://download.geofabrik.de/europe/britain-and-ireland.html (accessed 2022-03-11). DEM: Digital Elevation Model. GHSL: Global Human Settlement Layer; pop: population. SMOD: Settlement Model layers. DMSP: Defense Meteorological Satellite Program; VIIRS: Visible/Infrared Imager/Radiometer Suite. The links to the MERRA-2 data, DEM, slope, pop, SMOD and nighttime light data can be found in the reference list of the SI.

Text S1 Data sources and preparations of auxiliary predictors

Meteorological variables that played important roles in models in previous studies[15, 16] were obtained from three climate reanalysis data sources: the fifth generation of European ReAnalysis (ERA5), the land component of ERA5 (ERA5-Land) and HadUK-Grid. The ERA5[1] and ERA5-Land[2] datasets produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) provide spatiotemporal-resolved data on a wide range of meteorological variables. HadUK-Grid is a series of datasets for daily meteorological variables at 1 km × 1 km horizontal resolution across the British Isles derived from interpolation of in-situ observations[3, 4]. Hourly aerosol diagnostics data of 5 types of $PM_{2.5}$ composition were obtained from tavg1_2d_aer_Nx dataset (M2T1NXAER) [17] in the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). MERRA-2 reanalysis data were assimilated from multiple sources like model simulations, ground measurements, and satellite observations[5, 18]. Monthly anthropogenic source emission data were obtained from the Community Emissions Data System (CEDS) [6, 7] from the Pacific Northwest National Laboratory (PNNL). Pollutants selected included ammonia (NH_3), nitrogen oxides (NO_x), SO_2, non-methane volatile organic compounds (NMVOC), and components of PM: black carbon (BC) and organic carbon (OC). Hourly predictors from ERA5, ERA5-Land, MERRA-2, and CEDS were aggregated to daily average values and then interpolated to the grid cells. Specifically, the bilinear interpolation algorithm, which has been widely used in previous studies[19-21], was used for the ERA5, MERRA-2 and CEDS data. Since the spatial coverage of the ERA5-Land grid cells were slightly smaller than our modeling grids, we used another widely used algorithm[22, 23], the inverse distance weighting interpolation for ERA5-Land data.

Land cover classification gridded maps[8] were obtained from the Copernicus Climate Change Service (C3S). Version 2.0.7 provides the maps from 1992 to 2015, while version 2.1.1 provides data from 2016 to 2019. The 6 types of Intergovernmental Panel on Climate Change (IPCC) classes considered for the change detection were used to aggregate the original land cover classification system[24]. The area proportion of each class was calculated in each 1-km grid cell. We used the land cover data in 1992 for pre-1992 years. Road network data were downloaded from OpenStreetMap, whose information was collected by participants[9, 25]. The length of different types of roads in each grid cell was calculated. All the years in this study used the same road density data due to data availability. Although road networks in the UK could have changed over time, we used the data in 2022 to represent the overall spatial patterns of roads. Terrain including elevation[10] and slope[11] was downloaded from NASADEM and then aggregated respectively to averages in each 1 km grid cell. Gridded population and the degree of urbanization data were downloaded from the Global Human Settlement Layer (GHSL) in a 5-year time interval from 1980 to 2020 and then resampled to the modeling grid cells. The data of years without GHSL data were obtained by linear interpolation using data from the adjacent 5-year time interval. Stable nighttime light (NTL) data version 5[14] were obtained from a previous study[26] and then aggregated to averages of every 1 km grid cells. Some of the data sources used in this study went back as far as 1980, which led to the decision to limit the time span of this study.

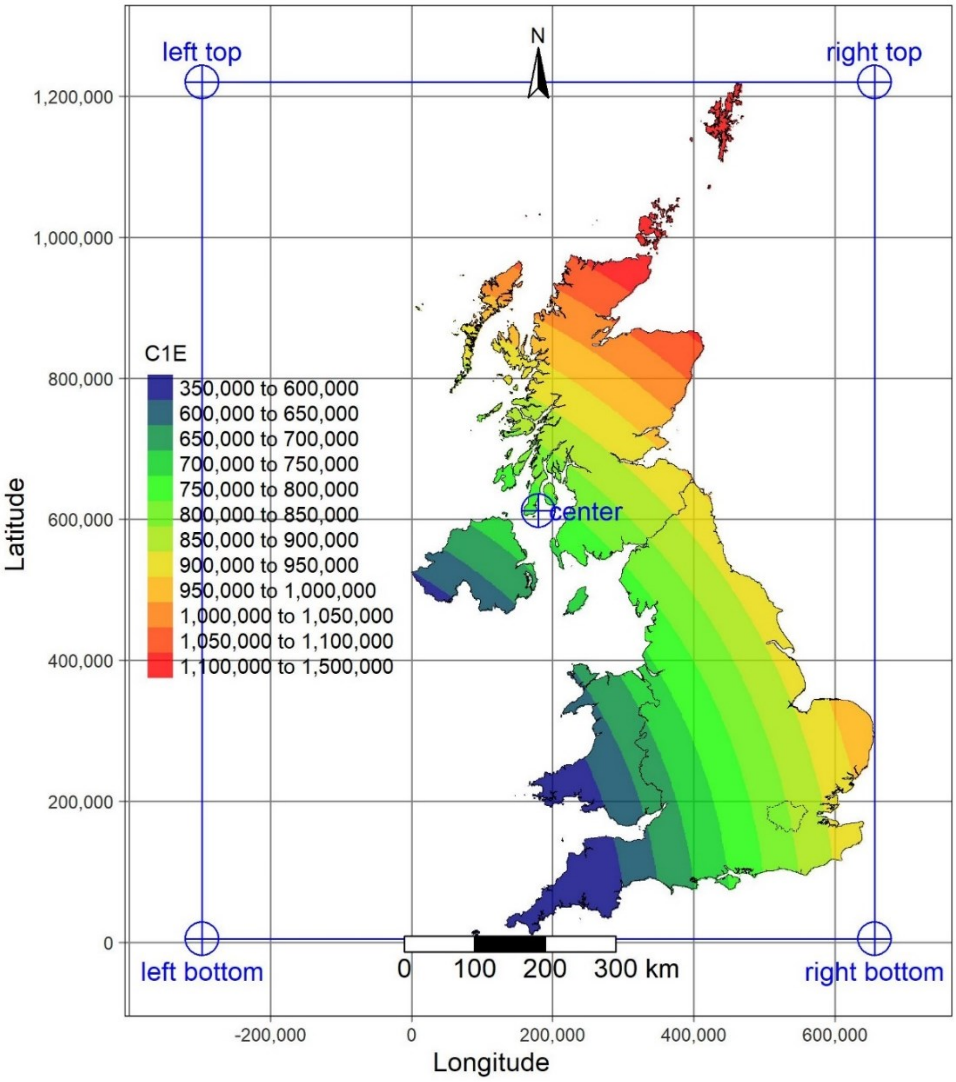


Figure S4. The spatial variation of the Euclidean distance from each grid to the left bottom corner of the study area (C1E). The blue rectangle is the rectangle around our study area. The points are the corners and the center of the rectangle.

Text S2 The formulas of spatiotemporal weights

The formula of spatial weights is shown as follows:

$$C_jE_i = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad \text{(Equation S1)}$$

Where C_jE_i represents the Euclidean distance from a grid cell i to a corner or the center j in the study region, x_i, y_i represents the longitude and latitude of the grid cell i, x_j, y_j represents the longitude and latitude of the corner or the center j.

The formula of temporal weights is shown as follows:

$$IDt_{mn} = \frac{1}{|DOY_m - DOY_n| + 1} \quad \text{(Equation S2)}$$

Where IDt_{mn} represents the inverse time interval from a day n to the middle day of a season n, DOYrepresents the order of a day in a year. To avoid 0 in the denominator, add 1 to the absolute value of the difference between the two days.

Table S2. List of spatiotemporal weights

Variable name	**Description**
C1E	The Euclidean distance from.a grid to the left bottom corner of the study area
C2E	The Euclidean distance from a grid to the left top corner of the study area
C3E	The Euclidean distance from a grid to the right bottom corner of the study area
C4E	The Euclidean distance from a grid to the right top corner of the study area
CCE	The Euclidean distance from a grid to the center of the study area
dow	The order of a day in a week
IDt1	The inverse time interval from a day to the spring equinox (21 March)
IDt2	The inverse time interval from a day to the summer solstice (21 June)
IDt3	The inverse time interval from a day to the autumn equinox (22 September)
IDt4	The inverse time interval from a day to the winter solstice (22 December)

Text S3 The LightGBM Algorithm

LightGBM is a novel implementation of the gradient boosting decision trees (GBDT) algorithm. LightGBM has three main optimization features to reduce complexity in finding the best split points in decision trees, as is shown in the right bottom panel of Figure 1. The histogram-based algorithm, which transforms continuous numeric features into discrete bins, is used to reduce the potential split points. Gradient-based one-side sampling (GOSS) is used to reduce the sample size without changing the data distribution by much. Exclusive feature bundling (EFB) is used to reduce the number of features without hurting the accuracy[27]. Therefore, LightGBM has strength in faster computation speed, lower memory consumption, and capability of handling big data when compared with other advanced algorithms like extreme gradient boosting (XGBoost) [27] and has been used in previous studies[28, 29]. GOSS was not used in this study due to our moderate sample size.

Text S4 Supplementary details about the stage 1 and stage 2 models

For the stage 1 model, since the model development and prediction were in grids where monitors were available, we created a few predictors in addition to those introduced in section 2.1.3. One-hot encoding was used to transform monitor types into new predictors. Specifically, monitor types were defined according to whether they were in rural or suburban areas, which nation they were in (England/Scotland), and whether they were near emission sources (background/industrial/traffic).

Since the stage 2 model needs to predict at locations where monitors were not available, predictors derived from monitor types were excluded.

Table S3. Hyperparameters used in this study

Name	**Long name**	**Values**	**The value selected in stage 1**	**The Value Selected in stage 2**
learning_rate	shrinkage rate	0.05, 0.1	0.05	0.1
num_leaves	maximum number of leaves in one tree	31, 63, 127, 255, 511, 1023, 2047, 4095	4095	1023
max_depth	maximum depth for the tree model	4-12	12	12
min_data_in_leaf	minimal number of data in one leaf	10,20	20	10
bagging_fraction	the ratio of the randomly selected subset of data without resampling	0.6-1	0.85	0.90
bagging_freq	frequency for bagging	3-5	4	3
feature_fraction	the ratio of the randomly selected subset of features on each iteration	0.5-1	0.94	0.57
lambda_l1	L1 regularization	0.5,1	0.5	0.5
lambda_l2	L2 regularization	0.5,1	0.5	0.5
max_bin	Maximum number of discrete bins per feature	63,255,511	511	63

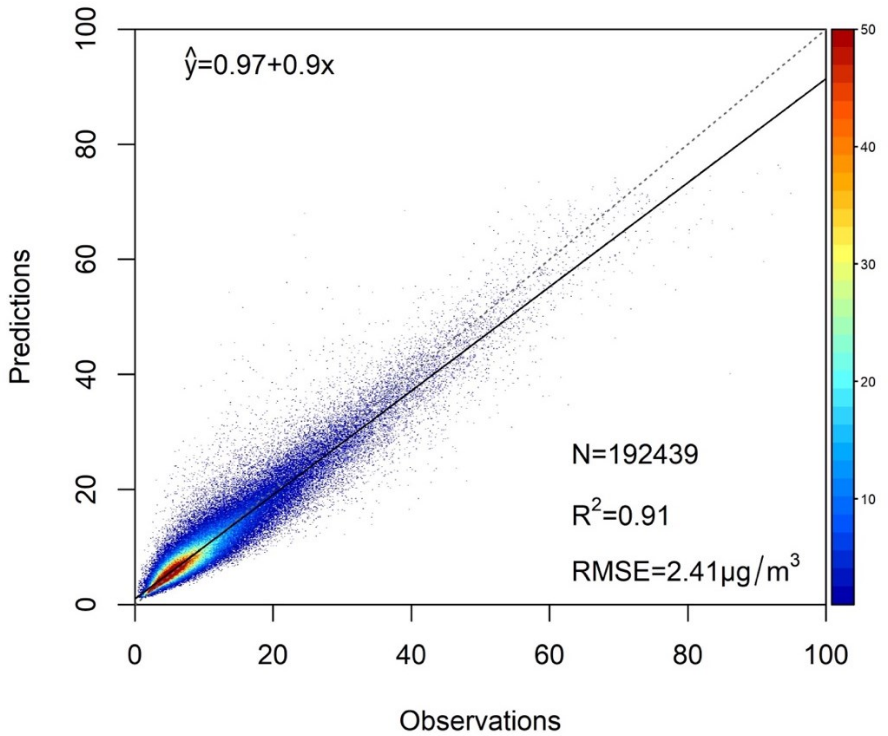


Figure S5. Density scatterplots of the 10-fold grid-based CV results for the stage 1 model

Table S4. The CV results of the stage 1 model from 2010 to 2019 at the daily level

Year	Sample size	R^2	RMSE	MAE
2010	10053	0.88	3.05	2.11
2011	12606	0.92	3.05	2.20
2012	13369	0.91	2.81	2.01
2013	13174	0.90	2.77	2.00
2014	15032	0.91	2.73	1.92
2015	16927	0.89	2.43	1.57
2016	19540	0.88	2.62	1.67
2017	24435	0.91	2.05	1.35
2018	29125	0.90	2.01	1.35
2019	38178	0.93	1.88	1.18

Note. The unit for RMSE and MAE is μg/m^3.

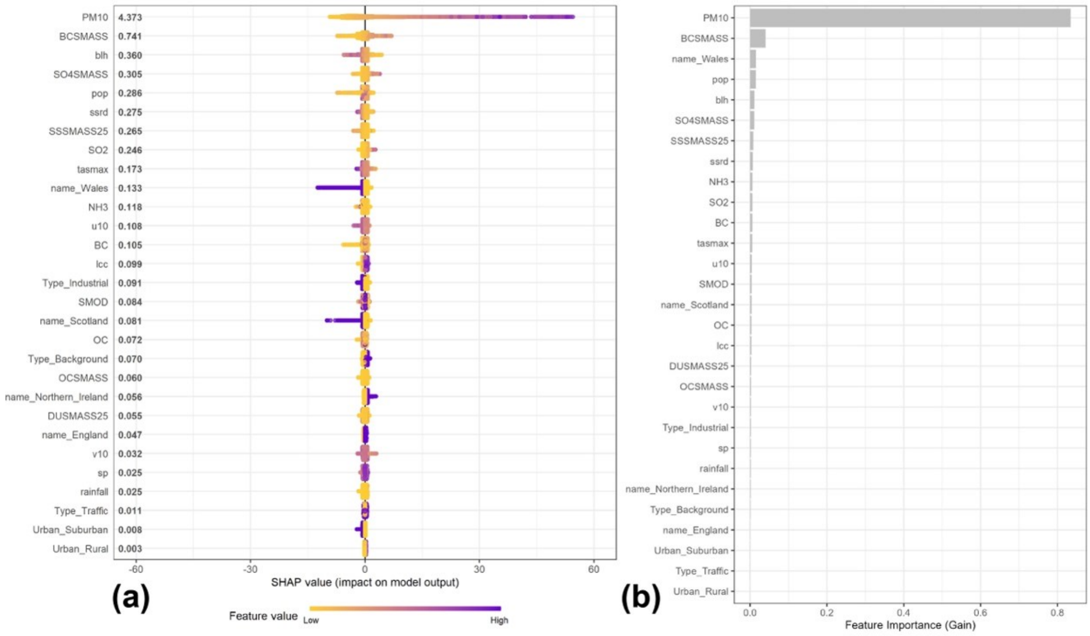


Figure S6. The interpretation of the stage 1 model with SHAP summary plot for $PM_{2.5}$ predictions in the development set (a) and feature importance of the predictors in relative percentage (b). The numbers next to the vertical axis of (a) represent mean absolute SHAP value by predictor variable. In (a), each dot in each row represents a data sample, the x position of each dot is the effect of a predictor variable on a model's prediction, and the color of the dot represents the value of that predictor variable. Dots that don't fit on the row are stacked to show density. Thirty-six predictions with PM_{10}>100 μg/m^3 were removed for better visualization in (a).

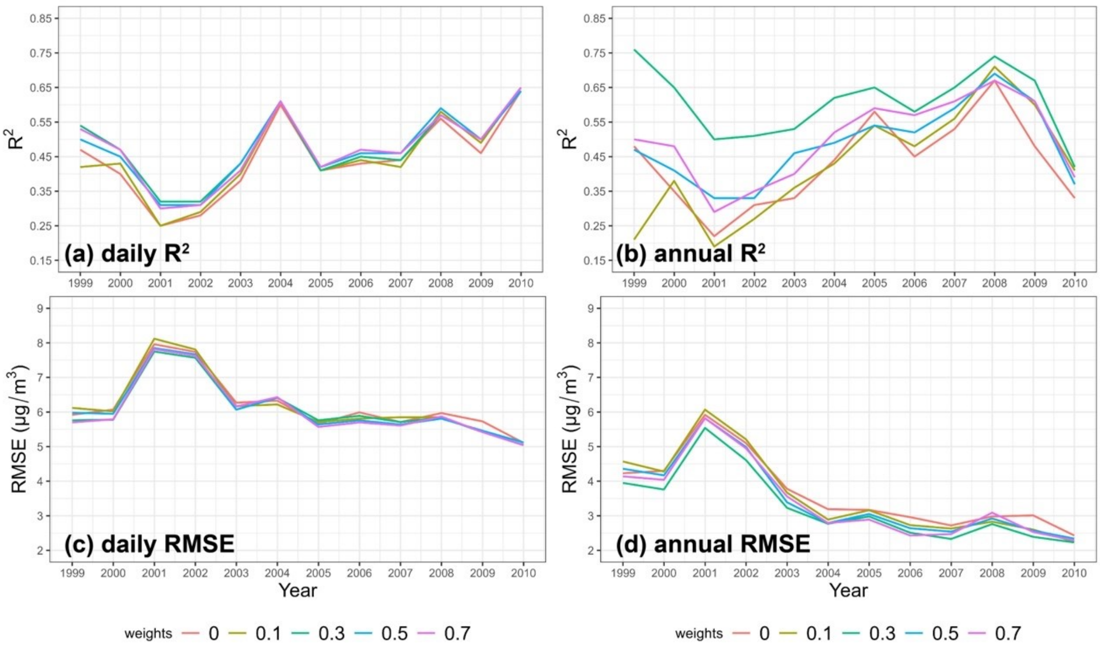


Figure S7. The comparison of stage 2 model testing results based on different weights in terms of R^2 (a) and RMSE (c) values at the daily level and R^2 (b) and RMSE (d) values at the annual level. The values for different years were linked by lines for better visual display.

Table S5. The by-year CV results of the stage 2 model from 2010 to 2019

Year	Daily				Monthly average				Annual average			
	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE
2010	17660	0.63	5.30	3.35	641	0.73	2.77	2.00	55	0.78	1.85	1.34
2011	18061	0.78	5.36	3.32	650	0.89	2.77	2.03	56	0.84	1.92	1.51
2012	20158	0.77	4.53	2.99	735	0.85	2.13	1.62	63	0.82	1.38	1.06
2013	19868	0.73	4.67	3.06	718	0.81	1.93	1.53	62	0.90	1.00	0.80
2014	21021	0.71	4.93	3.08	778	0.80	2.48	1.76	67	0.84	1.39	1.08
2015	24931	0.76	3.73	2.34	909	0.79	1.76	1.26	78	0.79	1.06	0.79
2016	27818	0.73	3.86	2.57	999	0.81	1.75	1.33	85	0.86	1.03	0.82
2017	33100	0.70	4.07	2.55	1163	0.84	1.70	1.30	99	0.83	1.13	0.90
2018	40737	0.67	3.84	2.59	1429	0.79	1.78	1.40	121	0.87	1.22	1.02
2019	48862	0.72	4.13	2.90	1697	0.83	2.37	1.95	145	0.79	1.83	1.53

Note. The unit for RMSE and MAE is μg/m^3.

Table S6. The testing results of the stage 2 model from 1998 to 2009 at the daily, monthly, and annual levels

Year	Daily				Monthly average				Annual average			
	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE
1998	793	0.55	5.66	4.19	28	0.57	4.22	3.31	3	0.76	3.95	3.08
1999	1331	0.48	5.71	3.88	47	0.55	4.12	2.65	4	0.65	3.76	2.29
2000	1379	0.32	7.73	5.06	48	0.31	5.96	3.53	4	0.50	5.54	3.13
2001	1383	0.32	7.60	4.89	48	0.30	5.53	3.43	4	0.51	4.61	3.27
2002	1379	0.43	6.09	4.45	47	0.43	3.75	2.95	4	0.53	3.23	2.77
2003	1406	0.60	6.50	4.60	48	0.57	3.98	3.27	4	0.62	2.77	2.34
2004	1633	0.44	5.64	4.23	58	0.41	3.83	3.07	5	0.65	2.98	2.75
2005	2017	0.46	5.82	4.05	71	0.51	3.27	2.63	6	0.58	2.51	2.20
2006	2072	0.45	5.71	4.07	72	0.51	3.14	2.39	6	0.65	2.33	1.97
2007	2260	0.58	5.82	4.02	81	0.66	3.59	2.67	7	0.74	2.76	2.14
2008	3410	0.50	5.43	3.69	127	0.62	3.18	2.35	11	0.67	2.39	2.08
2009	15751	0.65	5.05	3.45	577	0.65	2.95	2.21	50	0.42	2.23	1.78

Note. The unit for RMSE and MAE is μg/m^3.

Table S7. The testing results of the stage 2 model from 1998 to 2009 at the daily, monthly, and annual levels using the 100 km grid-based CV strategy

Year	Daily				Monthly average				Annual average			
	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE
1998	793	0.48	6.12	4.54	28	0.49	4.59	3.51	3	0.71	4.26	3.28
1999	1331	0.43	6.00	4.06	47	0.46	4.42	2.99	4	0.55	4.09	2.74
2000	1379	0.30	7.82	5.14	48	0.29	6.01	3.73	4	0.46	5.68	3.34
2001	1383	0.34	7.46	4.78	48	0.29	5.63	3.55	4	0.50	4.78	3.29
2002	1379	0.41	6.08	4.32	47	0.34	4.04	3.02	4	0.50	3.43	2.81
2003	1406	0.60	6.04	4.32	48	0.51	3.79	3.13	4	0.58	2.54	2.35
2004	1633	0.41	5.63	4.06	58	0.41	3.70	2.93	5	0.67	2.88	2.64
2005	2017	0.44	5.83	4.10	71	0.50	3.20	2.54	6	0.54	2.52	2.33
2006	2072	0.40	5.93	4.22	72	0.48	3.20	2.56	6	0.61	2.48	2.14
2007	2260	0.55	6.07	4.14	81	0.62	3.82	2.89	7	0.70	2.97	2.33
2008	3410	0.44	5.86	4.02	127	0.53	3.57	2.80	11	0.53	2.82	2.42
2009	15751	0.62	5.29	3.68	577	0.60	3.16	2.45	50	0.33	2.48	2.06

Note. The unit for RMSE and MAE is $\mu g/m^3$.

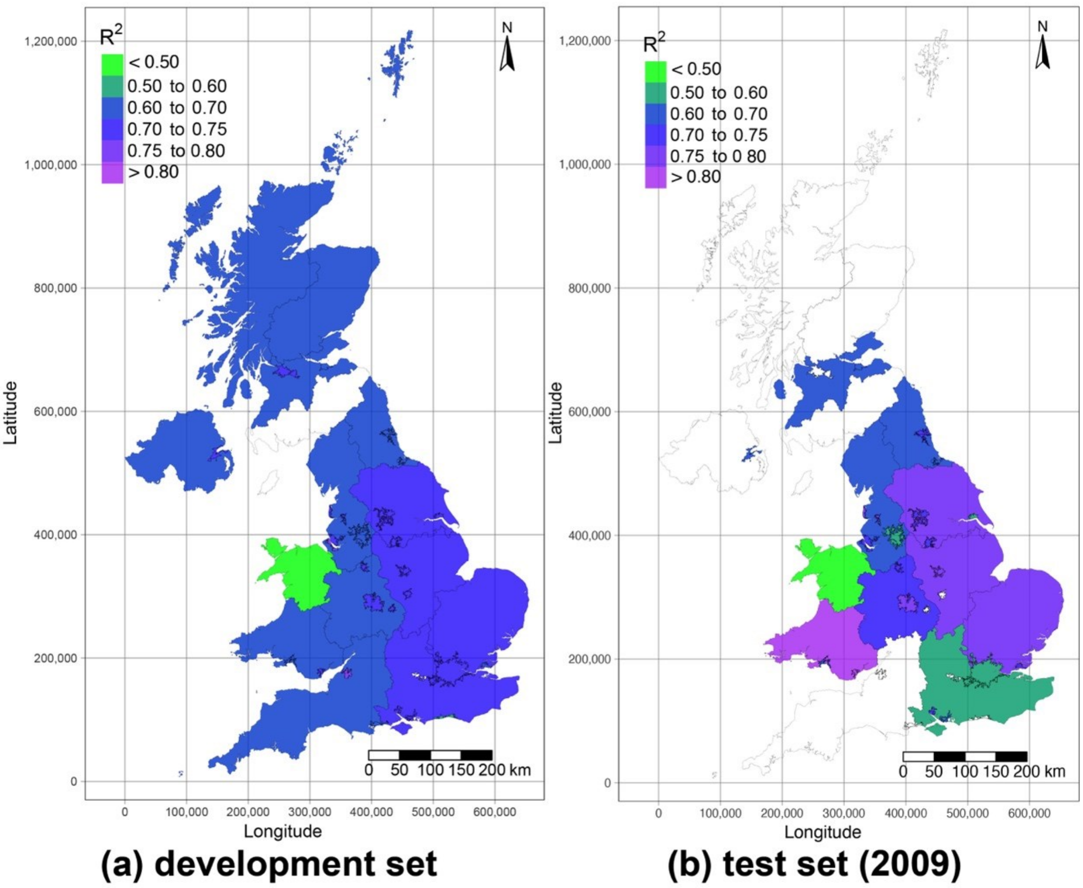


Figure S8 Spatial variances in the stage 2 model performance in different air quality zones and agglomerations. This figure visualizes the R^2 values between observed and estimated $PM_{2.5}$ concentrations in the development set from 2010 to 2019 (a) and the testing set in 2009 (b).

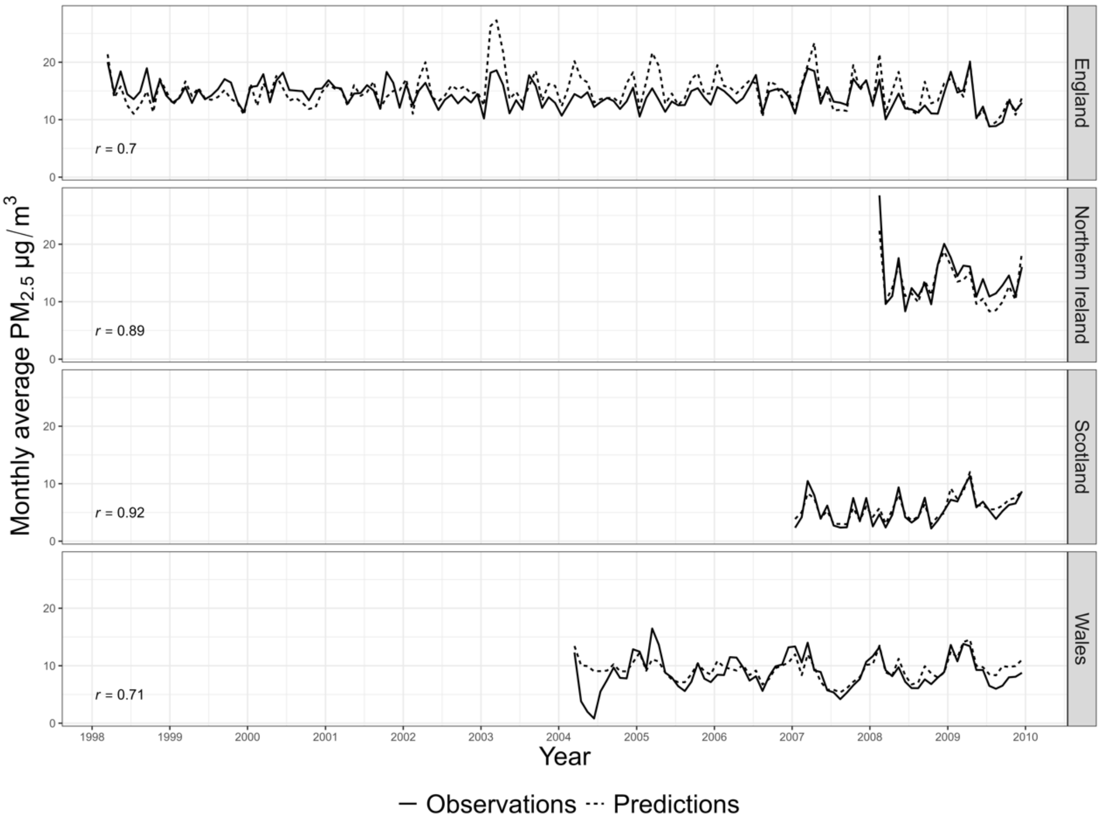


Figure S9. Time series in estimated (dashed) and observed (solid) monthly mean $PM_{2.5}$ concentrations in 4 subregions from 1998 to 2009. The correlation coefficients (r) between the observations and the predictions are shown at the bottom left of each facet.

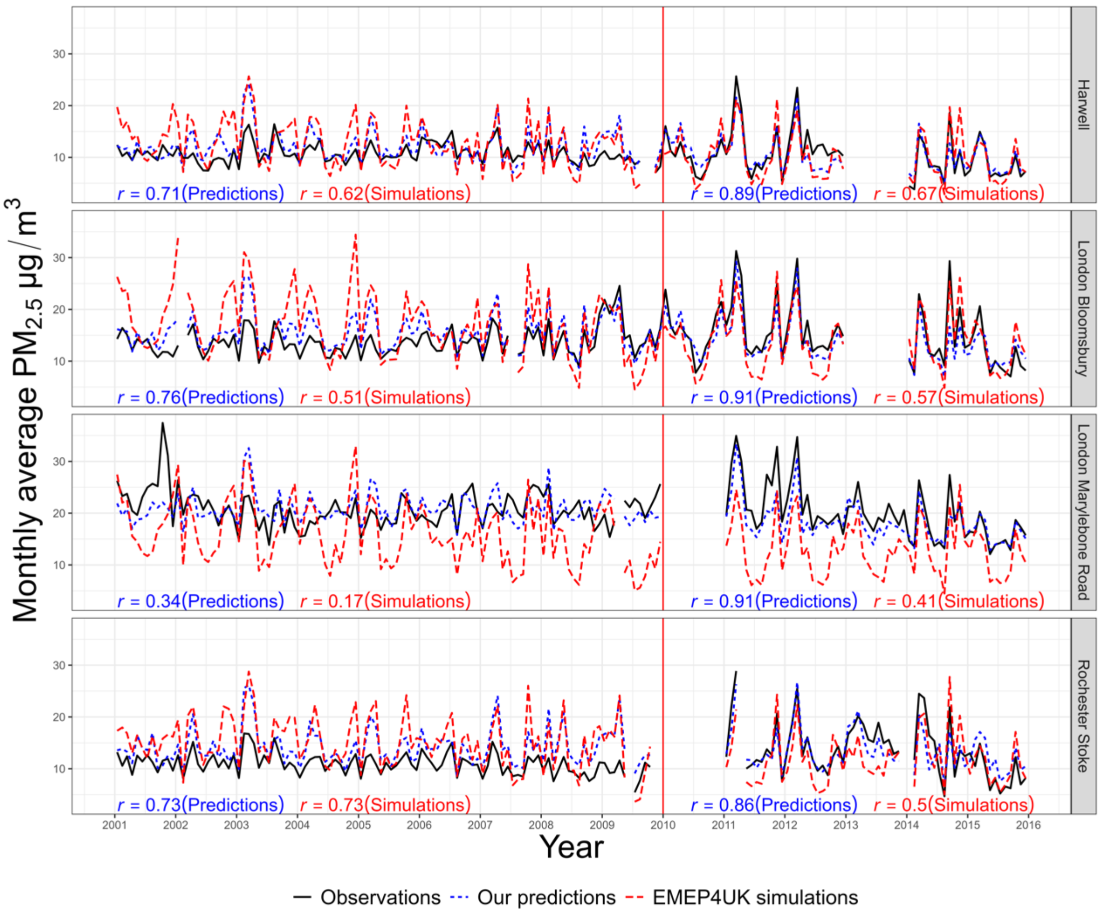


Figure S10. Time series in observed (solid black), our model estimated (dashed blue), and EMEP4UK-simulated (longdash red) monthly mean $PM_{2.5}$ concentrations from 2001 to 2019. The red vertical solid line is used to split the modeling years (after 2010) and the back extrapolation years (before 2010). The correlation coefficients (r) with the notation "(Predictions)" in blue shown at the bottom of each facet were calculated between the observations and our model predictions, while the correlation coefficients with the notation "(Simulations)" in red were calculated between the observations and the EMEP4UK simulations.

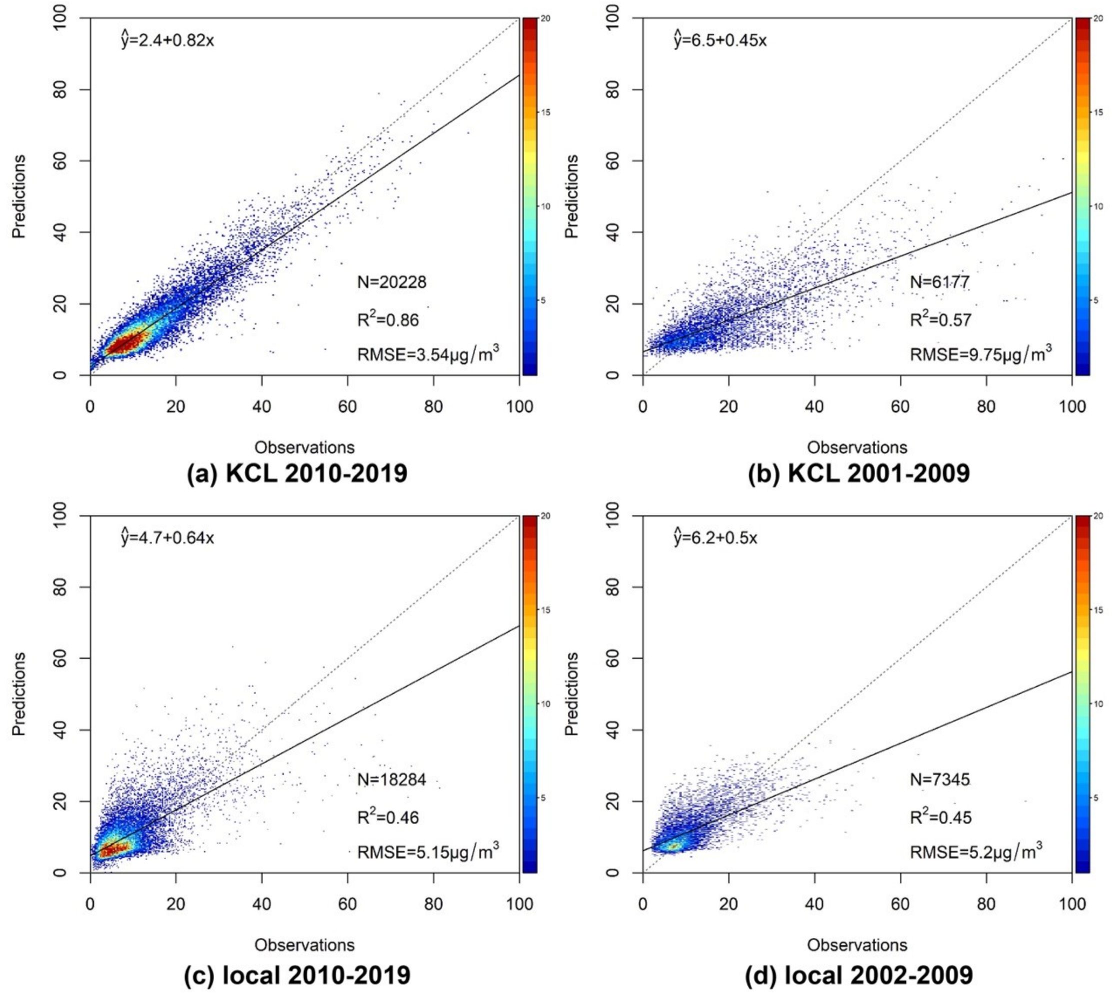


Figure S11. Density scatterplots of the testing results based on KCL and local networks for the stage 1 model (2010-2019) and the stage 2 model (before 2010)

Table S8. The testing results based on KCL and local networks for the stage 1 model (2010-2019) and the stage 2 model (2001-2009)

Year	KCL				local			
	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE
2001	277	0.58	7.58	4.56	-	-	-	-
2002	509	0.63	9.22	6.79	476	0.65	3.51	2.59
2003	845	0.65	11.46	8.08	708	0.66	4.97	3.30
2004	581	0.53	11.51	8.00	698	0.63	3.48	2.50
2005	538	0.62	12.38	9.05	1269	0.43	6.52	3.30
2006	940	0.42	13.22	9.57	982	0.42	5.09	3.75
2007	855	0.77	6.36	4.40	649	0.60	4.62	3.24
2008	1632	0.66	5.98	4.56	1269	0.31	5.50	3.70
2009	277	0.58	7.58	4.56	1294	0.32	5.21	3.73
2010	1918	0.84	3.59	2.62	1836	0.41	4.47	3.49
2011	1798	0.92	3.50	2.49	1601	0.53	6.36	3.64
2012	1890	0.80	5.15	3.54	1626	0.66	4.26	2.54
2013	1470	0.87	3.80	2.87	1664	0.52	4.88	3.58
2014	1293	0.90	3.45	2.68	1565	0.46	5.23	4.12
2015	1894	0.86	2.96	2.35	1905	0.29	5.35	3.66
2016	2444	0.86	3.51	2.44	1687	0.60	4.33	3.08
2017	2476	0.88	2.99	2.29	2223	0.43	5.35	3.77
2018	2949	0.81	3.09	2.27	2520	0.44	5.00	3.73
2019	2096	0.84	3.31	2.46	1657	0.42	5.94	4.41

Note. The unit for RMSE and MAE is $\mu g/m^3$.

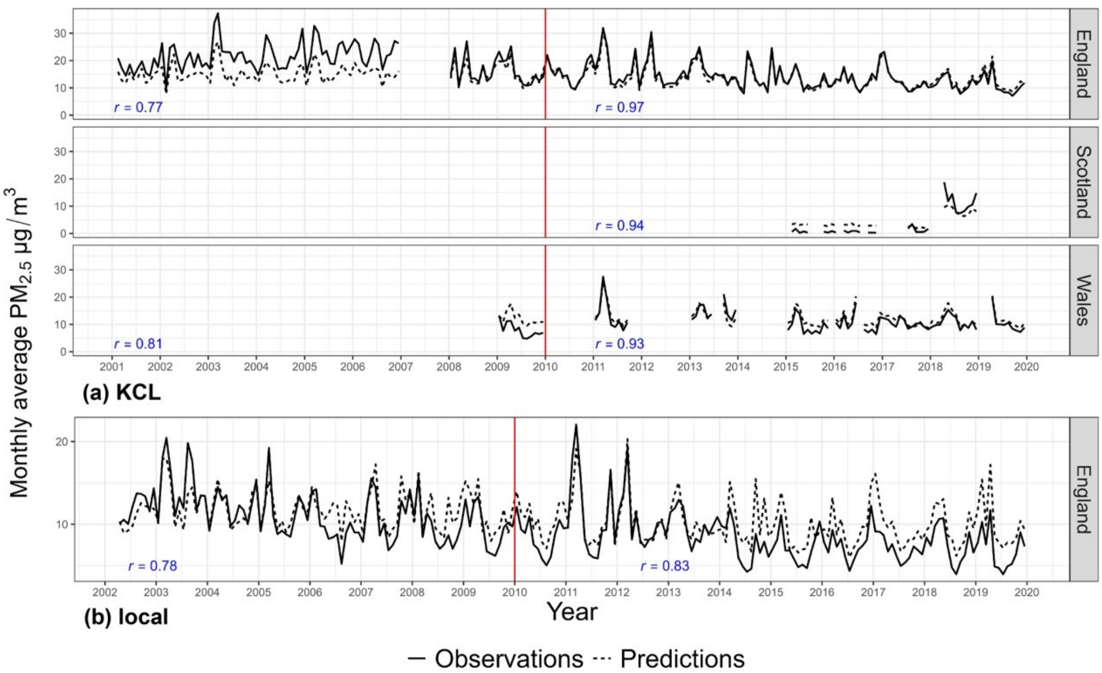


Figure S12. Time series in estimated (dashed) and observed (solid) monthly mean $PM_{2.5}$ concentrations from 2001 to 2019 based on observations from KCL (a) and local networks (b). The red vertical solid line is used to split the modeling years (after 2010) and the back extrapolation years (before 2010). The correlation coefficients (r) between the observations and the predictions over the 2 periods are shown at the bottom of each facet.

Table S9. Comparisons with observations measured before 2000 from previous literature

Location	Description	Period	Metric	Obs.	Est.	Refs
Haverah Park in Leeds (a moorland at Haverah Park, far from urbanized areas)	Instrument: Sierra Model 245 automatic dichotomous samplers; Frequency: daily, 24-h average (midday to midday)	1982/06/13-1982/09-28	period average	17.2	8.3	30
		1982/08/02	daily average	61.5	15.9	
		1982/07/31-1982/08/06	period average	46.2	11.6	
		1982/09/17	daily average	76.6	33.6	
		1982/09/16-1982/09/19	period average	68.6	28.0	
Leeds University in Leeds (A roof-top site in Leeds University, 2 km north of the city center)		1982/06/13-1982/09-28	period average	22.2	13.7	
		1982/08/02	daily average	59.1	21.1	
		1982/07/31-1982/08/06	period average	49.7	18.5	
		1982/09/17	daily average	142.1	44.4	
		1982/09/16-1982/09/19	period average	107.2	37.3	
Hodge Hill in Birmingham (70 meters south of an elevated section of the M6 motorway)	Instrument: The Ruprecht and Patashnick TEOM Frequency: hourly	Jan 1995	monthly average	11.0	10.1	31
		Feb 1995		11.0	9.6	
		March 1995		12.0	11.5	
		April 1995		15.0	14.5	
		May 1995		15.0	14.4	
		June 1995		11.0	12.3	
		Jan 1995-June 1995	period min	3.0	5.4	
			period max	43.0	32.1	
			period average	13.0	12.1	
		1995/04/01- 1995/07/31	period average	13.1	13.5	32
		1994/10/01-1995/9/31	period average	15.7	13.4	

Location	Description	Period	Metric	Obs.	Est.	Refs
		1994/10/01-1996/12/31	period average	14.5	14.7	33
			period min	2.1	5.4	
			10^{th} percentile	6.0	7.3	
			median	11.7	12.3	
			90th percentile	25.8	25.6	
			period max	82.8	48.0	
London Marylebone Road (an urban kerbside/roadside site, around 1 m from the kerbside of a major arterial route)	Instrument: TEOM Frequency: hourly	1997/06/20-1997/08/31	period average	26.40	18.12	34
		1997/09/01-1997/11/30	period average	23.30	20.02	
		1997/12/01-1998/02/28	period average	22.80	20.15	
		1997 (Jun-Dec)	50^{th} percentile	21.40	17.52	
			90^{th} percentile	38.10	27.48	
			95^{th} percentile	44.60	31.38	
			98^{th} percentile	46.90	35.52	
			99^{th} percentile	50.70	38.37	
			99.90^{th} percentile	55.40	42.38	
		1997/08/05-1997/08/21	period average	36.00	25.29	
			period max	81.60	36.27	
		1997/10/29-1997/11/14	period average	41.80	25.41	
			period max	97.10	39.92	
London Bloomsbury (an urban background site, within the south east corner of a small park in central London)	Instrument: TEOM Frequency: hourly	1997/06/20-1997/08/31	period average	18.90	13.76	
		1997/09/01-1997/11/30	period average	19.30	15.71	
		1997/12/01-1998/02/28	period average	15.90	16.49	

Location	Description	Period	Metric	Obs.	Est.	Refs
		1997 (Jun-Dec)	50th percentile	15.50	12.57	
			90th percentile	29.50	24.29	
			95th percentile	36.90	27.54	
			98th percentile	40.50	31.51	
			99th percentile	53.20	34.67	
			99.90th percentile	70.70	36.52	
		1997/08/05-1997/08/21	period average	28.20	21.32	
			period max	60.20	32.92	
		1997/10/29-1997/11/14	period average	27.90	21.11	
			period max	155.90	36.29	
Rochester (a rural site, on the western boundary of a rural primary school on the outskirts of the village of Lower Stoke, Rochester, Kent)	Instrument: TEOM Frequency: hourly	1997/06/20-1997/08/31	period average	14.20	14.19	
		1997/09/01-1997/11/30	period average	13.20	13.24	
		1997/12/01-1998/02/28	period average	13.00	13.86	
		1997 (Jun-Dec)	50th percentile	11.20	11.80	
			90th percentile	25.30	20.22	
			95th percentile	29.10	24.10	
			98th percentile	34.30	29.68	
			99th percentile	36.20	32.20	
			99.90th percentile	53.30	34.97	
		1997/08/05-1997/08/21	period average	25.30	22.89	
			period max	57.30	32.28	
		1997/08/05-1997/08/21	period average	18.30	15.30	
			period max	101.70	23.63	

Location	Description	Period	Metric	Obs.	Est.	Refs
Harwell Science Centre, Didcot, Oxfordshire (in the middle of an unfarmed field and surrounded by predominantly agricultural land)	Instrument: TEOM Frequency: hourly	1997/09/28-1997/11/30	period average	13.40	13.61	
		1997/12/01-1998/02/28	period average	12.00	12.50	
		1997 (Jun-Dec)	50th percentile	9.70	10.75	
			90th percentile	22.50	19.97	
			95th percentile	26.40	23.90	
			98th percentile	29.40	28.42	
			99th percentile	35.50	29.81	
			99.90th percentile	37.80	31.28	
		1997/08/05-1997/08/21	period average	23.90	15.79	
			period max	51.70	29.83	
The Archway Road (a roadside site in North London)	Instrument: Partisol Starnet 2000 system Frequency: 0.5 hour	1998/06/29-1998/08/08	period average	16.00	10.38	35
			period max	26.00	15.23	
			period min	7.00	6.93	
			Weekdays mean	16.00	10.21	
			Weekdays max	26.00	13.65	
			Weekdays min	9.00	6.93	
			Weekends mean	15.00	10.85	
			Weekends max	24.00	15.23	
			Weekends min	7.00	7.23	
		1999/03/01-1999/03/28	period average	27.00	17.11	
			period max	74.00	35.25	
			period min	7.00	7.53	
			Weekdays mean	28.00	17.79	
			Weekdays max	74.00	35.25	

Location	Description	Period	Metric	Obs.	Est.	Refs
			Weekdays min	10.00	8.39	
			Weekends mean	19.00	15.18	
			Weekends max	38.00	22.89	
			Weekends min	10.00	7.53	
University Old College in Edinburgh (on the roof, an urban background site, in central Edinburgh)	Instrument: The Ruprecht and Patashnick Partisol 2025 samplers Frequency: daily (midnight to midnight)	1999/09/16- 2000/09/15	annual average	8.5	8.7	[36, 37]
			90^{th} percentile	15.3	13.1	
			98th percentile	21.1	17.3	

Note. The time in the column “period” is shown in the form of “year/month/day”; The unit for observations and predictions is μg/m^3. TEOM: Tapered element oscillating microbalance, Obs: observations, Est: estimates.

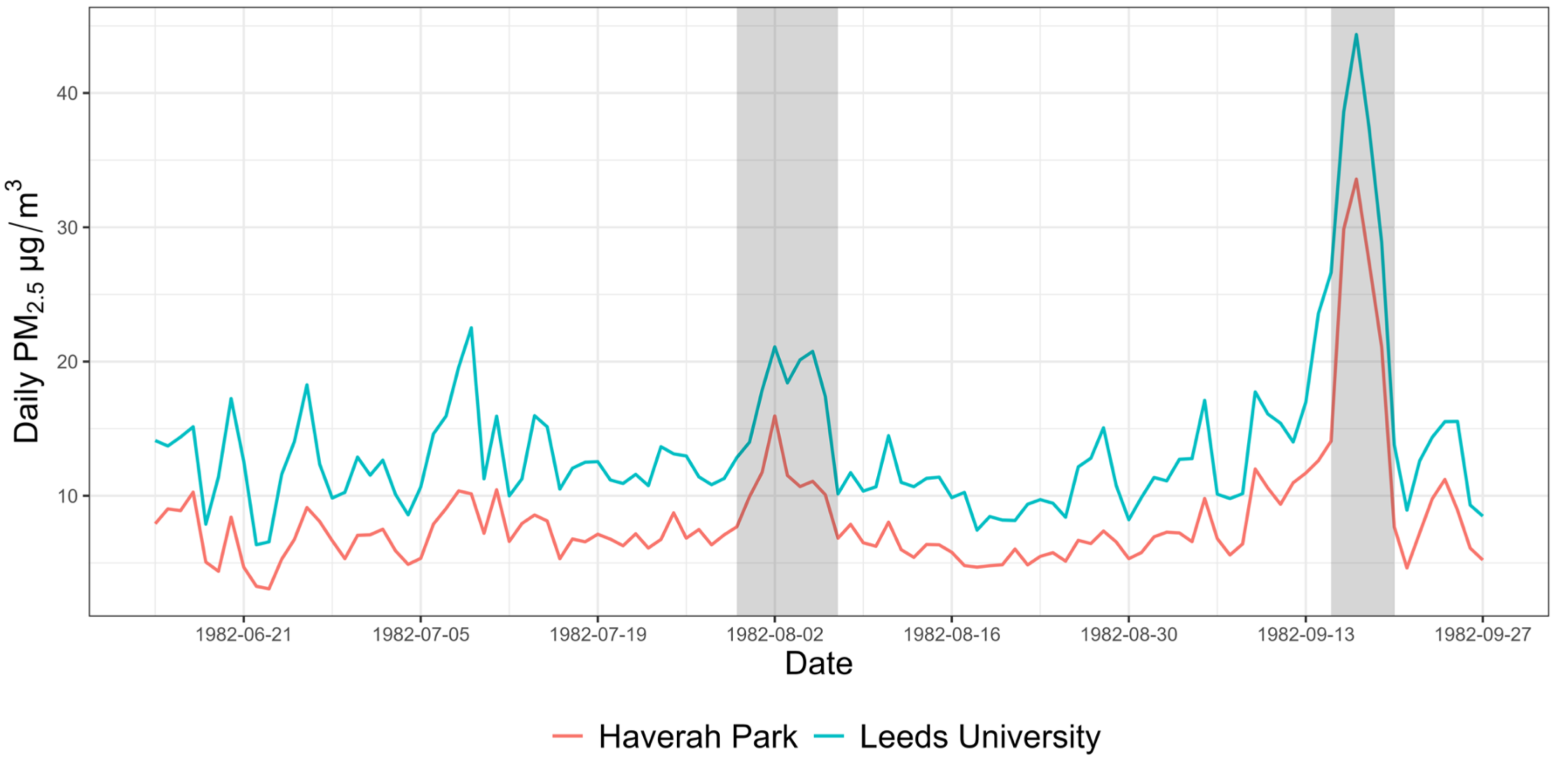


Figure S13. Time series of estimated daily $PM_{2.5}$ concentrations from June 13, 1982 to September 28, 1982 at Haverah Park (red) and Leeds University (blue). The 2 pollution episodes defined in the reference study[30] were highlighted in grey.

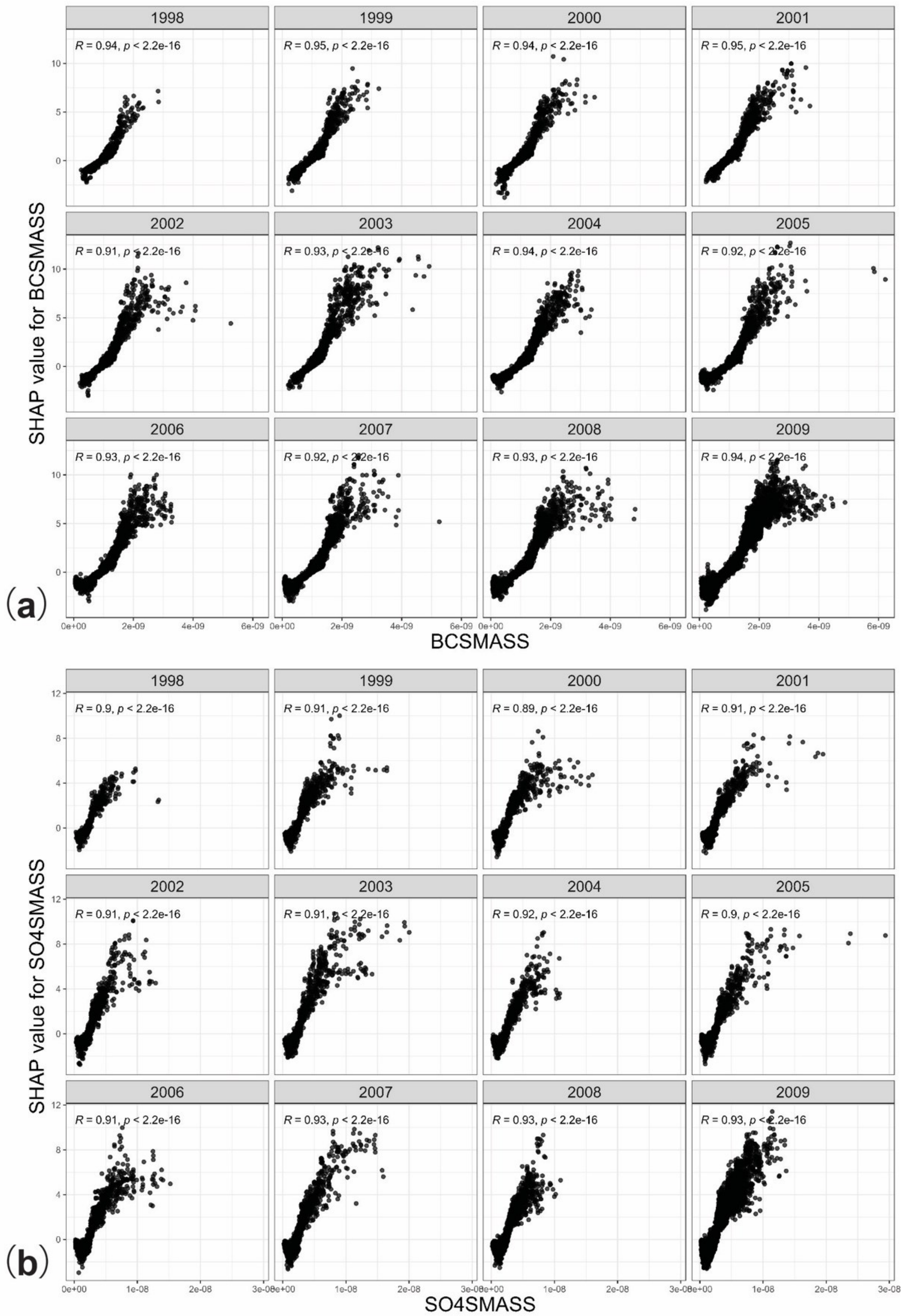


Figure S14. Effects of black carbon surface mass concentration (BCSMASS) (a) and sulfate surface mass concentration (SO4SMASS) (b) on the stage 2 model predictions in the testing set by year. The Pearson correlation coefficients (R) between the predictor variables and their SHAP values are shown in the upper left of each facet.

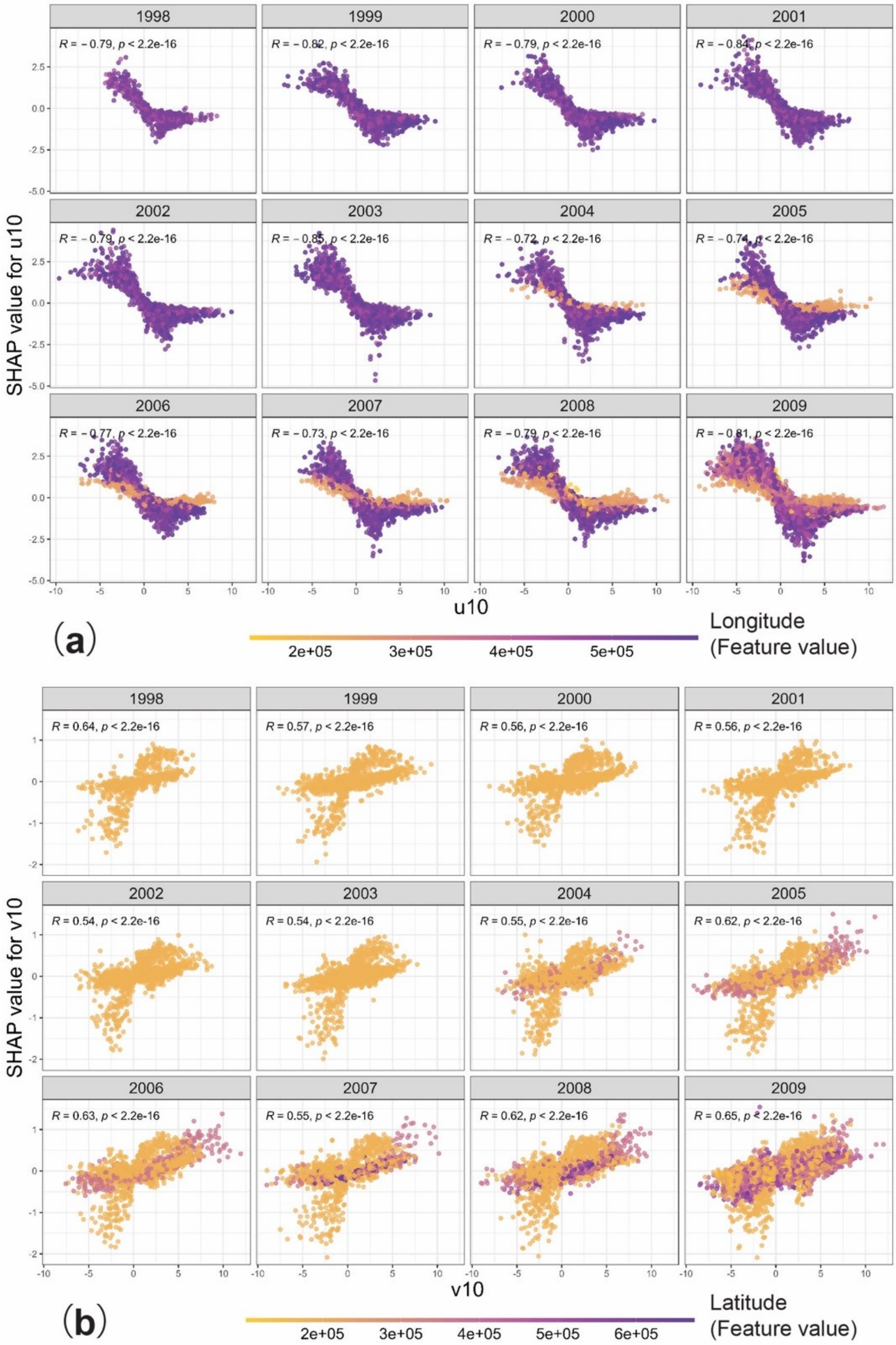
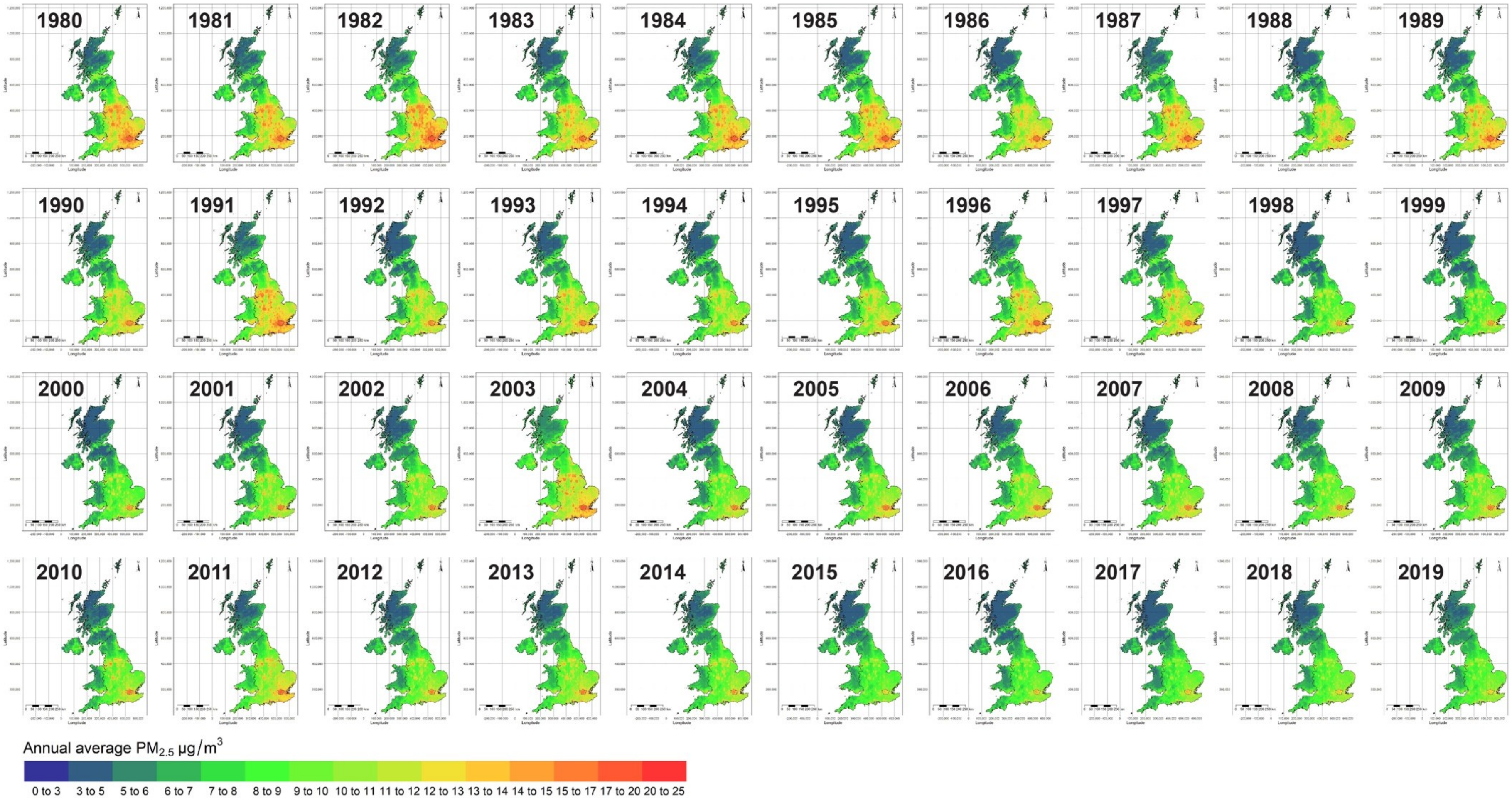


Figure S15. Effects of 10-m u-component of wind (u10, parallel to longitude) (a) and 10-m v-component of wind (v10, parallel to latitude) (b) on the stage 2 model predictions in the testing set by year. A positive u-component of wind is from the west, while a positive v-component of wind is from the south. The vertical distribution of the data in the dependence plot indicates the interaction effects between wind direction and other predictors. Although longitude and latitude were not directly used as predictors in our study, we use the color of the dot to represent the corresponding value of longitude and latitude in (a) and (b), respectively, to show how the effects of wind vary at different locations. The Pearson correlation coefficients (R) between the predictor variables and their SHAP values are shown in the upper left of each facet.



306

307 **Figure S16.** Spatial distribution of annual average estimated $PM_{2.5}$ concentrations in the UK from 1980 to 2019

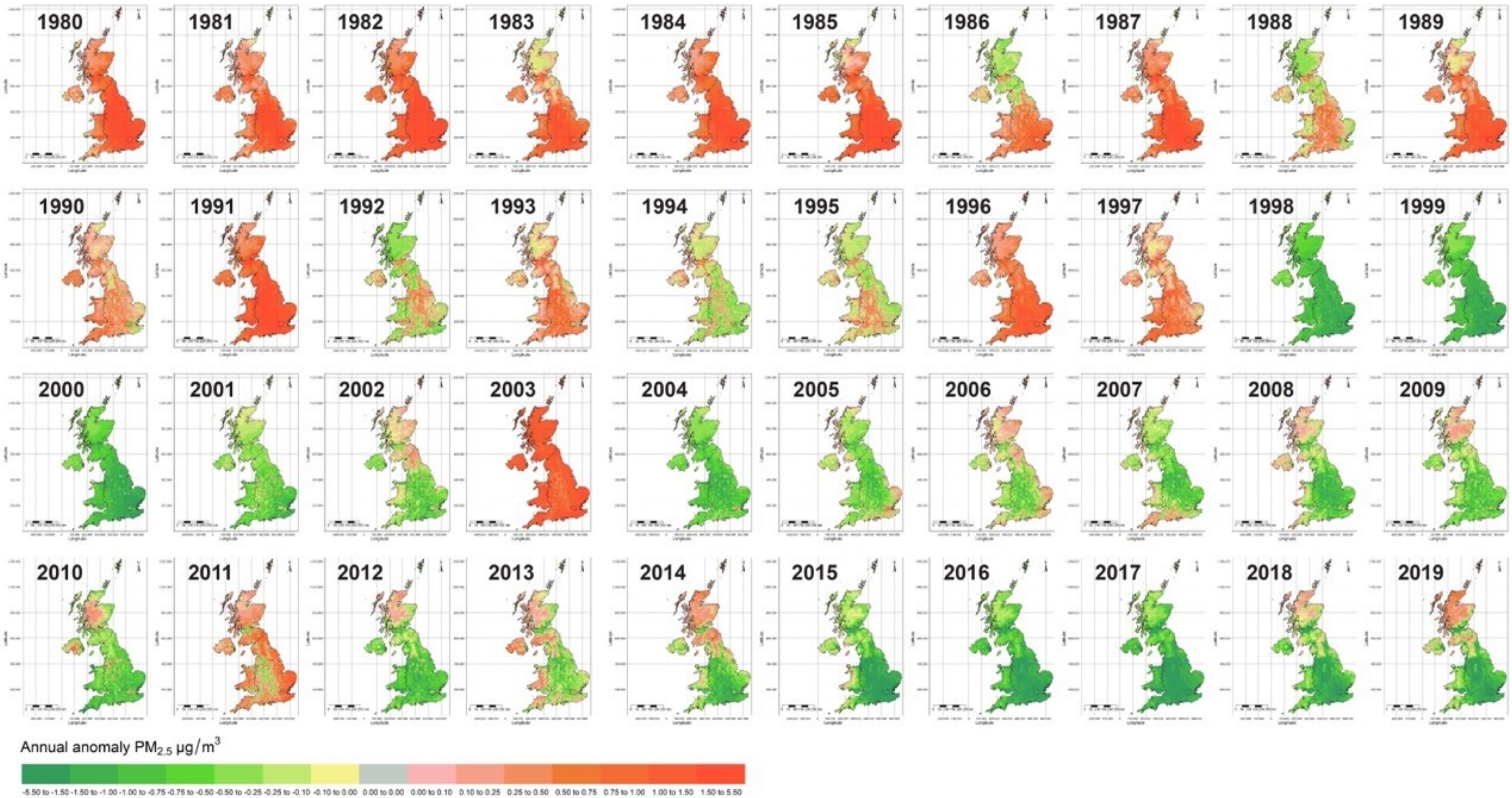


Figure S17. Spatial distribution of annual mean $PM_{2.5}$ anomalies in the UK from 1980 to 2019. The base line was the averages in each grid over the entire period

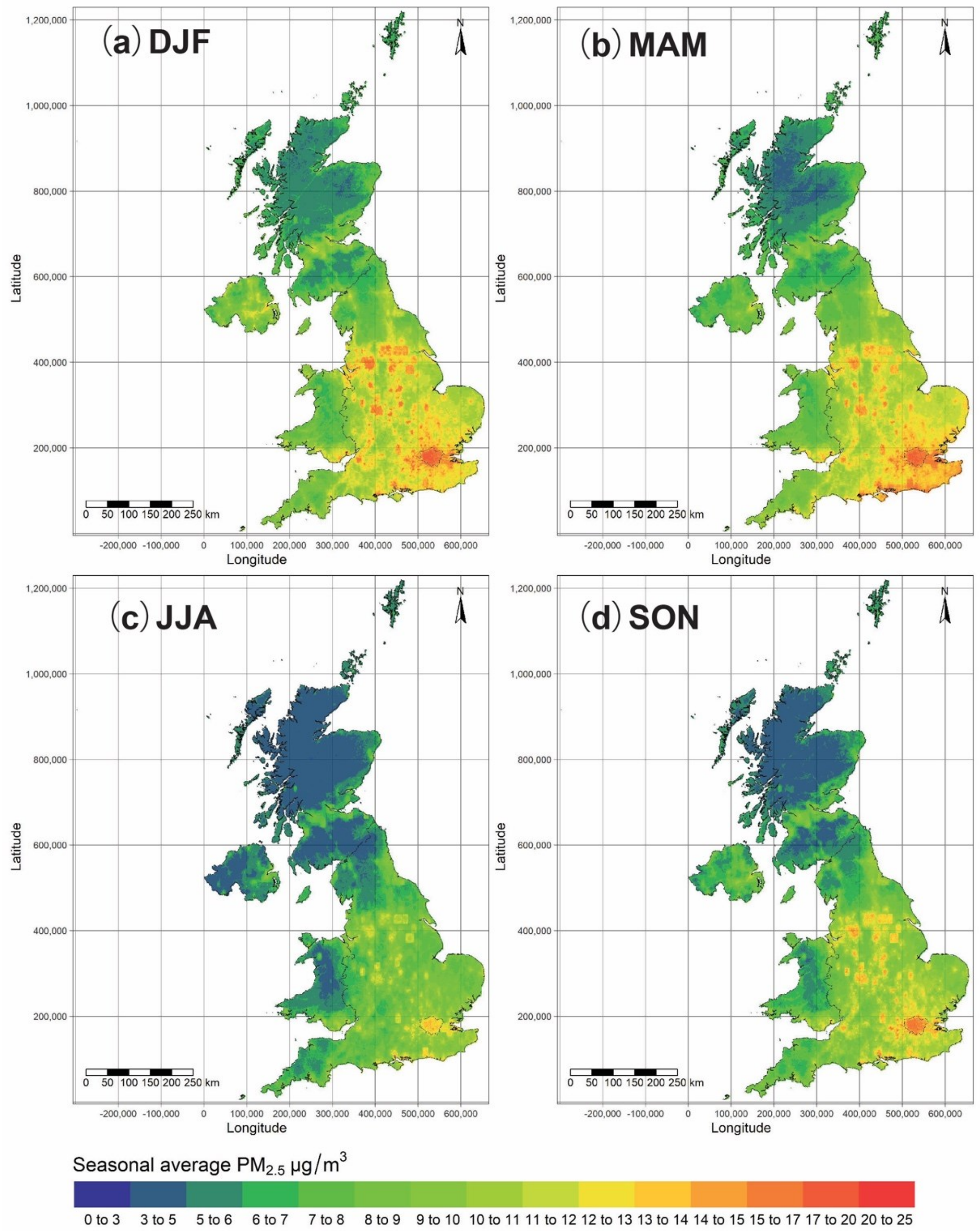


Figure S18. Spatial distribution of 4-decade seasonal average $PM_{2.5}$ estimates in the UK. DJF: Dec, Jan, Feb; MAM: Mar, Apr, May; JJA: June, July, Aug; SON: Sept, Oct, Nov.

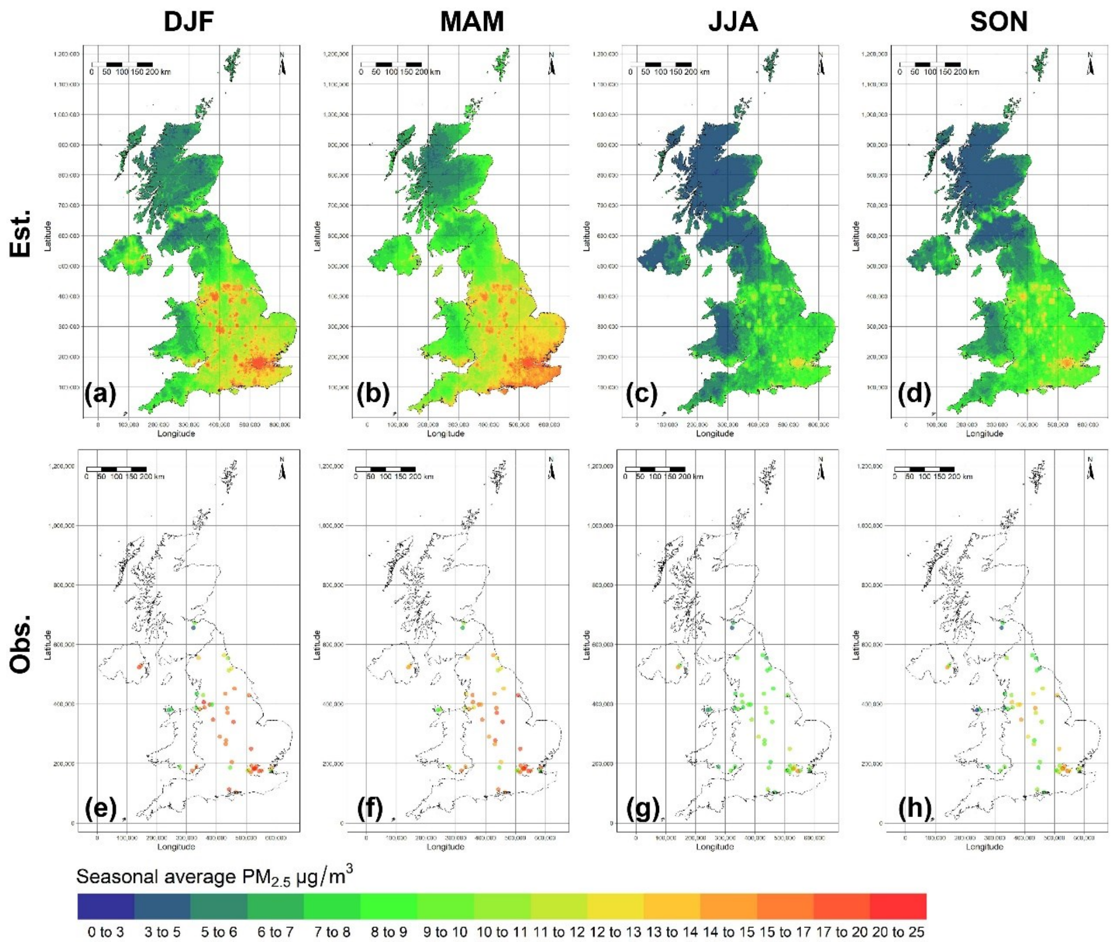


Figure S19. Comparisons of seasonal mean $PM_{2.5}$ and ground measured $PM_{2.5}$ concentrations in 2009. DJF: Dec, Jan, Feb; MAM: Mar, Apr, May; JJA: June, July, Aug; SON: Sept, Oct, Nov. Obs: observations, Est: estimates.

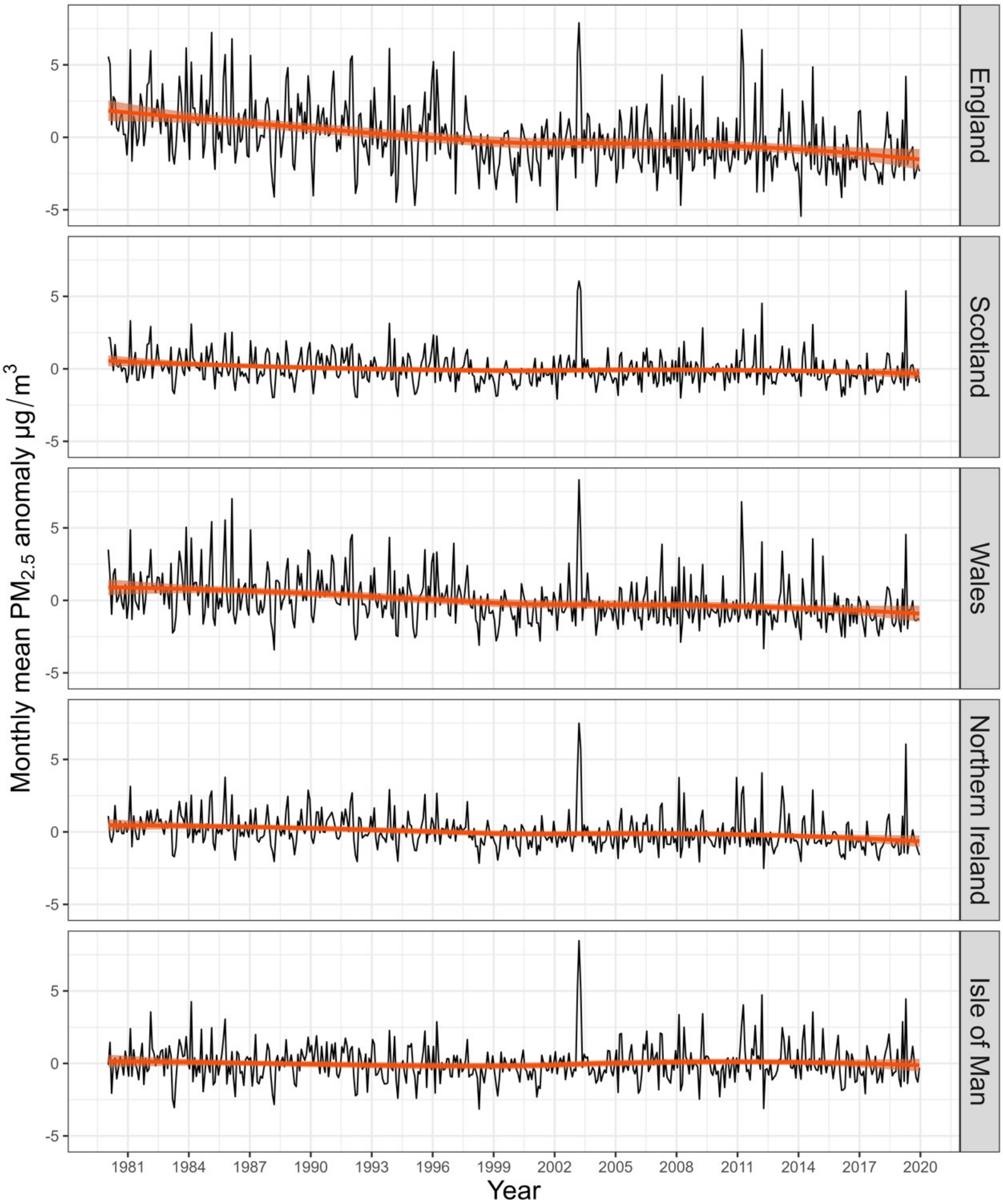


Figure S20. Time series of the monthly mean $PM_{2.5}$ anomalies from 1980 to 2019 in different subregions. The red lines with 95% confidence intervals (CIs) were derived with the locally estimated scatterplot smoothing (LOESS) approach.

Table S10. Trends and 95% confidence intervals (CIs) of the monthly mean $PM_{2.5}$ anomalies in the different subregions from 1980 to 2019

Region	Period	Trend ($\mu g/m^3$/year)	95% CI ($\mu g/m^3$/year)	Significance
	1980-1999	-0.12	(-0.17,-0.07)	p<0.05
England	2000-2019	-0.05	(-0.1,-0.01)	p<0.05
	1980-2019	-0.07	(-0.09,-0.06)	p<0.05
	1980-1999	-0.04	(-0.06,-0.02)	p<0.05
Scotland	2000-2019	-0.01	(-0.04,0.01)	p=0.4
	1980-2019	-0.02	(-0.02,-0.01)	p<0.05
	1980-1999	-0.06	(-0.1,-0.03)	p<0.05
Wales	2000-2019	-0.03	(-0.06,0)	p=0.09
	1980-2019	-0.04	(-0.06,-0.03)	p<0.05
	1980-1999	-0.04	(-0.06,-0.01)	p<0.05
Northern Ireland	2000-2019	-0.02	(-0.05,0.01)	p=0.12
	1980-2019	-0.02	(-0.03,-0.01)	p<0.05
	1980-1999	-0.02	(-0.05,0)	p=0.07
Isle of Man	2000-2019	0	(-0.03,0.03)	p=0.91
	1980-2019	0	(-0.01,0.01)	p=0.98

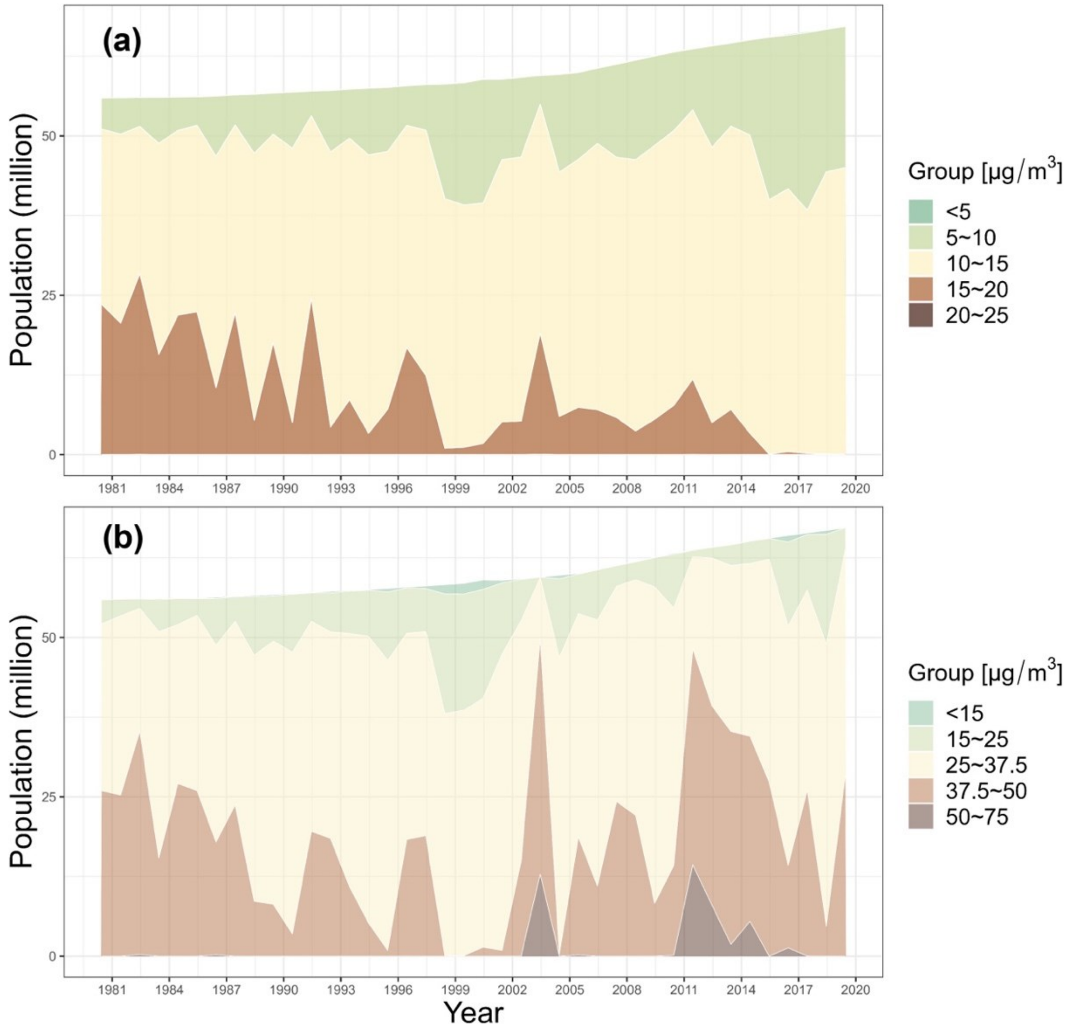


Figure S21. Time series of populations exposed to $PM_{2.5}$ pollution from 1980 to 2019 based on two annual metrics (a) annual average and (b) the 99th percentile of the annual distribution of 24-hour average.

Table S11. The grid-based CV results of the stage 2 model from 2010 to 2019

Year	Daily				Monthly average				Annual average			
	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE	Sample size	R^2	RMSE	MAE
2010	17660	0.71	4.63	3.14	641	0.65	3.02	2.19	55	0.48	2.42	1.84
2011	18061	0.81	4.73	3.20	650	0.81	3.03	2.27	56	0.51	2.36	1.82
2012	20158	0.79	4.29	3.05	735	0.73	2.78	2.21	63	0.45	2.21	1.77
2013	19868	0.81	3.91	2.78	718	0.70	2.38	1.77	62	0.58	1.83	1.41
2014	21021	0.81	3.84	2.64	778	0.79	2.30	1.71	67	0.58	1.71	1.25
2015	24931	0.77	3.72	2.46	909	0.68	2.26	1.76	78	0.50	1.77	1.45
2016	27818	0.80	3.48	2.53	999	0.77	2.09	1.69	85	0.70	1.71	1.38
2017	33100	0.83	3.19	2.34	1163	0.82	2.02	1.65	99	0.74	1.70	1.44
2018	40737	0.81	3.05	2.27	1429	0.76	1.99	1.63	121	0.73	1.68	1.38
2019	48862	0.85	3.04	2.26	1697	0.86	2.07	1.71	145	0.74	1.72	1.47

Note. The unit for RMSE and MAE is μg/m^3.

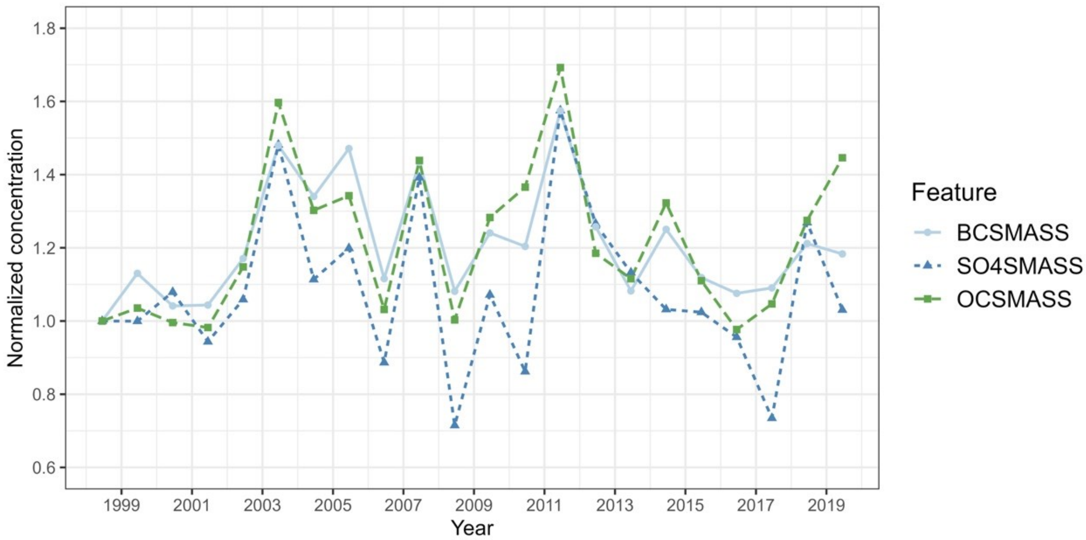


Figure S22. Time series of normalized average concentrations of 3 types of aerosols from MERRA-2 (normalized to 1998=1) from March to May in England in the develop set of the stage 2 model

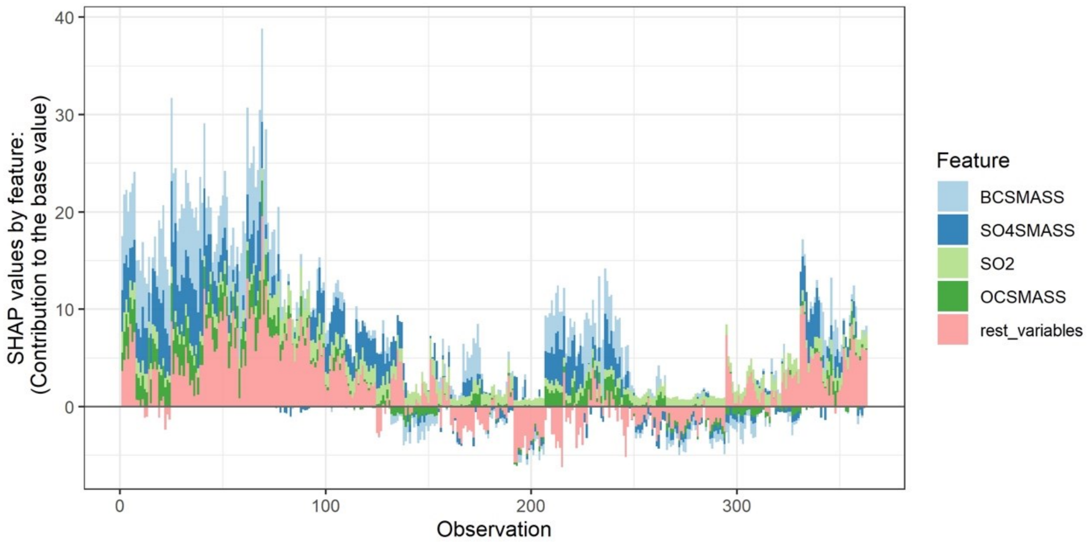


Figure S23. Effects of the stage 2 model predictors on $PM_{2.5}$ predictions from March to May 2003 in England in the testing set. Only the top 4 predictors are shown separately, other predictors are aggregated. The x-axis is the ID of predictions. The y-axis is the stacked SHAP values of the predictors for each prediction.

References

(1) Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **2020**, *146* (730), 1999-2049. DOI: https://doi.org/10.1002/qj.3803.

(2) Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* **2021**, *13* (9), 4349-4383. DOI: https://doi.org/10.5194/essd-13-4349-2021.

(3) Perry, M.; Hollis, D. The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology* **2005**, *25* (8), 1041-1054. DOI: https://doi.org/10.1002/joc.1161.

(4) Hollis, D.; McCarthy, M.; Kendon, M.; Legg, T.; Simpson, I. HadUK-Grid—A new UK dataset of gridded climate observations. *Geoscience Data Journal* **2019**, *6* (2), 151-159. DOI: https://doi.org/10.1002/gdj3.78.

(5) Gelaro, R.; McCarty, W.; Suárez, M. J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C. A.; Darmenov, A.; Bosilovich, M. G.; Reichle, R.; et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of climate* **2017**, *30* (14), 5419-5454. DOI: https://doi.org/10.1175/JCLI-D-16-0758.1.

(6) Hoesly, R. M.; Smith, S. J.; Feng, L.; Klimont, Z.; Janssens-Maenhout, G.; Pitkanen, T.; Seibert, J. J.; Vu, L.; Andres, R. J.; Bolt, R. M.; et al. Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development* **2018**, *11* (1), 369-408. DOI: https://doi.org/10.5194/gmd-11-369-2018.

(7) Feng, L.; Smith, S. J.; Braun, C.; Crippa, M.; Gidden, M. J.; Hoesly, R.; Klimont, Z.; van Marle, M.; van den Berg, M.; van der Werf, G. R. The generation of gridded emissions data for CMIP6. *Geoscientific Model Development* **2020**, *13* (2), 461-482. DOI: https://doi.org/10.5194/gmd-13-461-2020.

(8) Land cover classification gridded maps from 1992 to present derived from satellite observations. Copernicus Climate Data. https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.006f2c9a?tab=overview (accessed 2022-02-21).

(9) Haklay, M.; Weber, P. Openstreetmap: User-generated street maps. *IEEE Pervasive computing* **2008**, *7* (4), 12-18. DOI: https://doi.org/10.1109/MPRV.2008.80.

(10) NASADEM Merged DEM Global 1 arc second V001. NASA EOSDIS Land Processes Distributed Active Archive Center. https://lpdaac.usgs.gov/products/nasadem_hgtv001/ (accessed 2022-04-09).

(11) NASADEM Slope and Curvature Global 1 arc second V001. NASA EOSDIS Land Processes Distributed Active Archive Center. https://lpdaac.usgs.gov/products/nasadem_scv001/ (accessed 2022-04-09).

(12) GHS-POP R2022A - GHS population grid multitemporal (1975-2030) - OBSOLETE RELEASE. European Commission, Joint Research Centre (JRC). https://data.jrc.ec.europa.eu/dataset/d6d86a90-4351-4508-99c1-cb074b022c4a (accessed 2022-07-14).

(13) GHS-SMOD R2022A - GHS settlement layers, application of the Degree of Urbanisation methodology (stage I) to GHS-POP R2022A and GHS-BUILT-S R2022A, multitemporal (1975-2030) - OBSOLETE RELEASE. European Commission, Joint Research Centre (JRC). https://data.jrc.ec.europa.eu/dataset/4606d58a-dc08-463c-86a9-d49ef461c47f (accessed 2022-07-14).

(14) Harmonization of DMSP and VIIRS nighttime light data from 1992-2021 at the global scale.

figshare. https://figshare.com/articles/dataset/Harmonization_of_DMSP_and_VIIRS_nighttime_light_data_from_1992-2018_at_the_global_scale/9828827 (accessed 2022-07-14).

(15) Schneider, R.; Vicedo-Cabrera, A. M.; Sera, F.; Masselot, P.; Stafoggia, M.; de Hoogh, K.; Kloog, I.; Reis, S.; Vieno, M.; Gasparrini, A. A Satellite-Based Spatio-Temporal Machine Learning Model to Reconstruct Daily PM2.5 Concentrations across Great Britain. *Remote Sensing* **2020**, *12* (22), 3803. DOI: https://doi.org/10.3390/rs12223803.

(16) Danesh Yazdi, M.; Kuang, Z.; Dimakopoulou, K.; Barratt, B.; Suel, E.; Amini, H.; Lyapustin, A.; Katsouyanni, K.; Schwartz, J. Predicting Fine Particulate Matter (PM2.5) in the Greater London Area: An Ensemble Approach using Machine Learning Methods. *Remote Sensing* **2020**, *12* (6), 914. DOI: https://doi.org/10.3390/rs12060914.

(17) MERRA-2 tavg1_2d_aer_Nx: 2d,1-Hourly,Time-averaged,Single-Level,Assimilation,Aerosol Diagnostics, V5.12.4. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://disc.gsfc.nasa.gov/datasets/M2T1NXAER_5.12.4/summary (accessed 2022-04-26).

(18) Koster, R. D.; McCarty, W.; Coy, L.; Gelaro, R.; Huang, A.; Merkova, D.; Smith, E. B.; Sienkiewicz, M.; Wargan, K. *MERRA-2 input observations: Summary and assessment*; No. GSFC-E-DAA-TN37524; 2016. (accessed 2022-04-26).

(19) Yu, W.; Ye, T.; Zhang, Y.; Xu, R.; Lei, Y.; Chen, Z.; Yang, Z.; Zhang, Y.; Song, J.; Yue, X.; et al. Global estimates of daily ambient fine particulate matter concentrations and unequal spatiotemporal distribution of population exposure: a machine learning modelling study. *The Lancet Planetary Health* **2023**, *7* (3), e209-e218. DOI: https://doi.org/10.1016/S2542-5196(23)00008-6.

(20) Huang, C.; Hu, J.; Xue, T.; Xu, H.; Wang, M. High-Resolution Spatiotemporal Modeling for Ambient PM2.5 Exposure Assessment in China from 2013 to 2019. *Environmental Science & Technology* **2021**, *55* (3), 2152-2162. DOI: https://doi.org/10.1021/acs.est.0c05815.

(21) Wei, J.; Li, Z.; Guo, J.; Sun, L.; Huang, W.; Xue, W.; Fan, T.; Cribb, M. Satellite-Derived 1-km-Resolution PM1 Concentrations from 2014 to 2018 across China. *Environmental Science & Technology* **2019**, *53* (22), 13265-13274. DOI: https://doi.org/10.1021/acs.est.9b03258.

(22) Xue, T.; Zheng, Y.; Tong, D.; Zheng, B.; Li, X.; Zhu, T.; Zhang, Q. Spatiotemporal continuous estimates of PM2.5 concentrations in China, 2000–2016: A machine learning method with inputs from satellites, chemical transport model, and ground observations. *Environment International* **2019**, *123*, 345-357. DOI: https://doi.org/10.1016/j.envint.2018.11.075.

(23) Ma, Z.; Hu, X.; Sayer, A. M.; Levy, R.; Zhang, Q.; Xue, Y.; Tong, S.; Bi, J.; Huang, L.; Liu, Y. Satellite-Based Spatiotemporal Trends in PM2.5 Concentrations: China, 2004-2013. *Environmental Health Perspectives* **2016**, *124* (2), 184-192. DOI: https://doi.org/10.1289/ehp.1409481.

(24) Defourny, P.; Lamarche, C.; Marissiaux, Q.; Brockmann, C.; Boettcher, M.; Kirches, G. *Product User Guide and Specification: ICDR Land Cover 2016-2020*; 2021. https://datastore.copernicus-climate.eu/documents/satellite-land-cover/D5.3.1_PUGS_ICDR_LC_v2.1.x_PRODUCTS_v1.1.pdf (accessed 2022-02-21).

(25) Haklay, M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design* **2010**, *37* (4), 682-703. DOI: https://doi.org/10.1068/b35097.

(26) Li, X.; Zhou, Y. A Stepwise Calibration of Global DMSP/OLS Stable Nighttime Light Data (1992–2013). *Remote Sensing* **2017**, *9* (6), 637. DOI: https://doi.org/10.3390/rs9060637.

(27) Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A highly

efficient gradient boosting decision tree. In *Advances in neural information processing systems*, Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. Eds.; Vol. 30; Curran Associates, Inc., 2017.

(28) Wei, J.; Li, Z.; Pinker, R. T.; Wang, J.; Sun, L.; Xue, W.; Li, R.; Cribb, M. Himawari-8-derived diurnal variations in ground-level PM2.5 pollution across China using the fast space-time Light Gradient Boosting Machine (LightGBM). *Atmospheric Chemistry and Physics* **2021**, *21* (10), 7863-7880. DOI: https://doi.org/10.5194/acp-21-7863-2021.

(29) Zhong, J.; Zhang, X.; Gui, K.; Wang, Y.; Che, H.; Shen, X.; Zhang, L.; Zhang, Y.; Sun, J.; Zhang, W. Robust prediction of hourly PM2.5 from meteorological data using LightGBM. *National Science Review* **2021**, *8* (10). DOI: https://doi.org/10.1093/nsr/nwaa307.

(30) Clarke, A. G.; Willison, M. J.; Zeki, E. M. A comparison of urban and rural aerosol composition using dichotomous samplers. *Atmospheric Environment (1967)* **1984**, *18* (9), 1767-1775. DOI: https://doi.org/10.1016/0004-6981(84)90352-4.

(31) Quality of Urban Air Review Group. *Airborne Particulate Matter in the United Kingdom*; Department of Environment, UK, 1996. (accessed 2023/02/23).

(32) Harrison, R. M.; Deacon, A. R.; Jones, M. R.; Appleby, R. S. Sources and processes affecting concentrations of PM10 and PM2.5 particulate matter in Birmingham (U.K.). *Atmospheric Environment* **1997**, *31* (24), 4103-4117. DOI: https://doi.org/10.1016/S1352-2310(97)00296-3.

(33) Anderson, H. R.; Bremner, S. A.; Atkinson, R. W.; Harrison, R. M.; Walters, S. Particulate matter and daily mortality and hospital admissions in the west midlands conurbation of the United Kingdom: associations with fine and coarse particles, black smoke and sulphate. *Occupational and Environmental Medicine* **2001**, *58* (8), 504-510. DOI: https://doi.org/10.1136/oem.58.8.504.

(34) Yin, J. Monitoring of airborne particulate mass and number concentrations in the UK atmosphere. University of Birmingham, Birmingham, 2002. https://etheses.bham.ac.uk/id/eprint/10133/ (accessed 2023-02-23).

(35) Moore, M. An investigation into particulate air pollution in London, at a site on the Archway Road (A1) London. Middlesex University, London, 2001. https://eprints.mdx.ac.uk/13463/ (accessed 2023-02-23).

(36) Heal, M. R.; Hibbs, L. R.; Agius, R. M.; Beverland, I. J. Interpretation of variations in fine, coarse and black smoke particulate matter concentrations in a northern European city. *Atmospheric Environment* **2005**, *39* (20), 3711-3718. DOI: https://doi.org/10.1016/j.atmosenv.2005.03.007.

(37) Heal, M. R.; Hibbs, L. R.; Agius, R. M.; Beverland, I. J. Total and water-soluble trace metal content of urban background PM10, PM2.5 and black smoke in Edinburgh, UK. *Atmospheric Environment* **2005**, *39* (8), 1417-1430. DOI: https://doi.org/10.1016/j.atmosenv.2004.11.026.