Supplement of Atmos. Chem. Phys., 19, 11303–11314, 2019
https://doi.org/10.5194/acp-19-11303-2019-supplement

Atmospheric Chemistry and Physics
Open Access

Supplement of

Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique

Tuan V. Vu et al.

Correspondence to: Zongbo Shi (z.shi@bham.ac.uk) and Roy M. Harrison (r.m.harrison@bham.ac.uk)

Number of pages : 11
Number of tables : 3
Number of figures : 5

CONTENTS

Methods

Figures: Figure S1 to Figure S5

Tables

Section S1. Data collection and overview of air quality

Hourly air quality data for six air pollutants was collected in Beijing from 17/01/2013 to 31/12/2017 across 12 national air quality monitoring stations which were classified in three categories (urban, suburban, and rural areas) based on hierarchical clustering (Figure S1, Table 1). Specifically, $PM_{2.5}$ levels at urban, suburban and rural sites decreased from 89.8, 78.3, and 67.8 μg m^{-3} in 2013 to 59.6, 54.6, and 47.8 μg m^{-3} in 2017, respectively. In 2017, 23 % of days still exceeded the NAAQS-II. A higher decrease in PM_{10} levels by 20.2 % was found at urban sites compared to those at suburban sites (17.2 %). PM_{10} also shows exceedances of NAAQS-II standards both for daily averages (150 μg m^{-3}) and annual averages (70 μg m^{-3}). It suggests that particulate matter, especially $PM_{2.5}$ is still a critical air pollutant in Beijing. In 2017, SO_2 does not show exceedance of the NAAQS-II standards either for daily averages (150 μg m^{-3}) and annual averages (60 μg m^{-3}). For CO, only 12 days do not meet NAAQS-II standards of 4 μg m^{-3}. In contrast, the annual average concentration of NO_2 in 2017 was slightly higher than the NAAQS-II standard of 40 μg m^{-3}, with 18 days exceeding the NAAQS-II standard for daily averages (80 μg m^{-3}).

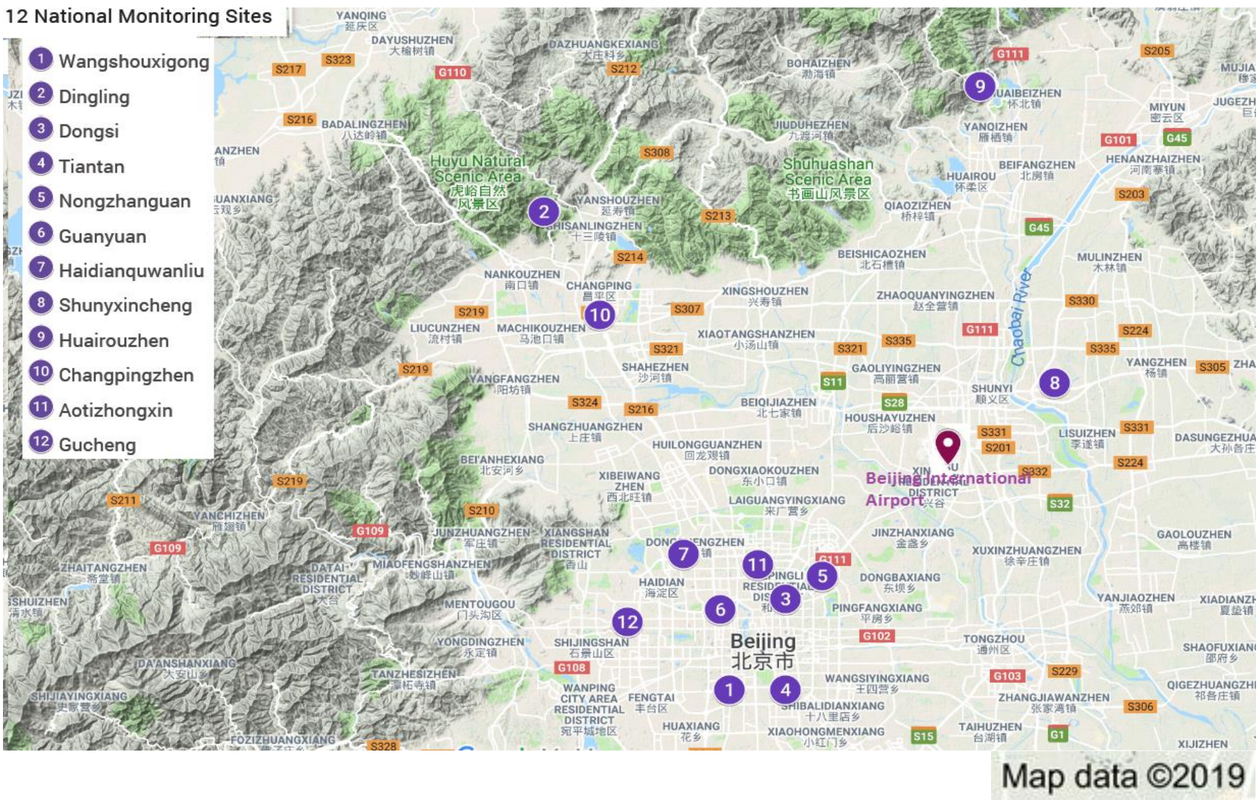


Figure S1. Map of the 12 monitoring stations (Source: © Google Maps topographic background imagery)

Section S2. Notices, regulation and policies for air pollution control in Beijing

Regulation and policies on energy system re-structuring:

- In October 2013, the government of Huairou district enforced a policy to replace anthracite stoves from 3000 rural households, change coal heating to electricity for 1170 households, supply liquefied petroleum to the countryside for 20,000 households, construct energy-saving residential housing and implement district heating; this reduced the consumption of 47,000 tons of poor quality coal.
- In Oct 2013, the government of Shijingshan, an urban district of Beijing, planned to cut 2800 tons of coal usage from coal-fired boilers in 2013, and reduce coal usage by more than 4500 tons in 2014, and eliminate coal-fired boilers in 2015.
- In November 2013, Miyun government issued an action plan to "Reduce coal for clean air" with a focus on urban transformation, conversion to natural gas, replacement with high quality coal, relocation of mountain communities, conservation of household energy, and removal of illegal constructions.
- In September 2014, the China State government released an important regulation on the "Reform and upgrade Action Plan for coal energy conservation and emission reduction (2014-2020)" that requires Beijing to place strict controls upon energy efficiency. Following that Action Plan, stack gas emissions of SO_2, NO_x, and PM from coal-fired power plants must be limited to below 10, 35, and 50 mg m^{-3} respectively.
- In March 2017, the Ministry of Environmental Protection issued the "2017 Air Pollution Prevention and Control Work Plan for Beijing-Tianjin-Hebei". According to this plan, before the end of October 2017, Beijing, Tianjin, Langfang and Baoding City of Hebei will become the "no-coal zone".

Regulations and policies on vehicle emission control: In order to control air pollution from vehicle emissions, during 2013-2017 the city announced a series of policies and regulations focusing on the implementation of stricter standards for new vehicles and vehicle fuels, elimination of yellow-label vehicles (which do not meet basic emission standards), and promotion of public transport. Consequently, Beijing led the nation in improving the fuel quality standards by adopting the desulfurization of gasoline and diesel fuels (sulfur content <10 ppm) in 2012, three years ahead of the surrounding regions (Tianijin and Hebei) and five years before the national deadline. Major policies for air pollution from transportation management:

- In February 2013, Beijing implemented the fifth phase emission standards for new light-duty gasoline vehicles (LDVs) and heavy-duty diesel vehicles (HDVs) for public transport.
- In June 2013, another notice from the Beijing government emphasized that all heavy-duty vehicles sold and registered in Beijing must meet the national fourth-phase emission standards
- In August 2014, a notice from Beijing's government declared that all spark ignition light vehicles must meet the national five phase standard from 1st January 2015.
- In 2014, Beijing Municipal Commission of Transport (BMCT) expanded traffic restrictions to certain vehicles, particularly yellow-label and non-local vehicles to enter the city within the sixth ring road during daytime since 2015.
- In November 2014, the governments of Yanquing and Miyun, two rural districts of Beijing, released regulations to prohibit yellow-label gasoline vehicles entering certain roads.
- In February 2015, the Beijing Municipal government issued a notice to promote elimination and replacement of old motor vehicles with an expectation of 1 million old vehicles/year phased out.
- Other policies which may have contributed to the enhancement of air quality during 2013-2017 included a ban of outdoor biomass burning and improved suppression of dust discharges from construction sites.

Section S3. Model performance and explanation

Variables and hyperparameters: The input variables contain time and MET variables. Time variables: day_unix (or t_{trend}) represents the emission trend of a pollutant; Julian_day (t_{JD}: the day of the years) represents for the seasonal variation; weekday/weekend represents the difference of pollution between the week and weekend days.

MET variables: wind speed (m s^{-1}), wind direction (^{o}), temperature (^{o}C), relative humidity (%), and atmospheric pressure (mbar). The back-trajectories can be used as a predictor feature, but it does not increase the performance of the model in this case.

Selected parameters in a random forest:

- Mtry=4: variables randomly sampled for splitting the decision tree
- Nodesize=3: minimum size of terminal nodes for model
- Ntree=200, the number of trees to grow. Figure S2 shows the dependence of model performance on the number of trees.

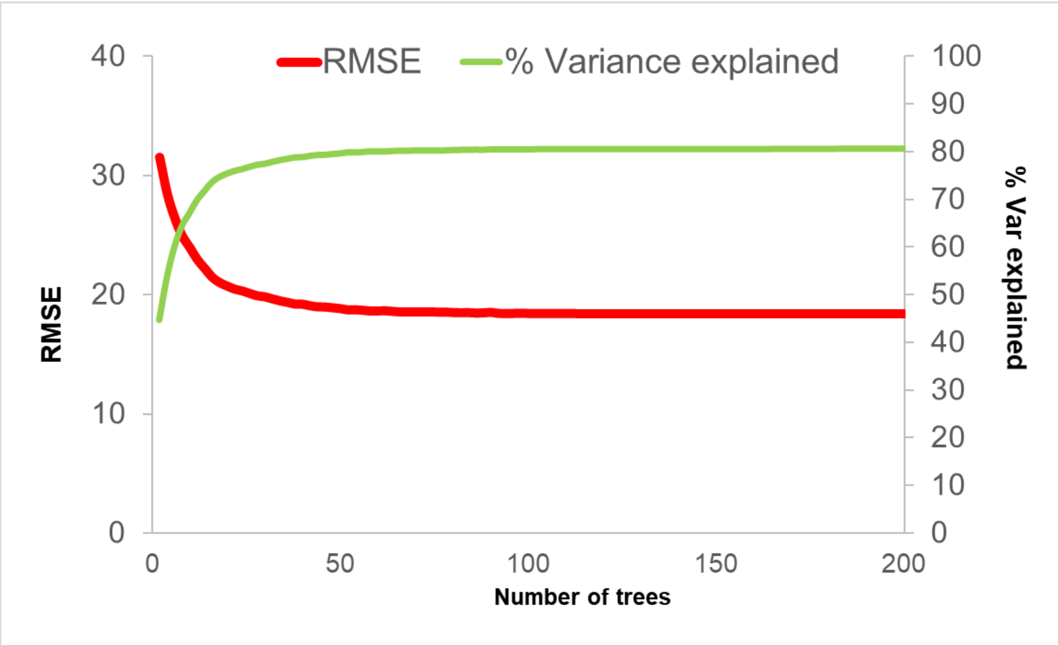


Figure S2. The influence of number of trees on the model performance for $PM_{2.5}$.

Model performance's evaluation

A random forest shows a good performance with the correlation (r^2) between hourly predicted and observed data for both training and testing data sets. In particular, r^2 value ranged 0.81-0.83, 0.75-0.79, 0.80-0.83, 0.88-0.90, 0.85-0.87, and 0.89-0.90 for $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO and O_3, respectively. Figure S3 shows the hourly correlation between observed and predicted data for a testing data. Other model evaluation metrics are shown in Table S2.

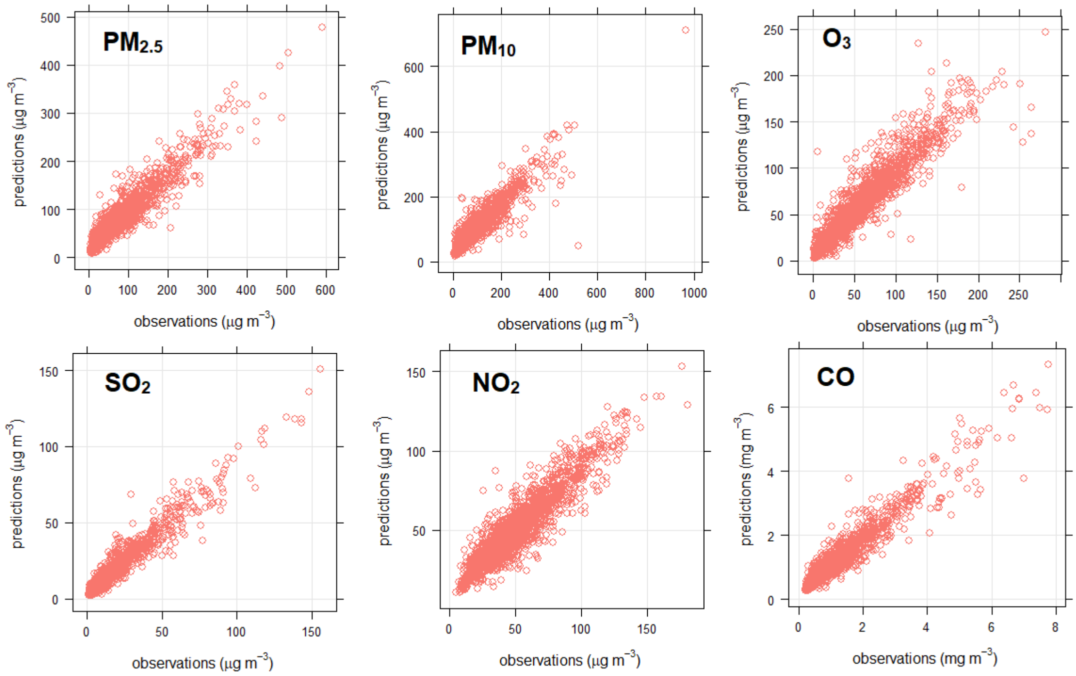


Figure S3. Correlations between daily observed and predicted data from testing data sets

As shown in Figure S3, it is likely that the model underestimates hourly concentration of air pollutants at the extremely high levels. These errors are reduced when we compare the weekly averaged concentration as shown in Figure S4.

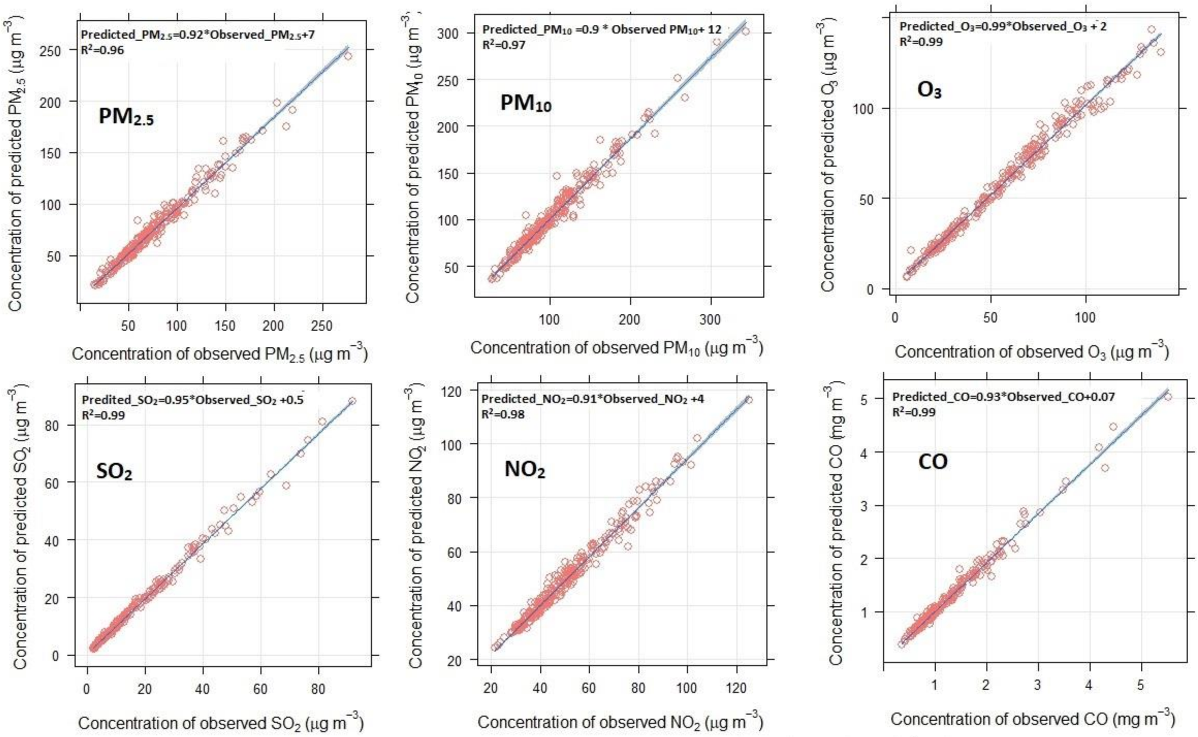
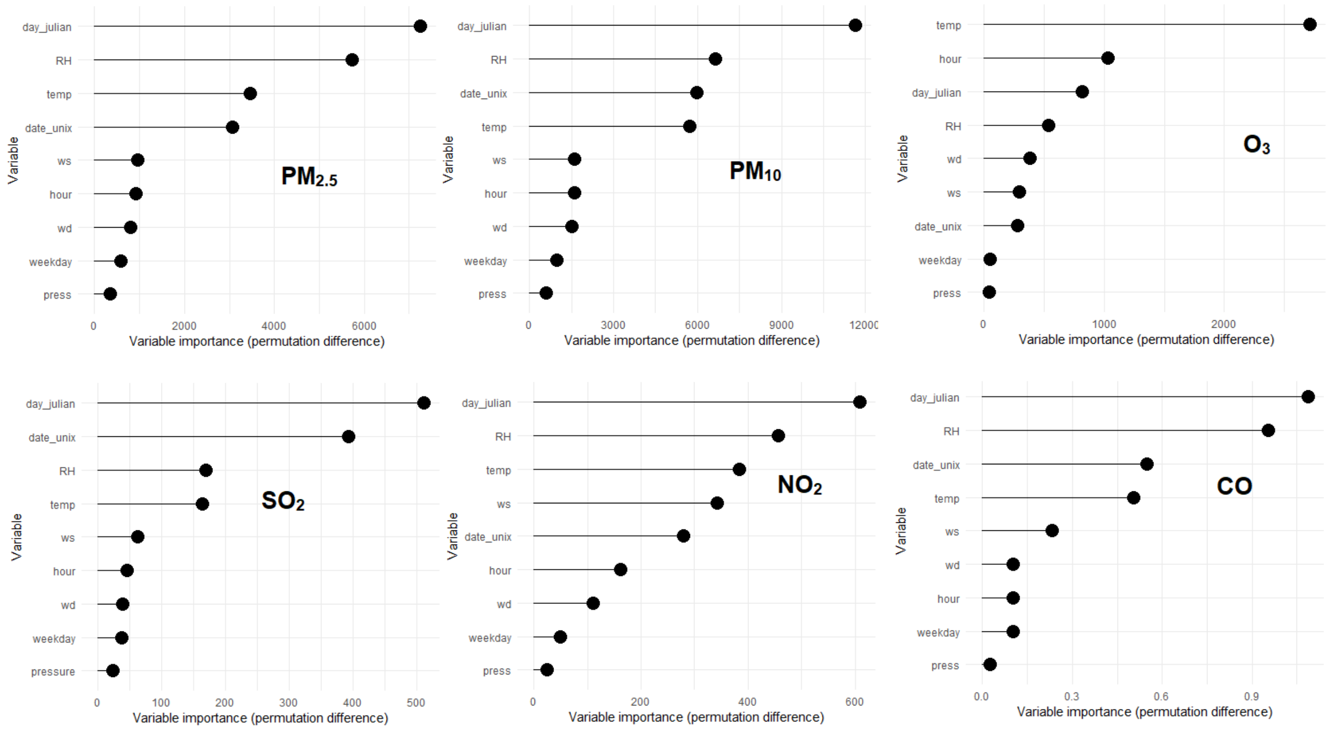


Figure S4. The correlation between observed and modelled concentrations is approximately 0.9-0.99 for weekly averaged data. In our study, a RF forest model was trained using a fraction of 0.7 from the datasets.

Variable importance and interactions:

As shown in Figure S4, seasonal variations (day_julian) play the most important variable in the model, except for ozone when temperature and diurnal pattern (hour) mainly control the predicted values. The trend (day_unix) shows more important role in the model of SO_2 and CO, indicating emission control shows most effectiveness on the decrease of SO_2 and CO. Regarding MET variables, humidity and temperature play a more important role in the model of PM while wind speed has a larger impact in the model of NO_2. The variable interaction is shown in Figure S5.



Figuer S4. Importance of predictor features: date_unix, day of the year (day_julian), hour of day (hour), week/weekend, temperature (temp), RH, pressure (press), wind speed (ws), wind direction (wd) in the random forest model. Figure 4 shows the day of the year (seasonal variable) is the most important variable controlling the concentration of the pollutant (except for ozone: the most important is the temperature variable). The trend (date_unix) has a larger effect on SO_2, than CO and PM, less effect on the NO2 and no significant effect on O3 concentration.

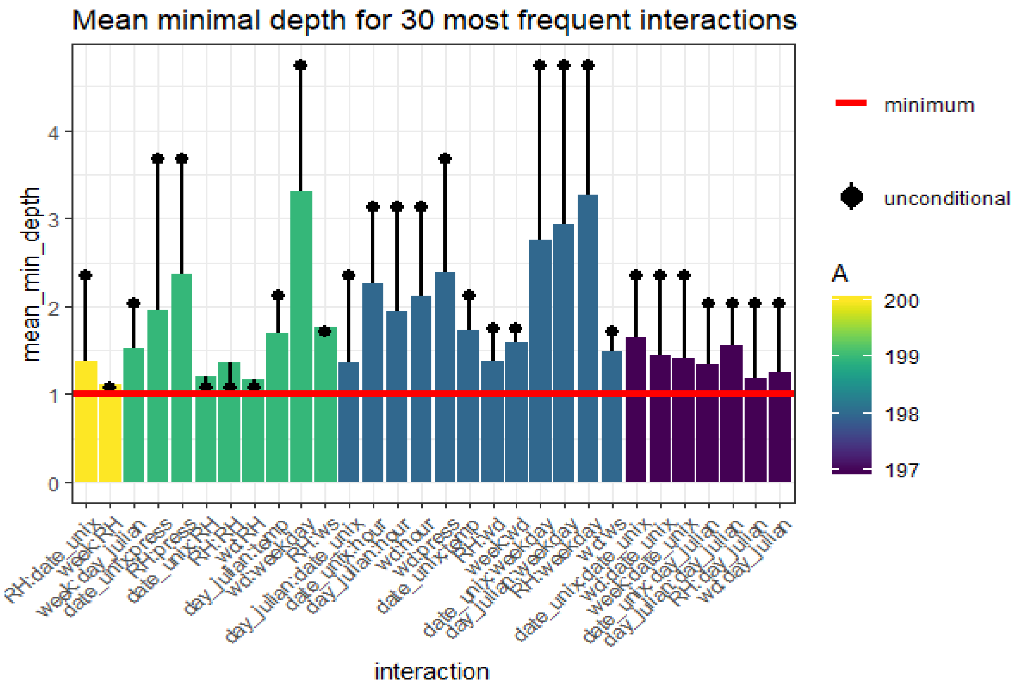


Figure S6. Features interactions in a random forest model for $PM_{2.5}$. This figure shows the co-occurrence of a pair of variables in a similar tree. For example, in the first node of the tree, RH and date_unix is the most frequent occurrence.

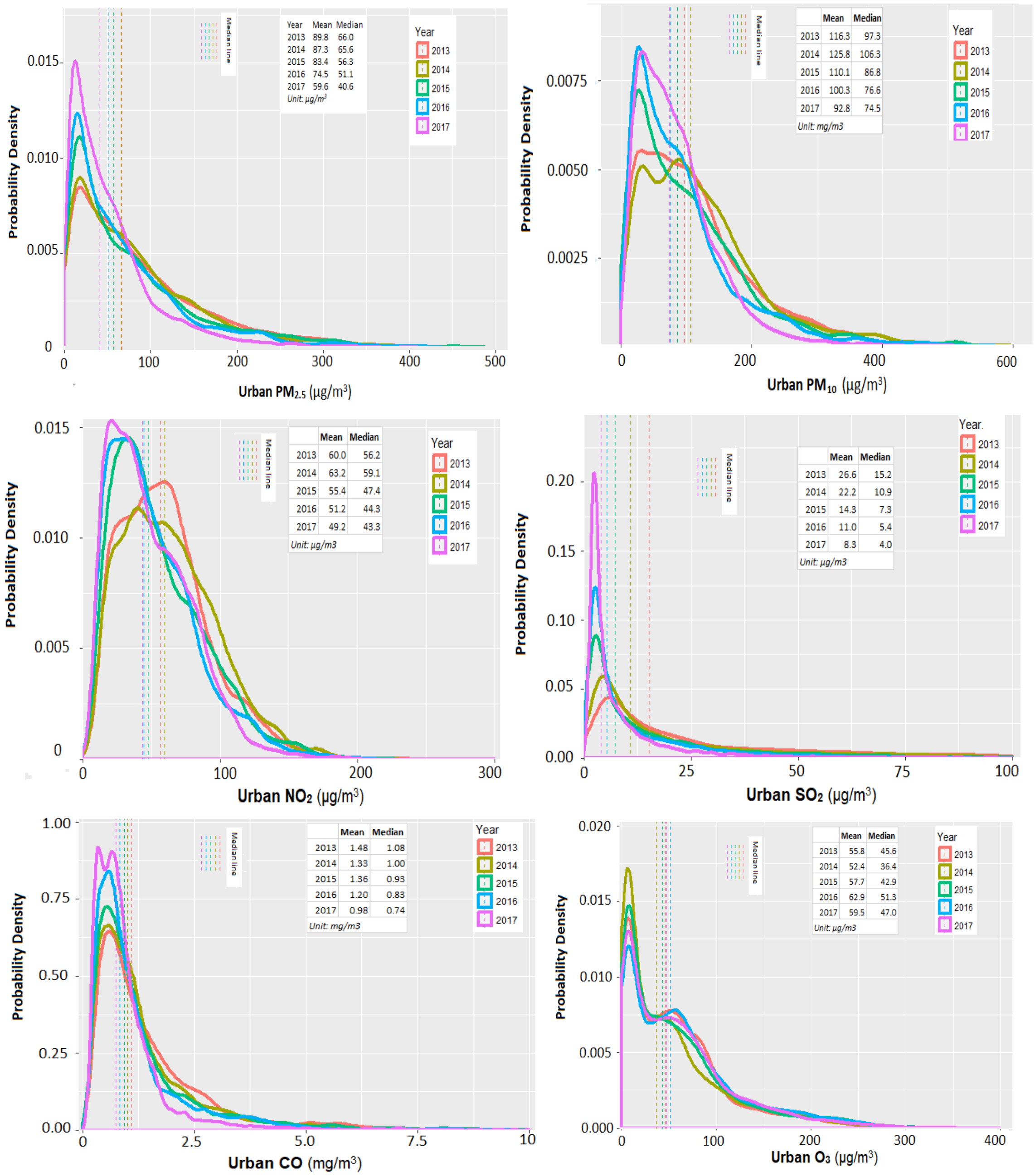


Figure S7. Probability density of urban air pollutant concentrations during 2013-2017. Number of heavy polluted events decreases from 2013 to 2017 for all pollutants, except ozone.

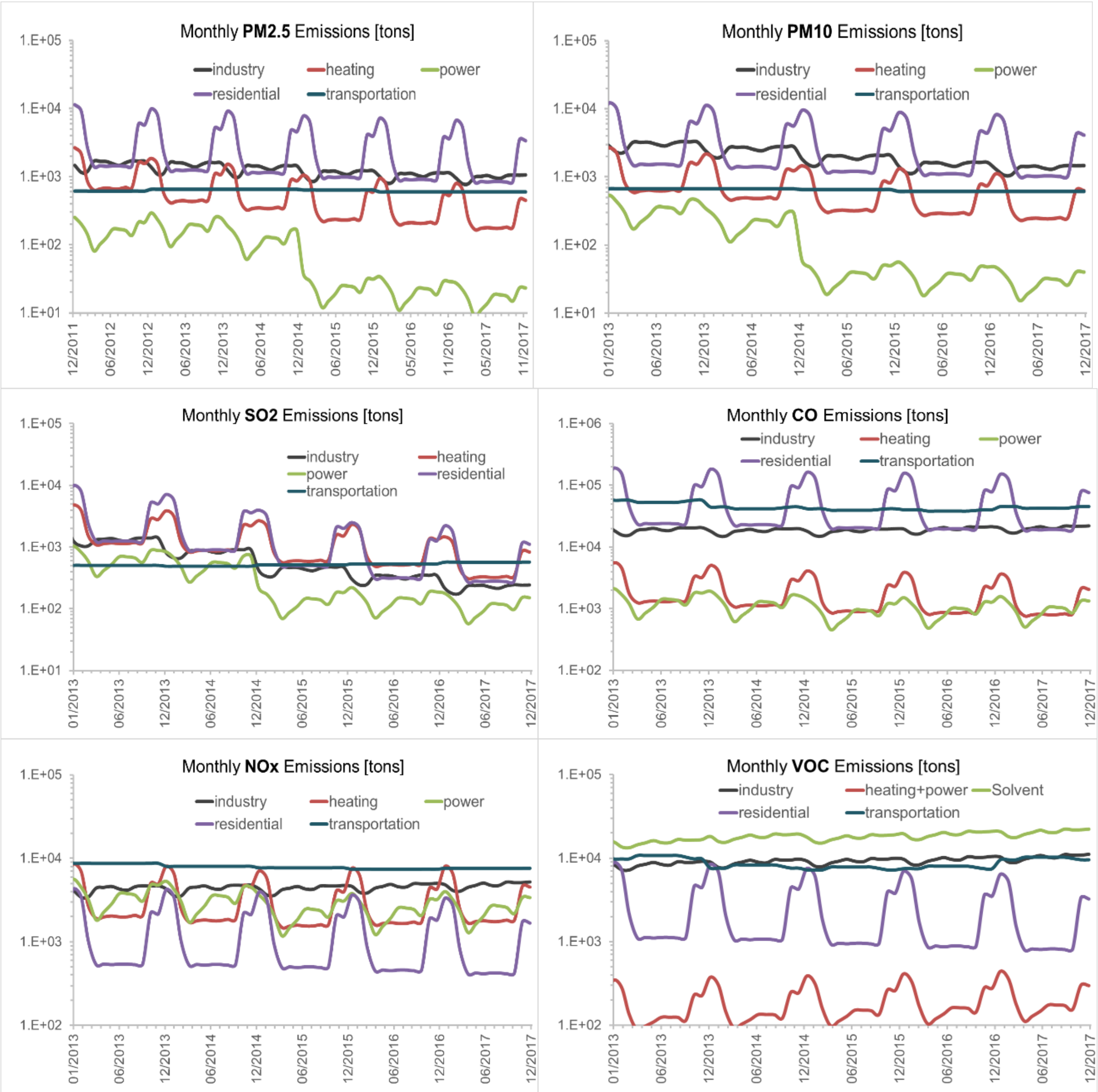


Figure S8. Monthly emission inventories of air pollutants in Beijing during 2013-2017. The emissions of $PM_{2.5}$, PM_{10}, NO_x, SO_2, CO in Beijing dropped by 35 %, 44 %, 11 %, 71 %, 17% from 76, 109, 260, 93, 1.7 Gg in 2013 to 49, 61, 231, 27, 1.4 Gg in 2017, respectively. Power sector represents the coal-fired, gas-fired and oil-fired power plants; industry sector includes two subsectors as industrial process and industrial boilers (to offer the mechanical energy); heating includes both industrial heating (to offer the thermal energy) and domestic heating (refers to centralized heating); residential sources are the urban and rural burning with traditional stoves with coal or biomass fuels; transportation includes both on-road and off-road traffic; solvent use contains all the subsectors which would use solvent during production processes, such as paint, ink, pharmaceutical production and household solvent use.

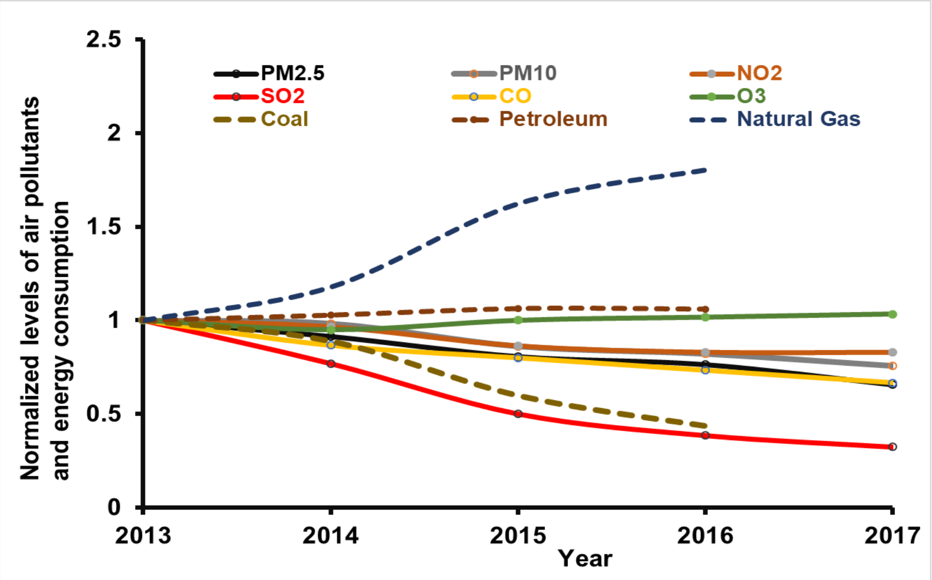


Figure S9. Normalized levels of air pollutants and energy consumption. The trend of SO_2 was very close to the normalized trend of coal consumption, but showed a faster decrease than trends of $PM_{2.5}$ and NO_2.

Table S1. Locations and categories of monitoring site

Station ID	Name	Category	Longtitude	Latitude
01	Wangshouxigong	Urban	116.37	39.87
02	Dingling	Rural	116.17	40.29
03	Dongsi	Urban	116.43	39.95
04	Tiantan	Urban	116.43	39.87
05	Nongzhanguan	Urban	116.47	39.97
06	Guanyuan	Urban	116.36	39.94
07	Haidianquwanliu	Urban	116.32	39.99
08	Shunyixincheng	Urban	116.72	40.14
09	Huairouzhen	Suburban	116.64	40.40
10	Changpingzhen	Suburban	116.23	40.20
11	Aotizhongxin	Urban	116.40	39.98
12	Gucheng	Suburban	116.26	39.93

Table S2: RF model performance for testing data set (in hourly time resolution).

Pollutants	RMSE	r2	FAC2	MB	MGE	NMB	NMGE	COE	IOA
PM2.5	17.9	0.95	0.94	0.62	10.00	0.01	0.14	0.81	0.91
PM10	43.1	0.79	0.87	1.46	27.10	0.01	0.26	0.57	0.79
NO2	14.3	0.78	0.95	-0.01	10.16	0.00	0.20	0.59	0.79
SO2	7.0	0.89	0.89	0.22	3.70	0.02	0.25	0.73	0.87
CO	0.4	0.86	0.96	0.01	0.24	0.01	0.21	0.67	0.84
O3	18.4	0.89	0.82	0.50	12.90	0.01	0.21	0.70	0.85

Note: FAC2 (fraction of predictions with a factor of two), MB (mean bias), MGE (mean gross error), NMB (normalised mean bias), NMGE (normalised mean gross error), COE (Coefficient of Efficiency), IOA (Index of Agreement) (Emery et al. 2017).

Table S3. Air Quality Standards. China's Air Quality Standards: GB 3095-2012, phase-in 2012-2016; WHO Air Quality Guidelines (2005). The Class 2 standards apply to urban areas.

Pollutants	Averaging time	China standards		WHO	unit
		Class 1	Class 2		
$PM_{2.5}$	annual	15	35	10	μg m^{-3}
	24 hours	35	75	25	μg m^{-3}
PM_{10}	annual	40	70	20	μg m^{-3}
	24 hours	50	150	50	μg m^{-3}
NO_2	annual	40	40	40	μg m^{-3}
	24 hours	80	80	-	μg m^{-3}
	hourly	200	200	200	μg m^{-3}
SO_2	annual	20	60	-	μg m^{-3}
	24 hours	50	150	20	μg m^{-3}
	hourly	150	500	-	μg m^{-3}
	10 min	-	-	500	μg m^{-3}
CO	annual	4	4	-	mg m^{-3}
	24 hours	10	10	-	mg m^{-3}
O_3	8-hour mean, daily max	100	160	100	μg m^{-3}
	hour	160	200	-	μg m^{-3}