Supplementary Material for Assessing Beijing's $PM_{2.5}$ Pollution: Severity, Weather Impact, APEC and Winter Heating

Xuan Liang, Tao Zou, Bin Guo, Shuo Li,
Haozhe Zhang, Shuyi Zhang, Hui Huang, Song Xi Chen

A.1 Seasonal Characteristics of the Wind in Beijing

Using the data of 60 months since January 2010, we summarize the distribution of wind directions and average speed for the four seasons in Fig. S1. It shows that the general wind distribution in Beijing was dominated by NW and SE, with winter by NW and summer by SE. The average speeds of NW in spring, autumn and winter were significantly higher than those of other directions. In summer, winds of all directions were calmer. In spring and summer, SE was much stronger than that in autumn and winter.

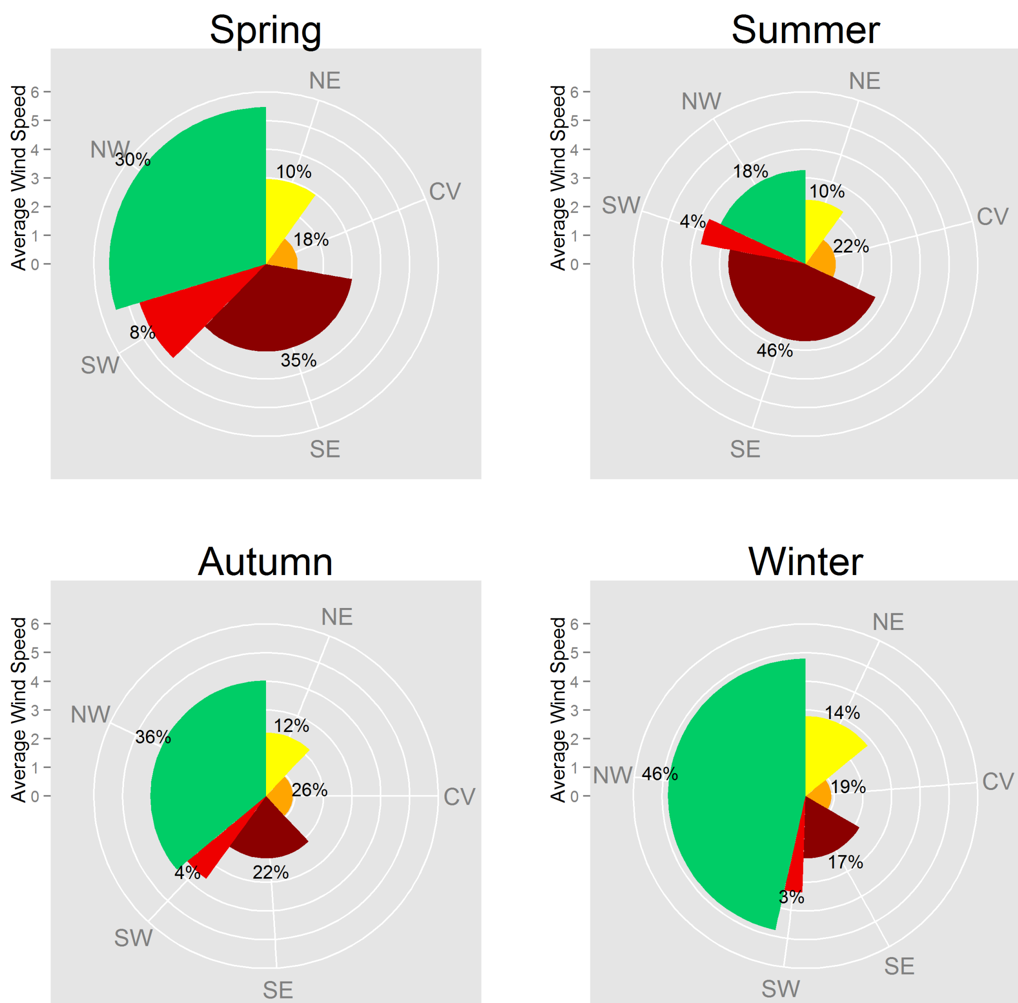

Fig. S1: **Seasonal wind patterns in Beijing.** The distribution of wind directions (shown via the width of angles) and the average speed (via the length of radius) for the four seasons.

A.2 Map of the Northern Part of the North China Plain

Fig. S2: **Map of the northern part of the North China Plain.** Beijing, in the north-west corner of the North China Plain, is hemmed in by Taihang Mountains to the west and Yanshan Mountains to the north. Source: Google Earth.

A.3 Influence of the Northerly Wind

Given the benefit of northerly wind, a natural question is "How long can Beijing's $PM_{2.5}$ remain below $35\mu g/m^3$ without substantial northerly wind?". Fig. S3 reports the lengths of Low PM periods, after excluding the time under northerly wind above $1.5m/s$ and $3.3m/s$, respectively. The medians and averages were dramatically shortened to 3 and 5 hours, respectively, if excluding northerly $> 1.5m/s$. They became 7 and 10 hours, respectively after excluding northerly $> 3.3m/s$. These were sharply smaller than 15 and 21 hours respectively for the entire Low PM period. These statistics indicate that Beijing can hold up for only 3 to 5 hours without the beneficial northerly wind.

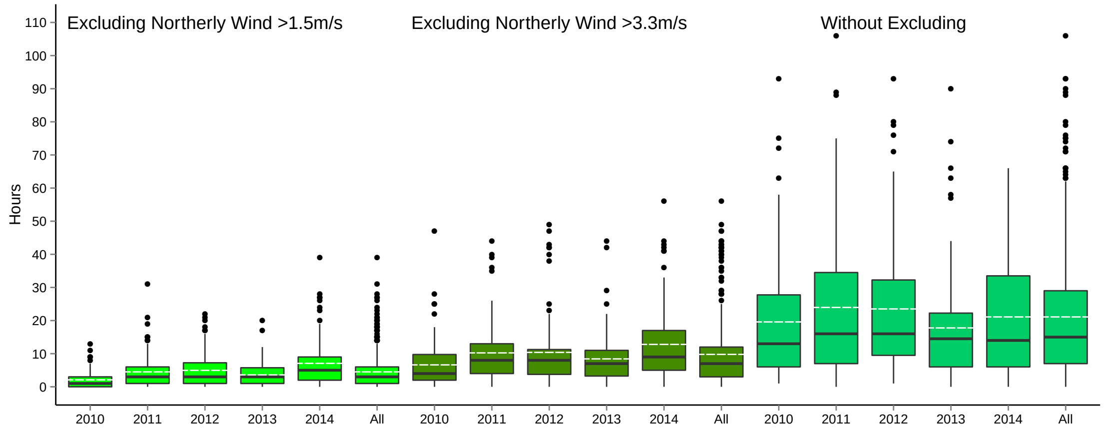

Fig. S3: Box-plots for the hours in Low PM periods excluding northerly wind $> 1.5m/s$ (left) and $3.3m/s$ (middle), and without excluding (right).

A.4 Model Assumptions

We provide the assumptions for the nonparametric estimation of Model (4.2).

We first define some notations.

(i) Denote f_X the density of a random variable X.

(ii) For a real valued function f, its $\|\cdot\|_p$-norm is defined as $\|f\|_p = \left(\int |f(x)|^p dx\right)^{1/p}$ and its $\|\cdot\|_\infty$-norm is defined as $\|f\|_\infty = sup|f(\cdot)|$. if f is continuous, then $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$.

(iii) Define $\sigma(X_t, t \in A)$ as the σ-Algebra generated by the random variable $X_t, t \in A$, where A is an index set of time.

(iv) The α-mixing coefficient of the strictly stationary process $V_t = (Y_t, X_t)$ for $t \in \mathbb{Z}$ is defined as

$$\alpha(k) = \sup_{\substack{B\in\sigma(V_s, s\le t)\\ C\in\sigma(V_s, s\ge t+k)}} |P(B\cap C) - P(B)P(C)| \text{ for } k \ge 1.$$

The process $\{V_t\}$ is said to be α-mixing if $\lim_{k\to\infty} \alpha(k) = 0$.

(v) Denote by $\mathcal{C}_{2,d}(b)$ the set of twice continuously differentiable real valued functions f on $\mathbb{R}^d$ such that $\|f\|_\infty \le b$ and $\|f^{(2)}\|_\infty \le b$, where $f^{(2)}$ denotes any partial derivative of order 2 for f.

For weakly dependent data, Bosq (1998) gives the assumptions for the asymptotic behaviour of the nonparametric regression estimator for estimating a general conditional mean $r(x) = E(g(Y)|X = x)$. In our case, $g(Y) = Y$ and $g(Y) = I(Y < y)$ for estimating $E(Y|X)$ and $F(Y < y|X)$, respectively.

Following Bosq (1998), for each month j, year i, the assumptions are given as follows.

(i) The joint density $f_{(X_{ijs},X_{ijt})}$ exists for any $s \neq t$ and belongs to $\mathcal{C}_{2,2q}(b)$ for a positive constant b, and

$$\sup_{s\neq t} \|f_{(X_{ijs},X_{ijt})} - f_{X_{ijs}} f_{X_{ijt}}\|_p < \infty,$$

for some $p \geq 2$, and q is the dimension of X_{ijt}.

(ii) The residuals $\{e_{ijt}\}$ satisfy $E(e_{ijt}|X_{ijt}, W_{ijt}) = 0$ and the second moment $E(e_{ijt}^2|X_{ijt} = x, W_{ijt} = w) = \sigma^2(x, w) < \infty$ for any x, w in their respective supports.

(iii) At each wind direction w, $V_{ijt} = (Y_{ijt}, X_{ijt})$ is strictly stationary and α-mixing, and the α-mixing coefficient given i and j satisfies $\alpha(k) \leq \gamma k^{-\beta}$, for $k \geq 1$ and constants $\gamma > 0$ and $\beta > \max(\frac{2(p-1)}{p-2}, q + 2)$.

(iv) At each wind direction w, $f_{X_{ijt}}(x)$ and $\phi(x) = \int y f_{V_{ijt}}(y, x)dy$ belong to $\mathcal{C}_{2,q}(b)$ for a positive constant b. And $f_{V_{ijt}}$ belongs to $\mathcal{C}_{2,q+1}(b)$. The density $f_{X_{ijt}}(x) > 0$ in the support of X_{ijt}.

(v) There exists a positive constant $a > 0$ such that $E\{\exp(a|Y_{ijt}|)\} < \infty$.

(vi) The kernel used for smoothing is a product kernel of a univariate kernel $k(\cdot)$ which satisfies $\int k(v)dv = 1$, $k(v) = k(-v)$ and $0 < \int v^2 k(v)dv = \kappa_2 < \infty$. And the bandwidth $h_s = O(n_{ij}^{-1/(q+4)})$ for $s = 1, \cdots, q$.

A.5 Model Diagnostics

We estimated Y_{ijt} by $\hat{Y}_{ijt}^{NP} = \hat{m}_{ij}(X_{ijt}, W_{ijt})$ and $\hat{Y}_{ijt}^{PL} = \hat{\beta} Y_{ij,t-1} + \hat{g}_{ij}(X_{ijt}, W_{ijt})$, where the superscript NP means using the model in (4.2) and PL means the partial linear model in (4.3). Therefore, we obtained the estimated residuals $\hat{e}_{ijt}^{NP} = Y_{ijt} - \hat{Y}_{ijt}^{NP}$ and $\hat{e}_{ijt}^{PL} = Y_{ijt} - \hat{Y}_{ijt}^{PL}$ for each observation t at year i and month j using the two models.

In order for the kernel regression estimation to work, the residuals $\{e_{ijt}\}$ in (4.2) are required to be stationary and weakly dependent. Hence, in the following we conducted two diagnostic checks on the stationarity and one diagnostic on the weak dependence of e_{ijt}. The diagnostics were based on the estimated residuals $\hat{e}_{ijt} = \hat{e}_{ijt}^{NP}$ and $\hat{e}_{ijt}^{PL}$ under the NP Model (4.2) and PL Model (4.3), respectively.

For the NP Model (4.2), we first diagnosed on the stationarity of the residuals by the augmented Dickey-Fuller (DF) test [45] and Phillips-Perron unit root test [46] to test on the null hypothesis

$$\text{H}_0\text{: } \{e_{ijt}\} \text{ is a unit root process versus } \text{H}_a\text{: } \{e_{ijt}\} \text{ is a stationary process.}$$

Both the augmented DF test and the Phillips-Perron unit root test are commonly used to test the stationarity of the residuals of regression models. That conducting both tests is to have extra insurance.

We performed the two tests on each of the 60 months from January, 2010 to December, 2014 and computed the p-values based on the asymptotic null distributions. For the augmented DF test, the null hypothesis of unit root was rejected in all the months but April, 2011 at 5% significance level. At the same time, the Phillips-Perron unit root test rejected the null hypothesis in all the months at the level of 1%. Both tests produced very consistent results, which constitute a strong evidence for $\{e_{ijt}\}$ in all the 60 months being stationary.

To check the weak dependence among the residuals, we used the adjusted rescaled range analysis [47] which is designed to test

$$\text{H}_0\text{: } \{e_{ijt}\} \text{ is a weakly dependent process versus } \text{H}_a\text{: } \{e_{ijt}\} \text{ is a long-range dependent process.}$$

The lag parameter chosen to estimate the long-run covariance in the test was $q = 10$ and $q = 20$ to ensure the robustness. Almost all the p-values for $q = 10$ exceeded 0.05 except those of March 2010 and May 2012. While in the situation $q = 20$, all p-values were larger than 0.05. These results suggest that we can not reject H_0 and the residuals were largely weakly dependent.

For the PL Model (4.3), we repeated the above diagnostic testings and reached the same conclusion as those under the NP model. In conclusion, the diagnostics above show that the residuals under both models were likely to be weakly stationary series, which is the basic condition needed for the kernel regression estimation adopted in our study.

A.6 Selection Frequencies of the Meteorological Variables

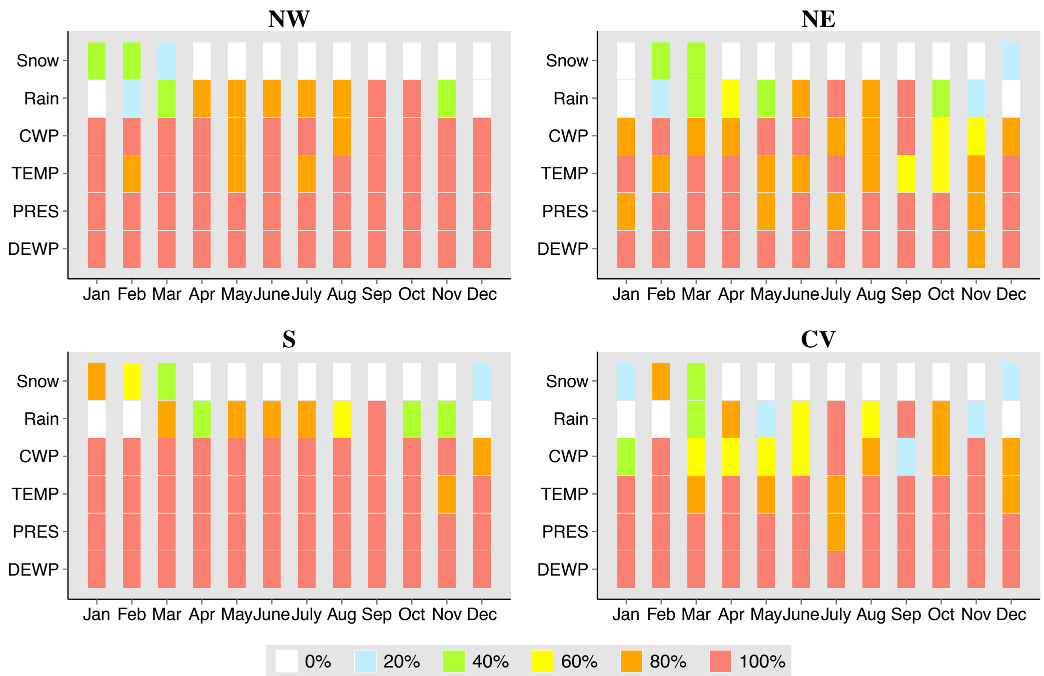

Fig. S4: The frequency of the meteorological variables selected for being useful of each month in five replications (five years from 2010 to 2014) under each of the wind direction

The bandwidths obtained by the cross-validation for each monthly model (4.2) contain information on the importance of variables in explaining the $PM_{2.5}$ concentration. If a covariate is redundant, the bandwidth selected by the cross-validation will diverge to the upper bound of the allowable range with probability tending to one; see [37] for details. As all the meteorological variables, except the wind direction, are continuous, the upper bound is the infinity.

We checked on the cross-validation bandwidths selected under each of the four wind directions (NW, NE, S, CV) for each month, and employed 15000 as the threshold to judge if a variable is redundant or not. The use of 15000 was made by the observation that the CV bandwidths were either small, in the range of 0.2 to 100 (mostly from 0.2 to 10), or above 15000. Fig.S4 reports the frequency of each variable which was selected for being useful among the 60 months under each wind direction. It shows that dew point and pressure were the most influential, followed by temperature and CWP. Rain and snow were significant in the summer and winter, respectively.It is not surprising to see CWP was less influential under the CV (calm and Variable wind) than the other wind directions, as it is hard to accumulate for this wind type.

A.7 Technical Details on the Adjusted Averages

The estimator of μ_{ij} is given in (5.2) . Taking expectation on both sides of (5.2), we get

$$\begin{aligned} E(\hat{\mu}_{ij}) &= (\sum_{a=1}^{n_{\cdot j}} n_{aj})^{-1} \left\{ \sum_{w=1}^{4} n_{ij} E\left(\hat{m}_{ij}(X_{ijt}, W_{ijt}) I(W_{ijt}=w)\right) \right. \\ & \left. + \sum_{w=1}^{4} \sum_{a=1, a\neq i}^{n_{\cdot j}} n_{aj} E\left(\hat{m}_{ij}(X_{ajt}, W_{ajt}) I(W_{ajt}=w)\right) \right\}. \end{aligned} \tag{A.1}$$

By the assumption of weak stationarity, we have

$$\begin{aligned} \sum_{w=1}^{4} E\left(\hat{m}_{ij}(X_{ijt}, W_{ijt}) I(W_{ijt}=w)\right) &= \sum_{w=1, a\neq i}^{4} E\left(\hat{m}_{ij}(X_{ajt}, W_{ajt}) I(W_{ajt}=w)\right) \\ &= \sum_{w=1}^{4} \int \hat{m}_{ij}(x, w) f_{\cdot j}(x, w) dx. \end{aligned} \tag{A.2}$$

Under the assumptions (i)-(vi) in section A.4, it was proved in [34] that

$$\hat{m}_{ij}(x, w) = m_{ij}(x, w) + O_p(n_{ij}^{-2/(q+4)}), \tag{A.3}$$

where q is the dimension of X_{ijt}.

Plugging (A.3) into (A.2) and (A.1), we get

$$E(\hat{\mu}_{ij}) = \sum_{w=1}^{4} \int m_{ij}(x, w) f_{\cdot j}(d) dx + O(n_{ij}^{-2/(q+4)}), \tag{A.4}$$

which implies

$$E(\hat{\mu}_{ij}) = \mu_{ij} + O(n_{ij}^{-2/(q+4)}). \tag{A.5}$$

This means that the proposed $\hat{\mu}_{ij}$ is an asymptotically unbiased estimator of the true μ_{ij}. For the variance of $\hat{\mu}_{ij}$, we can use similar techniques in [37] to prove $\hat{\mu}_{ij}$ is a consistent estimator.

A.8 The Adjusted Comparison of $PM_{2.5}$ between Two Years

In order to make a fair comparison of the average pollution levels by excluding the weather effects between year i_1, i_2, we need to compare the average and the percentiles of $PM_{2.5}$ via controlling the weather variables. In this section, we concentrate on the averages and the approach for percentiles is similar.

To control weather conditions, we need to compare the averages of $PM_{2.5}$ given the same weather variables $(X_{ijt}, W_{ijt}) = (x, w)$ for year i_1 and i_2, namely,

$$E(Y_{i_1jt}|X_{i_1jt} = x, W_{i_1jt} = w) - E(Y_{i_2jt}|X_{i_2jt} = x, W_{i_2jt} = w). \tag{A.6}$$

Then an aggregated version of the comparison (A.6) concerning all the possible realizations of the weather variables is

$$\sum_{w=1}^{4} \int \{E(Y_{i_1jt}|X_{i_1jt} = x, W_{i_1jt} = w) - E(Y_{i_2jt}|X_{i_2jt} = x, W_{i_2jt} = w)\} f_{\cdot j}(x, w)dx = \mu_{i_1j} - \mu_{i_2j},$$

say, the difference of the adjusted averages of $PM_{2.5}$ in the two years. By plugging in the estimators we propose, we can get the adjusted comparison as $\triangle\hat{\mu}_{i_1-i_2,j} = \hat{\mu}_{i_1j} - \hat{\mu}_{i_2j}$.

To quantitatively evaluate the adjusted differences between two months, we need to check if $\Delta\mu_{i_1-i_2,j}$ is significantly different from zero. Specifically, when $\Delta\hat{\mu}_{i_1-i_2j} > 0$, we construct an hypothesis testing problem

$$\text{H}_0\text{: } \Delta\mu_{i_1-i_2,j} = 0 \text{ vs } \text{H}_a\text{: } \Delta\mu_{i_1-i_2,j} > 0.$$

While in the situation $\Delta\hat{\mu}_{i_1-i_2,j} < 0$, we use the following test,

$$\text{H}_0\text{: } \Delta\mu_{i_1-i_2,j} = 0 \text{ vs } \text{H}_a\text{: } \Delta\mu_{i_1-i_2,j} < 0.$$

The test statistic is $T = \frac{\Delta\hat{\mu}_{i_1-i_2,j} - \Delta\mu_{i_1-i_2,j}}{\hat{\sigma}_{i_1-i_2,j}}$ where $\hat{\sigma}_{i_1-i_2,j}$ is obtained by the block bootstrap estimation method. The asymptotic null distribution of T is a standard normal distribution $Z \backsim N(0, 1)$. We reject the null hypothesis when $|T| > Z_\alpha$, where α is the significance level of the test and Z_α is the upper α percentile of the standard normal distribution.

References

46. Said, S.E., Dickey, D.A. (1984) Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599-607. (`doi:10.1093/biomet/71.3.599`)

47. Phillips, P.C., Perron, P. (1988) Testing for a unit root in time series regression. *Biometrika*, 75(2):335-346. (`doi:10.1093/biomet/75.2.335`)

48. Lo, A.W. (1991) Long-term memory in stock market prices. *Econometrica*, 59(5):1279-1313. (`doi:10.3386/w2984`)

Table S1: Differences in the adjusted averages, medians and 90th percentiles of $PM_{2.5}$ concentration ($\mu g/m^3$) between year 2013 and year 2012, and between year 2014 and year 2012. The quantities inside the parentheses are p-values for testing the null hypothesis of no underlying difference versus the underlying difference being positive or negative depending on the sign of the difference, based on the block size $l = 12$ in the bootstrap estimation of the standard error. The zero in the parentheses means the corresponding p-value being less than 0.01.

(a) **2013-2012**

Month	1	2	3	4	5	6
Average	0.7(0.47)	23.8(0)	6.2(0.18)	10.7(0.07)	-5.5(0.1)	8.1(0.12)
Median	-14.9(0.06)	-4.9(0.34)	8.2(0.15)	14.8(0.06)	-4.9(0.16)	13.4(0.15)
90th Percentile	10.1(0.39)	96.3(0)	-5.1(0.42)	-5(0.4)	-18.5(0.08)	-6.7(0.28)
Month	**7**	**8**	**9**	**10**	**11**	**12**
Average	1.2(0.38)	-12(0.02)	20.4(0)	11.2(0.16)	17.7(0.01)	34.3(0)
Median	5(0.11)	-3.3(0.3)	23.5(0)	26.6(0.01)	13.3(0.19)	14.6(0.08)
90th Percentile	-11.7(0.21)	-41.9(0)	36.9(0.02)	5.1(0.43)	65.9(0)	60.8(0)

(b) **2014-2012**

Month	1	2	3	4	5	6
Average	-8.8(0.24)	47.1(0)	20.3(0.01)	1.2(0.42)	-7.9(0.01)	-30.3(0)
Median	5(0.35)	31.2(0.01)	23(0.02)	26.3(0)	0(0.5)	-26.8(0.01)
90th Percentile	-91.2(0.04)	177.4(0)	15.2(0.28)	33(0.14)	-30.2(0)	-60.4(0)
Month	**7**	**8**	**9**	**10**	**11**	**12**
Average	34.3(0)	-6.6(0.11)	-2.7(0.21)	17.3(0.01)	-0.9(0.46)	2.6(0.35)
Median	37.6(0)	1.7(0.39)	-1.7(0.39)	33.2(0)	-28.2(0)	-4.9(0.32)
90th Percentile	63.8(0)	-21.8(0.01)	-6.7(0.15)	60.8(0)	70.9(0.01)	-5.1(0.44)

Table S2: Annual adjusted averages and percentiles of $PM_{2.5}$ concentration ($\mu g/m^3$) with the block size $l = 12$ in the bootstrap estimation of the standard errors. The "Original" and "Adjusted" are the averages or percentiles before and after the adjustment, respectively. "SE" is the estimated standard error of the adjusted $PM_{2.5}$ concentration using the bootstrap.

		Average	10%	25%	50%	75%	90%
	Original	104.7	21.6	39.1	78.0	147.6	233.1
2010	Adjusted	101.3	20.3	38.9	77.8	146.1	223.4
	SE	2.0	0.8	1.6	2.5	3.8	6.1
	Original	99.0	18.1	33.4	74.0	145.9	219.8
2011	Adjusted	97.6	15.1	31.1	75.9	147.4	236.2
	SE	2.1	1.2	2.0	3.5	3.9	6.5
	Original	90.7	13.8	27.1	68.1	135.6	205.2
2012	Adjusted	91.5	14.3	28.3	71.8	136.9	213.8
	SE	1.7	0.7	1.4	2.4	3.3	5.2
	Original	101.7	16.8	32.6	77.8	150.0	223.8
2013	Adjusted	101.2	16.5	32.7	79.4	151.6	229.4
	SE	1.9	0.7	1.4	2.5	4.2	4.7
	Original	98.0	15.0	32.5	74.7	145.0	214.9
2014	Adjusted	96.9	14.8	32.3	79.8	149.3	231.0
	SE	2.1	0.7	1.9	2.6	5.0	6.7

Table S3: APEC effect: the original and the adjusted averages and percentiles (standard deviation) of $PM_{2.5}$ concentration ($\mu g/m^3$) with the block size $l = 12$ in the bootstrap estimation of the standard errors. The "Original" and "Adjusted" are the averages or percentiles before and after the adjustment, respectively.

		Original				Adjusted			
Period	year	Average	50%	75%	90%	Average	50%	75%	90%
Nov3-12	2010	95.3	67.0	129.0	217.8	70.2(7.76)	53.5(17.17)	122.8(13.13)	184.9(49.79)
	2011	68.5	55.0	90.2	123.0	61.3(5.53)	56.8(5.35)	104.4(11.86)	188.3(47.46)
	2012	61.2	39.0	98.2	134.1	70.5(6.19)	65.1(10.25)	96.7(9.49)	134.6(15.1)
	2013	85.6	46.5	159.2	190.1	112.2(10.65)	102.7(24.93)	173.2(14.12)	234.2(28.08)
	2014	57.7	42.0	80.0	144.2	63.6(8.04)	43.5(12.39)	96.7(15.79)	158.1(26.09)
Nov6-12	2010	97.8	60.5	138.5	241.3	74.3(14.16)	28.6(37.42)	121.1(20.57)	234.2(53.98)
	2011	69.8	49.0	90.0	140.5	62.3(6.92)	51.8(5.09)	88.4(13.95)	163.1(43.52)
	2012	59.5	48.5	98.2	116.6	62.3(6.34)	58.5(11.73)	90(10.62)	117.8(7.56)
	2013	78.6	47.0	148.2	178.3	85.5(9.59)	63.5(25.11)	163.1(15.99)	186.6(29.33)
	2014	49.6	41.0	77.0	99.2	52.2(6.72)	40.2(13.71)	88.4(11.39)	112.7(18.33)

Table S4: Heating effect: the original and the adjusted averages and percentiles (standard deviation) of $PM_{2.5}$ concentration ($\mu g/m^3$) in non-heating and heating period in November with the block size $l = 12$ in the bootstrap estimation of the standard errors. The "Original" and "Adjusted" are the averages or percentiles before and after the adjustment, respectively.

		Original				Adjusted			
	time	Average	50%	75%	90%	Average	50%	75%	90%
2010	Nov1-14	81.5	49.5	117.5	186.7	66(5.02)	53.5(7.63)	112.7(10.99)	163.1(21.65)
	Nov15-30	179.9	154.0	249.0	399.8	186.9(17.58)	163.1(22.7)	327.7(43.55)	494.9(78.65)
2011	Nov1-14	78.0	56.0	102.0	153.0	71.7(7.66)	70.1(11.59)	124.5(15.05)	201(31.22)
	Nov15-30	137.4	123.0	205.0	285.0	129.7(11.65)	129.5(22.8)	221.3(15.2)	302.4(16.18)
2012	Oct20-Nov2	105.6	63.0	171.5	267.7	63.2(7.26)	36.9(7.54)	83.4(19.59)	231.4(29.63)
	Nov3-18	69.9	42.5	114.0	169.4	99.2(8.18)	96.7(12.26)	159.7(15.25)	201(13.09)
2013	Nov1-14	115.8	89.5	185.0	266.0	104.1(6.83)	88.4(14.09)	178.2(14.42)	266.9(20.54)
	Nov15-30	69.3	33.0	93.8	190.7	136.6(13.37)	119.5(14.53)	231.4(29.83)	348(48.1)
2014	Nov1-14	50.2	34.0	74.0	117.4	50.9(4.94)	33.6(5.98)	81.7(9.05)	139.6(19.33)
	Nov15-30	153.0	117.0	254.2	333.9	130.8(10.77)	93.4(22.02)	251.7(30.5)	363.2(20.11)

Table S5: Heating effect: the original and the adjusted averages and percentiles (standard deviation) of $PM_{2.5}$ concentration ($\mu g/m^3$) in non-heating and heating period in March with the block size $l = 12$ in the bootstrap estimation of the standard errors. The "Original" and "Adjusted" are the averages or percentiles before and after the adjustment, respectively.

		Original				Adjusted			
	time	Average	50%	75%	90%	Average	50%	75%	90%
2010	Mar8-22	96.0	67.0	116.0	219.4	115.7(11.75)	70.4(6.58)	116.1(20.28)	216.2(23.91)
	Mar23-Apr7	83.7	63.0	110.0	184.0	88.1(5.47)	73.7(6.14)	134.6(10.64)	184.9(10.35)
2011	Mar1-15	62.3	23.0	67.0	209.0	107.5(13.38)	67.1(7.71)	121.1(25.98)	277(50.98)
	Mar16-31	52.4	30.0	76.0	121.2	72.8(4.54)	67.1(6.8)	112.7(5.07)	164.8(16.17)
2012	Mar4-18	89.7	70.5	141.5	189.6	113.3(9.58)	93.4(8.25)	166.4(11.4)	282.1(50.24)
	Mar19-Apr3	81.4	51.5	108.8	193.0	66(6.48)	44.1(8.9)	86.8(9.95)	149.7(23.62)
2013	Mar3-17	163.5	144.0	235.0	329.5	130.4(9.67)	116.1(12.4)	201(18.36)	292.2(23.79)
	Mar18-Apr2	95.7	86.5	128.0	210.0	79.7(5.65)	72.1(8.79)	121.1(9.68)	198.3(19.05)
2014	Mar1-15	94.9	74.0	154.0	225.1	126(10.52)	119.5(19.26)	206.1(18.16)	256.7(13.13)
	Mar16-31	125.1	96.0	189.5	277.0	76(9.18)	75.3(13.47)	149.7(15.31)	216.2(29.1)

Table S6: The Coefficients (p-value) of the linear part in Model (4.3). The zero in the parentheses means the corresponding p-value being less than 0.001.

Year-Month	pm2.5lag
2014-1	0.85(0)
2014-2	0.78(0)
2014-3	0.38(0)
2014-4	0.94(0)
2014-5	0.62(0)
2014-6	0.6(0)
2014-7	0.81(0)
2014-8	0.81(0)
2014-9	0.77(0)
2014-10	0.95(0)
2014-11	0.9(0)
2014-12	1.02(0)

Table S7: The Coefficients (p-value) of the linear part in Model (8.1). The units of linear covariates are $\mu g/m^3$ except CO with unit being $100\mu g/m^3$. The zero in the parentheses means the corresponding p-value being less than 0.001.

Year-Month	SO2.lag	NO2.lag	CO.lag
2014-1	-0.15(0.598)	1.24(0)	3.93(0)
2014-2	1.21(0)	1.13(0)	3.52(0)
2014-3	0.17(0)	0.48(0)	0.99(0)
2014-4	-0.01(0.948)	0.07(0.15)	5.74(0)
2014-5	0.82(0.25)	0.18(0.016)	4.28(0.033)
2014-6	0.15(0.201)	0.12(0)	2.39(0)
2014-7	1.18(0.072)	0.03(0.756)	4.46(0)
2014-8	1.1(0.007)	0.16(0.008)	7.19(0)
2014-9	0.08(0.741)	0.31(0)	0.23(0)
2014-10	-0.08(0.827)	0.77(0)	5.64(0)
2014-11	0.91(0)	0.82(0)	3.29(0)
2014-12	1.22(0)	0.98(0)	1.64(0.01)

Table S8: The Coefficients (p-value) of the linear part in Model (8.2). The units of linear covariates are $\mu g/m^3$ except CO with unit being $100\mu g/m^3$. The zero in the parentheses means the corresponding p-value being less than 0.001.

Year-Month	pm2.5lag	SO2.lag	NO2.lag	CO.lag
2014-1	0.55(0)	0.11(0.485)	0.82(0)	1.39(0)
2014-2	0.77(0)	0.25(0.118)	0.41(0)	1.31(0)
2014-3	0.22(0)	-0.02(0.61)	0.28(0)	1.06(0)
2014-4	0.67(0)	0.03(0.749)	0.1(0.009)	1.86(0)
2014-5	0.52(0)	0.11(0.687)	0.11(0.006)	2.37(0.062)
2014-6	0.61(0)	-0.2(0.002)	0.08(0)	0.98(0)
2014-7	0.69(0)	0.28(0.248)	0(0.925)	1.82(0)
2014-8	0.51(0)	0.39(0.229)	0.09(0.163)	4.11(0)
2014-9	0.7(0)	-0.08(0.579)	0.18(0)	0.06(0)
2014-10	0.86(0)	0.2(0.034)	0.27(0)	0.34(0.221)
2014-11	0.56(0)	0.4(0)	0.36(0)	1.42(0)
2014-12	0.58(0)	0.78(0.003)	0.61(0)	0.6(0.195)