

# BMJ Open Bayesian model averaging method for evaluating associations between air pollution and respiratory mortality: a time-series study

Xin Fang,<sup>1</sup> Runkui Li,<sup>2,3</sup> Haidong Kan,<sup>4,5</sup> Matteo Bottai,<sup>1</sup> Fang Fang,<sup>6</sup> Yang Cao<sup>1,7</sup>

**To cite:** Fang X, Li R, Kan H, *et al.* Bayesian model averaging method for evaluating associations between air pollution and respiratory mortality: a time-series study. *BMJ Open* 2016;**6**:e011487. doi:10.1136/bmjopen-2016-011487

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-011487>).

Received 15 February 2016  
Revised 2 May 2016  
Accepted 10 June 2016

## ABSTRACT

**Objective:** To demonstrate an application of Bayesian model averaging (BMA) with generalised additive mixed models (GAMM) and provide a novel modelling technique to assess the association between inhalable coarse particles (PM<sub>10</sub>) and respiratory mortality in time-series studies.

**Design:** A time-series study using regional death registry between 2009 and 2010.

**Setting:** 8 districts in a large metropolitan area in Northern China.

**Participants:** 9559 permanent residents of the 8 districts who died of respiratory diseases between 2009 and 2010.

**Main outcome measures:** Per cent increase in daily respiratory mortality rate (MR) per interquartile range (IQR) increase of PM<sub>10</sub> concentration and corresponding 95% confidence interval (CI) in single-pollutant and multipollutant (including NO<sub>x</sub>, CO) models.

**Results:** The Bayesian model averaged GAMM (GAMM+BMA) and the optimal GAMM of PM<sub>10</sub>, multipollutants and principal components (PCs) of multipollutants showed comparable results for the effect of PM<sub>10</sub> on daily respiratory MR, that is, one IQR increase in PM<sub>10</sub> concentration corresponded to 1.38% vs 1.39%, 1.81% vs 1.83% and 0.87% vs 0.88% increase, respectively, in daily respiratory MR. However, GAMM+BMA gave slightly but noticeable wider CIs for the single-pollutant model (−1.09 to 4.28 vs −1.08 to 3.93) and the PCs-based model (−2.23 to 4.07 vs −2.03 to 3.88). The CIs of the multiple-pollutant model from two methods are similar, that is, −1.12 to 4.85 versus −1.11 versus 4.83.

**Conclusions:** The BMA method may represent a useful tool for modelling uncertainty in time-series studies when evaluating the effect of air pollution on fatal health outcomes.

## INTRODUCTION

Numerous time-series studies have indicated a positive association of ambient inhalable coarse particles, including particulate matters with diameters >2.5 µm and <10 µm

## Strengths and limitations of this study

- Provide a novel modelling technique allowing for the modelling uncertainty derived from knots selection for conventional GAMM to assess the association between air pollutants and adverse health outcomes.
- Provide robust effect estimation from time-series studies on PM<sub>10</sub> and fatal health outcomes.
- Uncertainty from variable selection and other sources was not investigated.
- No lag effect of PM<sub>10</sub> on respiratory mortality was examined.

(PM<sub>2.5–10</sub>) and PM<sub>10</sub> with daily respiratory death counts.<sup>1–6</sup> In time-series studies, one major methodology concern is potential confounding due to factors that vary on similar timescales as the pollutant concentrations or outcome. Although time-series studies have substantially strengthened the evidence base for the adverse health effect of PM<sub>10</sub>, methodological development of time-series studies to better adjust for confounding is still fully justified.<sup>7</sup> The impact of potential confounders, for example, weather and time, on the association of PM<sub>10</sub> with mortality or other health outcomes can be non-linear and may vary with season. Thus, a wide variety of approaches have been developed and applied in modelling and estimating the non-linear functions of continuous confounders in recent years. Prominent examples are smoothing splines,<sup>8</sup> penalised basis splines,<sup>9</sup> adaptive regression splines<sup>10–11</sup> and local polynomials.<sup>12</sup> These methods allow for greater flexibility in data modelling, because they relax the linearity assumption traditionally required in standard parametric methods.

The generalised additive mixed model (GAMM), an extension of the generalised additive model (GAM), has become a widely used method for evaluating short-term effect of air pollution. The fact that it allows for



CrossMark

For numbered affiliations see end of article.

## Correspondence to

Dr Xin Fang; [xin.fang@ki.se](mailto:xin.fang@ki.se)

serial correlations and spatial designs makes it a popular method in environmental epidemiological studies.<sup>13</sup> However, often there are several competing models to select from (ie, using different combinations of confounders), and the process of selecting the best/optimal model often varies and evaluating one's ultimate selection among the others is difficult. Model uncertainty can be significant, as selection of models might lead to largely different conclusions, but often the classical approach conditioning on a single presumed model ignores or underestimates such uncertainty.<sup>14</sup>

One common way to comprehend the problem is to conduct a 'sensitivity analysis' using a range of different plausible models to investigate the robustness of the estimates.<sup>15</sup> However, this still does not incorporate model uncertainty into effect estimates because it still requires selection of a single final model. To address this, researchers have proposed the Bayesian model averaging (BMA) method when assessing the triggering effect of air pollution on mortality.<sup>16</sup> BMA is a technique designed to account for the uncertainty of the model selection process. By averaging over many different competing models, BMA incorporates model uncertainty into the estimation of parameters and prediction. BMA has been applied successfully in many statistical model classes including linear regression, generalised linear models, Cox regression models and discrete graphical models, in all cases improving predictive performance.<sup>17</sup>

Our study re-examined the association between the concentration of PM<sub>10</sub> (including PM<sub>2.5-10</sub>) and the daily respiratory mortality rate (MR) from a time-series study design.<sup>1</sup> The goal of this paper is to demonstrate application of BMA within the GAMM frame and provide novel modelling techniques for time-series studies. We first demonstrated the application of the BMA method in GAMM. Second, we compared the estimates of three modelling techniques, that is, generalised linear mixed model (GLMM), optimal GAMM and Bayesian model averaged GAMM (GAMM+BMA). The study was approved by the Institutional Review Board of Basic Medical Sciences, Chinese Academy of Medical Sciences, China.

## MATERIALS AND METHODS

The data used in this study included the number of daily respiratory deaths, air quality and meteorological conditions from 1 January 2009 to 31 December 2010 in eight districts having air quality monitoring stations of a metropolitan area in Northern China. The geographic location of the eight districts was published elsewhere.<sup>1</sup> Respiratory mortality data were obtained from the regional Causes of Death Registry (CDR). All deaths in CDR were coded according to the 10th version of the International Classification of Diseases (ICD-10). The data collection was described in detail elsewhere.<sup>1 18</sup> In this study, ICD-10 codes J00–J98 were used to identify deaths due to respiratory diseases. In total, 10.38 million

permanent residents and 9559 respiratory deaths were included in the study. Air quality data included the concentrations of PM<sub>10</sub>, nitrogen oxides (NO<sub>x</sub>) and carbon monoxide (CO). The daily concentrations of these pollutants were presented as an average of 24 hourly measurements. To adjust for the effect of weather conditions, data on meteorological conditions, including mean daily temperature, relative humidity, wind speed and barometric pressure during the study period, were obtained from the local meteorological administration. Previous studies conducted in the same area mostly used citywide average pollutant concentrations,<sup>19 20</sup> whereas our study used pollutant concentrations measured in 11 stations of the study area and used station-specific pollutant concentrations. The spatial distribution of these monitoring stations over the districts was described elsewhere.<sup>1</sup>

The daily numbers of respiratory deaths in the eight districts were assumed to follow quasi-Poisson distribution to account for overdispersion by relaxing the distribution assumption that the variance equals the mean.<sup>21</sup> Given the non-linear relationship of the daily number of respiratory deaths to calendar day, temperature and barometric pressure, we used GAMM to account for these non-parametric components and district-level random effect. Natural splines were used to fit the non-linear trend of the mortality, adjusting for potential confounders, that is, meteorological conditions and day of the week (DOW). The full GAMM for single-pollutant included PM<sub>10</sub>, relative humidity, wind speed, DOW, smoothing functions for calendar day, temperature and barometric pressure, as well as random effect of districts, and can be expressed as:

$$\begin{aligned} \text{Log}(E(y_{i,t})) = & \beta_0 + \beta_1 \times \text{PM}_{10,t} + \beta_2 \\ & \times \text{relative humidity}_t + \beta_3 \\ & \times \text{wind speed}_t + \beta_4 \times \text{DOW}_t \\ & + s(\text{day}_t, n_1/\text{year}) \\ & + s(\text{temperature}_t, n_2) \\ & + s(\text{Barometric pressure}_t, n_3) + \beta \\ & \times \text{district}_i + Z_i\mu + \log(\text{population}_i) \quad (1) \end{aligned}$$

where  $E(y_{i,t})$  is the expected number of deaths in district  $i$  on  $t$ -th day, DOW is a dummy variable for day of week,  $\text{District}_i$  is a dummy variable for the eight districts and  $Z_i$  is a random intercept for districts  $i$ .  $s(\cdot)$ s are the smoothing functions realised by natural cubic spline with  $n_1$  knots per year to adjust for long-term temporal trend,  $n_2$  knots for temperature and  $n_3$  knots for barometric pressure.<sup>22</sup> Although natural cubic spline offers less flexibility at the limits where the second derivatives are zero, it presents a larger variance around the limits.<sup>23</sup> We used the annual average population size of each district as the offset in the Poisson regression model.<sup>24</sup> We also used the multipollutant model that included NO<sub>x</sub>

and CO to adjust for potential confounding from other pollutants. The optimal GAMM with the most appropriate number of knots for calendar day, temperature and relative humidity was determined by minimising Akaike information criterion (AIC).<sup>25</sup>

However, because we usually have rather limited knowledge about seasonal or longer time trend in the mortality time series, knot selection in GAMM might be complicated and oversmoothing or undersmoothing the series may potentially attenuate a true pollution effect.<sup>7</sup> In such a situation, the BMA method provides a plausible solution to incorporate the model uncertainty derived from the knot selection. The basic idea behind BMA is summarised as follows.<sup>17</sup>

Considering model  $M_k$  having the structure given by equation 1, if  $\beta$  is the coefficient of interest, then its posterior distribution given data  $D$  is

$$\text{pr}(\beta|D) = \sum_{k=1}^K \text{pr}(\beta|M_k, D) \text{pr}(M_k|D) \quad (2)$$

This is an average of the posterior distributions under each of the GAMM models considered, weighted by their posterior model probability. In equation (2),  $M_1, \dots, M_K$  are the models considered. The posterior probability for the model  $M_k$  is given by

$$\text{pr}(M_k|D) = \frac{\text{pr}(D|M_k) \text{pr}(M_k)}{\sum_{l=1}^K \text{pr}(D|M_l) \text{pr}(M_l)} \quad (3)$$

where  $M_k$  is one of the potential underlying models for data  $D$  with a prior probability  $\text{pr}(M_k)$  that it is true, and

$$\text{pr}(D|M_k) = \int \text{pr}(D|\beta_k, M_k) \text{pr}(\beta_k|M_k) d\beta_k \quad (4)$$

is the integrated likelihood of model  $M_k$ . In equation (4),  $\beta_k$  is the coefficient of model  $M_k$ ,  $\text{pr}(\beta_k|M_k)$  is the prior density of  $\beta_k$  under model  $M_k$ ,  $\text{pr}(D|\beta_k, M_k)$  is the likelihood and all probabilities are implicitly conditional on all models being considered.

The posterior mean and variance of  $\beta$  are defined as:

$$E[\beta|D] = \sum_{k=1}^K \hat{\beta}_k \text{pr}(M_k|D) \quad (5)$$

and

$$\text{Var}[\beta|D] = \sum_{k=1}^K (\text{Var}[\beta|D, M_k] + (\hat{\beta}_k - E[\beta|D])^2) \text{pr}(M_k|D), \quad (6)$$

where  $\hat{\beta}_k = E[\beta|D, M_k]$ .

The 95% Bayesian credible interval (CI) of  $\beta$  is

$$E[\beta|D] \pm 1.96\sqrt{\text{Var}[\beta|D]} \quad (7)$$

The posterior probability  $\text{pr}(M_k|D)$  in the above formula for each model was estimated by:<sup>26</sup>

$$\text{pr}(M_k|D) = \frac{\text{pr}(M_k) \cdot e^{-0.5(\text{BIC}_{M_k} - \overline{\text{BIC}})}}{\sum_{l=1}^K \text{pr}(M_l) \cdot e^{-0.5(\text{BIC}_{M_l} - \overline{\text{BIC}})}} \quad (8)$$

where  $\text{BIC}_{M_k}$  is the Bayesian information criterion (BIC) of model  $k$ , which extracts a penalty according to the number of terms in the model, and  $\overline{\text{BIC}}$  is the average of  $\text{BIC}_{M_l}$  ( $l=1, \dots, K$ ).

$\hat{\beta}_k$  in equation 5 and  $\text{Var}[\beta, M_k]$  in equation 6 are the posterior mean and variance of the parameter of interest, which derived from existing functions to estimate GAMMs may not correspond to the posterior mean and variance. In addition, we use BIC in equation 8 to approximate Bayes factors. Thus, the method presented here is an empirical approximation to a fully Bayesian form of model averaging, bridging classical frequentist and Bayesian estimation methods.<sup>27 28</sup>

**Table 1** Daily respiratory mortality rate and PM<sub>10</sub> concentrations by districts in the study area, 2009–2010

Districts	Population (in 1000)	Mortality rate (1/100 000 persons)		PM <sub>10</sub> (µg/m <sup>3</sup> )		NO <sub>x</sub> (µg/m <sup>3</sup> )		CO (mg/m <sup>3</sup> )	
		Median	P25–P75	Median	P25–P75	Median	P25–P75	Median	P25–P75
District 1	896	0.11	0–0.22	94.0	57–138	52.0	33–78	1.20	0.8–1.7
District 2	3001	0.10	0.06–0.13	106.5	67–151	72.0	50.5–109.5	1.30	0.85–1.9
District 3	851	0.24	0.12–0.35	110.3	73.5–159	70.5	50.5–107.5	1.38	1.0–2.1
District 4	2814	0.07	0.04–0.14	112.0	71–154	79.0	52–116	1.20	0.8–2.0
District 5	316	0.00	0–0.32	82.5	49–124	33.0	23–53	1.00	0.6–1.4
District 6	546	0.18	0–0.18	129.0	83–174	60.0	44–88	1.40	1.0–2.0
District 7	736	0.00	0–0.14	108.5	66–154	52.0	37–75	0.90	0.6–1.4
District 8	1218	0.25	0.08–0.33	105.5	68.5–150.5	73.0	53–107.5	1.35	0.95–2.0
Total	10 378	0.11	0–0.22	106.0	66–150	61.0	41–93	1.20	0.8–1.8

P25, the 25th percentile; P75, the 75th percentile.

In our study, the prior probability  $pr(M_k)$  was assumed to be from the uniform distribution, thus:

$$pr(M_k) = \frac{1}{K} \quad (9)$$

According to equation 5, the posteriors mean is averaged over all models, applying weights that depend on the degree to which data support each model. The weight

given by equation 8 incorporates the BIC component that penalises for dimensionality of the model. Therefore, the best models are weighted heavily in equation 5 by equation 8, which in turn heavily favours parsimonious models as well.

To consider correlations between  $PM_{10}$  and CO and  $NO_x$ , we introduced principal components (PCs) derived from principal component analysis (PCA) into the multipollutant models to exclude the impacts of

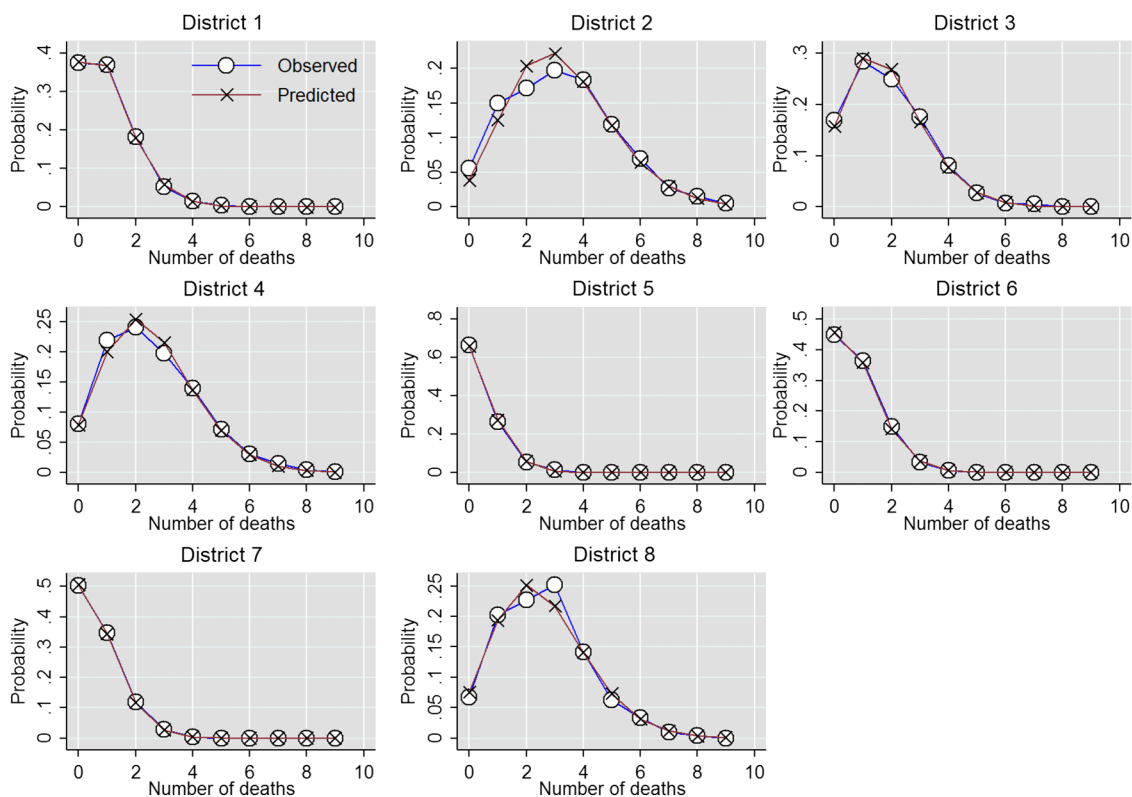
**Table 2** Meteorological conditions in the study area 2009–2010

	Mean	SD	Min	Q1	Median	Q3	Max
Air temperature (°C)	13.0	11.7	-12.5	1.7	14.7	24.3	34.5
Wind speed (m/s)	2.2	1.0	0.5	1.5	2.1	2.7	6.4
Relative humid (%)	51.0	19.2	13.0	35.0	52.0	67.0	92.0
Barometric pressure (kPa)	101.2	1.0	99.0	100.4	101.1	102.0	103.7

**Table 3** Pairwise Pearson correlation coefficients between pollutants and meteorological conditions

	$PM_{10}$	$NO_x$	CO	Temperature	Barometric pressure	Relative humidity
$NO_x$	0.4780*					
CO	0.5532*	0.8210*				
Temperature	-0.0157*	-0.3206*	-0.2939*			
Barometric pressure	-0.1845*	0.1535*	0.0892*	-0.8266*		
Humidity	0.2178*	0.1699*	0.3215*	0.3258*	-0.3121*	
Wind speed	-0.1413*	-0.4626*	-0.4800*	-0.0668*	0.0509*	-0.4859*

\* $p < 0.05$ .



**Figure 1** Observed proportion and predicted probability based on Poisson distribution of number of daily respiratory deaths in the study area between 2009 and 2010.

collinearity between the three pollutants.<sup>18</sup> We then transformed the regression coefficients of the PCs back to the regression coefficients of the original pollutants.

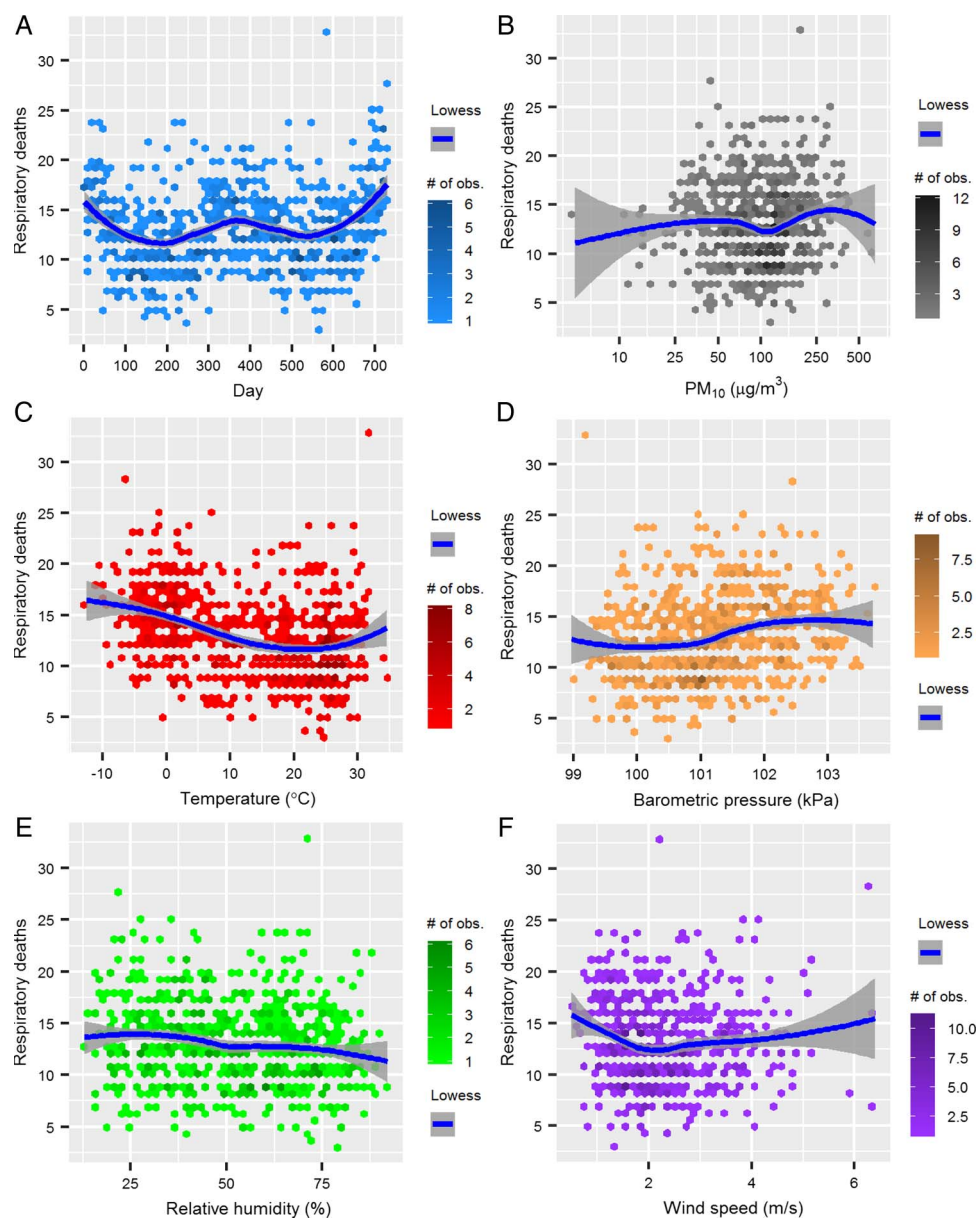
All analyses in our study were conducted in the statistical software Stata (V.14.1; StataCorp LP, College Station, Texas, USA) and using R software (V.3.2.3) packages 'mgcv', 'splines' and 'lme4'. Since <2% of the observations in the data set were incomplete, the listwise deletion method was used to handle missing values.

## RESULTS

The daily respiratory mortality rate (per 100 000 persons) and PM<sub>10</sub>, NO<sub>x</sub> and CO concentrations of the eight districts are shown in table 1. The highest daily mortality rates were found in districts 3 (median=0.24;

IQR=0.24) and 8 (median=0.25; IQR=0.25). During the 2-year study period, the annual median concentrations for PM<sub>10</sub>, NO<sub>x</sub> and CO were 106.0 µg/m<sup>3</sup>, 61.0 µg/m<sup>3</sup> and 1.20 mg/m<sup>3</sup>, respectively. The annual median concentrations of PM<sub>10</sub> and NO<sub>x</sub> were above the limits of Class II of the National Ambient Air Quality Standards of China (70 µg/m<sup>3</sup> for PM<sub>10</sub> and 50 µg/m<sup>3</sup> for NO<sub>x</sub>), but that for CO was below the national limit (4 mg/m<sup>3</sup>).<sup>29</sup>

The meteorological conditions of the same period in the study area are shown in table 2. The mean temperature was 13.0°C ranging from -12.5°C to 34.5°C, the mean relative humidity was 51.0% ranging from 13.0% to 92.0% and the mean barometric pressure was 101.2 kPa ranging from 99.0 to 103.7 kPa, a characteristic of a typical subhumid warm continental monsoon climate.



**Figure 2** Relationship between number of daily respiratory deaths and (A) days; (B) PM<sub>10</sub> concentrations and (C–F) meteorological conditions. Lowess; locally weighted scatterplot smoothing.

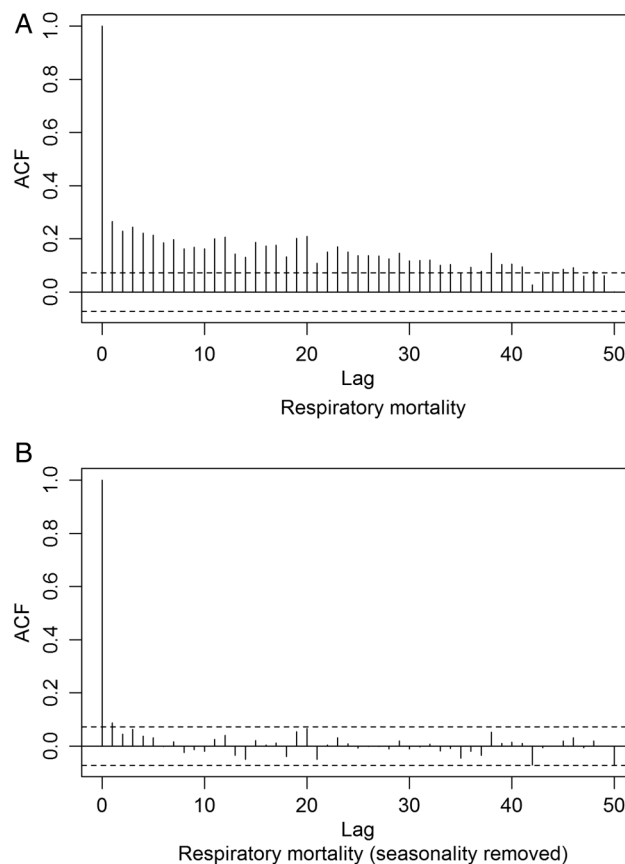
The pairwise Pearson's correlation coefficients between pollutants and meteorological conditions are shown in table 3. We observed strong linear correlation between temperature and barometric pressure ( $r=-0.83$ ,  $p<0.001$ ). To control for the collinearity, we included temperature, relative humidity and wind speed but not barometric pressure in the conventional GLMM analysis.

The observed and predicted (based on quasi-Poisson distribution) numbers of daily respiratory deaths of the eight districts between 2009 and 2010 corresponded well with no zero-inflation observed, supporting the distributional assumptions in GAMM analysis (figure 1). There were a clear temporal trend of daily number of respiratory deaths (figure 2A) and a non-linear relationship between daily number of respiratory deaths and temperature as well as barometric pressure (figure 2C, D), supporting the choice of including them as smoothing functions in GAMM. There was a moderate linear relationship of daily average relative humidity to daily number of respiratory deaths (figure 2E), and we therefore included daily average relative humidity as a linear component in the model. However, we did not observe a clear linear or non-linear relationship between daily number of respiratory deaths and wind speed (figure 2F). As a result, we included wind speed as a linear component first and as a smoothing function later in a sensitivity analysis.

We also examined the seasonality of respiratory mortality using autocorrelation function (ACF). The slow decrease in autocorrelation from lag 1 to lag 50 shown in the correlogram (figure 3A) indicates that there is some temporal trend in the mortality series. We removed this trend by regressing the mortality against a smoothing function of time. The correlogram of the residuals after removing the seasonality (figure 3B) shows substantially less autocorrelation. Thus, season is an important factor related strongly to  $PM_{10}$ , meteorological conditions and mortality as shown in figures 2 and 3.

There were moderate to strong correlations between  $PM_{10}$  and CO as well as  $NO_x$  (table 3), and we introduced the first and the second PCs derived from PCA into the multipollutant models. They accounted for 94.22% of variation of the three pollutants.

For GAMM analyses, we used different combinations of number of knots (from 3 or 4 to 24 knots for each smoothing function of calendar day, temperature and barometric pressure) and showed the results of combinations with the relative large posterior probability around the best model, that is, 12, 14 and 16 knots calendar day; 5, 6 and 7 knots for temperature and 4, 5 and 6 knots for barometric pressure, respectively. The knot combinations with convergence problem or extreme small posterior probability were excluded. The estimated coefficients, their corresponding SEs as well as AICs and BICs of 27 considered versions of single-pollutant (ie,  $PM_{10}$ ) GAMMs are shown in table 4. The estimated regression coefficients of  $PM_{10}$  changed little



**Figure 3** ACF for respiratory mortality for (A) raw data and (B) residuals after removing seasonality. ACF, autocorrelation functions.

with different knots for barometric pressure but increased with the increasing knots for temperature. However, when the number of knots of calendar day changed from 12 to 14 and from 14 to 16, the regression coefficients of  $PM_{10}$  showed a slight U-shape (table 4).

Estimated increases in respiratory MR for single-pollutant, multipollutant and PCA-based multipollutant models are presented in figure 4. The same knot combination for temperature, etc, was optimal across the single-pollutant and the multipollutant models. The results of the GLMMs, optimal GAMMs and GAMM+BMA are presented of in table 5. Only GLMM of the single-pollutant model confirmed a statistically significant association between  $PM_{10}$  and daily number of respiratory deaths, with the largest effect of  $PM_{10}$  3.07 (95% CI 0.91 to 5.27) per cent increase in daily respiratory MR per IQR increase in  $PM_{10}$  concentration. GAMM+BMA and the optimal GAMM of single-pollutant, multipollutants and PCA-based multipollutant showed comparable results for the effect of  $PM_{10}$  on daily respiratory MR, that is, one IQR increase in  $PM_{10}$  concentration corresponded to 1.38% vs 1.39%, 1.81% vs 1.83% and 0.87% vs 0.88% increase, respectively, in daily respiratory MR (table 5). However, by incorporating the uncertainty in knots selection, GAMM+BMA gave slightly but noticeable wider CIs for the single-pollutant

**Table 4** Coefficients of PM<sub>10</sub> of GAMMs for single-pollutant with different knots

Model	Number of knots (D, T, P)	$\hat{\beta}$	SE	AIC	BIC	Posterior probability
1	12, 5, 4	0.0001609643	0.0001499442	15063.93	15230.75	0.028997
2	12, 5, 5	0.0001609382	0.0001499458	15063.93	15230.75	0.028997
3	12, 5, 6	0.0001609357	0.0001499460	15063.93	15230.75	0.028997
4	12, 6, 4	0.0001611031	0.0001497939	15063.70	15230.51	0.032694
5	12, 6, 5	0.0001611031	0.0001497939	15063.70	15230.51	0.032694
6	12, 6, 6	0.0001611030	0.0001497939	15063.70	15230.51	0.032694
7	12, 7, 4	0.0001649193	0.0001497870	15063.35	15230.16	0.038947
8	12, 7, 5	0.0001649172	0.0001497871	15063.35	15230.16	0.038947
9	12, 7, 6	0.0001649211	0.0001497870	15063.35	15230.16	0.038947
10	14, 5, 4	0.0001592525	0.0001503313	15063.58	15230.39	0.034716
11	14, 5, 5	0.0001592566	0.0001503310	15063.58	15230.39	0.034716
12	14, 5, 6	0.0001592535	0.0001503312	15063.58	15230.39	0.034716
13	14, 6, 4	0.0001607594	0.0001500942	15063.40	15230.21	0.037985
14	14, 6, 5	0.0001607662	0.0001500939	15063.40	15230.21	0.037985
15	14, 6, 6	0.0001607661	0.0001500939	15063.40	15230.21	0.037985
16	14, 7, 4	0.0001646505	0.0001500780	15063.04	15229.85	0.045477
17	14, 7, 5	0.0001646507	0.0001500780	15063.04	15229.85	0.045477
18	14, 7, 6	0.0001646507	0.0001500780	15063.04	15229.85	0.045477
19	16, 5, 4	0.0001633165	0.0001502163	15063.64	15230.46	0.033522
20	16, 5, 5	0.0001633165	0.0001502164	15063.65	15230.46	0.033522
21	16, 5, 6	0.0001633165	0.0001502164	15063.65	15230.46	0.033522
22	16, 6, 4	0.0001650914	0.0001499675	15063.46	15230.27	0.036863
23	16, 6, 5	0.0001650913	0.0001499675	15063.46	15230.27	0.036863
24	16, 6, 6	0.0001650912	0.0001499675	15063.46	15230.27	0.036863
25	16, 7, 4	0.0001690050	0.0001499511	15063.10	15229.91	0.044133
26	16, 7, 5	0.0001690088	0.0001499509	15063.10	15229.91	0.044133
27	16, 7, 6	0.0001690044	0.0001499512	15063.10	15229.91	0.044133

AIC, Akaike information criterion; BIC, Bayesian information criterion; D, day; GAMMs, generalised additive mixed models; P, barometric pressure; T, temperature.

model (−1.09 to 4.28 vs −1.08 to 3.93) and the PCA-based model (−2.23 to 4.07 vs −2.03 vs 3.88). The CIs of the multiple-pollutant model from the two methods are similar, that is, −1.12 to 4.85 versus −1.11 versus 4.83. The results indicate that BMA provides inference about parameters taking account of the different knot selection strategies, this sometimes being an important source of uncertainty in additive model analysis, and in our example we have found that single model-based CIs tend to be narrow, which might increase the probability of false-positive finding.

In addition, the effect of the first PC in GAMM and GAMM+BMA was statistically significant (data not shown), potentially indicating a joint effect of PM<sub>10</sub>, NO<sub>x</sub> and CO on respiratory mortality. For sensitivity analysis, including wind speed as a linear component or a smoothing function in GAMM changed the results little (data not shown).

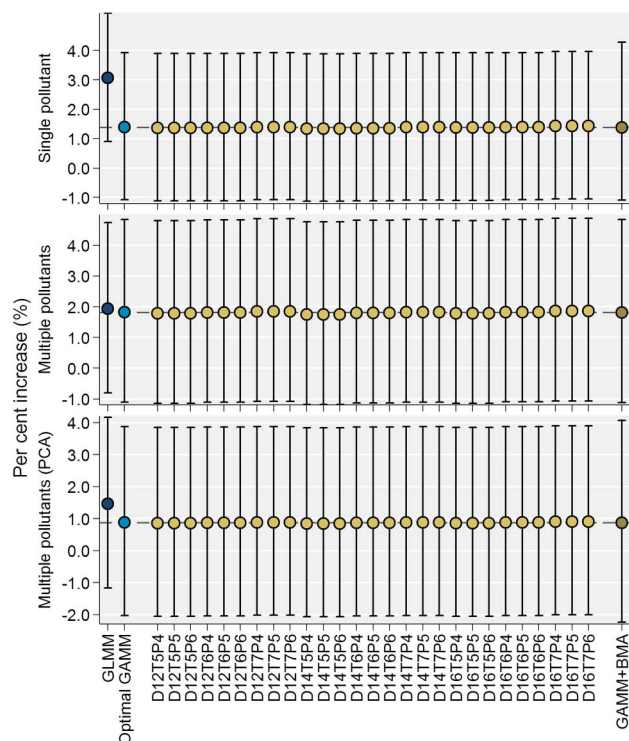
## DISCUSSION

Yang *et al*<sup>1</sup> and Zhang *et al*<sup>30</sup> had examined the triggering effect of PM<sub>10</sub> on respiratory mortality in the same area previously. However, the studies were based on GAM. Although GAM is a powerful method for modelling non-linear effects of continuous covariates in

regression models with non-Gaussian response, it cannot account for the between-cluster heterogeneity and within-cluster correlation of the pollutant concentrations. In recent years, smoothing based mixed model, that is, GAMM and its extensions have gained popularity in part for its ability to account for such limitations.<sup>13 31–33</sup>

Using a subset of multisite time-series data from Yang *et al*'s<sup>1</sup> study, we reinvestigated the associations between short-term exposure to ambient inhalable coarse particles PM<sub>10</sub> and daily number of deaths from respiratory diseases. By adding the random effect of districts to the additive predictors, we used GAMM to provide a unified likelihood framework for non-parametric regression for potentially correlated exposures.<sup>13</sup>

Standard statistical methods for estimating the association between air pollutants and adverse health outcomes often fail to incorporate the model uncertainty in effect estimates. For example, knot selection for splines used in GAMM is critical, because it determines the degree of smoothness in the smoothing function of time as well as the amount of residual temporal variation in mortality. In studies using GAMM to examine effect of pollutants, we usually have rather limited knowledge about the complexity of the seasonal and long-term trends in the mortality time series or in the pollution time series. Although there are often biological or



**Figure 4** Estimated per cent increase in daily respiratory deaths per IQR increase in PM<sub>10</sub> concentration in GLMM, optimal GAMM, GAMMs with different knots in day, temperature and pressure (indicated by D, T and P) and GAMM+BMA for single pollutant, multiple pollutants and PCA. BMA, Bayesian model averaging; GAMM, generalised additive mixed model; GLMM, generalised linear mixed model; PCA, principal component analysis.

mechanistic information that is applied, current approaches in choosing the number and location of knots are mainly data-driven<sup>34 35</sup> and are based on prior knowledge of the timescales where confounding is more likely to occur. More importantly, the single model selection at the end that ignores the entire process of getting to the final model and the failure to account for this uncertainty might bias our judgment.

Thus, there is concern about underestimated uncertainty in model selection in time-series air pollution studies. If a single ‘best’ model was used, the variance estimates for its coefficients will not fully reflect their

true uncertainties.<sup>36</sup> A coherent and conceptually simple way to take into account model uncertainty when making reference is BMA. In theory, BMA provides better average predictive performance than any single model, and this theoretical result has now been supported in practice in a range of applications involving different model classes and types of data.<sup>17</sup> While BMA is an attractive solution to the problem of model uncertainty, it is not yet part of the standard data analysis tools in epidemiological studies due to several practical difficulties, including impractical exhaustive summation of posterior distribution, the computational difficulty of integrals and the challenge of specifying prior distribution over competing models. In addition, owing to lack of official solution for GAMM in common commercial statistical software, model fit criteria such as the BIC are not handily available,<sup>26</sup> making it difficult to construct the model weight to weigh estimated coefficients of the individual models. Although Whitney and Ngo<sup>26</sup> demonstrated the application of BMA in the GAM frame, they only examined the piece-wise linear relationship between air pollutants and mortality, essentially limiting their algorithm to a GLMM.

In our study, we demonstrated the feasibility of implementing BMA in GAMM frame in R software environment. For single-pollutant and multipollutant models, our optimal GAMM and GAMM+BMA gave equivalent point estimates for effect of PM<sub>10</sub> on respiratory MR. Compared with previously published studies, our GAMM+BMA method for single-pollutant gave comparable results (per cent increases ranging from 0.87 to 1.38 vs 1.01 to 2.07<sup>1 18 30 37 38</sup>). However, our multipollutant models showed smaller effect of PM<sub>10</sub>, which was consistent with previous findings suggesting that the effect of PM<sub>10</sub> in multipollutant models was about two to three times smaller<sup>1 18</sup> or slightly reversed.<sup>30</sup> Although the relative increase of MR is rather small, taking into account the huge population (>10 million) in the study area, it still raises a severe public health challenge.

However, our main interest was not the estimated coefficients but the uncertainty of the estimation from different modelling strategies. The GAMM+BMA gave slightly but noticeable wider confidence even when only considering the model uncertainty from the knot. The averaged estimate across a set of potential valid models

**Table 5** Per cent increase in daily respiratory MR associated with an IQR increase in PM<sub>10</sub> concentration from GLMM, optimal GAMM and GAMM+BMA

Model	Single-pollutant		Multipollutant		Multipollutant (PCA)	
	Per cent	95% CI	Per cent	95% CI	Per cent	95% CI
GLMM	3.07	(0.91 to 5.27)	1.94	(−0.80 to 4.75)	1.47	(−1.17 to 4.17)
Optimal GAMM*	1.39	(−1.08 to 3.93)	1.83	(−1.11 to 4.83)	0.88	(−2.03 to 3.88)
GAMM+BMA	1.38	(−1.09 to 4.28)	1.81	(−1.12 to 4.85)	0.87	(−2.23 to 4.07)

\*Knots for days, temperature and barometric pressure are 14, 7 and 4, respectively.

BMA, Bayesian model averaging; GAMM, generalised additive mixed model; GLMM, generalised linear mixed model; MR, mortality rate; PCA, principal component analysis.



would derive more robust interval estimation for reducing the type I error. Future studies may use our technique to build a model-averaged dose–response relationship between  $PM_{10}$  and daily mortality to account for model uncertainty with respect to location and number of knots. Our method allows for a fully parametric characterisation of the effect of air pollutants on adverse health outcomes as well as adjusts for the non-parametric effects of other covariates.

There are also limitations in our study. First, we only considered the uncertainty from the knots selection but not from the selection of covariates and confounders, which is another major source of the uncertainty in estimating the effect of  $PM_{10}$ . However, this limitation can be easily overcome by including different covariates and smoothing functions in the GAMM. Second, GAMM sometimes demonstrates problems in convergence and SE estimation, which might result in the failure of coefficient estimation. The data with many zeros can cause such problems in a model with a log link, because a mean of zero corresponds to an infinite range of linear predictor values. Fundamentally, it is due to lack of identifiability, and it is the case in our study. However, the problem can be addressed using stricter convergence criteria and panelised splines.<sup>39</sup> Third, in our study, we did not investigate the lag effect of  $PM_{10}$  on respiratory mortality, because our interest lies on the uncertainty of the estimated association but not the association itself. However, previous studies in the same area indicated that the strongest effect of air pollutants on respiratory mortality was in day 0 (lag 0) and day 1 (lag 1), and the strongest cumulative effect was in 2-day moving average of day 0 and day 1.<sup>18</sup> We incorporated lag 1 of  $PM_{10}$  concentration in the single-pollutant and multipollutant models as a sensitivity analysis and estimated effects of lag 0  $PM_{10}$  on respiratory MR increase per IQR increase in lag 0  $PM_{10}$  reduced to 0.79 (−1.91 to 3.55) and 1.20 (95% CI −1.93 to 4.42), respectively. Similar results were found for GAMM+BAM. Thus, the effects of  $PM_{10}$  need to be investigated in further GAMM+BAM with different lag structures.

In our study, we selected a uniform distribution for the prior distribution. In Bayesian statistics, the choice of the prior distribution is often controversial. Different rules for selecting priors have been suggested in the literature.<sup>40</sup> The most intuitive solution is to use the uniform distribution as a non-informative prior when no information is available, although it does not integrate to 1 in most of the cases (this does not pose a major problem for Bayesian analyses). In view of the large sample size in our study, the data will dominate the posterior distribution (they will overwhelm the prior), so the selection of prior distribution would not pose much effect on our parameter estimation. Actually, we were informed by previous studies on the knot and variable selection, but we pretended to be uninformative in the current study in order to incorporate uncertainty (including inappropriate models) as more as possible. We would

like to use informative prior distribution to narrow the range of the included models and examine performance of different prior distributions in future studies.

In conclusion, there is an increasing interest in the use of GAMM to investigate the association between short-term exposure to  $PM_{10}$  and adverse health outcomes in time-series studies. Epidemiology studies incorporate different modelling strategies to adjust for confounding, making it difficult to compare results across studies. Furthermore, the uncertainty of model selection has rarely been considered and quantified in these studies. Using BMA in the GAMM frame is a promising approach to investigate the association of air pollution with adverse health using time-series data, as well as other applications. Naturally, BMA tends to produce larger SE than the models that ignore model uncertainty do. However, the conclusions of BMA are more robust than those derived from analyses dependent upon a particularly selected model. Future work should aim to extend the model averaging over more extensive families of models, employing the method to explore heterogeneity in other areas (such as confounder selection), developing corresponding package in R for widespread use and exploring implications of utilising these resulting estimates of effect and corresponding uncertainty on policy-making and decision-making.

#### Author affiliations

<sup>1</sup>Unit of Biostatistics, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

<sup>2</sup>College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China

<sup>3</sup>State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

<sup>4</sup>Key Laboratory of Public Health Safety of the Ministry of Education and Key Laboratory of Health Technology Assessment of the Ministry of Health, School of Public Health, Fudan University, Shanghai, China

<sup>5</sup>Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP<sup>3</sup>), Fudan University, Shanghai, China

<sup>6</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

<sup>7</sup>Clinical Epidemiology and Biostatistics, School of Medical Sciences, Örebro University, Örebro, Sweden

**Contributors** XF designed the study, did data analysis, interpreted the results and drafted the article; RL prepared the data, drafted the article and revised it critically; HK revised the article critically for important intellectual content; MB provided statistical consultation and revised the article critically; FF and YC are the guarantors of the study, and they monitored the study implementation and revised the article critically for important intellectual content. All authors contributed to further drafts and approved the version to be published.

**Funding** The authors gratefully acknowledge research grants to YC from the Junior Faculty Research Grants (C62412022) of the Institute of Environmental Medicine, Karolinska Institutet, and from the fund for PhD research (KID-funds) and travel (KI-foundations and funds) of Karolinska Institutet, Sweden, and research grants to RL from the Public Welfare Research Program of National Health and Family Planning Commission of China (201402022) and from the Opening Project of Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP<sup>3</sup>).

**Disclaimer** The funding sponsors had no role in the design of the study, in the collection, analysis or interpretation of data, in the writing of the manuscript or in the decision to publish the research results.

**Competing interests** None declared.

**Ethics approval** Institutional Review Board of Basic Medical Sciences, Chinese Academy of Medical Science, China.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- Yang Y, Cao Y, Li W, *et al.* Multi-site time series analysis of acute effects of multiple air pollutants on respiratory mortality: a population-based study in Beijing, China. *Sci Total Environ* 2015;508:178–87.
- Shang Y, Sun Z, Cao J, *et al.* Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. *Environ Int* 2013;54:100–11.
- Télez-Rojo MM, Romieu I, Ruiz-Velasco S, *et al.* Daily respiratory mortality and PM<sub>10</sub> pollution in Mexico City: importance of considering place of death. *Eur Respir J* 2000;16:391–6.
- Ostro BD, Hurley S, Lipsett MJ. Air pollution and daily mortality in the Coachella Valley, California: a study of PM<sub>10</sub> dominated by coarse particles. *Environ Res* 1999;81:231–8.
- Analitis A, Katsouyanni K, Dimakopoulou K, *et al.* Short-term effects of ambient particles on cardiovascular and respiratory mortality. *Epidemiology* 2006;17:230–3.
- Katsouyanni K, Touloumi G, Spix C, *et al.* Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *Air Pollution and Health: a European Approach*. *BMJ* 1997;314:1658–63.
- Peng RD, Dominici F, Louis TA. Model choice in time series studies of air pollution and mortality. *J Roy Stat Soc Ser A Stat Soc* 2006;169:179–98.
- Dominici F, McDermott A, Zeger SL, *et al.* On the use of generalized additive models in time-series studies of air pollution and health. *Am J Epidemiol* 2002;156:193–203.
- von Klot S, Peters A, Aalto P, *et al.* Ambient air pollution is associated with increased risk of hospital cardiac readmissions of myocardial infarction survivors in five European cities. *Circulation* 2005;112:3073–9.
- Terzi Y, Cengiz MA. Using of generalized additive model for model selection in multiple Poisson regression for air pollution data. *Sci Res Essays* 2009;4:867–71.
- Duarte BPM, Saraiva PM. Hybrid models combining mechanistic models with adaptive regression splines and local stepwise regression. *Ind Eng Chem Res* 2003;42:99–107.
- Jerrett M, Arain A, Kanaroglou P, *et al.* A review and evaluation of intraurban air pollution exposure models. *J Expo Anal Environ Epidemiol* 2005;15:185–204.
- Lin X, Zhang D. Inference in generalized additive mixed models by using smoothing splines. *J Roy Stat Soc Ser B Stat Methodol* 1999;61:381–400.
- Raftery AE, Gneiting T, Balabdaoui F, *et al.* Using Bayesian model averaging to calibrate forecast ensembles. *Mon Weather Rev* 2005;133:1155–74.
- Saltelli A, Ratto M, Andres T, *et al.* *Global sensitivity analysis: the primer*. West Sussex: John Wiley & Sons, 2008.
- Koop G, Tole L. Measuring the health effects of air pollution: to what extent can we really say that people are dying from bad air? *J Environ Econ Manag* 2004;47:30–54.
- Hoeting JA, Madigan D, Raftery AE, *et al.* Bayesian model averaging: a tutorial. *Stat Sci* 1999;14:382–401.
- Yang Y, Li R, Li W, *et al.* The association between ambient air pollution and daily mortality in Beijing after the 2008 olympics: a time series study. *PLoS One* 2013;8:e76759.
- Zhang F, Wang W, Lv J, *et al.* Time-series studies on air pollution and daily outpatient visits for allergic rhinitis in Beijing, China. *Sci Total Environ* 2011;409:2486–92.
- Zhang F, Xu J, Zhang Z, *et al.* Ambient air quality and the effects of air pollutants on otolaryngology in Beijing. *Environ Monit Assess* 2015;187:495.
- Ver Hoef JM, Boveng PL. Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology* 2007;88:2766–72.
- Dominici F, Daniels M, Zeger SL, *et al.* Air pollution and mortality: estimating regional and national dose–response relationships. *J Am Stat Assoc* 2002;97:100–11.
- Jbilou J, El Adlouni S. Generalized additive models in environmental health: a literature review. INTECH Open Access Publisher, 2012.
- Koken PJ, Piver WT, Ye F, *et al.* Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver. *Environ Health Perspect* 2003;111:1312–17.
- Yang L, Qin G, Zhao N, *et al.* Using a generalized additive model with autoregressive terms to study the effects of daily temperature on mortality. *BMC Med Res Methodol* 2012;12:165.
- Whitney M, Ngo L. Bayesian model averaging using SAS software. *SUGI* 2004;29:9–12.
- Clyde M. *Model averaging*. New York: Wiley-Interscience, 2003.
- Clyde M, George EI. Model uncertainty. *Stat Sci* 2004;19:81–94.
- State Environmental Protection Agency of China. *China national ambient air quality standard (GB 3095-2012)*. Beijing: State Environmental Protection Agency of China, 2012.
- Zhang F, Li L, Krafft T, *et al.* Study on the association between ambient air pollution and daily cardiovascular and respiratory mortality in an urban district of Beijing. *Int J Environ Res Public Health* 2011;8:2109–23.
- Xu M, Guo Y, Zhang Y, *et al.* Spatiotemporal analysis of particulate air pollution and ischemic heart disease mortality in Beijing, China. *Environ Health* 2014;13:109.
- Sun Z, Mukherjee B, Brook RD, *et al.* Air-Pollution and Cardiometabolic Diseases (AIRCMD): a prospective study investigating the impact of air pollution exposure and propensity for type II diabetes. *Sci Total Environ* 2013;448:72–8.
- Puett RC, Hart JE, Yanosky JD, *et al.* Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. *Environ Health Perspect* 2009;117:1697–701.
- He X, Shen L, Shen Z. A data-adaptive knot selection scheme for fitting splines. *IEEE Signal Proc Lett* 2001;8:137–9.
- Miyata S, Shen X. Free-knot splines and adaptive knot selection. *J Jpn Stat Soc* 2005;35:303–24.
- Viallefond V, Raftery AE, Richardson S. Variable selection and Bayesian model averaging in case-control studies. *Stat Med* 2001;20:3215–30.
- Zhou M, He G, Liu Y, *et al.* The associations between ambient air pollution and adult respiratory mortality in 32 major Chinese cities, 2006–2010. *Environ Res* 2015;137:278–86.
- Zhu R, Chen Y, Wu S, *et al.* The relationship between particulate matter (PM<sub>10</sub>) and hospitalizations and mortality of chronic obstructive pulmonary disease: a meta-analysis. *J Chron Obstruct Pulmon Dis* 2013;10:307–15.
- Health Effects Institute. Revised analyses of time-series studies of air pollution and health: special report. Boston: Health Effects Institute, 2003.
- Kass RE, Wasserman L. The selection of prior distributions by formal rules. *J Am Stat Assoc* 1996;91:1343–70.