Author's response to reviews

Title: Spatiotemporal Analysis of Particulate Air Pollution and Ischemic Heart Disease Mortality in Beijing, China

Authors:

Meimei Xu (xumeimei2005@163.com)
Yuming Guo (y.guo1@uq.edu.au)
Yajuan Zhang (zhyj830515@126.com)
Dane Westerdahl (danewest03@gmail.com)
Yunzheng Mo (moyunzheng@gmail.com)
fengchao Liang (liangfengchao126@126.com)
Xiaochuan Pan (xcpan@bjmu.edu.cn)

Version: 2
Date: 14 November 2014

Author's response to reviews: see over
Dear Dr Francesco Forastiere,

Thank you for your letter and the attached comments.

Because we have put new variable in our model based on the reviewer’s comments, our results and discussions have been modified accordingly. The changes in the revision have been highlighted. In order to better show the results, we have used “Fig.4” instead of “Table 2”, “Fig.5” and “Fig. 6” instead of “Table 4”.

Please find our revised manuscript and the following responses to the reviewer’s concerns. Please don’t hesitate to contact us if you have any questions.

Thank you for considering this manuscript.

Best Regards.

Xiaochuan Pan

Dept. of Occupational and Environmental Health

Peking University School of Public Health

38 Xue Yuan Road, Beijing 100191, P.R. China
Response to Reviewer Comments:

# Reviewer 1:

1. Page 6 lines 4-7: Daily temperature and humidity were not computed on the township level. So how have they been computed?

   Answer: Daily temperature and humidity from one station were used in this study, because time series analysis using temperature data from a single site can be used well to estimate the city-wide temperature–mortality association [1].

2. Pages 7 lines 26-28: Temperature and humidity were considered at the same lag 0. But several studies demonstrated that low temperatures during winters have a delayed effects of 1 to 5 days on health, while elevated temperature have effects on the current day of exposure. It could be appropriated to consider these two different temporal pattern for the cold and hot temperatures in the models, by using piecewise linear splines or two separated splines. It is possible that part of the seasonal effect was actually due to the effect of temperature.

   Answer: Studies has showed low temperature has a delayed effect of more than 10 days on mortality and high temperature has a delayed effect of around 3 days but a harvesting effect occurred after that [2, 3]. So we controlled the 14-day moving average temperature in the model in order to completely control for the potential effects of ambient temperature on
mortality [4, 5].

3. Page 8 line 2: “the township”

Answer: We have corrected it in the revision.

4. Page 8 line 3: IHD

Answer: We have corrected it in the revision.

5. Page 8 line 11: Generalized additive model (GAM)

Answer: We have corrected it in the revision.

6. Page 9 line 24: 51.7±19.9%

Answer: We have corrected it in the revision.

7. Page 10 line 13: Has the reported value of correlation between estimated and observed PM10 been cross validated?

Answer: The reported value of correlation between estimated and observed PM10 was produced during cross-validation.

8. Page 12 lines 9-12: please express the sentence in a better way, it is possible that something is missing.

Answer: We have improved the statement in the revision.
9. Page 13 line 26: “time-series” model

Answer: We have corrected it in the revision.

10. Since age and season have different effects in their categories, it is necessary to provide a test for their heterogeneity to give strength to this point.

Answer: We have done a Z test for heterogeneity from stratified analyses.

Please find the changes in the results section.
##Reviewer 2:

1. Major Compulsory Revisions

1.1. Data analysis

- Page 6 Line 13-15: Please explain why you chose inverse distance weighting and ordinary kriging. I am not an expert in kriging methods, but to my knowledge ordinary kriging assumes no structural trend in air pollution concentration. However, in your case PM10 concentration was higher in the south than in the north of Beijing. Hence, universal kriging would be better.

Answer: Inverse distance weighting and kriging are the most common interpolation methods. Inverse distance weighting is a typical deterministic GIS function for spatial interpolation. Ordinary kriging is the most common stochastic interpolation method. About the performance of IDW and kriging, the findings have been mixed[6]. That is why we also chose IDW.

We chose ordinary kriging instead of universal kriging for several reasons. First, although it may seem more appropriate because of the “varying mean” concentration across the city, universal kriging requires a predetermined set of “exploratory variables” to explain the varying means [7]. The exploratory variable may include emissions, land use, population, road network distribution, altitude, climatology and so on. Because so far we cannot obtain these variables in our study, we believed that it was
better to apply the more conservative ordinary kriging rather than risk to extrapolating the data drift to areas where no drift exists [8].

Secondly, we tried to do the simplest universal kriging that just included longitude and latitude. But this did not perform better than ordinary kriging based on cross-validation and was even worse (Table A). The reason is that although the average daily PM$_{10}$ concentration during the study period is higher in the south than in the north, it is not true each day.

Based on the above reasons, we still keep ordinary kriging in our study.

Table A The comparison between the predicted and observed PM$_{10}$ concentrations using different kriging methods at 27 monitoring stations during 2008 and 2009.

<table>
<thead>
<tr>
<th>station</th>
<th>Ordinary kriging</th>
<th>Universal kriging</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation$^a$</td>
<td>RMSE</td>
</tr>
<tr>
<td>1</td>
<td>0.98</td>
<td>14.45</td>
</tr>
<tr>
<td>2</td>
<td>0.97</td>
<td>20.96</td>
</tr>
<tr>
<td>3</td>
<td>0.99</td>
<td>14.74</td>
</tr>
<tr>
<td>4</td>
<td>0.99</td>
<td>12.96</td>
</tr>
<tr>
<td>5</td>
<td>0.99</td>
<td>12.87</td>
</tr>
<tr>
<td>6</td>
<td>0.97</td>
<td>24.14</td>
</tr>
<tr>
<td>7</td>
<td>0.93</td>
<td>30.38</td>
</tr>
<tr>
<td>8</td>
<td>0.99</td>
<td>13.62</td>
</tr>
<tr>
<td>9</td>
<td>0.96</td>
<td>27.38</td>
</tr>
<tr>
<td>station</td>
<td>Ordinary kriging</td>
<td>Universal kriging</td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>Correlation$^a$</td>
<td>RMSE</td>
</tr>
<tr>
<td>10</td>
<td>0.97</td>
<td>22.26</td>
</tr>
<tr>
<td>11</td>
<td>0.97</td>
<td>20.22</td>
</tr>
<tr>
<td>12</td>
<td>0.97</td>
<td>22.66</td>
</tr>
<tr>
<td>13</td>
<td>0.96</td>
<td>29.31</td>
</tr>
<tr>
<td>14</td>
<td>0.98</td>
<td>15.73</td>
</tr>
<tr>
<td>15</td>
<td>0.96</td>
<td>24.67</td>
</tr>
<tr>
<td>16</td>
<td>0.97</td>
<td>20.83</td>
</tr>
<tr>
<td>17</td>
<td>0.98</td>
<td>16.91</td>
</tr>
<tr>
<td>18</td>
<td>0.90</td>
<td>42</td>
</tr>
<tr>
<td>19</td>
<td>0.98</td>
<td>15.25</td>
</tr>
<tr>
<td>20</td>
<td>0.94</td>
<td>26.8</td>
</tr>
<tr>
<td>21</td>
<td>0.97</td>
<td>20.13</td>
</tr>
<tr>
<td>22</td>
<td>0.98</td>
<td>20.82</td>
</tr>
<tr>
<td>23</td>
<td>0.95</td>
<td>25.11</td>
</tr>
<tr>
<td>24</td>
<td>0.99</td>
<td>12.14</td>
</tr>
<tr>
<td>25</td>
<td>0.99</td>
<td>12.96</td>
</tr>
<tr>
<td>26</td>
<td>0.93</td>
<td>28.26</td>
</tr>
<tr>
<td>27</td>
<td>0.92</td>
<td>26.95</td>
</tr>
<tr>
<td>overall</td>
<td>0.96</td>
<td>22.40</td>
</tr>
</tbody>
</table>

1.2. Results
• Page 11 Line 6/7 and Figure 3: The effect estimates of GAM were smaller than in GAMM, however, they are not significantly different according to Figure 3. This should be mentioned in the text.

Answer: We appreciated the reviewer’s comments. To our knowledge, the effect estimates using different exposure metrics cannot be statistically tested and just can be compared by numbers.

The lags should be reordered in Figure 3: single lags first (lag0…. Lag5) and then lag01….lag05.

Answer: The lags have be reordered in Figure 3.

A harvesting effect can be seen then; this should be discussed within the text.

Answer: We have added the discussion about harvesting effects in the revision.

1.3. Discussion

• Please add a section discussing potential biological mechanisms explaining the association between PM\textsubscript{10} and ischemic heart disease.

Answer: We have added the statement in discussion section.

2. Minor Essential Revisions
• Overall, the language needs to be revised.

Answer: We have improved the language in the revision according to the reviewer’s suggestions.

• I suggest using the terms ‘spatially resolved PM$_{10}$’ and ‘PM$_{10}$’ instead of ‘kriged PM$_{10}$’ and ‘non-spatial PM$_{10}$’ throughout the manuscript.

Answer: We have used “spatially resolved PM$_{10}$” and “averaged PM$_{10}$” instead of “kriged PM$_{10}$” and “non-spatial PM$_{10}$”.

2.1. Abstract

• Line 10: Please change ‘304 township-level areas’ to ‘287 township-level areas’. As health data was not available for 17 townships.

Answer: We have changed it in the revision.

• Line 11: I recommend changing ‘A spatiotemporal model was used…’ to ‘ A generalized additive mixed model was used…’.

Answer: We have changed it in the revision.

• Line 13 and 14: Please state that the co-effects of the seasons, gender and age were studied in a stratified analysis in order to make that clear to the reader. The word ‘adjusted’ should be deleted in this context.

Answer: We have changed it in the revision.
• Line 14: I suggest replacing ‘time series analysis’ with ‘generalized additive model’.

Answer: We have changed it in the revision.

• Line 19: The p-value could be excluded. I suggest writing ‘for the 2-day average’ instead of ‘at lag 0-1 day’.

Answer: We have changed it in the revision.

• Line 23: Please change ‘statistically larger effects’ to either ‘larger effects’ or ‘statistically stronger effects’.

Answer: We have changed it in the revision.

• Line 13 and 27: ‘of the exposed population’ should be deleted as everyone living in those townships is exposed.

Answer: We have changed it in the revision.

2.2. Key words

I recommend replacing ‘PM$_{10}$’ with ‘particulate matter’.

Answer: We have changed it in the revision.

2.3. Introduction

• Line 5: Please check if there was an update on the Global Burden of
Disease Program.

Answer: We have changed it in the revision.

• Line 12: PM10 should be defined at first mention (the definition can be deleted in line 15).
Answer: We have changed it in the revision.

• References are missing at the end of the sentence beginning in line 11 ‘Studies on the …’.
Answer: We have added the references in the revision.

• Please put a proper reference behind line 18-20.
Answer: We have added the references in the revision.

• References are missing at page 5 line 1-2 ‘However, few studies…’.
Answer: We think our previous statement is not right, so we have changed the statement in the revision.

2.4. Material and Methods

Study Area

• Please give some information on the size and population in the individual townships (e.g. range of square kilometers and population).
Answer: We have added the information in the revision.

• Please give some information on the population in Beijing during the study period (2008/2009).

Answer: The population data during the study period were not available. So we used the population data from the Sixth National Population Census in 2010. This data is authoritative.

Data collection

• Why was health data not available in Minyun and Yanqing? The missing 17 townships are all located in those two counties? Please elaborate.

Answer: We cannot obtain health data from the two counties. The Missing 17 townships are all located in those two counties.

• Please explain if PM_{10} data of 27 monitoring sites was available on a daily basis and give some information on missing data.

Answer: PM_{10} data of all the monitoring stations was available on a daily basis. The missing rate during the study period was from 0.4% to 6.7%. Imputation will produce error so we did not fill the missing value before interpolating.

We have added the related information in the revision.
• Figure 1: Please include a smaller map showing the location of Beijing in China.

‘Town’ should be replaced by ‘Township’.

Answer: We have made a new map including a smaller map showing the location of Beijing in China (Fig.1).

“Town” has been replaced by “Township” in the new map.

• Line 9: Please change ‘addictive’ to ‘additive’.

Answer: We have corrected it in the revision.

• Line 11/12: Please make clear that a random intercept for townships was included into the model.

Answer: Generalized additive mixed models are proposed for overdispersed and correlated data in studies involving clustered and spatial design [9, 10]. A random area-level intercept can be included in order to model those areas with higher death rates. We have improved the statement in the revision.

• Line 28: Degrees of freedom for temperature and relative humidity are based on a study from the US where meteorological conditions are different compared to China. I suggest checking the model fit with
several degrees of freedom.

Answer: Our previous study [11] has shown temperature and relative humidity were adjusted well using spline with three degrees of freedom. We also have checked the model fit from 3 to 6 and found temperature and relative humidity with three degrees of freedom had the best model fit.

We have improved the statement in the revision.

• Overall, I suggest including holidays and influenza epidemics also as potential confounders.

Answer: We have included public holidays in the model. But for influenza epidemics, we cannot get influenza morbidity data to define influenza epidemics. Some studies [12, 13] have shown respiratory mortality can be used to represent the influenza epidemics. We have tried to define influenza epidemics when the 7-day moving average of the respiratory mortality was greater than the 90th percentile of its city-specific distribution [12, 13], but our results showed that influenza epidemics just occurred during January, 2008. We don’t believe this method is right in our study, because H1N1 influenza was epidemic in 2009 and we did not find influenza epidemics in 2009 by this method. So this variable is not controlled in our model. About the potential confounder of influenza epidemics, studies from the United States have
typically not included it because the effects of influenza epidemics may be reduced after seasonality is adjusted in the model [14].

• Page 8 Line 15 GAM: I assume log(E(Yt)) is meant. As an alternative you could estimate also log(E(Yi,t)) and including the random intercept for township i here as well as a sensitivity analysis.

Answer: In order to compare the effects of spatially resolved PM$_{10}$ and averaged PM$_{10}$, we also have estimated the effects of averaged PM$_{10}$ using GAM and GAMM (that included the random intercept) in the revision (Fig. 3).

2.5. Results

• Text and Tables: Please verify that $\mu$g/m3 is written with superscript 3.

Answer: We have corrected it in the revision.

• Page 10 Line 23: The p-value can be excluded here.

Answer: We have corrected it in the revision.

• Figure 2: Categories of PM$_{10}$ concentrations should be consistent.

Answer: PM$_{10}$ concentrations were classified by percentile in our previous figure. We have made a new figure using the equal interval of PM$_{10}$ concentrations (Fig.2).
• Figure 3: I suggest choosing different symbols for GAM and GAMM in the figure. Please insert a reference line at y-Axis=0.

Answer: We have made a new figure based on the reviewer’s suggestion.

• Table 1: Please insert column headings also on page 23. Please explain in the footnote that PM$_{10}$ represents the average of the 27 monitoring stations. For kriged PM$_{10}$ I suggest using the 287 townships where health data was available. Footnote: Please replace ‘Number’ with ‘Number of deaths’. It should be mentioned that the maximum limit of detection for PM$_{10}$ concentration is 600 µg/m3.

Answer: We have made the changes based on the reviewer’s suggestion.

• Table 2 and table 3: CI: confidence interval should be defined in the footnote.

Answer: We have made the changes based on the reviewer’s suggestion.

2.6. Discussion

• Page 12: The sentence starting in line 13 should be moved to the end of the discussion.

Answer: We have moved the sentence to the end of the discussion in the revision.
• Line 24-26 misestimating exposure in rural areas: I think this is also a problem for kriging as for example in Beijing most monitoring stations are located around the center and only few in the northern or western part (Figure 1). It might be difficult to accurately estimate PM$_{10}$ concentrations in the northern or western rural area of Beijing in this case.

Answer: This may a problem but our cross-validation results showed that the sites in the rural area have good correlation between measurements and estimates (table S3). So this suggested that PM$_{10}$ concentrations in the northern or western rural area of Beijing are relatively reliable. At least, compared with using the averaged PM$_{10}$, exposure misclassification reduced using spatially resolved concentrations in these areas.

• Page 13: It is not necessary to repeat all the numbers again.

Answer: We have improved the statement in this section.

• Line 9-27: With regard to heterogeneity of results, the number of study years may also play a role here.

Answer: We have improved the statement in this section.

• The section starting at page 13 line 29 can be deleted as temperature is
not comparable to PM$_{10}$ with regard to spatial variability.

Answer: We have improved the statement in this section.

• Page 14 Line 5-13: The main problem with the two-pollutant model may be that for SO$_2$ and NO$_2$ only a single daily average for the whole city was available. This should be mentioned in this context.

Answer: We have improved the statement in this section.

• Page 14 Line 18: The study by Zeka et al was conducted in the US, therefore, differences in the results might be due to differences in meteorological conditions in the US and China.

Answer: We have improved the statement in this section.

• Line 27: Please verify if males possibly were older than females.

Answer: Our data has age group so we cannot say whether males were older than females.

About the effects in females, we apologized here because our previous code for estimating effect in females had a small error that we did not find. We have corrected the results in the revision.

2.7. Supplemental Material

• Table S3: I recommend adding information on the location of the
monitors (traffic-site, urban background).

Answer: As far as we know, the monitors during the study periods were not classified.

• Table S4: Please change ‘krigied’ to kriged or ‘spatially resolved’.

Answer: We have corrected it in the revision.

3. Discretionary Revisions

Modelling the association between PM$_{10}$ and IHD mortality

• GAMM: I recommend excluding the random intercept for township i as a sensitivity analysis and comparing the results with the model including the random intercept for township i.

Answer: We think it does not make sense to remove the random intercept for township i because GAMM cannot be built without including the random intercept.

• I suggest additional adjusting for barometric pressure as a sensitivity analysis.

Answer: We have tried to adjust for barometric pressure in the model, but the model fit become worse and barometric pressure had no effects on mortality, so we did not included barometric pressure in the model.
References:


