INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD • INDIA

Research and Publications

Air Pollution in Indian Cities: Short Term Mortality Impacts and Interactions with Temperature

Hem H. Dholakia
Dhiman Bhadra
Amit Garg

W.P. No.2014-04-01
April 2014

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 015
INDIA

Air Pollution in Indian Cities: Short Term Mortality Impacts and Interactions with Temperature

Hem H. Dholakia
Public Systems Group, Indian Institute of Management, Ahmedabad
Email: hemdholakia@iimahd.ernet.in

Dhiman Bhadra
Production and Quantitative Methods Group, Indian Institute of Management, Ahmedabad
Email: dhiman@iimahd.ernet.in

Amit Garg
Public Systems Group, Indian Institute of Management, Ahmedabad
Email: amitgarg@iimahd.ernet.in

Abstract

Background: Indian cities are among the most polluted globally, yet assessments of short term mortality impacts due to pollution have been limited. Furthermore, studies examining temperature – pollution interactions on mortality are largely absent. Addressing this gap remains important in providing research evidence to better link health outcomes and air quality standards for India.

Methods: Daily all-cause mortality, temperature, humidity and particulate matter less than 10 microns (PM_{10}) data were collected for five cities – Ahmedabad, Bangalore, Hyderabad, Mumbai and Shimla spanning 2005 - 2012. Poisson regression models were developed to study short term impacts of PM_{10} as well as temperature – pollution interactions on daily all-cause mortality.

Results: We find that mortality associated with a 10 $\mu g/m^3$ PM_{10} increase is highest for Shimla (1.36%, 95% confidence interval = -0.38% to 3.1%) and the least for Ahmedabad (0.16%, 95% CI = -0.31% to 0.62%). The corresponding values for Bangalore, Hyderabad and Mumbai are 0.22% (-0.04% - 0.49%), 0.85% (0.06% - 1.63%) and 0.2% (0.1% - 0.3%) respectively. The relative health benefits of reducing pollution are higher for cleaner cities (Shimla) as opposed to dirtier cities (Mumbai). Overall we find that temperature and pollution interactions do not significantly impact mortality for the cities studied.

Conclusions: This is one of the first multi-city studies that assess heterogeneity of air pollution impacts and possible modification due to temperature in Indian cities that are spread across climatic regions and topographies. Our findings highlight the need for pursuing stringent pollution control policies in Indian cities to minimize health impacts.

Keywords: particulate matter; PM_{10}; Health effect; Temperature-pollution interactions; time-series; GAM

Author contributions: H.H.D and A.G. were responsible for research design, methodology and interpretation of results. H.H.D and D.B. carried out statistical analysis. All authors contributed to writing the manuscript.

Financial interests' declaration: None declared

Air Pollution in Indian Cities: Short Term Mortality Impacts and Interactions with Temperature

1. Introduction

Short term health impacts of air pollution have been extensively studied for developed countries using time series and case-crossover studies (Lee et al., 2014; Li et al., 2013; Samet et al., 2000; Samoli et al., 2008; J. D. Schwartz, 2004). These findings have played an important role in determining air quality standards in the respective countries. For instance, the U.S. Environmental Protection Agency (USEPA) reviews health research every five years to recommend revisions to National Ambient Air Quality Standards, as mandated by the Clean Air Act (Bell, Samet, & Dominici, 2003; USEPA, 1970). However, epidemiological studies, to inform air pollution policy, are largely limited in the context of developing countries such as India (Balakrishnan et al., 2011).

Indian cities today are among the most polluted in the world and it is estimated that outdoor air pollution leads to approximately 670,000 deaths annually (Lim et al., 2013). In India, the Central Pollution Control Board (CPCB) set up under the Air Act of 1981(MoEF, 1981) , is mandated with setting and reviewing the National Ambient Air Quality Standards (NAAQS). Current standards, for particulate matter set by the CPCB (CPCB, 2009) are much higher than those recommended by the World Health Organization (Krzyzanowski & Cohen, 2008). In addition, unlike other countries (Bell et al., 2003; Dominici, Peng, Zeger, White, & Samet, 2007), the CPCB does not take into account findings from health literature when deciding on air quality standards (Balakrishnan et al., 2011). A periodic review of epidemiological evidence informs policy makers about current health risks associated with air pollution and sets the agenda towards finding a balance between reducing health impacts and the costs of implementing further air pollution controls (Dominici, McDermott, & Hastie, 2004).

One potential reason for the lack of tight coupling between ambient air quality standards and health outcomes may be limited epidemiological evidence in the Indian context. A comprehensive review of air pollution and health in Asia found only three time-series studies that examine the short term impacts of air pollution on mortality for the cities of Delhi and Chennai (Balakrishnan et al., 2011; HEI, 2010; Rajarathnam et al., 2011).

However, studies for other cities are needed for at least two important reasons. The first reason is that for a country like India, cities vary widely in terms of development pathways, sources and levels of pollution and policy responses to curb pollution. This presents challenges for generalization of findings from single city studies to the whole country. Second, a changing climate may likely alter pollution levels and subsequently modify health risks over time (Jacob & Winner, 2009; Ren, Williams, & Tong, 2006; Tagaris et al., 2009). Consequently, temperature and pollution interactions for cities that lie in different climatic regimes may be quite different. An understanding of these health risks would play an important role in shaping policy to thwart air pollution.

To address the aforementioned research gaps, we use a time-series approach using semi-parametric Poisson regression to study the short term mortality impacts of particulate matter (PM_{10}) as well as temperature – pollution interactions for five cities – Ahmedabad, Bangalore, Hyderabad, Mumbai and Shimla. Being situated in different climactic zones of India, we hope that the observations derived from our findings on these cities will give a fairly good idea about the environment – mortality interaction patterns prevalent in India as a whole.

2. Methods

2.1 Mortality data

Daily all-cause mortality data were collected from the birth and death registers of the municipal corporations of Ahmedabad, Bangalore, Hyderabad, Mumbai and Shimla. For most cities, information on age and cause of death were not available. Table 1 summarizes the climactic characteristics and topography of the above cities.

Table 1. Cities distributed by climate zone

Climate zone	Representative cities	Topography
Hot and dry	Ahmedabad	Plains
Cold	Shimla	Hilly regions
Temperate	Bangalore	Plateau
Composite	Hyderabad, Lucknow	Plains
Warm and humid	Mumbai	Coastal areas

India is divided into five climate zones namely – hot and dry, warm and humid, composite, temperate and cold. The rationale for choosing these cities was that they are each representative of a different climate zone. In addition to climate zone, these cities represent varied topography – plains, plateau, coastal areas and hilly regions. Air pollution levels vary from city to city based on sources of pollution and policy measures. Additionally, different weather patterns may modify pollution related health risks leading to wide spatial heterogeneity. Thus our choice of cities provides a snapshot of differential health risks across India.

2.2 Weather and PM_{10} data

Daily data on maximum and minimum temperature, relative humidity and dew point temperature were collected from the Indian Meteorological Department (IMD). The IMD has a record of daily weather variables since the year 1948. Daily measurements of PM_{10} were collected from the Central Pollution Control Board (CPCB) database. Under the National Ambient Air Quality Monitoring Program (NAMP) the CPCB monitors four criteria pollutants i.e. Sulphur Dioxide (SO_2), Oxides of Nitrogen (NO_x), Total suspended particles (TSP) and particulate matter less than 10 microns (PM_{10}) for 342 stations located in 127 cities across India.

Typically two measurements are taken per week for each station implying that 100 - 120 observations are available per year. These measurements are made available through the CPCB website and the values reported are a 24-hour average. Every city has a different number of air quality monitors that range from one in Shimla to nine in Hyderabad. For a given year, if any monitor had less than 75% of recorded observations (i.e. less than 90 observations), then it was not used in the analysis. Scatterplots of daily mortality, PM_{10} concentrations and temperature for the different cities are shown in the supplementary material.

To create a population level exposure series for particulate matter, we used the centring approach described by Schwartz (2000). For each monitor, the mean (over all observations) of that particular monitor was subtracted from each observation. This demeaned data was then divided by the standard deviation of that particular monitor to get a standardized series for that monitor. This process was repeated for all monitors in a given city. The standardized series across all monitors was averaged to get one single series. Finally, this single series was multiplied by the standard deviation of all monitors taken together and the mean of all monitors taken together was added back to each observation (Schwartz, 2000). The resultant series was the final exposure series used in the regression model.

2.3 Analytical models

We adopted a semi-parametric regression framework to develop the exposure – response relationship between air pollution and mortality for the sampled cities (Balakrishnan et al., 2011; Peng, Dominici, & Louis, 2006; Rajarathnam et al., 2011). The logarithm of daily expected deaths was modelled as a function of daily air pollution measurements in the presence of other confounding variables such as temperature and humidity. Smooth functions were used to control for effects of daily temperature, humidity and seasonal and long term trends as these are non-linearly related to mortality (see supplementary material figures). Thus the regression equation can be expressed as:

$$Log[E(Y_{ij})] = \beta PM_{10i,j-1} + \sum_{j=1}^{P} g(x_{ij}) + DOW_{ij} + \varepsilon_{ij} \qquad \ldots(1)\ldots$$

where Y_{ij} is the daily mortality count for the i^{th} city on the j^{th} day and is assumed to follow an over-dispersed Poisson distribution. The pollution (PM_{10}) measurement for the i^{th} city on the j^{th} day lagged by one observation is represented using $\beta PM_{10i,j-1}$. However, as PM_{10} is measured once every three days, today's deaths (*t*) are related to the PM_{10} value three days prior (*t*-3). The covariates x_{ij} represent daily temperature, relative humidity and time for the i^{th} city on the j^{th} day. The effects are expressed by an unknown smooth function $g(\bullet)$ constructed using natural cubic splines. Details about the structure of $g(\bullet)$ are given in the supplementary material. An indicator variable for each day of week is given by DOW_{ij}. The error term is modelled using ε_{ij}. The parameter of interest is β associated with the pollution variable (PM_{10}). This parameter gives an estimate of the increase in mortality associated with a unit change in PM_{10} concentrations.

For temperature, humidity and time, the amount of smoothness (i.e. optimal degrees of freedom) was determined based on the approach by Dominici et al. (2004). The underlying idea is that β is sensitive to degrees of freedom selected for temperature, humidity and time. The approach by Dominici et al. (2004), where optimal degrees of freedom are chosen such that they predict PM_{10} instead of daily mortality, provides asymptotically unbiased estimates of the β parameter. The details of the algorithm implemented to arrive at these optimal values have been provided as supplementary information.

Mortality impacts related to pollution may be delayed i.e. exposures on preceding days may determine current health outcomes (Bhaskaran, Gasparrini, Hajat, Smeeth, & Armstrong, 2013; Braga, Zanobetti, & Schwartz, 2001). To account for this, single lag models (lag 1) was embedded in Equation (1). The lack of daily PM_{10} measurements did not allow for use of distributed lag models as this may introduce large errors (Braga et al., 2001; Zanobetti, Wand, Schwartz, & Ryan, 2000).

2.4 Sensitivity analysis

In order to compare some plausible scenarios, a sensitivity analysis was undertaken where the estimates (β) were tested using (i) zero lags for the pollution variable; (ii) minimum temperature instead of maximum temperature and (iii) including other pollutants such as sulphur dioxide.

2.5 Temperature – pollution interactions

To study the interaction effects of temperature and pollution (PM_{10}) on mortality, a two step approach as suggested by Ren et al., (2006) was adopted. The first step involved fitting Equation (1) with an interaction term to capture the joint effects of pollution and temperature.

This model is given in Equation (2). The term $t_{ij}*PM_{10ij}$ expressed the interaction between daily temperature and pollution while its effect is quantified by the coefficient α.

$$Log[E(Y_{ij})] = \beta PM_{10i,j-1} + \sum_{j=1}^{P} g(x_{ij}) + \alpha(t_{ij}*PM_{10ij}) + DOW_{ij} + \varepsilon_{ij} \qquad \ldots(2)\ldots$$

If the interaction term (or α) is found to be significant, then a second model is used to understand if interaction effects are more significant during hotter or colder temperatures.

2.6 Software

All analysis was performed in the statistical environment R version 2.15.1. The package *mgcv* (version 1.7-24) was used to fit the models described in equations (1) and (2). The package *ggplot2* (version 0.9.3.1) was used for graphical representations.

3. Results

3.1 Summary statistics

As seen in Table 2, there is wide variation among different cities when it comes to daily pollution levels, mortality, temperature, as well as number of complete observations available for analysis. The highest PM_{10} levels are observed for Mumbai (174.4 ± 86.6) and the lowest for Shimla (54.4 ± 25.2). The daily number of deaths varies across cities and seems to be linked to population size. Shimla had the lowest number of daily deaths (4.2 ± 2.7) whereas Mumbai (225.6 ± 30.7) had the highest.

Table 2. Observed values for different variables across cities (mean ± standard deviation)

City	Temperature (°C)		Relative Humidity (%)	Daily Deaths	PM_{10} (μg/m³)	No of complete observations
	Maximum	Minimum				
Ahmedabad	34.3 ± 4.6	21.5 ± 5.6	56.7 ± 17.8	100 ± 18	93.9 ± 58.7	602
Bangalore	29.4 ± 2.7	18.7 ± 2.3	66.6 ± 15.5	120.7 ± 17	108.3 ± 69.8	307
Hyderabad	33.6 ± 3.8	20.7 ± 3.6	53.1 ± 17.6	74.7 ± 16.2	80.4 ± 21.9	498
Mumbai	32.3 ± 2.4	22.7 ± 4.05	69.2 ± 13.8	225.6 ± 30.7	174.4 ± 86.6	2012
Shimla	20 ± 5.1	11.2 ± 5.3	NA	4.2 ± 2.7	54.4 ± 25.2	962

*NA – No data available

Each city had a different number of air quality monitors. The number of air quality monitors ranged from one in Shimla to nine in Hyderabad. For every air quality monitor, percentage of missing data by year varied (see Table S1 in supplementary information). Thus, air pollution impacts were estimated for the period of 2008 - 09 for Hyderabad and Bangalore; from 2005-2009 for Ahmedabad; from 2005 - 2011 for Mumbai and 2006 – 2009 for Shimla. The implication is that since different periods are being studied for different cities, making comparison of results across cities somewhat difficult.

3.2 Exposure – response estimates

The percentage increase in mortality associated with a 10 μg/m³ increase in PM_{10} is reported in Table 3. The highest increase was seen for Shimla (1.36%) and the least for Ahmedabad (0.16%). Bangalore and Mumbai showed similar results with a 0.22% and 0.20% mortality increase respectively.

Table 3. Percentage increase in mortality for every 10 µg/m^3 PM_{10} increase

		Sensitivity analysis		
	Core model	**No Lag for PM**	**Minimum temperature**	**SO_2 included**
Ahmedabad	0.16%	0.06%	0.12%	-0.32%
(95% CI)	*(-0.31 to 0.62)*	*(-0.42 to 0.55)*	*(-0.36 to 0.61)*	*(-1.01 to 0.38)*
Bangalore	0.22%	0.16%	0.17%	0.23%
(95% CI)	*(-0.04 to 0.49)*	*(-0.11 to 0.43)*	*(-0.09 to 0.43)*	*(-0.18 to 0.64)*
Hyderabad	0.85%	0.48%	0.83%	0.41%
(95% CI)	*(0.06 to 1.63)*	*(-0.30 to 1.27)*	*(0.04 to 1.62)*	*(-0.69 to 1.51)*
Mumbai	0.20%	0.18%	0.16%	0.13%
(95% CI)	*(0.10 to 0.30)*	*(0.08 to 0.27)*	*(0.06 to 0.25)*	*(0.04 to 0.23)*
Shimla*	1.36%	0.97%	0.54%	-
(95% CI)	*(-0.38 to 3.1)*	*(-0.83 to 2.7)*	*(-1.17 to 2.27)*	

*No humidity and SO_2 measurements were available for Shimla; Values in the brackets represent 95% confidence intervals.
Negative values imply that the effect of pollution on mortality is not significant

The sensitivity analysis showed that mortality estimates were lower when no lag for pollution was used, across all cities. The estimates of the core model did not change significantly if minimum temperature was used as a confounding variable. The inclusion of SO_2 reduced the impact of PM_{10} on mortality for Hyderabad and Mumbai, although, these differences were not significant. In addition, no significant interaction effect (at a 5% level) between temperature and pollution on mortality was observed. Table 4 shows the estimates, standard errors and p-values for the interaction term between temperature and pollution.

Table 4. Interaction effects of temperature and pollution for all cities

City	**β co-efficient**	**Std. error**	**p – value**
Ahmedabad	0.00328	0.00202	0.11
Bangalore	0.00607	0.00473	0.20
Hyderabad	0.00759	0.00757	0.31
Mumbai	-0.00380	0.00281	0.16
Shimla	-0.00007	0.00054	0.89

3.3 Comparison with other studies

Our results are in close agreement with previous studies (Balakrishnan et al., 2011; Rajarathnam et al., 2011; Romieu et al., 2012), which find 0.44% (95% CI = 0.17 to 0.71) for 0.15% (95% CI = 0.07 to 0.23) increase in mortality for every 10 $\mu g/m^3$ PM_{10} increase for Chennai (Balakrishnan et al., 2011), and Delhi (Rajarathnam et al., 2011), respectively. Studies from cities in the United States have shown 0.3% to 0.5% increases in mortality with every 10 $\mu g/m^3$ PM_{10} increase (Samet et al., 2000). More recent studies for Latin America also corroborate these findings - 0.48% to 1.26% mortality increases for every 10 $\mu g/m^3$ PM_{10} increase were observed for cities across Brazil, Chile and Mexico (Romieu et al., 2012). These findings imply that the short term mortality risks associated with pollution are not very different across cities. However, exposure of the large Indian population in to outdoor pollution translates into a significant increase in mortality. Figure 1 compares the percentage change in mortality for every 10 $\mu g/m^3$ increase in PM_{10} across different studies.

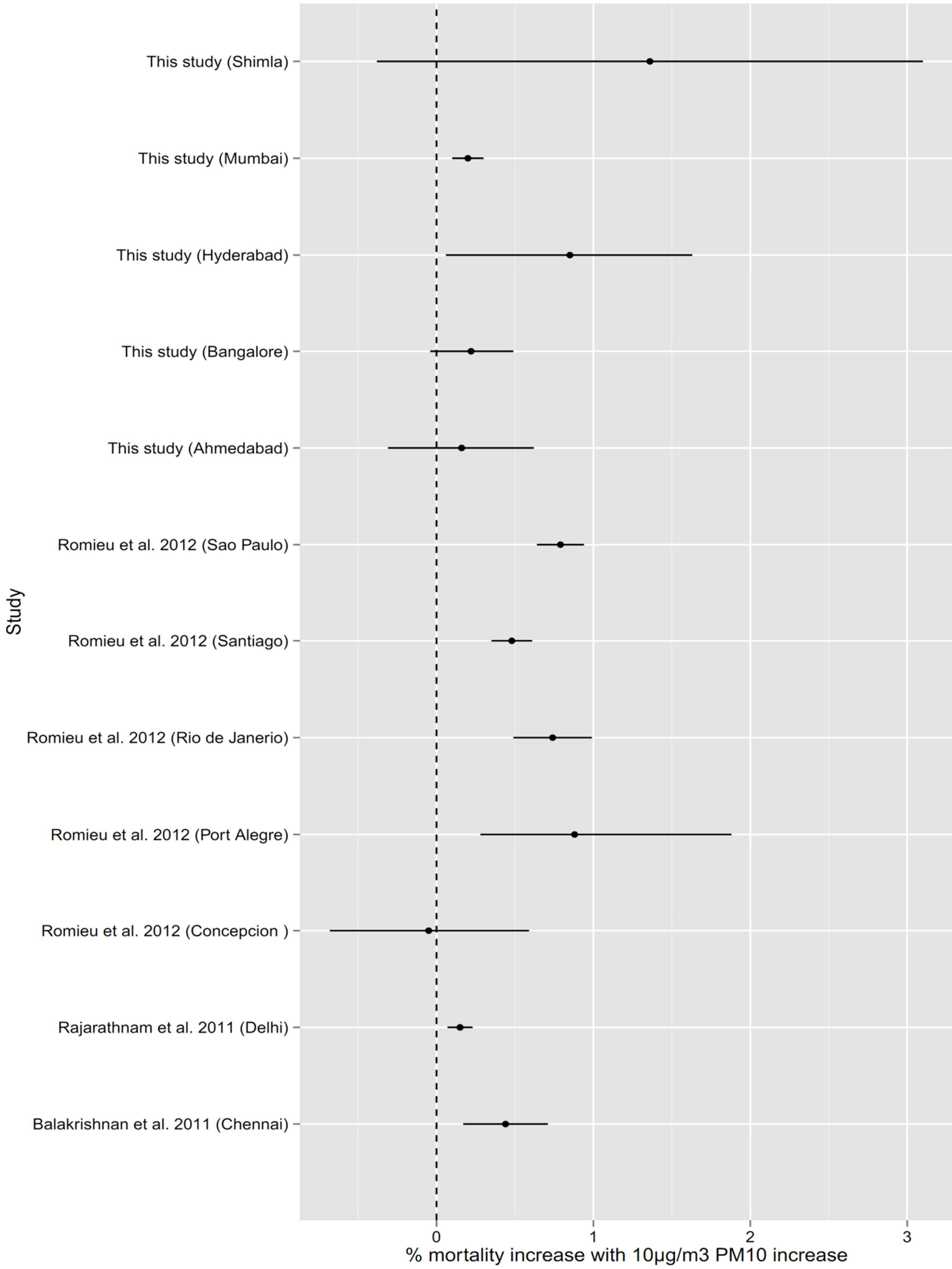

Figure 1 shows the central estimate and 95% confidence intervals for percentage increase in mortality with every 10µg/m^3 increase in PM_{10}. We compare estimates for five cities analyzed in this study and with those from selected previous studies.

4. Discussion

Our study was designed to explore the differences in short term impacts of air pollution on mortality and possible modification due to temperature in Indian cities that are spread across climactic zones and topographies. The results add to the existing body of epidemiological knowledge in the context of a developing country, like India.

An interesting finding of our analysis is that cities such as Ahmedabad and Mumbai that have higher levels of pollution do not experience a significant increase in mortality for every 10 $\mu g/m^3$ increase in PM_{10}. In contrast, the percentage increase in mortality is highest for Shimla which is among the cleanest cities. These results are corroborated by recent studies that have attempted to develop a new set of exposure – response relationships linking pollution and health (Lim et al., 2013; Smith et al., 2013). One of their main findings was that the impact of pollution on health may indeed be non-linear in nature i.e. when baseline pollution levels are high, health benefits associated with reduction in PM_{10} may be very small as compared to a situation where baseline pollution levels are low (Lim et al., 2013; Smith & Peel, 2010; Smith et al., 2013).

The import of these findings is that small reductions in pollution in cleaner cities will yield large health benefits, whereas in less cleaner cities, even large reduction in pollution may yield only modest health benefits in a relative sense. Although this seems counter-intuitive, it is not to suggest that the focus should be on reducing pollution in cleaner cities alone. On the contrary it underscores the need for rapid and aggressive policy measures in both types of cities to curb air pollution. Ambitious targets towards achieving ambient air quality standards should be set in highly polluted cities. On the other hand, cleaner cities could leverage significant health gains even by focussing on small reductions in pollution.

The pollutant of choice in our study was particulate matter less than ten micrograms in size (PM_{10}). This is because, it is the most routinely monitored air pollutant in India. Although studies show that fine particulate matter ($PM_{2.5}$) has more severe health impacts, this pollutant is not yet monitored in India on a regular basis.

The differential health impacts of single versus multiple pollutant models are of interest in epidemiology, although it is unclear whether including more than one pollutant in the analysis is necessarily more beneficial as opposed to single pollutant models (Tolbert, Klein, Peel, Sarnat, & Sarnat, 2007). We focussed primarily on the impacts of PM_{10} on mortality. Inclusion of sulphur dioxide (SO_2) along with PM_{10} did not change our estimates significantly, similar to previous findings (Rajarathnam et al., 2011). High percentage of missing data precluded incorporating nitrous oxides (NO_x) in our modelling framework. Thus, our results seem to be fairly robust and may not be influenced significantly if more pollutants are added.

In the context of our modelling framework, we did not find significant impacts of temperature-pollution interactions on mortality for the cities studied. This may be because temperature – pollution interactions are highly complex and non-linear and therefore may not have been captured adequately in the current model framework. Some studies e.g. Ren et al., (2006), have used more complex approaches such as modelling of the interaction using locally weighted smoothing functions (LOESS). However, a key limitation of such complex models is the inability to interpret estimates in an intuitive manner. Furthermore, the results from previous studies are varied implying that interaction effects may be city specific in nature. For instance, within the United States alone, Ren et al., (2008), found that while ozone modified the temperature mortality relationship in northern cities, no such effects were observed for southern cities. Further research is needed to better understand how temperature and pollution interactions influence health risks across cities in India.

One key limitation of our dataset is that there were significant missing data for the different air quality monitors across cities except Mumbai. This affected the parameterization and structure of the semi-parametric model used in our analysis. Furthermore, the fact that pollution estimates were not significant for Ahmedabad and Shimla may be a reflection of measurement error. Needless to say, better monitoring will help in developing more accurate exposure – response relationships across cities.

To create a consistent exposure series, Balakrishnan et al. (2011), developed a spatial model for Chennai. A 0.5 square kilometre grid was superimposed on a map of zones in the city. For each grid cell, PM values of the nearest AQM (measured as distance from centroid of grid to AQM) were assigned. For each zone, the PM exposure series was an average of the air quality reading on a particular day weighted by the number of grid cells it was assigned to in the zone. This approach was preferred to a simple average or centering approach used in this study. Whereas, a spatial model has distinct advantages, it requires a large amount of disaggregated information such as daily number of deaths in different zones of the city. Since that information was not available for cities in the current study, a centering approach was adopted. The advantage of the centering approach is that although a difference in measurements across monitors may influence variability of the exposure series and lead to underestimates, the slope co-efficient (i.e. β) corresponding to the pollution parameter (i.e. PM_{10}) remains unchanged if one or several AQM's are used (Balakrishnan et al., 2011; Rajarathnam et al., 2011; Wong, Ma, Hedley, & Lam, 2001).

Though it has been pointed out that impacts of air pollution is primarily linked to cardio-respiratory mortality(Pope III et al., 2002; Samoli et al., 2014; USEPA, 2009) , the present study only examined all-cause mortality. This was because information on cause-of-death and age groups was not available for the cities which were considered. Mortality in India is underreported and on an average only 67% of all deaths gets registered, with high variability across different states (Dhar, 2013). Of these, it is only institutional deaths that contain information on cause of death. It is reasonable to suggest that enhancing the quality of mortality registration data is vital for future studies.

In conclusion, the study of air pollution on mortality remains an important area of research in the Indian context. Clearly there remains a need to strengthen data quality and carry out similar studies for many more cities. In addition to the time-series approach used in this study, cohort studies are required to understand air pollution related health risks in India. Epidemiological evidence can help guide policy by providing evidence to tightly couple health outcomes and air quality standards, thereby minimizing the impacts of outdoor air pollution in India.

References

Balakrishnan, K., Ganguli, B., Ghosh, S., Sankar, S., Thanasekaran, V., Rayadu, V. N., & Caussy, H. (2011). *Short term effects of air pollution on mortality: results of a time-series analysis in Chennai, India* (Research Report No. 157). Health Effects Institute.

Bell, M. L., Samet, J., & Dominici, F. (2003, November). *Time series studies of particulate matter*. Johns Hopkins University, Department of Biostatistics Working Paper 10, Johns Hopkins, Bloomberg School of Public Health.

Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L., & Armstrong, B. (2013). Time series regression studies in environmental epidemiology. *International Journal of Epidemiology*. doi:10.1093/ije/dyt092

Braga, A. L. F., Zanobetti, A., & Schwartz, J. (2001). The time course of weather-related deaths. *Epidemiology*, *12*(6), 662–667.

CPCB. (2009, November 18). National ambient air quality standards. *The Gazette of India*. New Delhi. Retrieved from http://cpcb.nic.in/National_Ambient_Air_Quality_Standards.php

Dhar, A. (2013, April 18). Birth and death registration still low in India. *The Hindu*. New Delhi. Retrieved from http://www.thehindu.com/news/national/births-and-deaths-registration-still-low-in-india/article4630425.ece

Dominici, F., McDermott, A., & Hastie, T. J. (2004). Improved semiparametric time series models of air pollution and mortality. *Journal of the American Statistical Association*, *99*(468), 938–948.

Dominici, F., Peng, R. D., Zeger, S. L., White, R. H., & Samet, J. M. (2007). Particulate Air Pollution and Mortality in the United States: Did the Risks Change from 1987 to 2000? *American Journal of Epidemiology*, *166*(8), 880–888. doi:10.1093/aje/kwm222

HEI. (2010). *Outdoor air pollution and health in the developing countries of Asia: a comprehensive review* (Special Report No. 18). Boston, MA: Health Effects Institute.

Jacob, D. J., & Winner, D. A. (2009). Effect of climate change on air quality. *Atmospheric Environment*, *43*(1), 51–63. doi:10.1016/j.atmosenv.2008.09.051

Krzyzanowski, M., & Cohen, A. (2008). Update of the WHO air quality guidelines. *Air Quality, Atmosphere & Health*, *1*, 7 – 13. doi:10.1007/s11869-008-0008-9

Lee, H., Honda, Y., Lim, Y.-H., Guo, Y. L., Hashizume, M., & Kim, H. (2014). Effect of Asian dust storms on mortality in three Asian cities. *Atmospheric Environment*, *89*, 309–317. doi:10.1016/j.atmosenv.2014.02.048

Li, P., Xin, J., Wang, Y., Wang, S., Shang, K., Liu, Z., … Wang, M. (2013). Time-series analysis of mortality effects from airborne particulate matter size fractions in Beijing. *Atmospheric Environment*, *81*, 253–262. doi:10.1016/j.atmosenv.2013.09.004

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., … Aryee, M. (2013). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260.

MoEF. Air (Prevention & Control of Pollution) Act (1981). Retrieved from http://www.envfor.nic.in/legis/air/air1.html

Peng, R. D., Dominici, F., & Louis, T. A. (2006). Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*(2), 179–203.

Pope III, C. A., Burnett, R., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality and long term exposure to fine particulate matter. *Journal of the American Medical Association*, *287*, 1132 – 1141.

Rajarathnam, U., Seghal, M., Nairy, S., Patnayak, R. C., Chhabra, S., Kilnani, K. V., & Ragavan, S. (2011). *Time series study on air pollution and mortality in Delhi* (Research report No. Report No. 157). Boston, MA: Health Effects Institute.

Ren, C., Williams, G. M., Morawska, L., Mengersen, K., & Tong, S. (2008). Ozone modifies associations between temperature and cardiovascular mortality: analysis of NMMAPS data. *Occupational and Environmental Medicine, 65,* 255 – 260.

Ren, C., Williams, G. M., & Tong, S. (2006). Does Particulate Matter Modify the Association between Temperature and Cardiorespiratory Diseases? *Environmental Health Perspectives*. doi:10.1289/ehp.9266

Romieu, I., Gouveia, N., Cifuentes, L. A., Ponce de Leon, A., Junger, W., Hurtado-Diaz, M., … Tzintzun-Cervantes, G. (2012). *Multicity Study of Air Pollution and Mortality in Latin America (the ESCALA Study).* (Research Report No.170). Boston, MA: Health Effects Institute.

Samet, J. M., Zeger, S. L., Dominici, F., Curriero, F., Coursac, I., Dockery, D. W., … Zanobetti, A. (2000). The national morbidity, mortality, and air pollution study Part II: morbidity and mortality from air pollution in the United States. *Research Reports Health Effects Institute, 94*(pt 2), 5–79.

Samoli, E., Peng, R., Ramsay, T., Pipikou, M., Touloumi, G., Dominici, F., … Katsouyanni, K. (2008). Acute Effects of Ambient Particulate Matter on Mortality in Europe and North America: Results from the APHENA Study. *Environmental Health Perspectives*, *116*(11), 1480–1486. doi:10.1289/ehp.11345

Samoli, E., Stafoggia, M., Rodopoulou, S., Ostro, B., Alessandrini, E., Basagaña, X., … Forastiere, F. (2014). Which specific causes of death are associated with short term exposure to fine and coarse particles in Southern Europe? Results from the MED-

PARTICLES project. *Environment International*, *67*, 54–61. doi:10.1016/j.envint.2014.02.013

Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology*, *11*(3), 320–326.

Schwartz, J. D. (2004). The effects of particulate air pollution on daily deaths: a multi-city case crossover analysis. *Occupational and Environmental Medicine*, *61*, 956 – 961.

Smith, K. R., Frumkin, H., Balakrishnan, K., Butler, C. D., Chafe, Z. A., Fairlie, I., … Schneider, M. (2013). Energy and Human Health. *Annual Review of Public Health*, *34*(1), 159–188. doi:10.1146/annurev-publhealth-031912-114404

Smith, K. R., & Peel, J. L. (2010). Mind the Gap. *Environmental Health Perspectives*, *118*(12), 1643–1645. doi:10.1289/ehp.1002517

Tagaris, E., Liao, K.-J., DeLucia, A. J., Deck, L., Amar, P., & Russell, A. G. (2009). Potential impact of climate change on air pollution-related human health effects. *Environmental Science & Technology*, *43*(13), 4979–4988.

Tolbert, P. E., Klein, M., Peel, J. L., Sarnat, S. E., & Sarnat, J. A. (2007). Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta. *Journal of Exposure Science and Environmental Epidemiology*, *17*, S29–S35. doi:10.1038/sj.jes.7500625

USEPA. (1970). The Clean Air Act (CAA); 42 U.S.C.s/s 7401 et seq. (1970), Clean Air Act and Amendments of 1970 (PL 91-604;42 USC 1857h-7 et seq.; amended 1970). United States Environmental Protection Agency. Retrieved from http://www2.epa.gov/laws-regulations/summary-clean-air-act

USEPA. (2009). *Integrated Science Assessment for Particulate Matter* (No. EPA/600/R-08/139F). United States Environment Protection Agency.

Wong, C. M., Ma, S., Hedley, A. J., & Lam, T.-H. (2001). Effect of air pollution on daily mortality in Hong Kong. *Environmental Health Perspectives*, *109*(4), 335.

Zanobetti, A., Wand, M. P., Schwartz, J. D., & Ryan, L. M. (2000). Generalized additive distributed lag models - quantifying mortality displacement. *Biostatistics*, *1*(3), 279 – 292.

eAppendix

Table eS1. Percentage of missing values for different air quality monitors (AQM)

	2005	2006	2007	2008	2009	2010	2011
Ahmedabad							
AQM1	13	22	18	70	na	na	na
AQM2	14	23	13	74	na	na	na
AQM3	63	26	20	18	na	na	na
AQM4	na	na	na	43	na	na	na
Bangalore							
AQM1	na	na	na	15	16	na	na
AQM2	na	na	na	35	19	na	na
AQM3	na	na	na	0	21	na	na
AQM4	na	na	na	23	48	na	na
Hyderabad							
AQM1	na	na	na	8	71	na	na
AQM2	na	na	na	20	29	na	na
AQM3	na	na	na	10	10	na	na
AQM4	na	na	na	22	28	na	na
AQM5	na	na	na	11	11	na	na
AQM6	na	na	na	10	10	na	na
AQM7	na	na	na	10	10	na	na
AQM8	na	na	na	18	27	na	na
AQM9	na	na	na	11	10	na	na
Mumbai*							
AQM1	26.3	32.1	23	63	49.9	24.7	47.7
AQM2	na	na	57	11.2	8.8	4.4	2.7
Shimla							
AQM 1	na	0	0	0	0	na	na

The percentage missing has been computed assuming that any monitor having 120 observations for a given year will imply 100% data.
*Mumbai AQM's had over 135 measurements for each year. Hence, percentage missing has been computed assuming that 360 observations for a year imply 100% data.
**na – implies that no data was available for those years

Algorithm for optimal degrees of freedom for confounding variables

This study arrived at the optimal degrees of freedom for each of the confounding variables by using the algorithm developed by Dominici et al., (2004). The idea behind the algorithm is to estimate the optimal degrees of freedom (i.e. smoothness) for each variable in the regression. The algorithm has four steps which are as follows:

1. A smooth function of time (or temperature or humidity as the case may be) was regressed on the daily values of particulate matter (PM_{10}) using different degrees of freedom that range from 1 to n. For each of these a generalized cross validation score (GCV) was obtained. The GCV score is a measure to assess the goodness of fit of a model. The lower the GCV better is the model fit.

   The above model is shown in equation (1), where x_{ij} represents the variable of interest (time, temperature or humidity) for the i^{th} city on the j^{th} day, and the effect is captured by the smooth function *s*. The smooth function used was a natural cubic spline whose structure has been described in equations (2) and (3). The pollution level for the i^{th} city on the j^{th} day is given by PM_{10ij}. The degree(s) of freedom that yielded the lowest GCV score was selected and this was labelled as *d*. The equation can be written as

$$PM_{10ij} = \alpha_0 + s\,(x_{ij}) \qquad \text{degree(s) of freedom} = \{1, 2, 3\ldots n\} \qquad \ldots(1)\ldots$$

   The expression for the natural cubic spline used for the function *s* is given by a piecewise polynomial of the form

$$S\,(x) = \begin{cases} s_1\,(x) \text{ if } & x_1 \leq x < x_2 \\ s_2\,(x) \text{ if } & x_2 \leq x < x_3 \\ \ldots & \end{cases} \qquad \ldots(2)\ldots$$

$s_{n-1}(x)$ if $x_{n-1} \leq x < x_n$

where $s_k(x) = a_k (x - x_k)^3 + b_k (x - x_k)^2 + c_k (x - x_k) + d_k$ …(3)…

for $k = 1, 2, \ldots n-1$

2. The above degrees of freedom (i.e. d) was multiplied by a factor $\lambda \geq 3$. Let this be denoted as d^*. It has been proved theoretically (Dominici et al., 2004) that if d^* degrees of freedom is used to model the PM_{10}-mortality relationship, then the estimate of β for the pollution variable is asymptotically unbiased (Dominici et al., 2004).

3. Daily mortality was regressed on particulate matter and a smooth function of time with degree(s) of freedom varying from 1 to d^*. The equation can be written as

$$Log\,[E(Y_{ij})] = \alpha_0 + \beta PM_{10ij} + s_{ij}(x_{ij}) \qquad \text{degrees of freedom} = \{1,\ldots,d^*\} \quad \ldots(4)\ldots$$

where Y_{ij} is the daily number of deaths for the i^{th} city on the j^{th} day and is assumed to follow an over-dispersed Poisson distribution. From the above analysis, we can obtain the asymptotic bias and variance of air pollution risk estimates for each of the degrees of freedom {of $s(x_{ij})$} from 1 to d^*.

4. Finally, a bandwidth selection procedure as described by Dominici et al. (2004) is used to select the optimal degrees of freedom. To do this, fitted values for deaths were obtained for equation (4). Using these fitted values, a bootstrap procedure was implemented to randomly generate 100 bootstrap samples ($\hat{\beta}_d^{b,i}$) of the air pollution risk estimate (β).

5. For each of the 100 samples generated, the regression model (4) was fitted using varying degrees of freedom for the smooth function of time starting from 1 all the way to $d*$. This process resulted in generating 100 values of β for each degree of freedom. The estimate of the bootstrap average of β was computed as:

$$\hat{\beta}_d^{o,i} = \frac{1}{B} \sum_1^B \hat{\beta}_d^{b,i} \qquad \ldots(5)\ldots.$$

6. Using these newly generated β values, the unconditional squared bias (USB) and unconditional variance (UV) was estimated as described by Dominici et al., (2004). The USB provides an estimate of the difference between the average of β generated in the bootstrap for every degree of freedom and that of the β value corresponding to $d*$ degrees of freedom ($\hat{\beta}_{\widehat{m}}^{i}$). The UV gives an indication of how precise these USB estimates are. The formulae for calculating USB and UV are respectively:

$$USB = \frac{1}{N} \sum_1^N \left(\widehat{\beta}_d^{o,i} - \hat{\beta}_{\widehat{m*}}^{i} \right)^2 \qquad \ldots(6)\ldots$$

$$UV = \frac{1}{N} \sum_1^N \frac{1}{B-1} \sum_1^B \left(\widehat{\beta}_d^{b,i} - \hat{\beta}_d^{o,i} \right)^2 \qquad \ldots(7)\ldots$$

The degrees of freedom that give the lowest USB was selected as these are closest to the unbiased estimate of β. Among these USB estimates, the one that had the lowest variance was chosen as it is the most precise estimate of β. The degrees of freedom thus selected, is the optimal degrees of freedom. The optimal degrees of freedom for the smooth functions of temperature, humidity and time considered in our study are shown in Table S.2.

Table eS2. Optimal degrees of freedom estimated for different cities via bootstrap

City	Temperature	Humidity	Time
Ahmedabad	7	5	13
Bangalore	4	5	5
Hyderabad	6	3	30
Mumbai	8	4	24
Shimla	9	-	28

*df/yr – Degrees of freedom

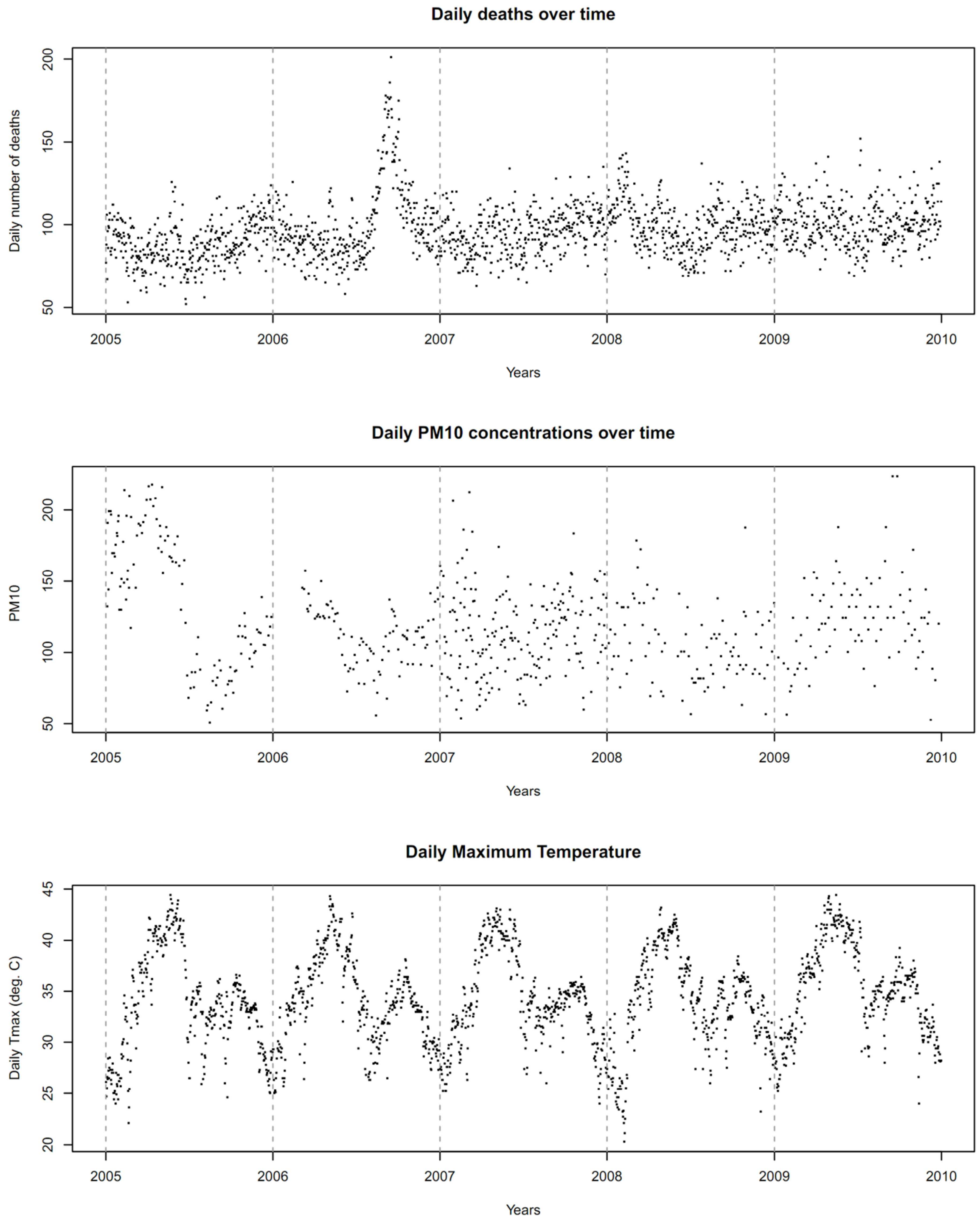

Figure eS.1 The scatter plots in these three panels show the association between daily number of deaths, daily pollution measurements and daily measurements of maximum temperature over time for Ahmedabad

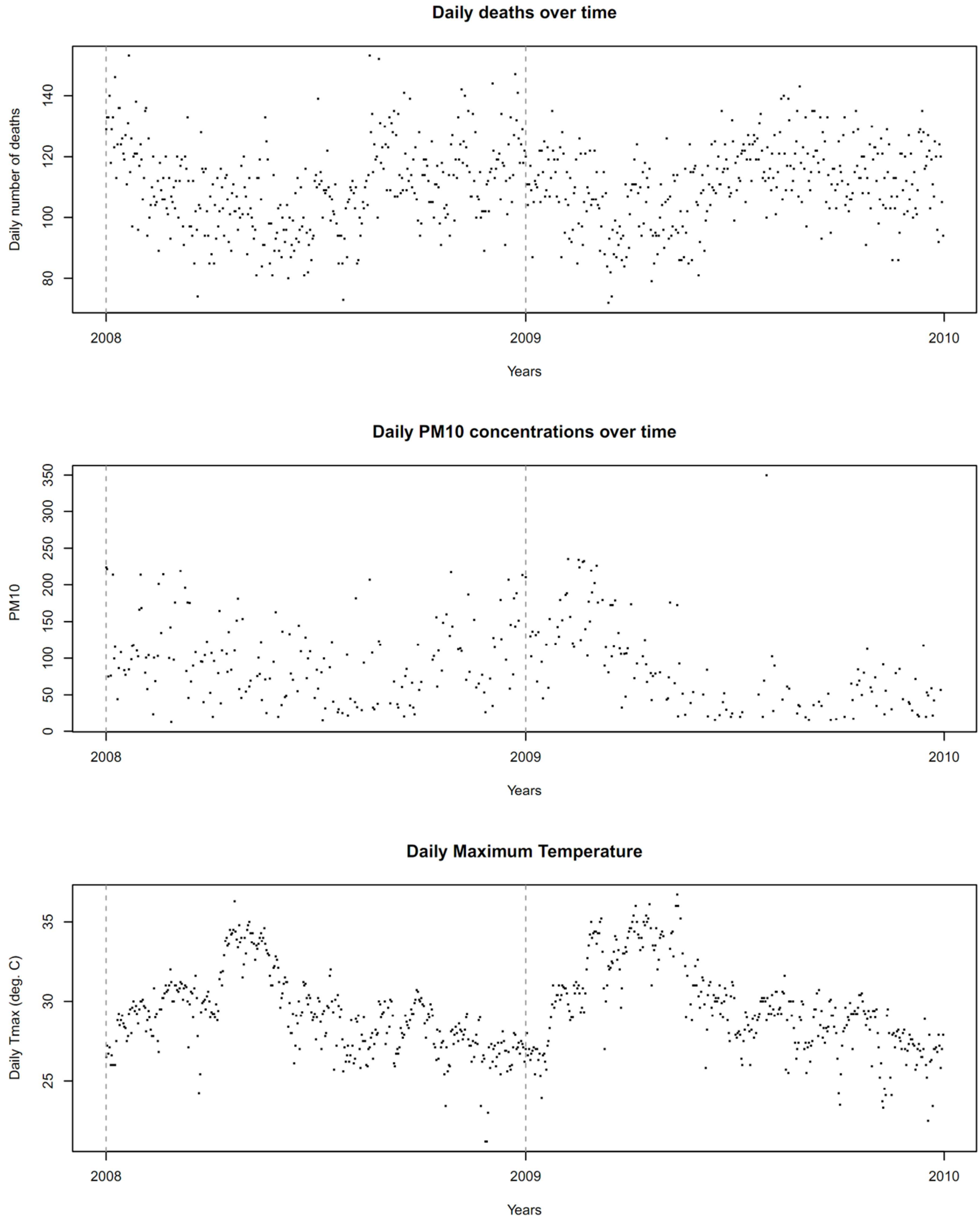

Figure eS.2 The scatter plots in these three panels show the association between daily number of deaths, daily pollution measurements and daily measurements of maximum temperature over time for Bangalore

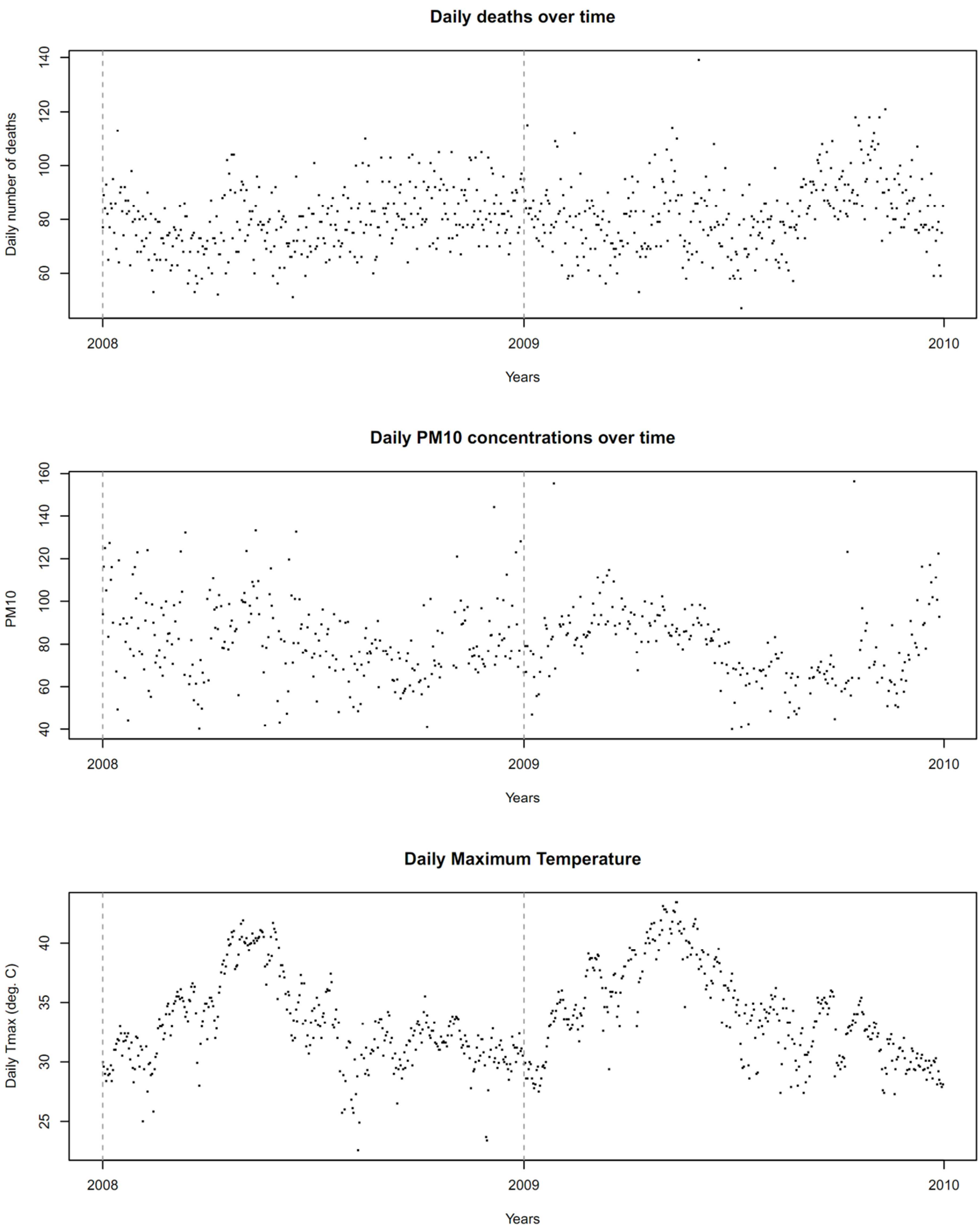

Figure eS.3 The scatter plots in these three panels show the association between daily number of deaths, daily pollution measurements and daily measurements of maximum temperature over time for Hyderabad

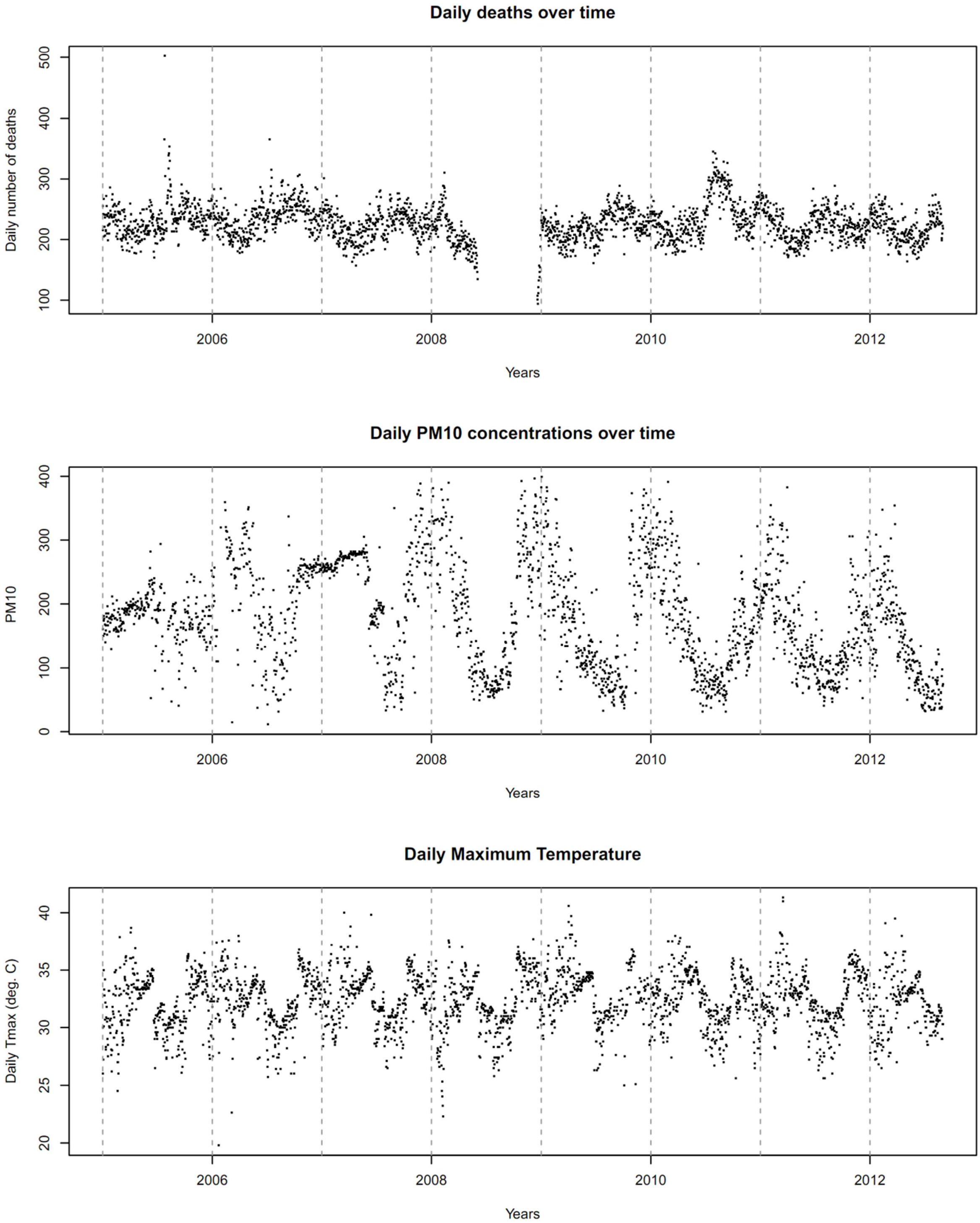

Figure eS.4 The scatter plots in these three panels show the association between daily number of deaths, daily pollution measurements and daily measurements of maximum temperature over time for Mumbai

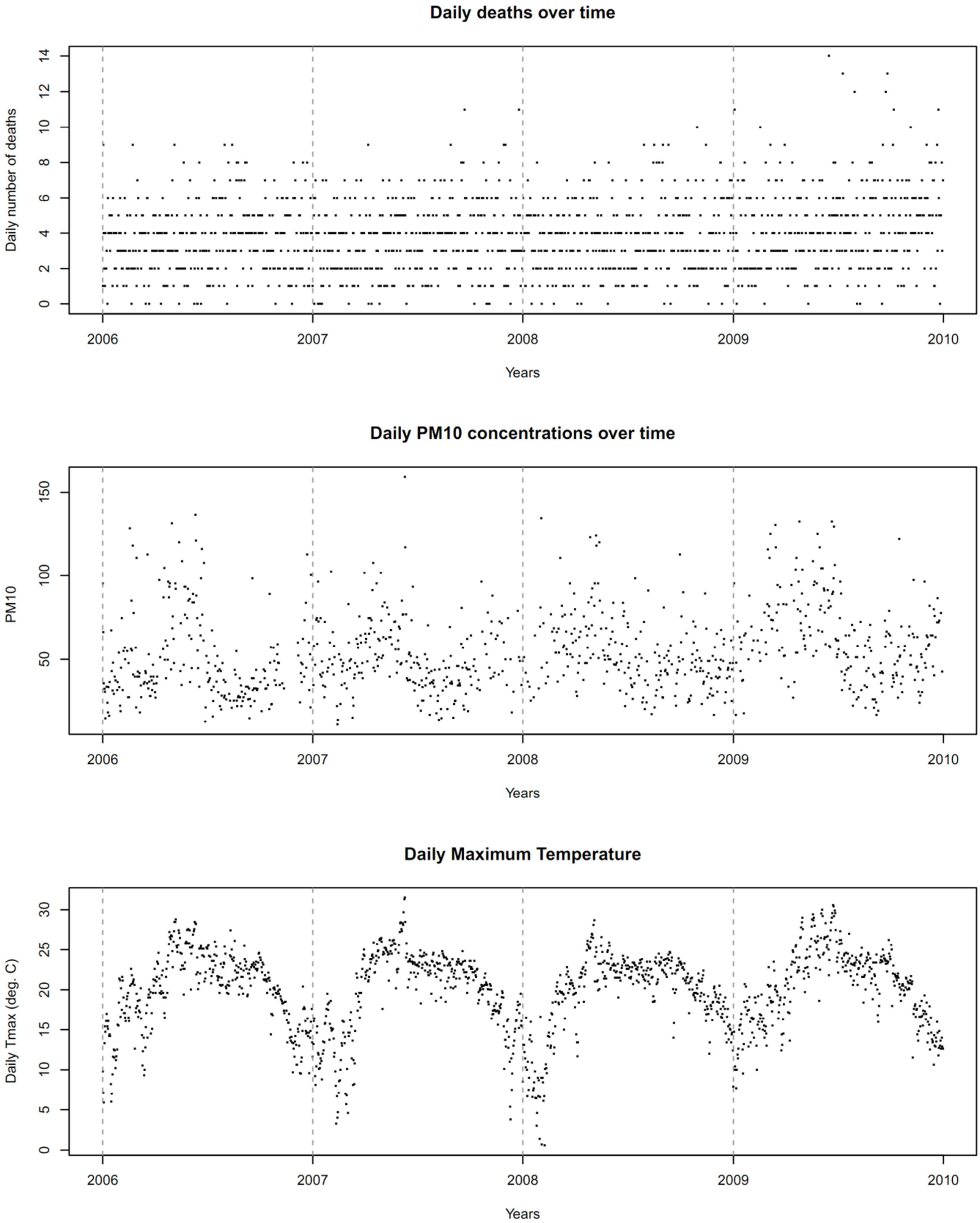

Figure eS.5 The scatter plots in these three panels show the association between daily number of deaths, daily pollution measurements and daily measurements of maximum temperature over time for Shimla