Cloudy with a Chance of Cardiovascular Disease:

Spatial Interpolation of Mortality Attributable to Air Pollution in China

Joseph Frostad

A thesis

submitted in partial fulfillment of the

requirements for the degree of

MASTER OF PUBLIC HEALTH

University of Washington

2015

Examining Committee:

Mohammad Hossein Forouzanfar, Chair

Haidong Wang, Member

Program authorized to offer degree: Global Health

University of Washington

Abstract

Cloudy with a Chance of Cardiovascular Disease:
Spatial Interpolation of Mortality Attributable to Air Pollution in China

Joseph Frostad

Chair of the Supervisory Committee:
Assistant Professor Mohammad Hossein Forouzanfar
Department of Global Health

Ambient air pollution, especially particulate matter smaller than 2.5 microns (PM2.5), is a significant contributor to premature mortality. This issue is particularly relevant in China due to rapid urbanization and population growth over the past several decades. Increased pollution in major population centers has resulted in massive loss of life via respiratory and cardiovascular disease. For the Global Burden of Disease (GBD) 2013 Risk Factors analysis, grid-level exposure estimates of annual PM2.5 and cause-specific integrated exposure response (IER) curves were used to predict the relative risk of mortality from a related cause for every pixel. These gridded relative risks were combined with cause-specific provincial mortality results in order to quantify the mortality attributable to air pollution for every province. To provide increased spatial granularity that can further inform the literature and policy-makers, cause-specific mortality was interpolated in order to generate predictions at the pixel level. This analysis demonstrated spatial and temporal trends in mortality attributable to air pollution in China over the GBD time series from 1990 to 2013. Mortality rates were shown to rise over time as a function of PM2.5 exposure, while total mortality increases were most prominent in several key population centers. The cause composition of total attributable mortality was found to have changed over the time series as well, with background rates of LRI and COPD decreasing so that cardiovascular disease makes up the bulk of mortality. As demonstrated by the non-linear concave shape of the IER curve, the distribution of mortality across levels of PM2.5 indicates that the majority of deaths are occurring in areas with moderate levels of air pollution and high population. Future interventions must focus on generating significant decreases in PM2.5 exposure across the Eastern seaboard, where Chinese urbanization is concentrated. This work is limited by the assumption of constant underlying mortality rates within provinces. As the spatial granularity of GBD mortality estimates continues to improve, these estimates will become more accurate and have increased utility to environmental health policy.

Introduction & Background

Exposure to ambient air pollution has been established by a robust variety of epidemiological studies as a risk factor for premature mortality (Pope III, 2002). Negative health outcomes are linked to both acute and chronic exposure across a range of endpoints. Of particular concern is fine particulate air pollution. As particles decrease in size, the extent of penetration increases – nanoparticles are capable of penetrating into the gas exchange area of the alveoli and even migrating to other organs through the bloodstream. Particles that are smaller than 2.5 micrometers have been studied extensively and linked to lung cancer, respiratory and cardiopulmonary disease (Turner 2011). Cardiopulmonary diseases make up the majority of all-cause mortality attributable to air pollution.

Two potential physiological pathways for the health impact of PM2.5 that have been proposed and supported by epidemiological evidence are pulmonary inflammation leading to accelerated atherosclerosis and altered cardiac autonomic function (Pope 2003). The risk factor analysis of the Global Burden of Disease 2010 study estimated that the deaths of 3.2 million people were attributable to ambient particulate matter air pollution in the year 2010 alone (Lim 2012). There is no baseline level of PM2.5 that has been determined to be safe and without negative health impact, meaning that all populations are in some way exposed. To take this into account, interventions against ambient air pollution must take a multi-pronged approach to mitigate the issue across an assortment of pathways. Some interventions need to be targeted at the population level, like increased emphasis on nutrition, exercise, and preemptive use of cardiovascular medication. Others are most effective on the community level, targeting specific

point sources of pollution and behaviors that increase risk based on the local context. (Giles 2010) There is a critical need to calibrate interventions according these circumstances in order to maximize efficiency when combating such a pervasive hazard. The inherently environmental nature of ambient air pollution exposure makes understanding spatial patterns particularly important for interventions. Such an analysis has notable applications for China – due to current sociopolitical context of air pollution there and the unique spatial distribution of exposure with regards to their demographic composition.

In some countries, simply viewing the spatial distribution of the PM2.5 exposure would be sufficient to provide significant insights on the most efficient way to target interventions. However, in China, the most extreme levels of air pollution were estimated to be in the Western provinces. This is an artifact of the intense dust storms that pummel the Gobi desert and the regions that surround it. While dust storms and the resulting increases to ambient PM2.5 are currently considered by experts to have major impacts on health, assumed in our models to be equivalent in effect to other sources of PM2.5, these areas of China are sparsely populated and therefore have marginal impact on the overall burden of disease from air pollution. This would make intervening based solely on the pixel-level exposure a highly inefficient strategy, since it would mean deprioritizing the areas where the bulk of disease burden occurs.

Subnational estimates of mortality attributable to air pollution in China were generated at the provincial level as a part of the Global Burden of Disease 2013. This level of subnational granularity was used because the cause-specific mortality estimates for China were generated using provincial-level datasets and statistical modelling. However, the spatial nature of the

exposure data used to calculate the population-attributable fraction of ambient air pollution disease burden is such that estimating at a much higher level of resolution is feasible. Given that the exposure to ambient air pollution of the Chinese population is generated at the pixel level and then aggregated to provinces for application mortality estimates, the same underlying statistical principles support the subsequent disaggregation of the calculated attributable mortality back down to pixels.

The current sociopolitical climate in China is very amenable to changes in policies relevant to air pollution as growing public concern has resulted in aggressive reporting on the issue from state-sponsored media. A Pew Research Center poll (2013) found air pollution to be a primary public concern, noting that from 2008 to 2013 the percentage of Chinese who rated it to be a very big problem had increased by sixteen to 47% and that eleven percent of that change had occurred in the past year alone. The Chinese government's response to air pollution and the resulting health burden has major implications when considering the substantial economic and demographic shifts taking place there as a part of the shift from developing nation to global superpower. Beyond China's emergence as a key global influence, the dynamics of global atmospheric circulation make pollution there relevant to the health of many surrounding populations (Zhang 1997). Due to the aforementioned factors, research that contributes to a more nuanced response to air pollution in China has the potential to translate into major population health gains by contributing to increased policy relevance of the mortality attributable to air pollution estimates generated as part of the Global Burden of Disease 2013 Risk Factors study.

Methodology

In order to estimate the mortality attributable to ambient particulate air pollution, a variety of epidemiological data sources collected across time and space are combined into a single coherent dataset that can be manipulated using computer software and statistical methodology. First, the exposure to PM2.5 must be derived at the most granular level of spatial resolution. Several data sources and methods are used in order to improve accuracy from GBD 2010. The initial input comes in the form of gridded estimates calculated from satellite data and chemical transport models. The former data is created using aerosol optical depth (AOD) readings from satellites.

AOD is a measurement of how much light extinction occurs as a solar beam travels through the earth's atmosphere. Particles in the atmosphere can absorb or scatter the energy in a ray of light, as a result of this process the amount of direct sunlight that does not reach the Earth's surface is recorded as the AOD. AOD is a function of the volume of particles in the vertical column of atmosphere directly over the location of a given observation (Augustine 2008). The conversion rate between AOD and PM2.5 were generated using the simulated relationship between the two variables in the Goddard Earth Observing System chemical transport model (GEOS-Chem). The conversion rate was also calibrated to reflect ground levels, of which exposure is more germane to human health impacts, by using information from the CALIOP satellite instrument to determine the relationship between AOD in the entire atmospheric column and the ground-level PM2.5 concentration (Winker 2003). These estimates

of gridded PM2.5 were generated at the resolution of 0.1° x 0.1° for 2000, 2005, 2010 and 2011. This would be roughly equivalent to 6.5 square miles at the equator.

The second input to the gridded dataset is provided by TM5-FASST, which is a nested 3-dimensional global atmospheric chemical transport model (Van Dingenen 2014). This model is used to simulate how atmospheric chemicals interact with each other on the global scale and in this case predict the reactions between species of interest in the life cycle of particles that are relevant to ambient air pollution. In order to predict PM2.5 concentrations, data on international aviation and shipping, forest fires, average meteorology, dust, sea salt, and industrial emissions inventory are input to the model. TM5-FASST produces estimates at the 1° x 1° resolution. These estimates are resampled to the 0.1° x 0.1° resolution to prepare for merging with the satellite data using population density provided by Gridded Population of the World, version 3 (GPWv3). GPWv3 is a population dataset that is created through the comprehensive collection of tabular population data, followed by spatial matching of these counts to geographic boundaries and transformation to grids with the assumption of proportional allocation within the boundary (SEDAC 2015).

The satellite data and TM5-FASST estimates are averaged at the 0.1° x 0.1° grid cell resolution in order to create a hybrid estimate. The annual rate of change between 2010 and 2011 was calculated and then used to extrapolate to 2013 using an exponential growth function. The major strength of the resulting dataset is that it provides information on the relative difference in PM2.5 between geographic areas through unified methodology, which is highly relevant to the underlying philosophy inherent to the Global Burden of Disease project. However, these estimates can be divergent from the “gold standard” of PM2.5 data: monitoring

stations that physically measure ground-level PM2.5. As such, the gridded PM2.5 data is further calibrated by estimating the relationship between estimated and measured PM2.5 in any grids that can be matched to a ground monitoring station.

In order to conduct this calibration, first a dataset of global ground monitoring data must be compiled. The combination of a literature review with WHO database extraction and input from an international network of GBD expert group collaborators yielded more than four thousand datapoints in 79 different GBD countries. These datapoints comprised 3,387 unique ground monitoring locations. Unfortunately, only 46% of the datapoints were direct measurements of PM2.5. The rest were measurements of PM10 that were converted to PM2.5 using several methods. If possible, they were converted using a ratio of PM2.5 to PM10 from other ground monitors that were within 50 km – this approach covered 1,151 datapoints, or 30% of the total. Another 899 datapoints were estimated using a ratio derived from monitors in the same country. Finally, a remaining 169 monitors that did not fit the first two criteria were simply estimated using the ratio of 0.5, from GBD 2010.

The finished ground monitor dataset was then spatially matched to the relevant grids. Regression techniques were used on the merged data in order to quantify the relationship between the estimated PM2.5 in a grid and the measured ground-level PM2.5. In order to improve model fit, interaction terms were used to account for measurement error in the ground-monitoring data. Available information on the measurement error included whether PM2.5 there was directly measured or estimated from PM10, whether the type of the monitoring site was known, and if the exact coordinates of the site were provided. The inclusion of these variables in the regression improved the RMSE and R^2 of our model. Regional

random effects were also tested in order to account for potential differences in the relationship between estimated and measured PM2.5 across geographies, with the rationale the standards of PM2.5 ground monitoring may not be the same for all countries. The existence of regional variation was demonstrated by this model, but overall the fits were not improved enough to justify this approach. The coefficients of the resulting calibration model were sampled using the variance-covariance matrix in order to create draws of uncertainty, and then applied to the global gridded PM2.5 dataset in order to predict ground-monitored PM2.5 exposure at the 0.1° x 0.1° resolution worldwide.

The next stage of the analysis is to estimate the increased risk of mortality caused by a given level of PM2.5 exposure. To accomplish this, integrated exposure response (IER) curves were calculated for the health outcomes that are linked to ambient air pollution exposure by epidemiological studies. The concept of IER curves was introduced in GBD 2010, positing that health impacts across several different kinds of exposure – including ambient air pollution, household air pollution, secondhand smoking, and direct smoking – are caused by varying levels of PM2.5 inhalation. This equation was estimated by assuming a power function and then using non-linear least squares (nls) to fit it according to epidemiological data across the full range of exposure for the different sources of PM2.5. Figure 1.1 displays the functional form and weakly informative priors used to the fit the IER curve.

Figure 1.1: IER Curve Equation and Relevant Assumptions

Data Likelihood

$$\log(RR_i) \sim \mathcal{N}(\mu_i, \sigma_i)$$

Model

$$\mu_i = \log\left(\frac{1 + \alpha \times \left(1 - e^{-\beta \times (exposure_i - TMRED)^\gamma}\right)}{1 + \alpha \times \left(1 - e^{-\beta \times (counterfactual_i - TMRED)^\gamma}\right)}\right)$$

Data

RR_i : measured relative risk for data point i
σ_i : variance of data point i based on study information
$TMRED$: theoretical minimum
$exposure_i$: measured exposure for data point i
$counterfactual_i$: counterfactual exposure for data point i

Priors

$$\alpha \sim \Gamma(1.0, 0.01)$$
$$\beta \sim \Gamma(1.0, 0.01)$$
$$\gamma \sim \Gamma(1.0, 0.01)$$

The data used to fit the equation came from studies estimating the relative risk of relevant health outcomes based on PM2.5 exposure. The level of PM2.5 for a given study was mapped using literature values or determined through collaboration with the study authors.

For the current iteration of GBD, newly published relative risks were added to the curve-fitting dataset such that convergence rates became inappropriately low. Given this issue, new methods were used to estimate the parameters of the equation. Stan, a Bayesian modelling approach utilizing Monte-carlo Markov Chain (MCMC) sampling methods, was employed with weakly informative priors guiding convergence of the parameters (Stan 2014). Similar to the previous model in GBD 2010, uncertainty in the RR estimates was reflected in the weight of

each datapoint in the curve-fitting approach. The mean and confidence interval of each of the three parameters was estimated by taking 1,000 draws during modelling. Four MCMC chains were used for each parameter, each running for 20,000 samples and dropping the first 10,000 warmup walks. Every 20th sample was kept thereafter in order to yield 1,000 draws of the parameters. Figures 2.1-5 show the MCMC chains for each cause as they converge for the three parameters being estimated by Stan.

Figure 2.1: 4-chain MCMC sampling for Cardiovascular Ischemic Heart Disease (age 50)

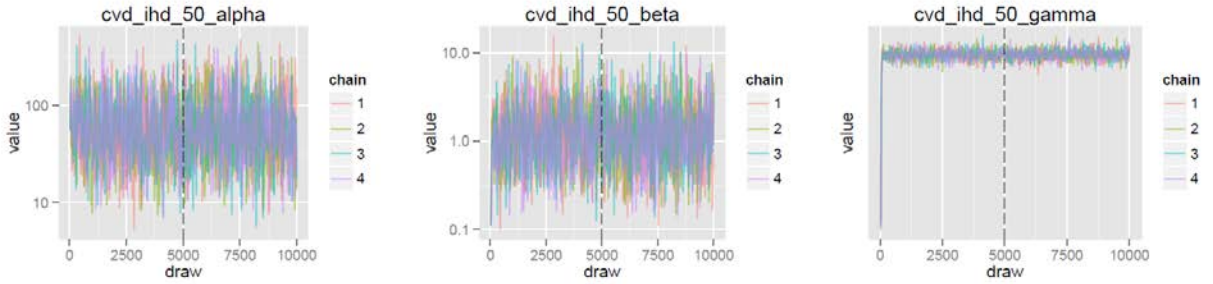


Figure 2.2: 4-chain MCMC sampling for Cardiovascular Stroke (age 75)

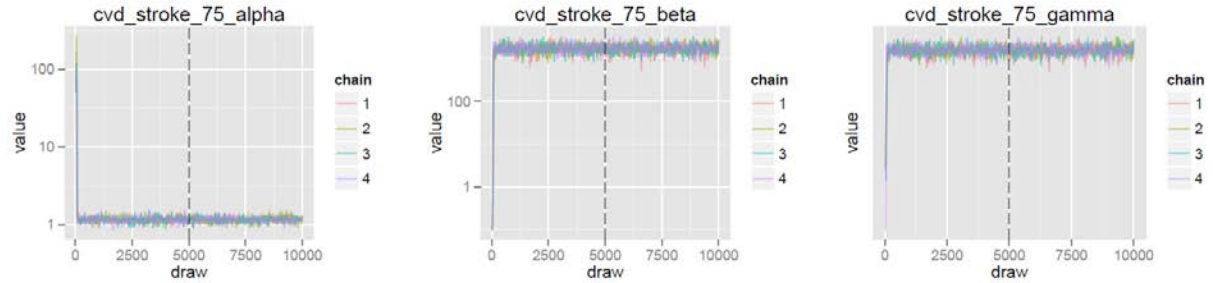


Figure 2.3: 4-chain MCMC sampling for Lower Respiratory Infection (All age)

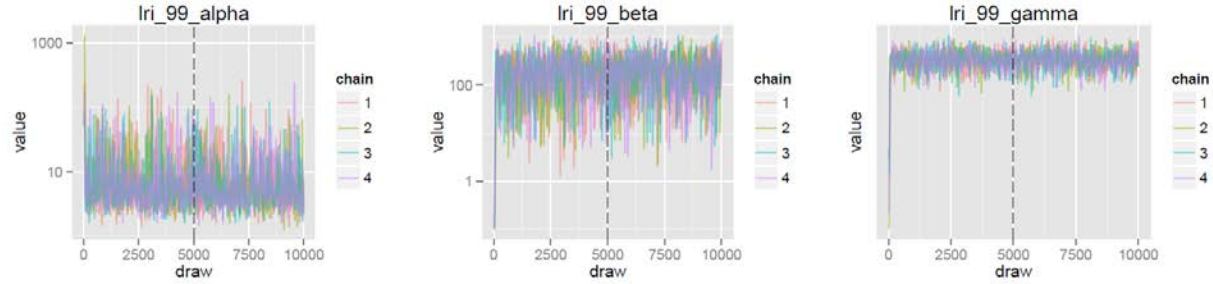


Figure 2.4: 4-chain MCMC sampling for Lung Cancer (All age)

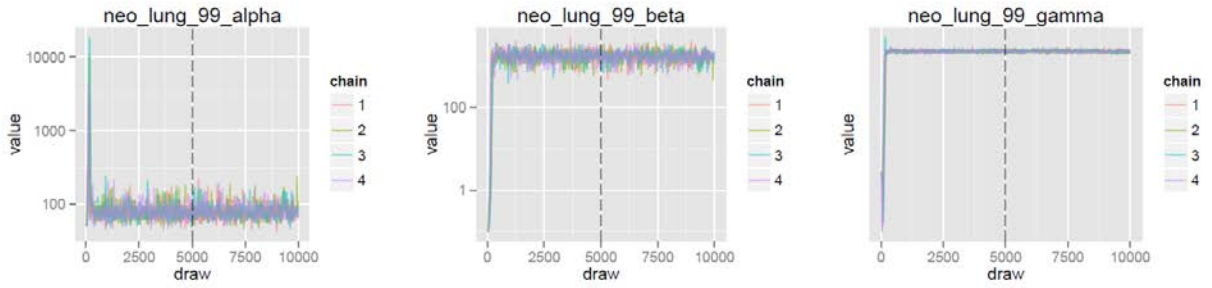
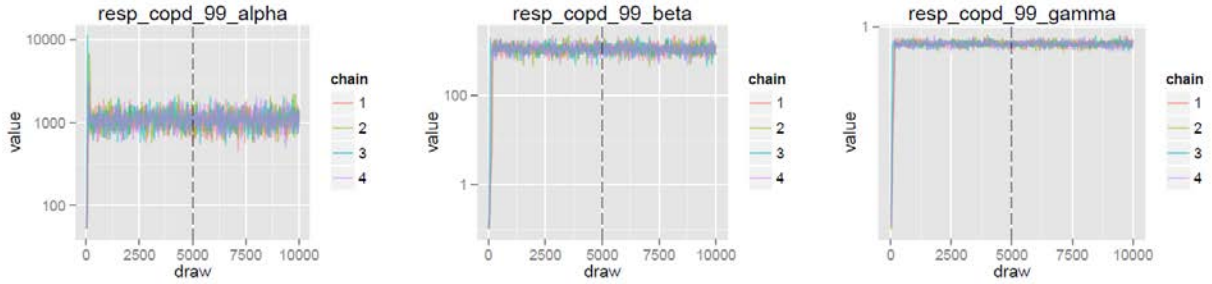


Figure 2.5: 4-chain MCMC sampling for Respiratory COPD (All age)



An IER curve was created for each of the relevant causes: cardiovascular ischemic heart disease (IHD), cardiovascular stroke, lower respiratory infection (LRI), lung cancer, and chronic obstructive pulmonary disease (COPD). For the latter three causes, a single curve was used for all age groups. For IHD and stroke, age specific curves were generated using the reported ages in each study and the GBD2010 assumption that 110 years old is the terminal age where death is imminent and all risks have a relative risk of 1. There was no granularity for sex in the IER curves. The resulting estimates of the three parameters for each cause could then be used to estimate cause-specific relative risks at the grid-level based on the predicted PM2.5. Figures 3.1-5 display the shape of IER curves and the relative risk data used to fit them, with the colors denoting different sources of exposure, the X axis denoting the intensity of exposure, and the Y axis denoting the relative risk.

Figure 3.1: IER curve for Cardiovascular Ischemic Heart Disease (age 50) – Left is all exposure, right is truncated to 0-300 ug/m^3, levels of exposure relevant to ambient air pollution

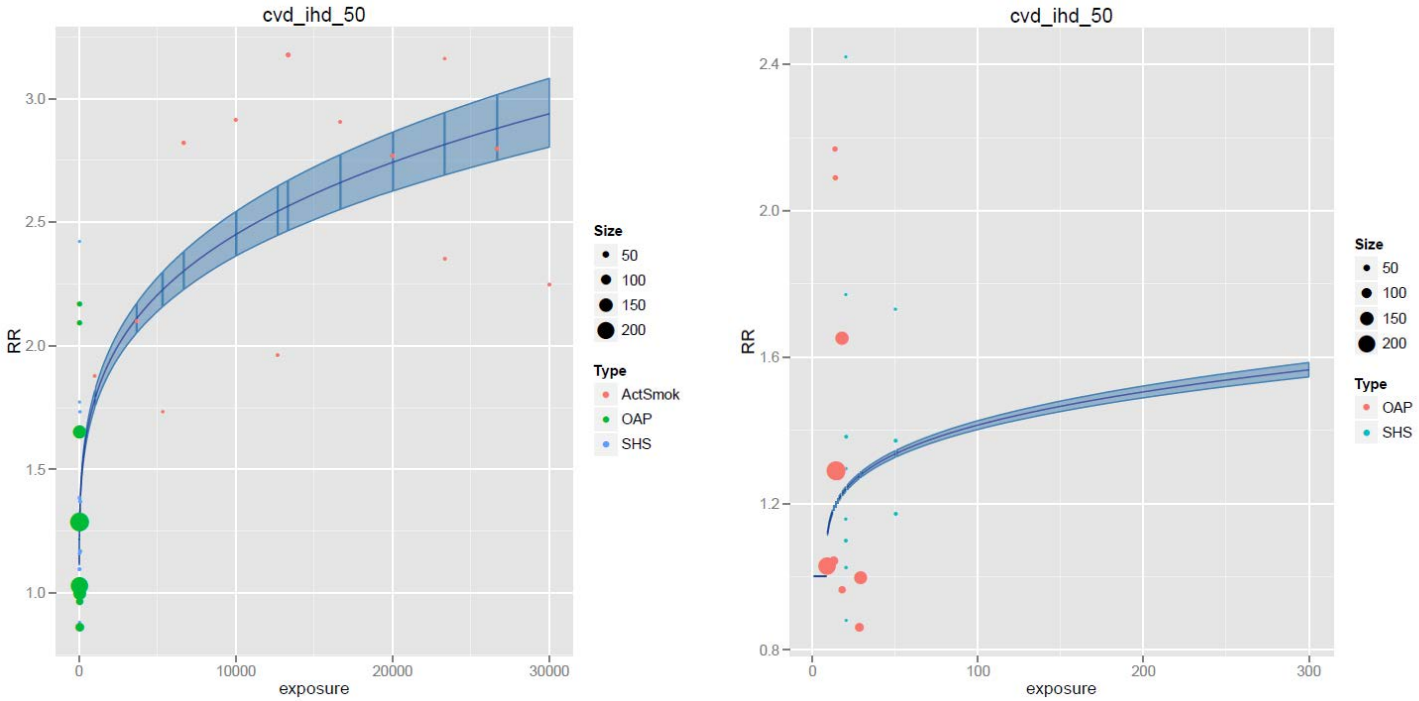


Figure 3.2: IER curve for Cardiovascular Stroke (age 50) – Left is all exposure, right is truncated to 0-300 ug/m^3, levels of exposure relevant to ambient air pollution

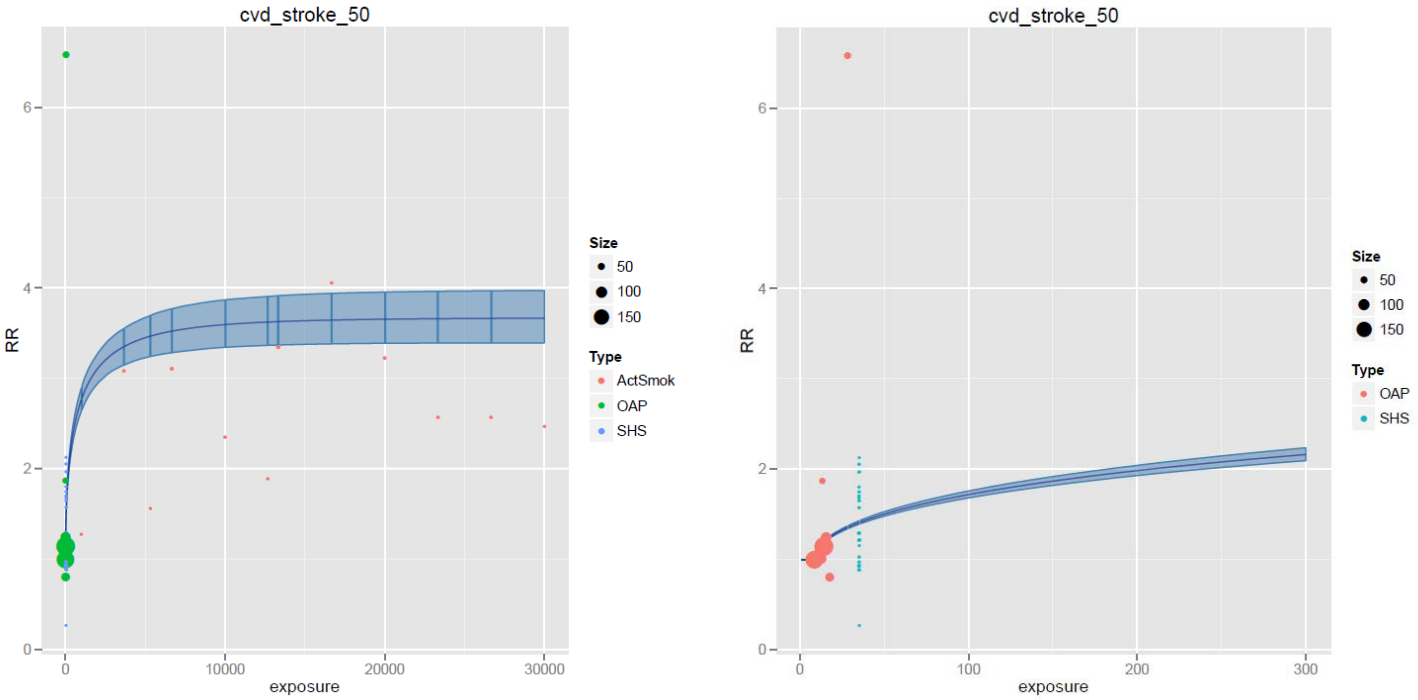


Figure 3.3: IER curve for Lower Respiratory Infection (all age) – Left is all exposure, right is truncated to 0-300 ug/m³: levels of exposure relevant to ambient air pollution

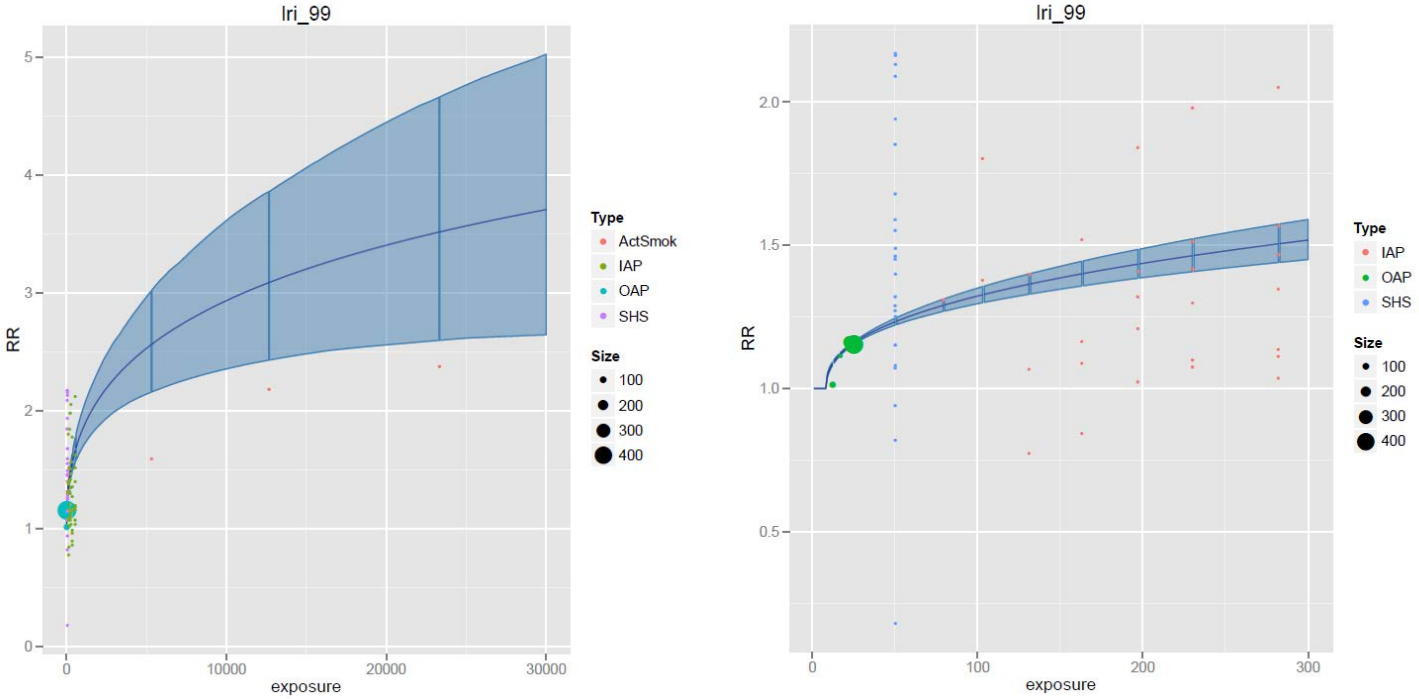


Figure 3.4: IER curve for Lung Cancer (all age) – Left is all exposure, right is truncated to 0-300 ug/m³: levels of exposure relevant to ambient air pollution

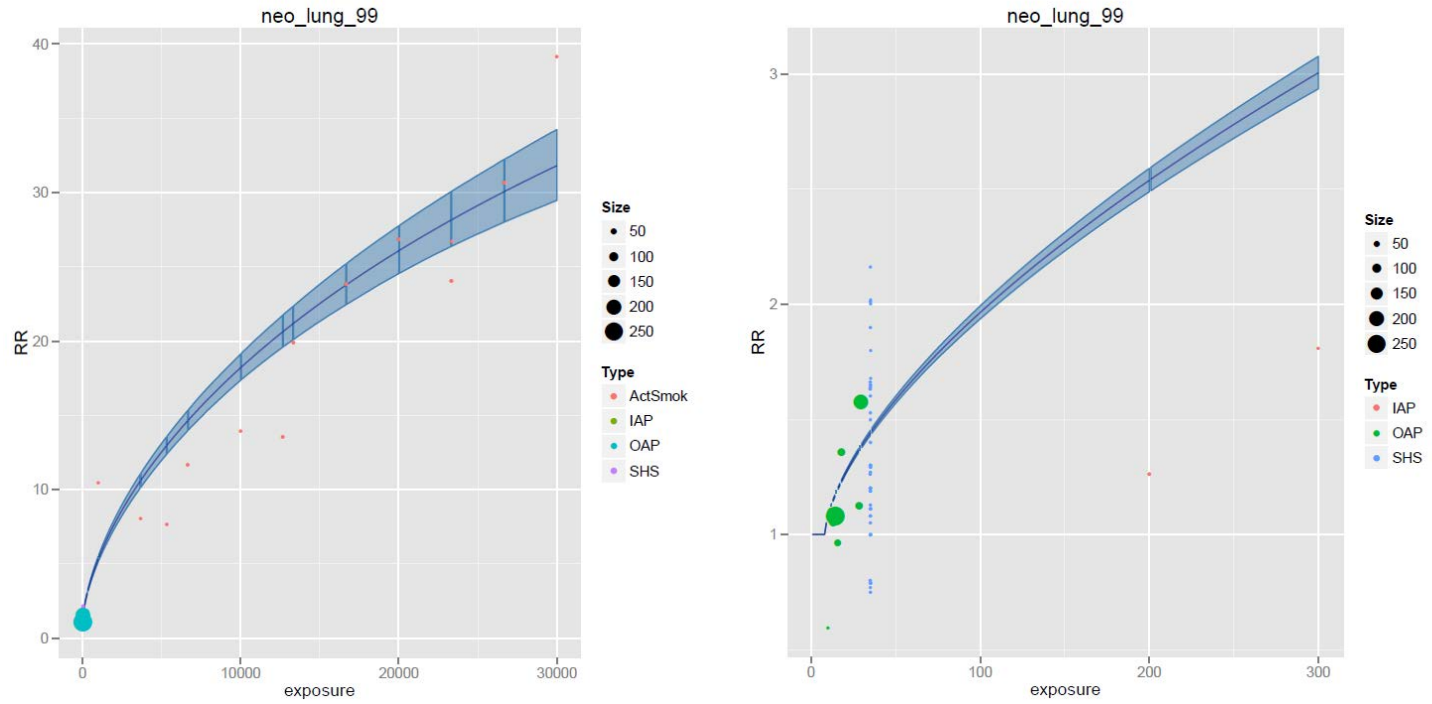
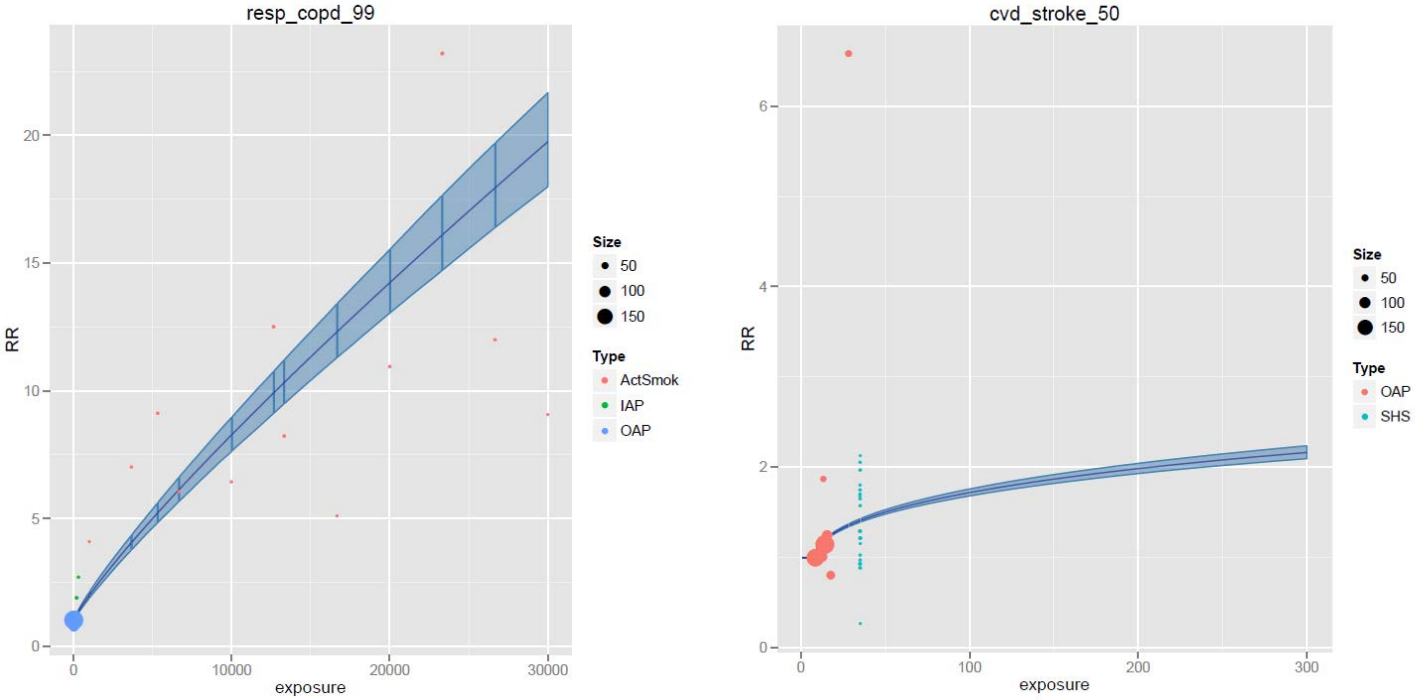


Figure 3.5: IER curve for Respiratory COPD (all age) – Left is all exposure, right is truncated to 0-300 ug/m³: levels of exposure relevant to ambient air pollution



To estimate the mortality attributable to air pollution in China, Population Attributable Fractions (PAFs) are calculated for each cause. Cause-specific mortality was estimated at the provincial level in China and published as a part of the Global Burden of Disease 2013 (Murray 2015). In order to be merged onto the mortality data, PAFs were aggregated to the provincial level using population weights averages determined by GPWv3. Provincial-level cause-specific PAFs were multiplied against the relevant cause-specific mortality in order to produce the burden attributable to ambient air pollution for every province in China.

As previously stated, the provincial mortality estimates provide new insight to the epidemiological distribution of disease burden that is attributable to air pollution. They can be used to inform policy-makers on interventions that can be conducted at the province level

based on the circumstances therein. However, the environmental nature of air pollution means that having estimates of attributable mortality at a higher level of spatial resolution would be very useful when considering interventions that could occur at a city or even neighborhood level, targeting areas of key importance when considering how air pollution disease burden is geographically distributed. To accomplish this, the provincial level estimates can be disaggregated back down to the level of exposure through the reversal of the process used to aggregate them. First, the cause-specific mortality rate in each grid must be calculated by using the provincial mortality rate and assumption that rates are constant across grids in a given province. By applying the relationship between the grid-level PAF and the provincial population average PAF to the provincial mortality rate, the mortality rate for a grid can be estimated. Then, by multiplying the grid-level PAF to the grid-level mortality rate, cause-specific attributable mortality rates can be estimated at the 0.1° x 0.1° resolution. The equations used for this estimation process are displayed in Figure 4.1.

Figure 4.1: Equations for spatial interpolation of provincial mortality back to grid level

$$MORT_{i,c} = \frac{MORT_c(1 - PAF_c)}{1 - PAF_{i,c}}$$

$$MORT_{i,c} * PAF_{i,c} = A.MORT_{i,c}$$

A.MORTi,c = Attributable mortality rate in each grid(i) for cause/age(c)

MORTi,c = Underlying mortality rate in each grid(i) for cause/age(c)

MORTc = Province-specific mortality rate for cause/age(c)

PAFc = Province-specific PAF for cause/age(c)

RRi,c = Grid-level RR for cause/age(c), based on grid-level calibrated exposure and IER curve for cause/age(c)

$PAF_{i,c} = (RR_{i,c} - 1)/RR_{i,c}$

Finally, the total attributable mortality in a grid can be calculated by multiplying the mortality rate by the population in a grid. However, GPWv3 is for all ages, and this process must be done iteratively across age groups in order to aggregate by cause. As such, the population composition of the province is assumed to be constant across grids and used to scale the grid-level population into age-specific population. The final output of this analysis is the mortality attributable to ambient particulate matter air pollution, by cause and total, for every pixel in China.

Results

The results of this analysis are best displayed in graphical format in order to portray the spatial and temporal distributions that are unveiled. Figures 5.1-2 indicate the change in mortality rate by grid over the entire GBD time series. It is clear that the mortality rates follow the trends in PM2.5 exposure during this time period. Of special note is the high mortality rate seen in Western china, which is driven by the dust storms of the Gobi desert. As previously mentioned, however, the population in this region is limited and therefore the levels of PM2.5 here have a negligible impact on the total burden. As such, maps showing the total mortality after conversion from the rate are useful to visualize the areas that make more significant contributions to the total picture of burden attributable to ambient air pollution. Figures 5.3-4 display the total mortality per pixel, giving us a better idea of how the spatial and temporal distribution here has changed over this time period.

Figure 5.1: Rate of mortality attributable to ambient air PM2.5 (per 100,000 population) for each pixel in China, estimated for the year 1990.

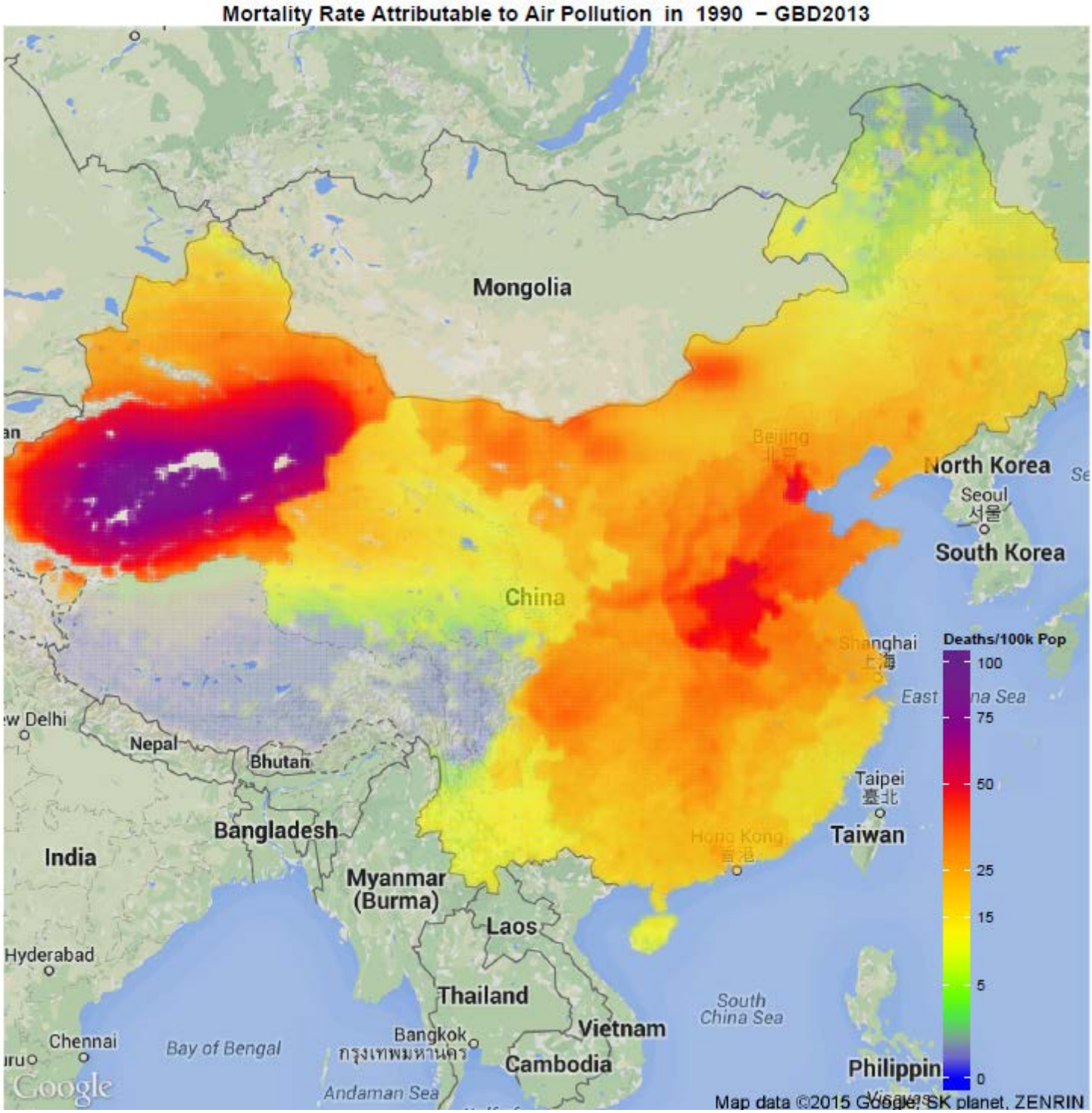


Figure 5.2: Rate of mortality attributable to ambient air PM2.5 (per 100,000 population) for each pixel in China, estimated for the year 2013

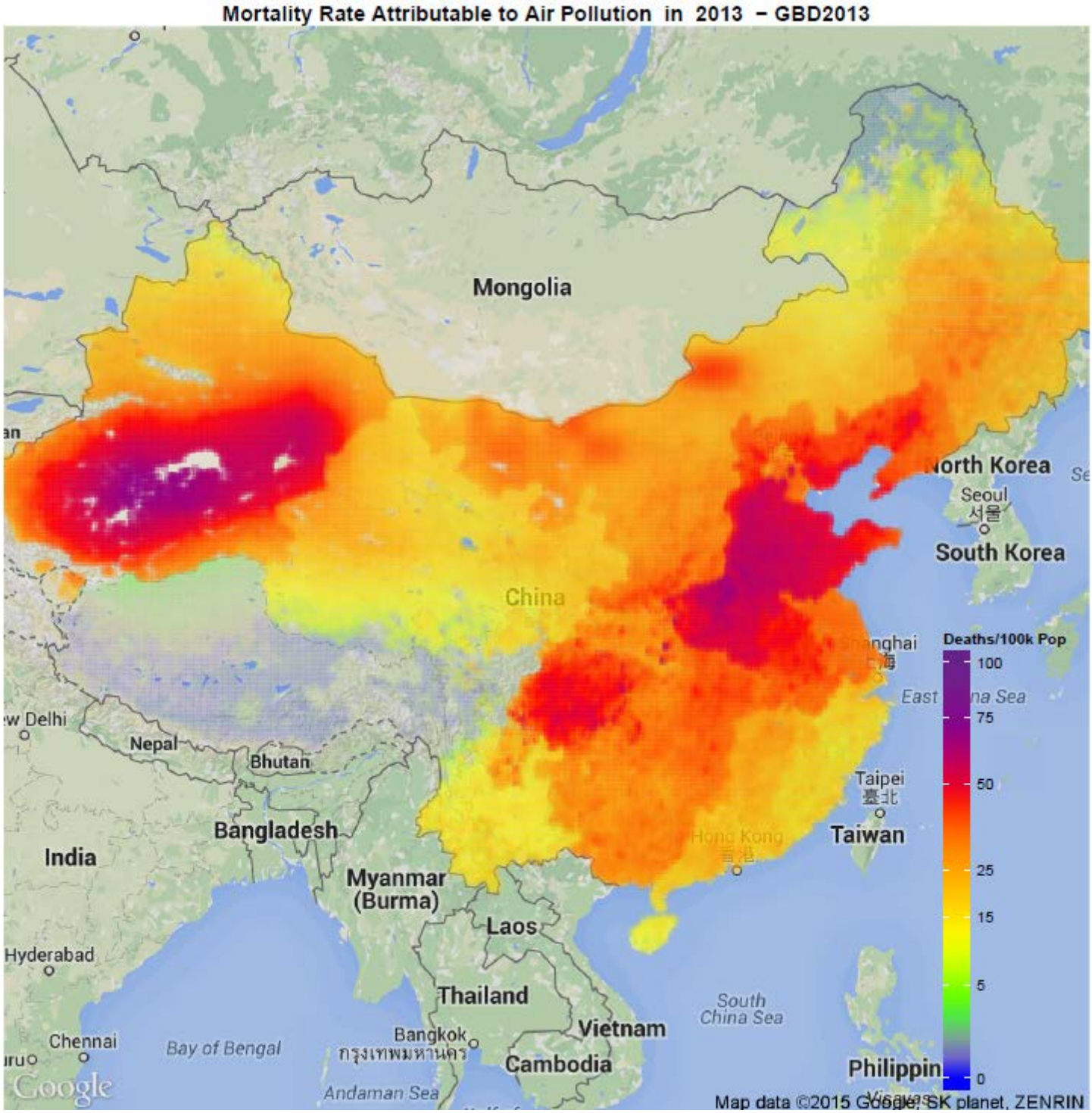


Figure 5.3: Total mortality attributable to ambient air PM2.5 (per 100,000 population) for each pixel in China, estimated for the year 1990

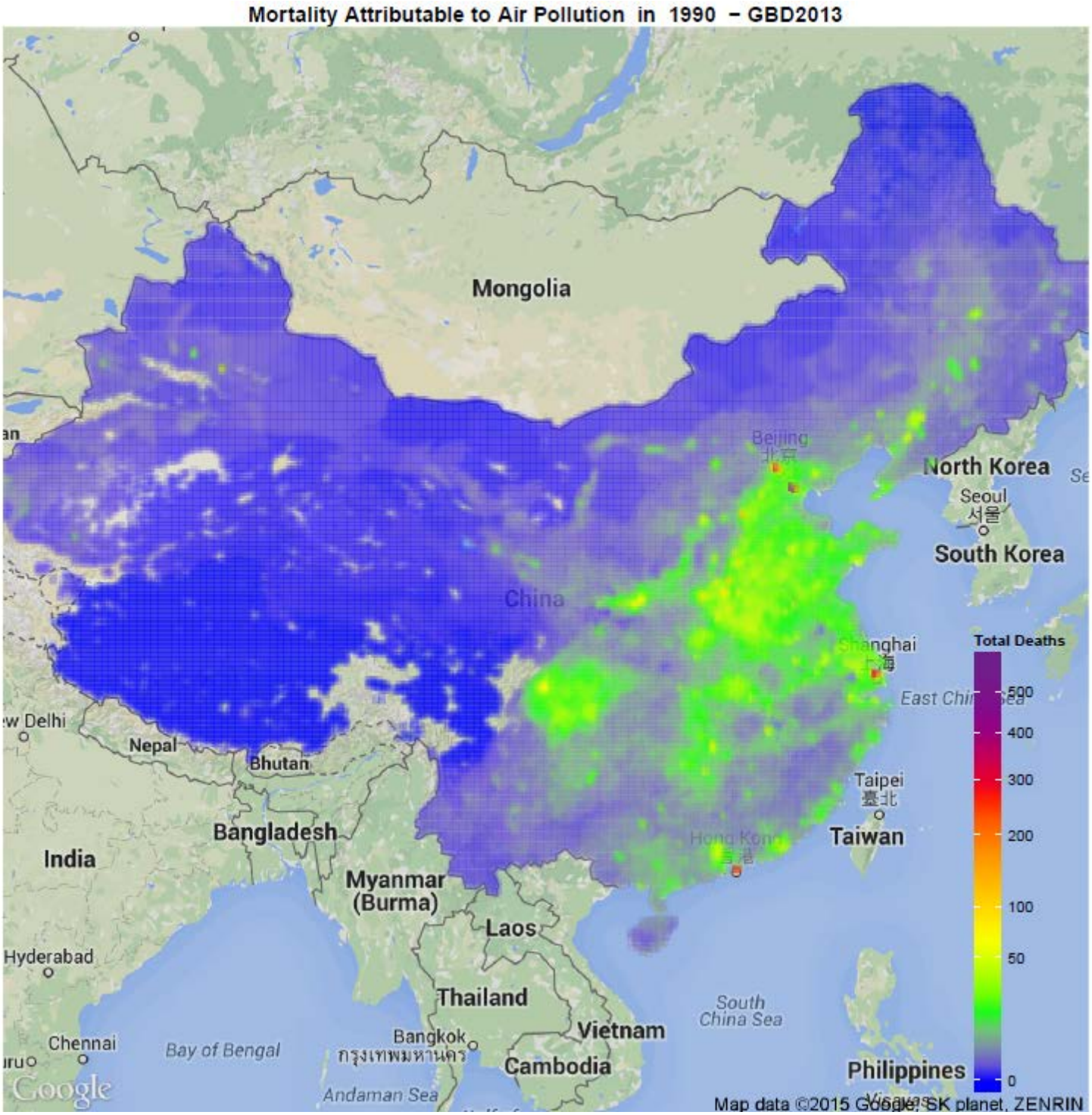
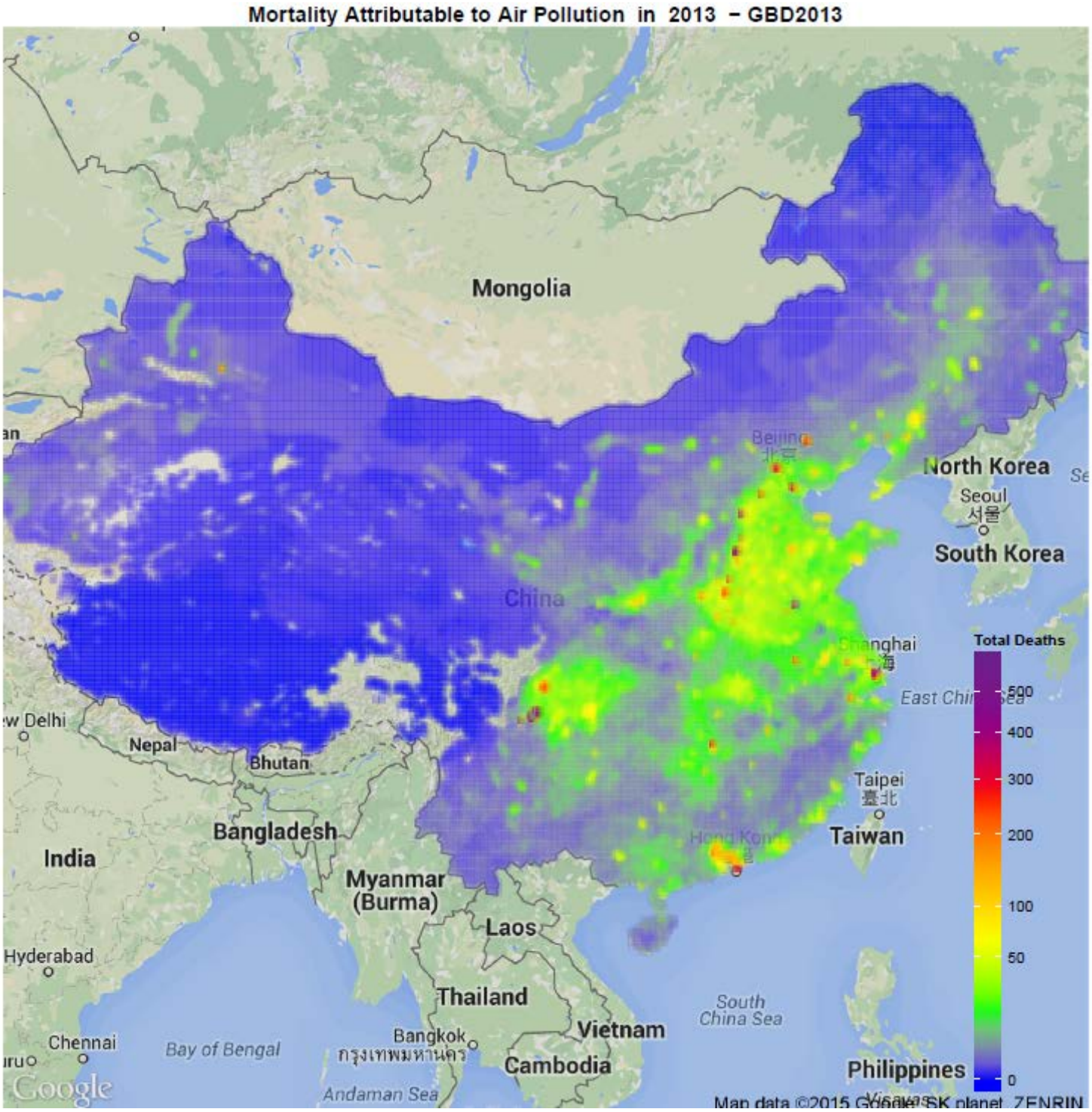


Figure 5.4: Total mortality attributable to ambient air PM2.5 (per 100,000 population) for each pixel in China, estimated for the year 2013

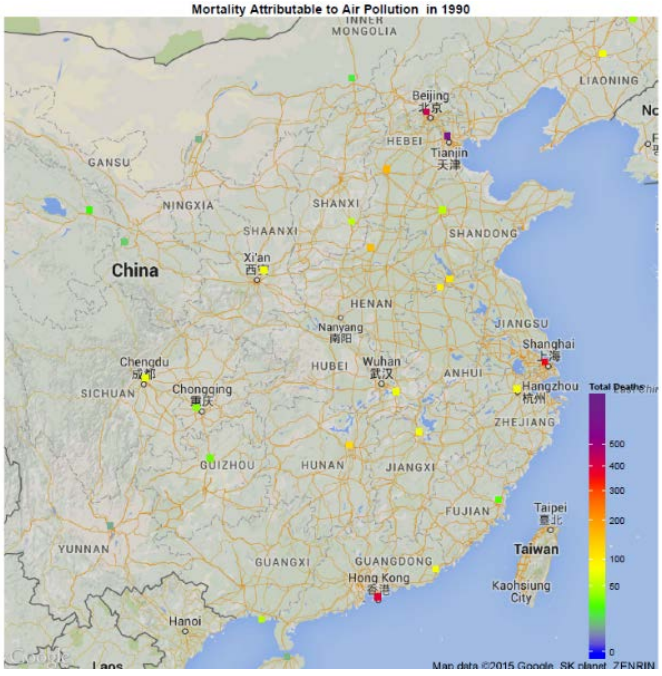


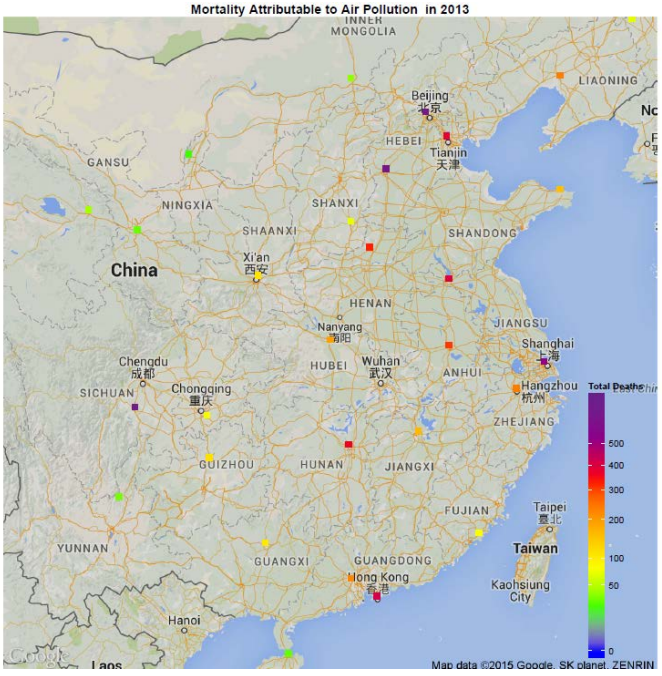
The maps of total mortality show that the regions of high exposure/mortality rate in Western China are largely a nonfactor when considering the total burden. Instead, the population centers in Eastern China are driving the majority of mortality that is being caused by ambient air pollution. The intense urbanization, industrialization, and population growth seen in China over the past two decades has a clear impact on the spatial distribution of mortality attributable to air pollution. Whereas burden was largely being driven by major National Central Cities like Shanghai, Beijing, and Tianjin, the rapid growth of other major cities as centers of population and industry has seen mortality counts rise as well. In particular, Sichuan province is now responsible for a much greater share of total mortality than in 1990. The cities and provinces directly southeast of Beijing have also seen their share of mortality increase, potentially as a result of pollution travelling from major industrial and population centers via atmospheric currents. Another interesting analysis that can be conducted as a result of this spatial interpolation is the ranking of grids by the amount of total mortality seen. Tables 1.1-2

ISO3	Name	Year	Pop	PM2.5	Deaths	Death Rate	Address
CHN_354	Hong Kong Special A	2013	1,646,592.0	28.6	411.7	25.0	Cheung Yuen Road, Kam Shan, Hong Kong
CHN_491	Anhui	2013	646,233.6	76.1	309.7	47.9	011 Xian Dao, Luyang Qu, Hefei Shi, Anhui Sheng, China
CHN_492	Beijing	2013	1,916,308.0	86.6	695.4	36.3	Xi Bei Wang Xi Lu, Haidian Qu, Beijing Shi, China, 100094
CHN_493	Chongqing	2013	138,151.3	91.5	78.4	56.7	001 Xiang Dao, Banan Qu, Chongqing Shi, China
CHN_494	Fujian	2013	444,369.0	29.9	81.8	18.4	Feng Chi Lu, Jinjiang Shi, Quanzhou Shi, Fujian Sheng, China, 362005
CHN_495	Gansu	2013	118,751.0	58.6	29.5	24.8	58 Jiu Da Lu, Chengguan Qu, Lanzhou Shi, Gansu Sheng, China, 730000
CHN_496	Guangdong	2013	865,937.3	56.6	222.2	25.7	21 Yan Jiang Nan Lu, Nanhai Qu, Foshan Shi, Guangdong Sheng, China
CHN_497	Guangxi	2013	284,545.2	49.3	102.7	36.1	San He Lu, Liubei Qu, Liuzhou Shi, Guangxi Zhuangzuzizhiqu, China
CHN_498	Guizhou	2013	304,935.0	53.9	113.6	37.3	311 Xian Dao, Huichuan Qu, Zunyi Shi, Guizhou Sheng, China
CHN_499	Hainan	2013	235,173.3	19.9	28.7	12.2	31 Hai Dian Er Dong Lu, Meilan Qu, Haikou Shi, Hainan Sheng, China, 570000
CHN_500	Hebei	2013	1,071,361.0	99.3	635.2	59.3	16 Che Zhan Nan Da Jie, Zhengding Xian, Shijiazhuang Shi, Hebei Sheng, Chi
CHN_501	Heilongjiang	2013	223,063.4	44.7	76.4	34.2	China, Heilongjiang Sheng, Haerbin Shi, Songbei Qu, 301569053
CHN_502	Henan	2013	489,669.1	124.2	324.0	66.2	Meng Dian Gong Ye Da Dao, Huixian Shi, Xinxiang Shi, Henan Sheng, China
CHN_503	Hubei	2013	498,026.3	62.5	193.4	38.8	Xi Wang Lu, Xiangzhou Qu, Xiangyang Shi, Hubei Sheng, China
CHN_504	Hunan	2013	921,552.4	63.2	361.2	39.2	559 Yin Shan Lu, Yuelu Qu, Changsha Shi, Hunan Sheng, China
CHN_505	Inner Mongolia	2013	169,142.7	25.3	39.2	23.2	310 Sheng Dao, Chahaeryouyiqian Qi, Wulanchabu Shi, Neimenggu Zizhiqu,
CHN_506	Jiangsu	2013	1,132,598.0	62.6	395.9	35.0	104 Guo Dao, Tongshan Qu, Xuzhou Shi, Jiangsu Sheng, China
CHN_507	Jiangxi	2013	509,039.5	63.9	160.2	31.5	Nan Jiang Lu, Qingshanhu Qu, Nanchang Shi, Jiangxi Sheng, China, 330029
CHN_508	Jilin	2013	222,657.6	36.2	70.3	31.6	Unnamed Road, Dongliao Xian, Liaoyuan Shi, Jilin Sheng, China
CHN_509	Liaoning	2013	558,392.4	42.3	236.3	42.3	Song Tong Xian, Panshan Xian, Panjin Shi, Liaoning Sheng, China
CHN_510	Ningxia	2013	98,490.2	48.7	23.7	24.1	Zhen Su Lu, Xixia Qu, Yinchuan Shi, Ningxia Huizuzizhiqu, China
CHN_511	Qinghai	2013	213,731.8	35.7	44.7	20.9	259 Ning Da Lu, Chengbei Qu, Xining Shi, Qinghai Sheng, China
CHN_512	Shaanxi	2013	277,090.2	78.3	109.4	39.5	Wei Fang Lu, Baqiao Qu, Xian Shi, Shaanxi Sheng, China
CHN_513	Shandong	2013	361,235.9	42.6	152.8	42.3	Chang Hua Lu, Huancui Qu, Weihai Shi, Shandong Sheng, China, 264203
CHN_514	Shanghai	2013	2,255,308.0	58.7	507.5	22.5	2145 Bao An Gong Lu, Baoshan Qu, Shanghai Shi, China
CHN_515	Shanxi	2013	188,163.1	74.8	72.4	38.5	Wen Ming Lu, Changzhi Shi, Shanxi Sheng, China
CHN_516	Sichuan	2013	1,440,650.0	77.1	701.7	48.7	Le Qing Lu, Shizhong Qu, Leshan Shi, Sichuan Sheng, China
CHN_517	Tianjin	2013	1,049,574.0	69.0	375.9	35.8	Tie Dong Bei Lu Chu Kou, Beichen Qu, Tianjin Shi, China
CHN_518	Tibet	2013	3,600.2	13.3	0.1	1.5	318 Guo Dao, Rikaze Shi, Rikaze Diqu, Xizang Zizhiqu, China
CHN_519	Xinjiang	2013	794,752.2	44.1	247.4	31.1	1197 Hong Miao Zi Lu, Shayibake Qu, Wulumuqi Shi, Xinjiang Weiwuerzizhi
CHN_520	Yunnan	2013	154,880.9	41.9	31.6	20.4	Sha Mu Duan, Dongchuan Qu, Kunming Shi, Yunnan Sheng, China
CHN_521	Zhejiang	2013	963,499.8	71.1	239.0	24.8	Jin Chang Lu, Gongshu Qu, Hangzhou Shi, Zhejiang Sheng, China, 310000

ISO3	Name	Year	Pop	PM2.5	Deaths	Death Rate	Address
CHN_354	Hong Kong Special	1990	1,601,659.0	31.0	386.6	24.1	Cheung Yuen Road, Kam Shan, Hong Kong
CHN_491	Anhui	1990	286,519.0	56.1	98.5	34.4	019 Xiang Dao, Duji Qu, Huaibei Shi, Anhui Sheng, China
CHN_492	Beijing	1990	1,048,020.0	64.5	420.9	40.2	Xi Bei Wang Xi Lu, Haidian Qu, Beijing Shi, China, 100094
CHN_493	Chongqing	1990	105,461.0	52.3	35.4	33.6	G 93 Cheng Yu Huan Xian Gao Su, Shapingba Qu, Chongqing Shi, China
CHN_494	Fujian	1990	153,440.0	27.5	25.7	16.7	192 Xian Dao, Jinan Qu, Fuzhou Shi, Fujian Sheng, China
CHN_495	Gansu	1990	76,563.0	34.1	12.4	16.3	501 Xiang Dao, Linxia Shi, Linxia Huizuzizhizhou, Gansu Sheng, China
CHN_496	Guangdong	1990	356,611.0	28.8	77.2	21.7	233 Sheng Dao, Chaoan Qu, Chaozhou Shi, Guangdong Sheng, China
CHN_497	Guangxi	1990	255,099.0	22.7	49.7	19.5	Chang Sha Lu, Yinhai Qu, Beihai Shi, Guangxi Zhuangzuzizhiqu, China
CHN_498	Guizhou	1990	115,361.0	31.8	30.3	26.3	311 Xian Dao, Huichuan Qu, Zunyi Shi, Guizhou Sheng, China
CHN_499	Hainan	1990	68,223.0	20.8	9.3	13.6	Unnamed Road, Xiuying Qu, Haikou Shi, Hainan Sheng, China
CHN_500	Hebei	1990	434,986.0	67.1	165.8	38.1	16 Che Zhan Nan Da Jie, Zhengding Xian, Shijiazhuang Shi, Hebei Sheng,
CHN_501	Heilongjiang	1990	172,587.0	27.8	41.7	24.2	1555 Zhong Yuan Da Dao, Songbei Qu, Haerbin Shi, Heilongjiang Sheng,
CHN_502	Henan	1990	311,462.0	73.0	158.0	50.7	Meng Dian Gong Ye Da Dao, Huixian Shi, Xinxiang Shi, Henan Sheng, Chi
CHN_503	Hubei	1990	216,503.0	48.0	72.6	33.6	Bi Huang Xian, Echeng Qu, Ezhou Shi, Hubei Sheng, China
CHN_504	Hunan	1990	369,513.0	49.4	120.3	32.6	559 Yin Shan Lu, Yuelu Qu, Changsha Shi, Hunan Sheng, China
CHN_505	Inner Mongolia	1990	86,212.0	22.7	18.0	20.9	310 Sheng Dao, Chahaeryouyiqian Qi, Wulanchabu Shi, Neimenggu Zizh
CHN_506	Jiangsu	1990	380,985.0	59.2	109.6	28.8	104 Guo Dao, Tongshan Qu, Xuzhou Shi, Jiangsu Sheng, China
CHN_507	Jiangxi	1990	281,129.0	46.3	75.9	27.0	Nan Jiang Lu, Qingshanhu Qu, Nanchang Shi, Jiangxi Sheng, China, 33002
CHN_508	Jilin	1990	183,660.0	21.3	40.9	22.3	Unnamed Road, Dongliao Xian, Liaoyuan Shi, Jilin Sheng, China
CHN_509	Liaoning	1990	293,110.0	29.1	75.8	25.9	435 Xian Dao, Shuncheng Qu, Fushun Shi, Liaoning Sheng, China
CHN_510	Ningxia	1990	38,413.0	35.9	8.6	22.4	301 Sheng Dao, Dawukou Qu, Shizuishan Shi, Ningxia Huizuzizhiqu, Chin
CHN_511	Qinghai	1990	144,704.0	27.3	23.0	15.9	259 Ning Da Lu, Chengbei Qu, Xining Shi, Qinghai Sheng, China
CHN_512	Shaanxi	1990	204,964.0	68.3	76.8	37.5	Nong Xing Nan Jie, Yanliang Qu, Xian Shi, Shaanxi Sheng, China, 710089
CHN_513	Shandong	1990	124,538.0	61.6	51.2	41.1	309 Guo Dao, Tianqiao Qu, Jinan Shi, Shandong Sheng, China
CHN_514	Shanghai	1990	1,608,711.0	38.8	353.5	22.0	2145 Bao An Gong Lu, Baoshan Qu, Shanghai Shi, China
CHN_515	Shanxi	1990	140,028.0	51.0	52.4	37.4	Wen Ming Lu, Changzhi Shi, Shanxi Sheng, China
CHN_516	Sichuan	1990	214,622.0	56.3	72.7	33.9	1980 Shu Long Da Dao Nan Duan, Xindu Qu, Chengdu Shi, Sichuan Sheng
CHN_517	Tianjin	1990	1,215,570.0	65.4	651.5	53.6	Tie Dong Bei Lu Chu Kou, Beichen Qu, Tianjin Shi, China
CHN_518	Tibet	1990	2,148.0	16.2	0.0	1.2	10 Tuan Jie Lu, Nielamu Xian, Rikaze Diqu, Xizang Zizhiqu, China
CHN_519	Xinjiang	1990	371,455.0	39.0	131.7	35.5	1197 Hong Miao Zi Lu, Shayibake Qu, Wulumuqi Shi, Xinjiang Weiwuerzi
CHN_520	Yunnan	1990	57,208.0	18.3	7.4	12.9	610 Xiang Dao, Wuhua Qu, Kunming Shi, Yunnan Sheng, China
CHN_521	Zhejiang	1990	333,890.0	46.9	75.6	22.6	Jin Chang Lu, Gongshu Qu, Hangzhou Shi, Zhejiang Sheng, China, 310000

present the grid with highest mortality for every province, along with the mortality rate per 100,000 population, the average annual PM2.5 exposure, the estimated population. The "Address" variable has been populated using geolocation based on the grid coordinates. Figures





6.1-2 show the location of these low-performing grids, indicating the intensity of mortality by color and displaying how the top grids have changed across the time series. Here, we see that some grids have remained the maximum for the province and just continued to grow, while in other provinces new grids entirely have jumped to the top of the mortality rankings. Using this geolocated data, policy-makers can determine where to target interventions, and potentially determine how or why things have changed in such a manner across the past two decades.

Another interesting result that can be drawn from this work is the change in cause composition over the time series. Figures 7.1-5 show how the cause profile of air pollution has shifted drastically over the past two decades, causing cardiovascular health to become much more important than LRI or COPD when considering the impact of ambient air pollution.

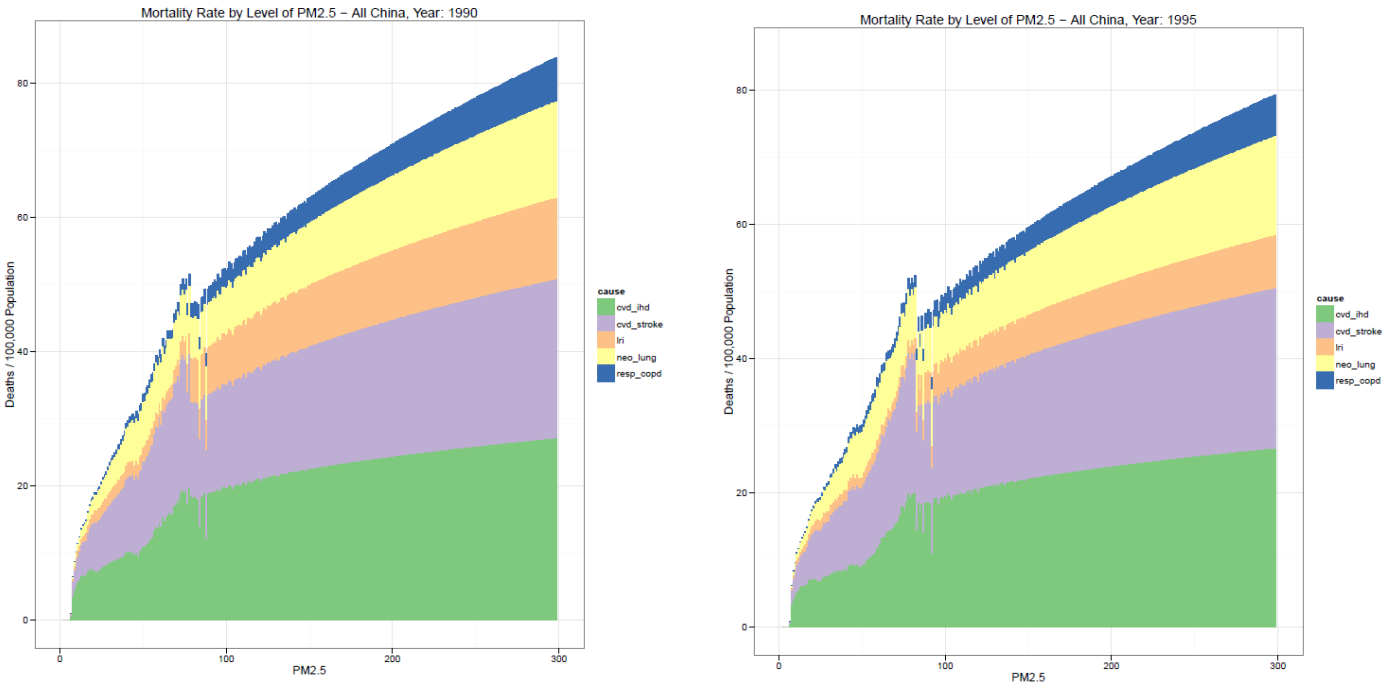


Figure 7.1-2: Average mortality rate at a given level of PM2.5 (rounded to the nearest integer to create even bins), from 1990-1995

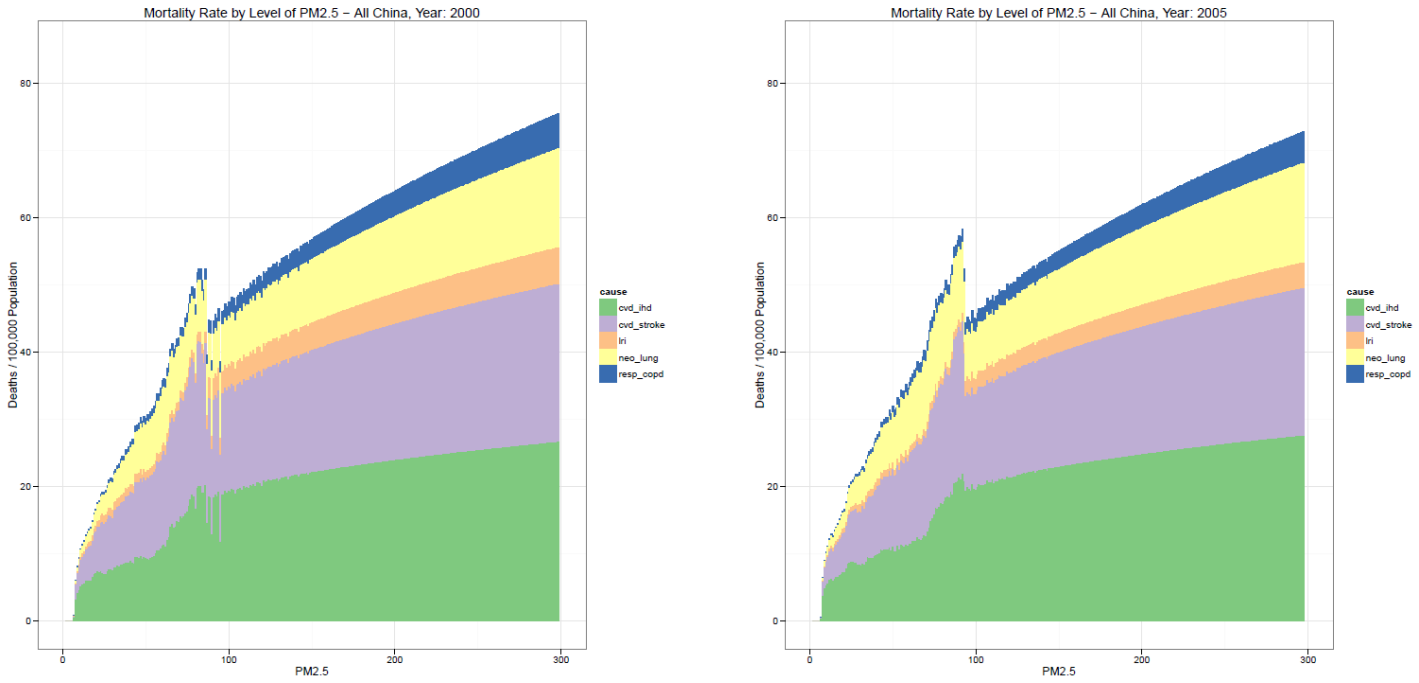
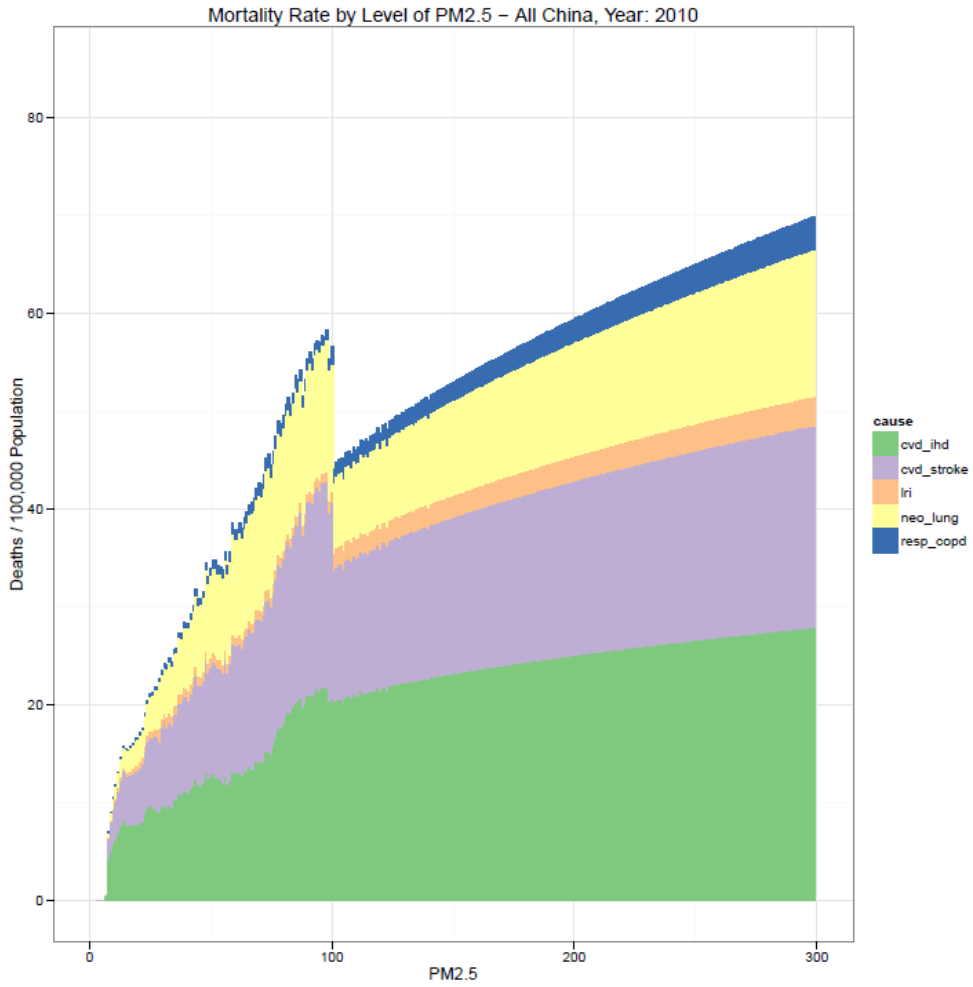
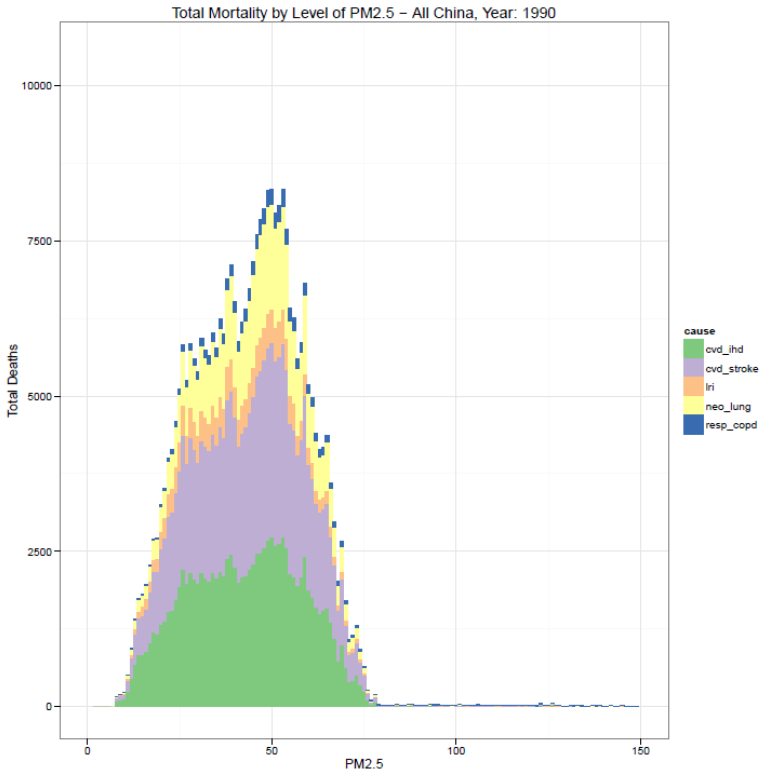
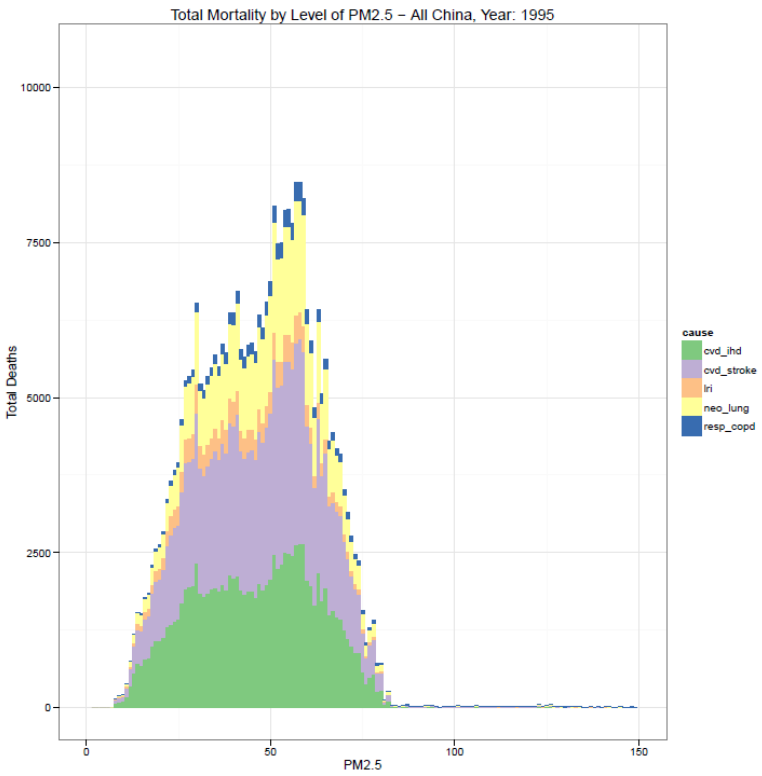


Figure 7.1-2: Average mortality rate at a given level of PM2.5 (rounded to the nearest integer to create even bins), from 2000-2010

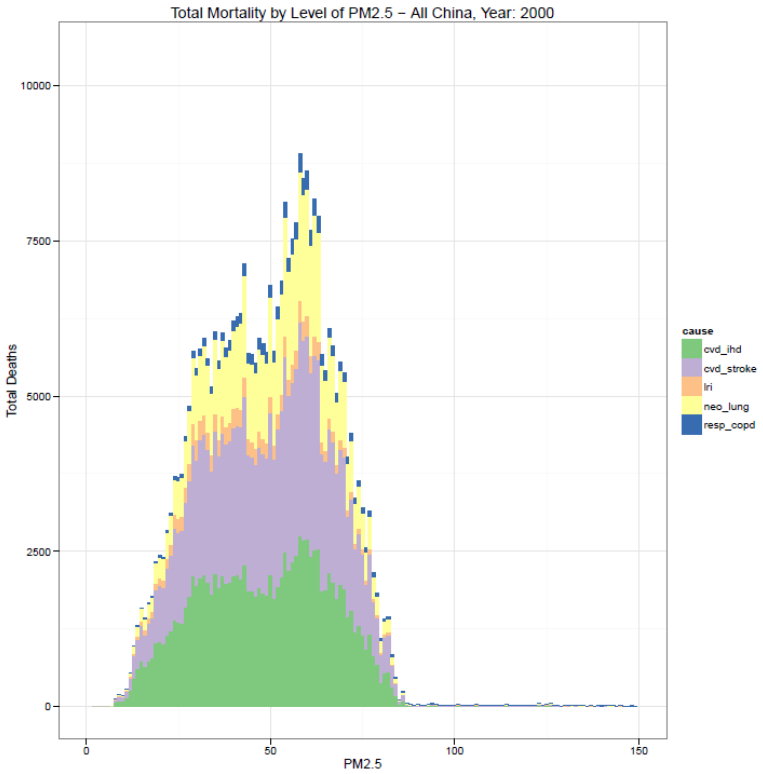


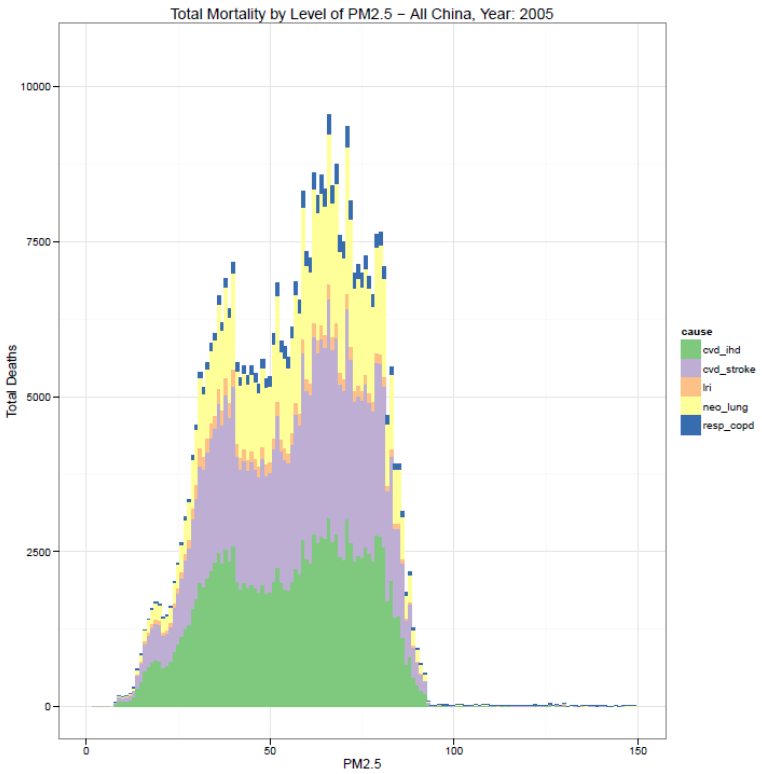
Likewise, notable trends are seen when conducting the same analysis but using the sum of total

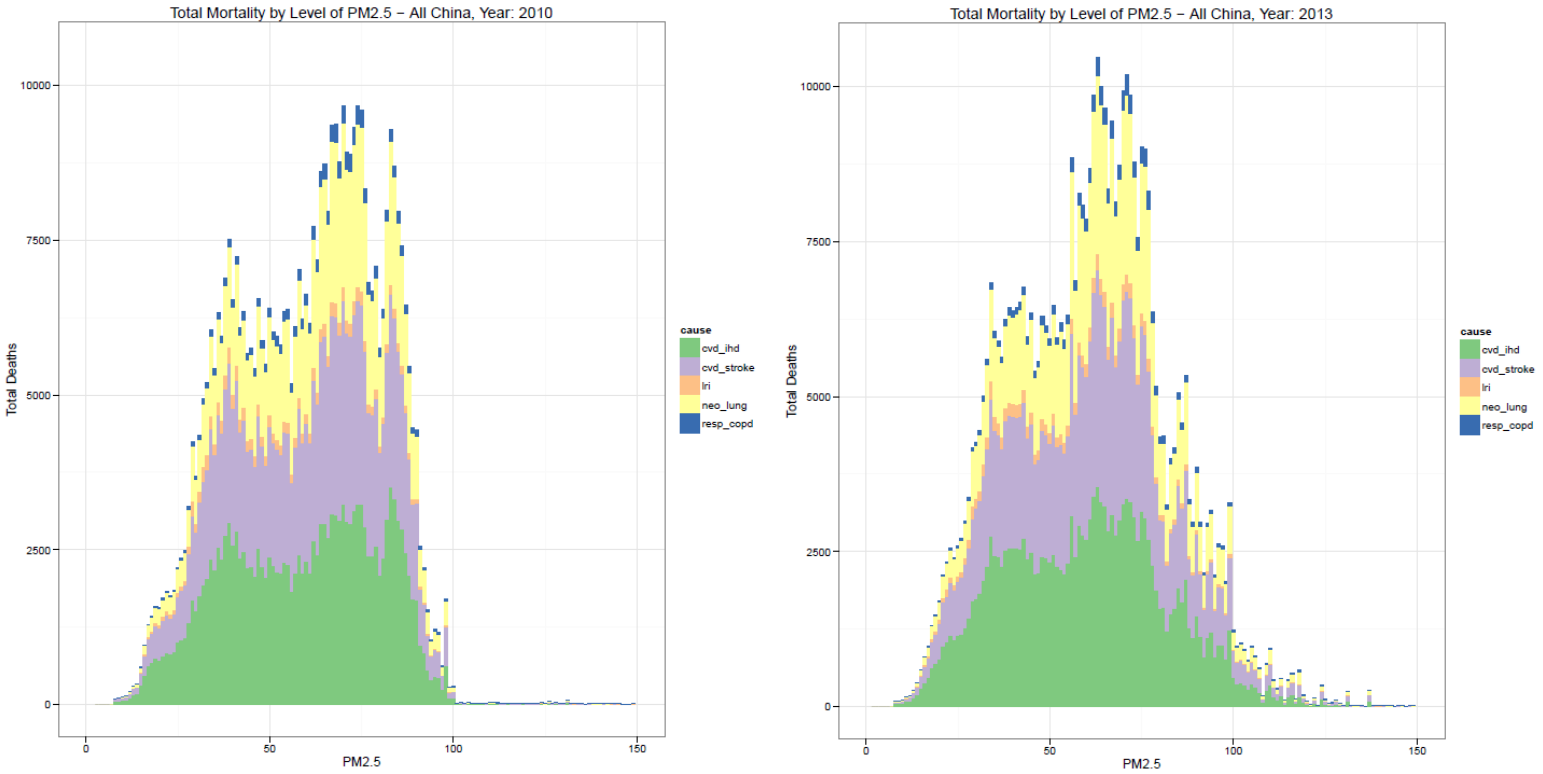




mortality at each PM2.5 level. Figures 8.1-5 show how the total mortality has grown and







continued to shift towards higher and higher levels of PM2.5, following the trajectory of population. These graphs reflect again that LRI and COPD have become non-factors in terms of total mortality and that cardiovascular health is driving the majority of burden. They also show the interesting conclusion that the majority of PM2.5 health impact is normally distributed and actually occurs at moderate levels of PM2.5.

Discussion

The increased understanding of spatial distribution and cause composition of the mortality attributable to air pollution in China provided by this work presents some very important implications. The ability to generate graphical displays of mortality, both by rate and the total, demonstrates narratives about how the impact of air pollution changes across time and space that could be very useful to policy-makers. The shift in the distribution of mortality

rates across the GBD time series shows that the rate is largely mirroring changes in exposure across the time period. If interventions are required to be provided in a manner that is derived from a philosophy grounded in equity, taking the strongest actions in places with high levels of PM2.5 exposure –and therefore high rates – might be the best course of action. However if efficiency and a strong cost-benefit ratio are paramount to the intervention strategy, looking at the distribution and change in total mortality across pixels might be more relevant. The narrative provided by such a display indicates that over the GBD time series rapid urbanization and the related increases to both population and PM2.5 exposure in major cities across Eastern China is the most important contributor to total burden. This trend would suggest that the most efficient way to intervene on this issue would be decisive actions in population clusters with high air pollution, impacting the largest amount of citizens at once.

However, the functional form and inherent shape of the integrated exposure response curve provides another important factor that must be considered when formulating an effective response to the health burden of air pollution. Given that the curves have a non-linear concave shape, the biggest improvements to population health will be made from shifting burden from moderate to low levels of PM2.5. This has some potentially dire implications when considering that the highest rates are found in places with high levels of air pollution. Intervening on these areas or the hotspot cities that host pixels with the highest total mortality is unlikely to be effective unless major improvements are being made. It is simply insufficient to decrease air pollution in a pixel from 100 ug/m^3 to 80 ug/m^3. While this is a sizable decrease, it would translate into only a minor change in the relative risk for the causes of interest for air pollution. This could suggest that many of the major cities in China are too far gone in terms of

low-cost interventions and there the focus should be on cities that are on the edge or only moderately exposed to PM2.5. In order to make any significant impact on the total burden from air pollution in China, major improvements would be necessary. A comprehensive strategy would necessitate a focus on decreasing pollution country-wide, with highly-intensive efforts to make massive decreases in ambient particulate matter in the major cities in East China. This would probably require sweeping changes to economic and environmental policy, as well as individual behavior. It is clear that this research indicates a sizable and growing issue that will be challenging to act on, demonstrating that immediate and decisive intervention is required.

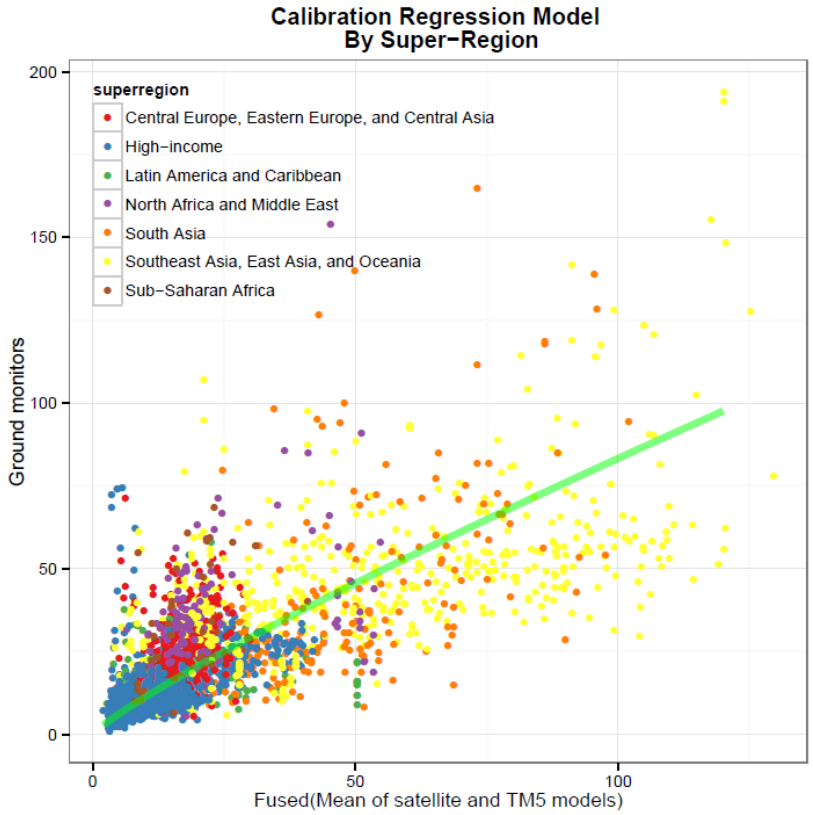
The exploration of trends in cause composition across the GBD time series also provides a pretty interesting story about urbanization and development in China. It can be observed that the total share of causes like LRI and COPD has decreased markedly, no longer making up a significant portion of the total mortality attributable to air pollution. Lung cancer and cardiovascular conditions like IHD and stroke have filled the vacancy and become the key drivers of health burden from PM2.5. This is likely driven by the rapid urbanization and industrialization of the country across the past two decades. Urbanization and increased wealth in China have allowed for better access to healthcare, presenting the opportunity for China to decrease the impact of communicable diseases and undergo the epidemiological transition seen in developed countries. At the same time, however, the Western lifestyle has its drawbacks. Urbanization and income improvements are also linked to behavioral risk factors like exercise, diet, and smoking that are very detrimental to cardiovascular health. Likewise the aging population of China makes non-communicable disease rates a major threat to future health outcomes. Along with the rising exposure to air pollution and its impact on mortality,

these lifestyle risks will be a very important outcome of urbanization that Chinese policy-makers must consider as they continue to shift from a developing country to a global super-power.

There are several key limitations to this work that must be considered when applying these results to public health policy and air pollution interventions. First, the exposure data has some documented flaws that cause both under and overestimation in certain scenarios. PM2.5 can be overestimated in Western China as a result of the Gobi desert dust storms. These storms are very real and do cause large changes in PM2.5 levels, resulting in negative health outcomes. However, there is not a scientific consensus on the extent of health impacts caused by dust or quantification of the differential impact between dust and other sources of PM2.5. This relationship needs to be further explored in order to improve prospective iterations of this analysis.

The gridded data can also underestimate ground level PM2.5, especially in small areas with high levels of pollution. We attempted to mitigate this using the calibration regression, but the assumption of a global relationship between gridded PM2.5 and ground level measurements may not accurate reflect the true situation. For example, GBD2013 air pollution exposure estimates were found to underestimate ground level PM2.5 measured in the nearby capitol of Mongola, Ulaanbaatar. Long considered to be one of the most polluted cities in the world due to severe wintertime and nighttime pollution, Ulaaanbaatar unfortunately falls victim to the fact that seasonal and diurnal variations in pollution were found to be underestimated by the satellite and TM5 models (Van Donkelaar 2015). Satellite measurements of AOD are unavailable during the night and spotty during the wintertime when changes in

cloud cover make readings difficult to attain. By improving the methodology behind calibration and endeavoring to make the addition of further ground monitoring datasets, perhaps these issues can be corrected in the next round of estimation. Another potential limitation lies in the extrapolation of data to 2013 based on the relationship between 2010-2011. This short time series may be unstable and not accurately reflect the true change in PM2.5 in the years since the final measurements. The 2013 estimates must be used carefully with respect to this potential source of error.



The calibration stage, in which ground monitor data is predicted from the gridded dataset by quantifying the relationship between both types of data in shared location is also a source of error. This error can be particularly troublesome for the case of China, as the bulk of datapoints are located in high-income Western countries and they tend to drive the results of the regression. Figure 9.1, where the yellow points show the relationship between ground monitor and gridded data for the super-region containing China, demonstrates this issue. It is likely that for a significant amount of the grids in China, the calibration regression over-estimates PM2.5 compared to the ground monitoring stations.

The addition of random effects was tested in some of the regression models tried for this analysis but did not end up providing sufficient improved fit across all regions. Efforts need to be made to improve the spatial range of the ground monitor database in order to provide a more robust dataset to regress across. In addition, new regression techniques such as spatial regression can be attempted in order to control for the regional variations and perhaps provide a situation where random effects become more useful. The rising importance of air pollution as a risk factor has resulted in policy-makers across the global to push for more monitoring stations, which will increase the richness of the available data for future iterations. This push has been especially notable in China where social and political forces have driven a markedly increased demand for accurate reporting on air pollution conditions. We hope that our work will demonstrate the utility of these estimates and the continued need for better data to help sustain this drive.

Finally, the spatial interpolation of mortality data that is estimated at the provincial level relies on assumptions that may not be fully accurate. In particular, the assumption that underlying mortality rates are constant across grids in a province is unlikely to truly reflect the epidemiological distribution. Rates of cardiovascular mortality in particular have been shown to vary in China across different levels of urbanicity. (Yusuf 2001) Beyond air pollution, a diverse array of behavioral risk factors are also linked to cardiopulmonary diseases. These risks are often dependent on lifestyles and habits that tend to be quite different in urban areas when compared to rural. Increasing urbanization in the country does somewhat mitigate this issue, but without improved spatial granularity of mortality estimates, spatial interpolation will always rely on assumptions. As the GBD moves towards increased resolution at all levels of analysis,

this process will become more and more accurate. Until then, it can provide policy-makers with a more robust understanding of the spatial distribution of air pollution mortality and allow for more informed intervention strategies to mitigate the substantial issues being faced by the growing Chinese population.

Conclusions & Future Direction

There is a significant body of evidence indicating that air pollution is a major contributor to the burden of disease in China. Even as the underlying mortality rates of linked causes like LRI and COPD decline, the rapid urbanization and demographic changes occurring in China mean that the total burden of air pollution continues to increase. This work provides new insight to decision-makers regarding the cause composition and spatial distribution of this burden. It also helps to provide a framework for forecasting analyses that can leverage the use of multiple scenarios to test policies before they are implemented and help ensure an efficient approach to intervention against this pervasive and almost omnipresent silent killer. As the Global Burden of Disease project continues to iterate, two planned areas of increasing focus dovetail nicely with the simultaneous trajectory of air pollution research. The first is the aforementioned forecasting. The design of various scenarios based on planned or theoretical changes in policy has a rich history in environmental science. Models such as the TM5 chemistry transport model used to generate the gridded dataset are very useful for this kind of projection analysis. Likewise, the Global Burden of Disease Forecasting team intends to provide future estimates for all causes, taking into account a variety of important covariates across the GBD landscape. As these estimates continue to improve and account for sources of uncertainty in

projection of relevant variables the utility of air pollution forecasts will increase and provide policy-makers with important understanding about the consequences of legislative action or inaction.

Another important goal for future iterations of the GBD is increased spatial granularity for estimated causes and parameters. Agreements with a growing list of countries to do subnational work at varying levels of enumeration will help to mitigate the key limitation of this work – assumption of constant mortality rate across grids. Waldo Tobler's oft-quoted first law of geography posits that, "Everything is related to everything else, but near things are more related than distant things." As such, improvements to the spatial resolution of mortality estimates will help to make the necessary assumptions of underlying mortality more reasonable and accurate. The incorporation of new gridded information will also help with spatial interpolation. One example is the potential use of gridded urbanicity data to leverage the extensive literature regarding differences in mortality, especially cardiovascular, between urban and rural locations. This relationship can then be used to improve disaggregation techniques. Likewise, general improvements to the gridded datasets must be undertaken in order to upgrade this work, especially with careful consideration to the aforementioned limitations. Key issues that have been raised such as the case of Ulaanbaatar can be used to drive methodological improvements in both the TM5 and satellite models in order to expand their utility.

Future rounds of the GBD will continue this kind of work in order to improve the accuracy and utility of these results. Even considering the limitations of this effort, several important contributions can be made to the understanding of ambient air particulate matter

pollution and how it impacts the health of the Chinese population. The changes in spatial distribution and cause composition across the two decades that the GBD seeks to explore provide interesting conclusions for policy-makers and academics working to mitigate the health impact of air pollution. Unfortunately, one frightening conclusion is that the bulk of air pollution health impact is occurring at moderate levels of air pollution. Given this and considering the non-linear concave shape of the Integrated Exposure Response curve, focusing on high pollution areas is not going to yield a major improvement in population health. Simply getting air pollution levels back to moderate levels is not going to be sufficient to avert the majority of burden. As such, major efforts need to be undertaken to reduce air pollution across the region. Work like this that explores the spatial distribution of mortality attributable to air pollution can help to identify areas that are on the edge of becoming moderately polluted and would be susceptible to cost-effective burden aversion by stopping them before it's too late. Likewise, we can also identify the areas of serious problem that will need heavy lifting in order to avert the loss of life and health due to PM2.5. Continued study of this phenomenon will help to provide critical information to key policy-makers who can help make a difference as China continues to develop into a global super-power and deal with the epidemiological transitions that accompany this change.

References

Augustine, J. A., G. B. Hodges, E. G. Dutton, J. J. Michalsky, and C. R. Cornwall (2008), An aerosol optical depth climatology for NOAA's national surface radiation budget network (SURFRAD), J. Geophys. Res., 113, D11204

Chien, L., Yang, C., & Yu, H. (2012). Estimated Effects of Asian Dust Storms on Spatiotemporal Distributions of Clinic Visits for Respiratory Diseases in Taipei Children (Taiwan). Environ Health Perspect Environmental Health Perspectives, 1215-122

Burnett, R. T., Pope, C. A., Ezzati, M., Olives, C., Lim, S. S., Mehta, S., et al. (2014). An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environmental health perspectives, 4, 397–403

Giles LV, Barn P, Künzli N, et al. From Good Intentions to Proven Interventions: Effectiveness of Actions to Reduce the Health Impacts of Air Pollution. Environmental Health Perspectives. 2011;119(1):29-36. doi:10.1289/ehp.1002246.

Murray, C., et al. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013 The Lancet, Volume 385 , Issue 9963 , 117 - 171

Lim SS, Vos T, Flaxman AD, et al. 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. Lancet. 2012;380(9859):2224-2260. doi:10.1016/S0140-6736(12)61766-8.

"Environmental Concerns on the Rise in China." Pew Research Center, Washington, D.C. (2013) <http://www.pewglobal.org/files/2013/09/Pew-Global-Attitudes-Project-China-Report-FINAL-9-19-132.pdf>

Pope III C, Burnett RT, Thun MJ, et al. Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. JAMA. 2002;287(9):1132-1141. doi:10.1001/jama.287.9.1132.

Pope, C. (2003). Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution: Epidemiological Evidence of General Pathophysiological Pathways of Disease. Circulation, 71-77.

Gridded Population of the World (GPW), v3 | SEDAC http://sedac.ciesin.columbia.edu/data/collection/gpw-v3 (accessed Jul 22, 2015).

Turner, M., Krewski, D., Pope, C., Chen, Y., Gapstur, S., & Thun, M. (2011). Long-term Ambient Fine Particulate Matter Air Pollution and Lung Cancer in a Large Cohort of Never-Smokers. American Journal of Respiratory and Critical Care Medicine, 1374-1381.

Singh, G. K.; Siahpush, M. Widening rural–uban disparities in life expectancy, U.S., 1969–2009. American Journal of Preventive Medicine 2014, 46, e19-e29

Stan Development Team. 2014. RStan: the R interface to Stan, Version 2.5. http://mc-stan.org/rstan.html

Van Donkelaar, A.; Martin, R. V; Brauer, M.; Boys, B. L. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environ. Health Perspect.* **2015**, *123* (2), 135–143.

Winker, D. M.; Pelon, J. R.; McCormick, M. P. The CALIPSO mission: spaceborne lidar for observation of aerosols and clouds. In Third International Asia-Pacific Environmental Remote Sensing Remote Sensing of the Atmosphere, Ocean, Environment, and Space; Singh, U. N., Itabe, T., Liu, Z., Eds.; International Society for Optics and Photonics, 2003; pp 1–11

Yusuf, S., Reddy, S., Ounpuu, S., & Anand, S. (2001). Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. Circulation, 23, 2855–2864.

Zhang, X., Arimoto, R., & An, Z. (1997). Dust emission from Chinese desert sources linked to variations in atmospheric circulation. J. Geophys. Res. Journal of Geophysical Research, 28041.