International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Spatial Analysis of The World Bank's Global Urban Air Pollution Dataset

Doll, C.N.H.

IIASA Interim Report
August 2009

International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-09-033**

Spatial analysis of the World Bank's global urban air pollution dataset

Christopher N.H. Doll (doll@iiasa.ac.at)

Approved by

Arnulf Grubler
Transition to New Technologies (TNT) Program

August 31, 2009

International Institute for Applied Systems Analysis
Registration number: ZVR 524808900

Contents

Abstract

This paper describes the method used to spatially render the World Bank's air pollution database of modelled PM_{10} concentrations for some 3,200 locations across the world for cities over 100,000 people and capital cities. The dataset has very good spatial coverage with each of the 11 GGI regions well populated by data points. Mapping these point concentrations with respect to population density reveals that most densely populated areas of the world are well accounted for with the exception of sub-Saharan Africa. In total, 1.96 billion people are accounted for, which is 2/3 of the total global urban population. South Asia and the Middle East/North Africa have the highest average concentration. Concentrations and exposures are then mapped according to World Health Organization guidelines. It is found that much of the urban populations in the world's most populous countries have concentrations that lie outside of even the lowest air quality targets. Finally, concentrations are compared to other spatially explicit datasets of pollutants to build a picture of spatial air pollution patterns and help to understand the characteristics of the model.

Acknowledgments

The author gratefully acknowledges the guidance of Arnulf Grubler in preparing this report and discussion and review comments of Zig Klimont and his colleagues.

About the Author

Christopher Doll produced this report as part of his research activities while at IIASA as an institute post-doctoral research scholar 2007-2009. He holds bachelor's, master's and Ph.D. degrees from the University of London (UK) in geography & mathematics and remote sensing. His primary research interests involve the use of night-time satellite data and other spatially explicit datasets to understand relationships between urbanization, environment and development.

Spatial analysis of the World Bank's global urban air pollution dataset

Christopher N.H. Doll

1 Introduction

Particulate matter less than 10μm in diameter (PM_{10}) is one of the standard metrics by which air quality is measured. It is of importance in both environmental and public health arenas due to its origin and its likelihood to be inhaled by humans into the respiratory tract. The World Bank's global air pollution dataset (World Bank, 2008) provides figures for the average annual PM_{10} concentration ($\mu g/m^3$) concentration based on a multiple regression model. The dataset contains values for 3,226 cities over 100,000 inhabitants along with capital cities in 180 countries along with the corresponding population for the year 2000.

Particulate matter can be distinguished between a number of sources and domains such as: indoor and outdoor, natural and anthropogenic, urban and rural. In urban areas, the primary sources arise from the combustion of fuels, construction activities, road dust re-suspension and wind (WHO, 2005) but can also be affected by proximity to deserts as well as longer range transport from other natural sources. Taken together, concentrations can vary widely according to a number of factors such as geographic situation, prevailing weather conditions and the time of day. Therefore it is important to make clear what is being considered here is the average ambient annual concentration for a given location (city).

This paper details the process used to produce this enhanced dataset along with the issues encountered its analysis. A preliminary analysis is performed to describe the spatial distribution of concentrations both in absolute terms and in relation to the underlying population along with its shortcomings and omissions original form the list. Comparisons are made with selected cities from other published studies as a limited form of validation before proceeding to consider PM_{10} concentrations in relation to the World Health Organization Air Quality Guidelines. The analysis then considers other spatially explicit emissions datasets to elucidate the broad picture of PM_{10} in relation to other pollutants. Articulating these spatial distributions provides a means of evaluating the dataset and identifies areas of interest, which may be verified to be due to the limitations of the model or point indeed to a policy implication. In doing so, this report aims to provide assistance to IIASA spatial modelling activities by helping to understand the assumptions one can draw from it, facilitating assessment of its overall usefulness for use in scenario work concerning integrating air pollution measures with other connected global change issues such as climate change and human health.

2 Model Specification

PM_{10} concentrations are derived from Bank's Global Model of Ambient Particulates (GMAPS) model. Measurements in of particulate matter (PM) from 304 cities (Cohen *et al.*, 2005), were used to calibrate the model; this represents just under 10% of the final dataset. Note that the term particulate matter is used here as some of these locations collected measurements on the total suspended particles (TSP) and not necessarily the PM_{10} fraction. Panel data from 1985-1999 was collected from residential and non-residential sites. Most monitoring locations were located in the US and Europe, with other locations including Central America, India and China and a handful in South America and Africa. Locations were divided into PM_{10} and TSP and subdivided into those which had historical data as well as current. China for example only had TSP data, India only had historical PM_{10} records.

Preliminary analysis found a systematic variation in size composition of PM based on income levels although TSP and PM_{10} were not perfectly correlated. There was also found to be significant variation year to year within cities, which was comparable to between country variations. A model was designed to accommodate the 15 years of panel data since 1985 using the observed data by monitoring site for each PM type as the unit of observation which would account for location differences and incorporate TSP information from poorer countries which have no PM_{10} data.

Determinants of ambient PM concentration were categorized into anthropogenic factors and geo-climatic factors. In all 48 parameters were used. These were subdivided into categories of emissions/economic activity (city population, per-capita GDP, national per-capita fuel consumption by class including transport fuels), abatement policies and technology/knowledge (income based) for the anthropogenic factors. In additional to situational parameters (elevation, distance to coast), 22 local climatic factors (means and frequencies of wind speed, precipitation, temperature, frosty and cloudy days) were used.

These were used in a two stage estimation of PM_{10} & TSP consisting of within country marginal effects over time and space and average country effects. This gives the model the trade-off that the precision in estimating the contribution of specific factors that are applicable to limited set of places is sacrificed for plausibility of predictions in the broader set of geographic locations. The secondary cross-sectional model of TSP was used to generate estimates of PM_{10}/TSP ratios to predict PM_{10}. The resultant output from the model accounted for 88% of the observed variation. Although the aggregated nature of the determinant data and the quality of initial monitoring data means there is significant uncertainty in the estimates (Pandey *et al.*, 2002).

3 Geolocation of Cities

The geolocation of cities was done using a variety of data sources and in such a way so as to iteratively identify and correct errors. The main data source was a database of 50,000 locations available from the Center for International Earth Science Information Network (CIESIN *et al.*, 2004a). After accounting for naming conventions (capitalization, removal of accents and other diacritical notation not used in the CIESIN dataset), some 2,400 could be located using this database by running a simultaneous matching of countries and then city names. The remaining cities were

geolocated using the online database from the World Gazetteer (World Gazetteer, 2008), the US Census Bureau's place name Gazeteer (U.S. Census Bureau, 2008) and corroborated with Google Earth. Ancillary information from Wikipedia was also used to help identify alternative names, locations and population. Twenty four cities could not be reliably located and were therefore excluded from the dataset. The dataset was then overlaid night-time lights (NOAA-NGDC, 2008) and inspected for coincidence of points within lit areas, the rationale being that all cities over 100,000 inhabitants should be co-located with visible light emissions. A further five cities in China were excluded from the final dataset because of their mismatch with night-time lights. These are listed in Appendix A.

To facilitate comparisons with other datasets, the UN-city codes for cities with a population greater than 750,000 has also been added for 564 cities around the world.

One issue discovered was that the CIESIN database is very detailed in China and many cities were not present because suburbs were listed rather than urban agglomerations. This meant that Chinese cities had to be assessed manually for this to ensure that the names were being matched to locations of equivalent populations. It should be noted that in the World Bank database cities in Taiwan are listed in China.

4 Spatial Distribution of Locations

The distribution of points in the dataset summarized by IIASA's GGI reporting regions is given in Table 1. Alongside is the corresponding total population of locations within each region along with the population weighted average concentration and its minimum and maximum concentration. The coverage of locations is well distributed over the 11 regions, even in the relatively lowly populated region of the Eastern Europe (EEU). There is also a large variation in concentrations varying by more than a factor of four between the highest and the lowest regions.

Table 1. Distribution of the locations of air pollution measurements summarized by GGI region with the total urban population covered and the population weighted average of concentration.

IIASA(11) Region	# locations	Population (millions)	Weighed PM_{10} Concentration	Minimum	Maximum
MEA	199	119.95	106.8	23.2	358.9
SAS	402	237.85	105.0	12.6	239.1
CPA	390	430.85	86.5	20.0	166.8
AFR	183	124.04	69.0	10	282.9
PAS	182	128.82	66.4	17.4	159.1
LAM	458	245.67	45.5	6.7	173.2
EEU	124	35.02	36.3	14.9	100.6
FSU	299	118.81	33.0	5.9	110.6
WEU	455	191.90	31.6	9.0	66.5
PAO	240	96.31	30.6	11.2	61.2
NAM	265	228.24	24.7	10.4	47.6
Total	**3,197**	**1,957.44**	**62.9**	**5.9**	**358.9**

MEA – Middle East-North Africa; SAS – South Asia; CPA – Centrally Planned Asia; AFR – Sub-Saharan Africa; PAS – Pacific Asia; LAM - Latin America; EEU – Eastern Europe; FSU – Former Soviet Union; WEU – Western Europe (incl. Turkey); PAO – Pacific OECD; NAM – North America.

These locations are plotted on a map of population density (Grubler *et al.*, 2007) in Figures 1 and 2, split into the Americas and the rest of the World respectively. Locations appear to coincide well with the areas of high population density in most regions of the world; China and India in particular are well covered by data points. The dataset covers 1.96 billion people, which is just over 2/3 of the total urban population (2.84bn in 2000; (UN, 2006)). Due to the explicitly urban nature of the dataset, populations in Africa are selectively accounted for given the general lower level of urbanization across the continent. Some populous countries may only have a few data points. This can be seen in Figure 2, where many areas of high population density have no coverage by the dataset.

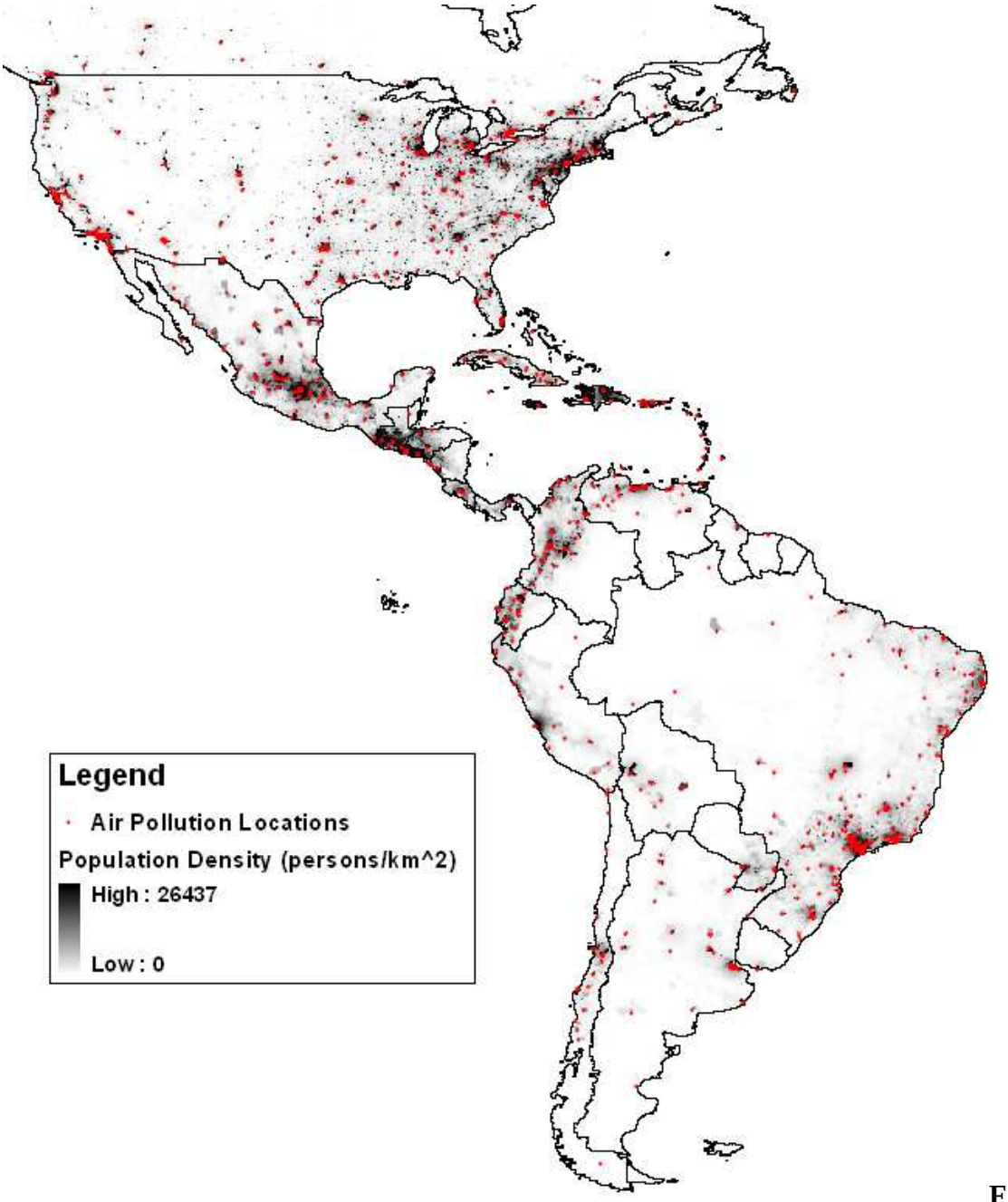

Figure 1. Locations of PM_{10} concentrations in the Americas overlaid on the IIASA population density map (A2r, 2000) with national boundaries (CIESIN *et al.*, 2004b).

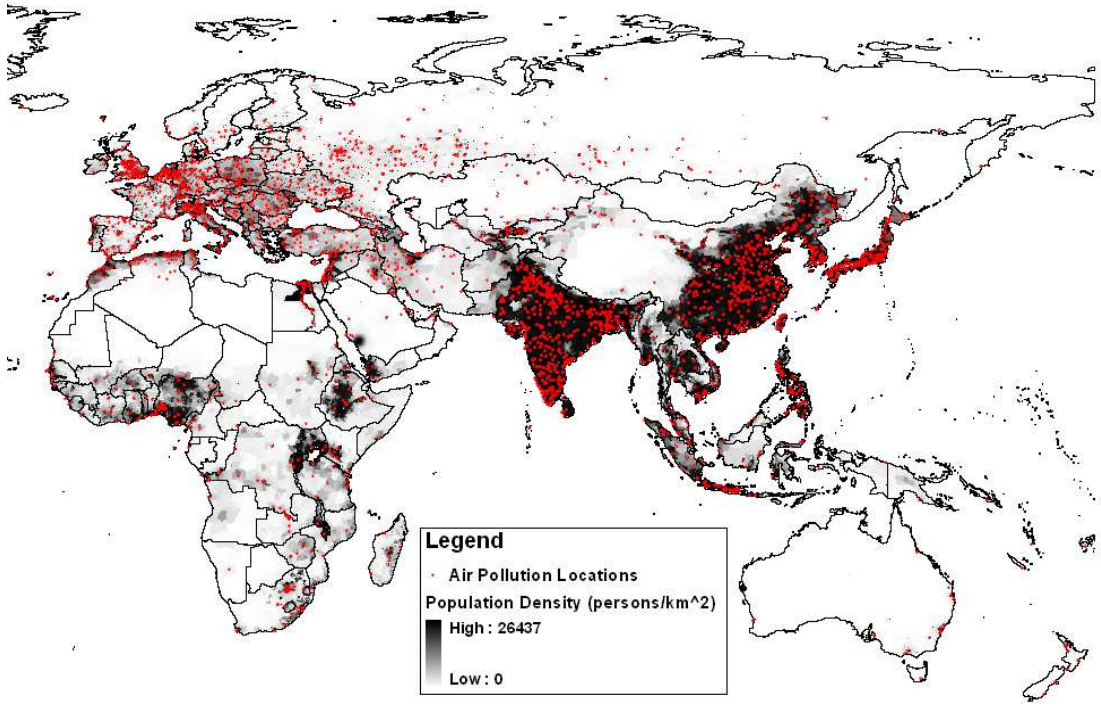

Figure 2. Locations of PM_{10} concentrations in Europe, Africa and Asia overlaid on the IIASA population density map (A2r, 2000) with national boundaries (CIESIN *et al.*, 2004b).

5 PM_{10} Concentrations around the World

The global distribution of concentrations is shown in Figure 4. Classes are displayed as deciles (an equal number of members in each class but classes do not have a fixed concentration range). The decile of lowest concentration values covers 12μg/m^3 and 260μg/m^3 in the highest with cities in Sudan and Nigeria occupying 9 of the top 10 most polluted cities in this dataset. Considered at the country level, Sudan has the highest PM_{10} concentrations with Mali second and Pakistan third. However Nigeria is 17th in the country ranking due to these highly polluted cities making up only a small proportion of the Nigerian sample. Pakistan, by contrast has many polluted cities and an urban population 7 times that of Sudan. Generally, concentrations cluster spatially into easily identifiable regions. Loci of highest concentrations are observed in Northern India (Southern India has appreciably lower concentrations), in China (especially around Northeast, throughout the Middle East and West Africa (Ghana being a notable exception). Bolivia is the other notable country outside of this set to have very high concentrations.

Although the reported range runs from 5 -360μg/m^3, as seen in the histogram of values in Figure 3, most values are in the lower part of this range; indeed 90% of the values are less than 100μg/m^3. Table 2 reports the population covered in each decile. These increase with concentration such that 700 million people are covered in the 5 lowest deciles compared to 1.26 billion in the 5 highest deciles.

Table 2. Range of concentrations covered in each decile with the associated population in each class.

Decile	Range	Population (millions)	% Population
1	5.9 - 18.3	123.64	6.3
2	18.3 - 21.9	114.95	5.9
3	21.9 - 25.5	174.02	8.9
4	25.5 - 30.3	148.95	7.6
5	30.3 - 36.9	138.78	7.1
6	36.9 - 45.5	189.76	9.7
7	45.5 - 60.5	210.51	10.7
8	60.5 - 78.8	262.52	13.4
9	78.8 - 100.6	279.99	14.3
10	100.6 - 359.9	318.72	16.3

With respect to economy, Figure 3 shows the relationship between mean weighted concentration and log per-capita GDP (World Bank, 2007). Generally, we observe that PM_{10} concentrations decline with increasing wealth. However there are notable outliers and the distribution does not exactly fit the hypothesized inverted-U relationship environmental Kuznets curve for an environmental pollutant.

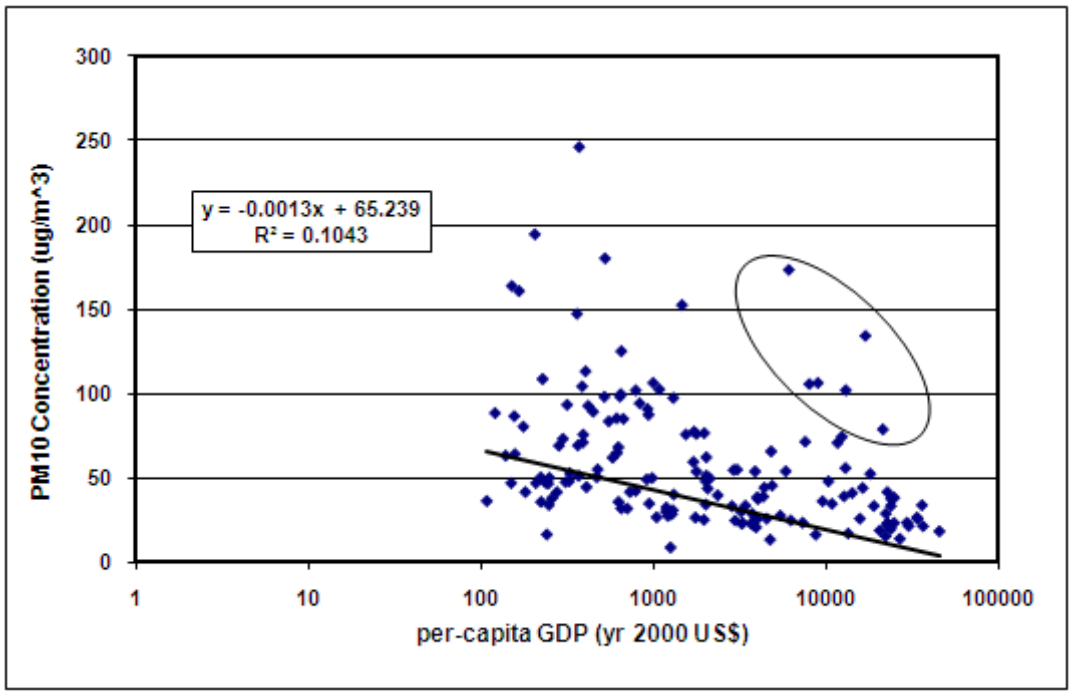

Figure 3. Relationship between PM10 concentrations and Log per-capita GDP (year 2000 US$) for 161 countries. Gulf states and Uruguay highlighted as small populations with high PM_{10} concentrations.

Acknowledging that we are not exactly comparing like with like (the World Bank dataset explicitly considers cities, and comparing against total per-capita GDP, rather than the urban component which will be higher) we may nevertheless make a number of inferences from the figure. It might be expected that urban environments in the

poorest countries are dirtier but the figure shows that income is no barrier to having low concentrations. However there are also countries throughout the income range with high PM_{10} concentrations. The group of countries encircled in Figure 3 has relatively high concentrations for GDP/capita and include Saudi Arabia and other small Gulf States. These outliers suggest that physical geography trumps economy as GDP is ineffective against the natural sources of PM_{10} from deserts, to which many of these countries find themselves in close proximity. Uruguay, whose capital Montevideo is the sole location for that country is also included in this group.

The persistence of high PM_{10} concentrations over GDP range also suggests that source substitution may be contributing to this effect, whereby economic progress eliminates one source only to be substituted for another. Transitions from public to increased private transportation being one example. Indeed the decline of concentrations maybe more indicative of policy interventions to limit emissions and drive incremental technological change rather than radical shifts to completely new technologies.

The lowest concentrations are found in Belarus, followed by the UK, France and Scandinavia for Europe. Colombia and the Northern parts of Brazil are also low along with selected cities in North America. Between these end points neighboring areas have concentrations in adjacent classes. There are however a number of locations were very low concentrations can be found compared to the surrounding area.

Valparai in Tamil Nadu is the focus for low emissions in South India. Srinagar (Northern India), Mendoza (Argentina), Kampala (Uganda) and Mbandaka (Democratic Republic of Congo) all have concentrations which are anomalously low compared to neighboring locations. Three of these cities share similar elevations (~1000m), a parameter which had a high significance in the model and this feature of the model combined with other similarities could conspire to produce this result.

Spatially comprehensive databases of air pollution are rare (hence the creation of this database) so validation of PM_{10} concentrations can be tricky not least because studies do not always give an annual average concentration. PM_{10} emissions can be highly variable, especially in regions where is a large natural source contribution and variable meteorological effects. A study of ambient air over Beijing under different conditions on April 2000 give values for three sites in the city of 66-128 $\mu g/m^3$ for a non-dust storm day; 259-317 $\mu g/m^3$ for a haze pollution day and 667-849 $\mu g/m^3$ for a dust storm day (Xie *et al.*, 2005). The annual average in World Bank dataset is 106 $\mu g/m^3$. Other studies also distinguish between wet and dry season. Kim Oanh *et al.*, (2006) publish average concentrations for 6 Asian cities in both the wet and dry season. Comparing these ranges to the annual average in the dataset reveals a large discrepancy for three of the six cities. Beijing and Chennai (Madras in the database) in particular have low values (Figure 5).

Chan and Yao (2007) tracked a decline in annual average PM_{10} concentration over 1999-2005 of 180-142 $\mu g/m^3$ but this is still considerably higher than the 106 $\mu g/m^3$ figure in the World Bank dataset. A 2003-2005 average of 95$\mu g/m^3$ compares more favorably to the dataset value of 87 $\mu g/m^3$ for Shanghai. Larsen *et al.* (2008) assembled average values for 20 cities around the globe for a subset of years ranging from 1997-2005. There is an overall agreement within +/- 15% with some major discrepancies most notably Bangkok, Bogota and Kolkata (Figure 6).

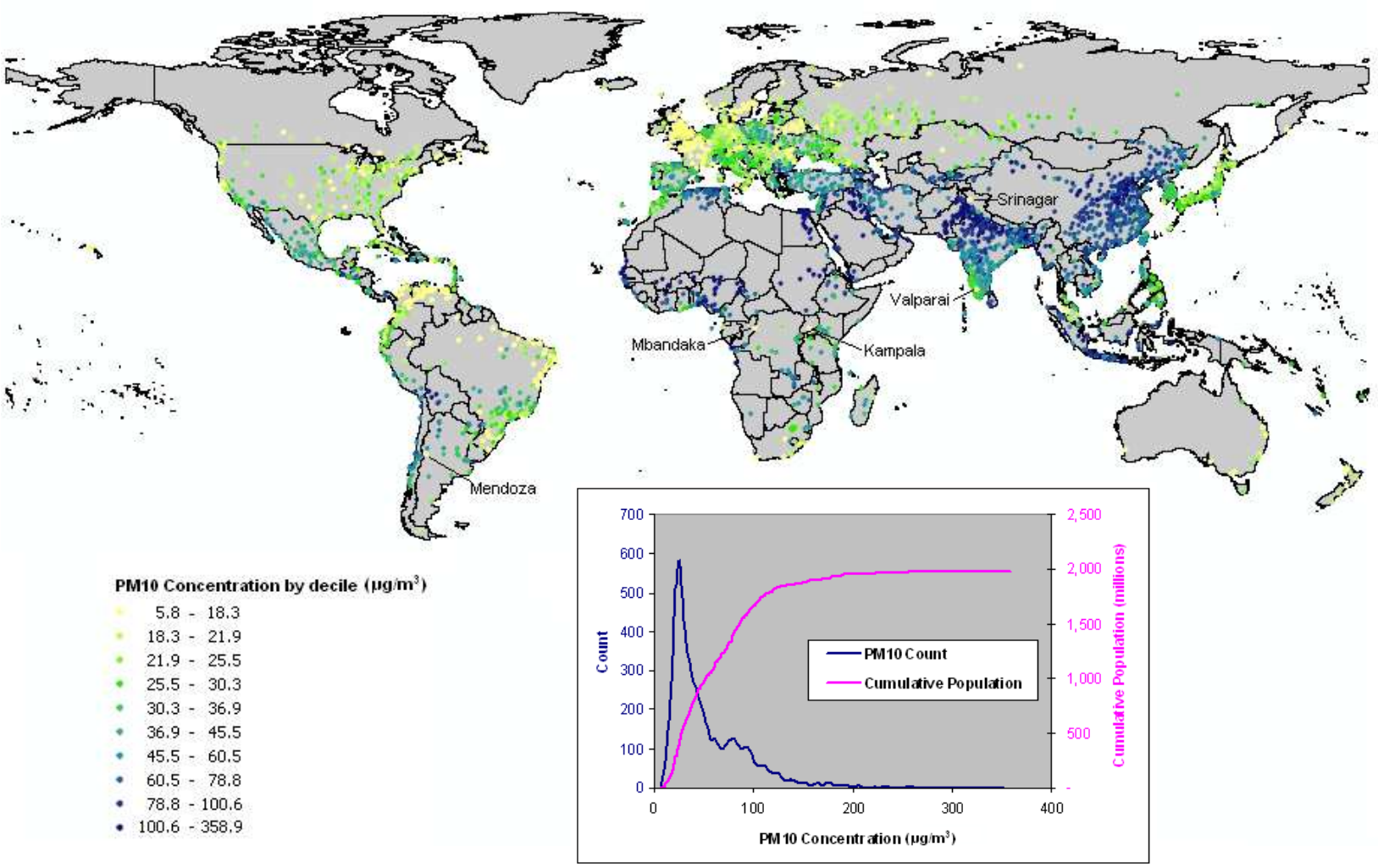

Figure 4. Locations of PM_{10} concentration classed by decile with the associated concentration and population distributions.

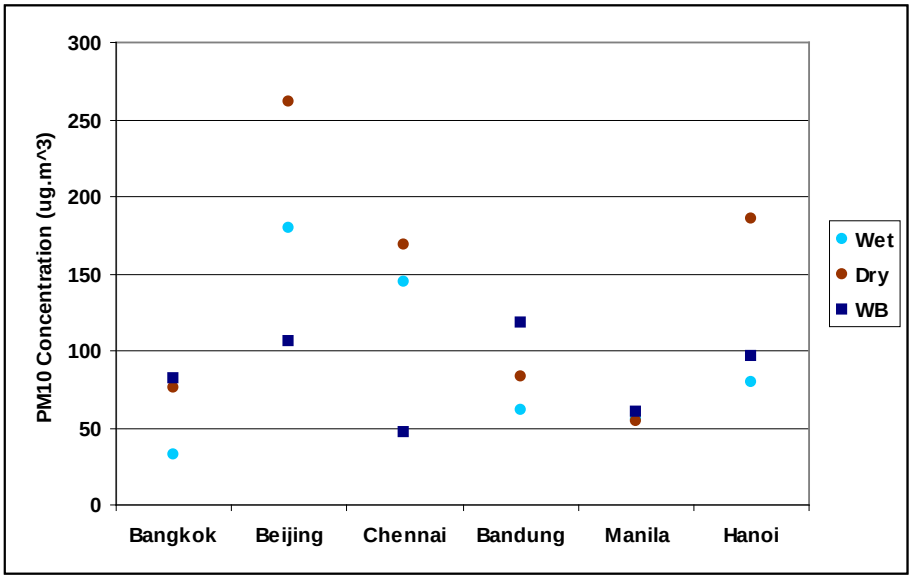

Figure 5. World Bank (WB) values relative to observations taken in the wet and dry season by Kim Oanh *et al.*, (2006) for six Asian cities.

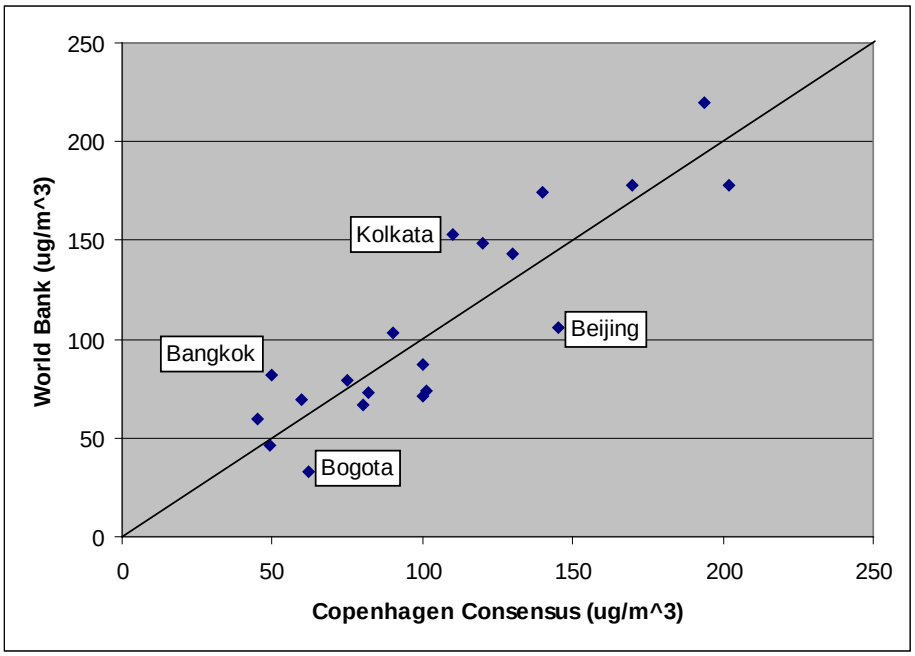

Figure 6. Level of agreement in annual average PM_{10} concentrations between World Bank dataset and the studies synthesized in the Copenhagen Consensus report (Larsen *et al.*, 2008) for 20 major cities around the world.

6 PM_{10} Exposures and WHO Targets

The World Health Organization (WHO) issues air quality guidelines (AQG) based concentration levels correlated to health risks. The AQG are the lowest levels which cardiopulmonary and lung cancer mortality has been shown to increase with long term exposure to $PM_{2.5}$ (particulate matter <= 2.5μm). Although PM_{10} concentrations are the more widely measured the health guidelines are based on studies of exposure to

$PM_{2.5}$. A $PM_{2.5}$/ PM_{10} ratio of 0.5 is used to calculate the equivalent PM_{10} concentration to match the health guideline to the observations. This ratio is deemed typical of developing country urban areas and is at the lower end of the range found in developed country urban areas (0.5-0.8; WHO, 2005). The concentrations for the AQG and interim levels are outlined in Table 3 and their spatial distribution is plotted in Figure 6.

Table 3. WHO Guidelines and rationale for annual mean concentrations for particulate matter. (WHO, 2005).

	PM_{10} ($\mu g/m^3$)	**$PM_{2.5}$ ($\mu g/m^3$)**	**Basis for Selection**
Interim Target 1 (IT-1)	70	35	15% long-term higher mortality risk than AQG
Interim Target 2 (IT-2)	50	25	~ 6% lower mortality risk than IT-1
Interim Target 3 (IT-3)	30	15	~ 6% lower mortality risk than IT-2
AQG	20	10	

The WHO is just one body which sets standards on air quality. It is noted that standards vary across the world not just between different countries but also in how the standard is denoted as an annual mean or the exceeding of a certain number of days/hours above a given level of concentration. The European Commission's standard is currently set a 40 $\mu g/m^3$ to be reduced to 20 $\mu g/m^3$ by 2010 (European Commission, 1999). There is as yet at the time of writing no standard for $PM_{2.5}$. By contrast the US environmental protection agency has no average annual target for PM_{10} instead using 150 $\mu g/m^3$ over a 24-hour period although it does mandate 15 $\mu g/m^3$ annual and 35 $\mu g/m^3$ over a 24-hour period for $PM_{2.5}$ (EPA, 2009). Given these regional and reporting differences, the WHO AQG is used here as a basis for comparison as it is the pre-eminent international body which publishes annual average PM_{10} guideline concentrations.

The sequence of maps in Figure 7 shows a gradual shift from less developed and rapidly developing countries to developed countries as target concentrations become increasingly more stringent. The corresponding exposed population is given alongside. The highest value in the dataset (359 $\mu g/m^3$) is more than 5 times the level of the first interim target with almost ¾ billion people outside of the most lenient interim target. Only 164 million people or 8.4% of the dataset's population reside in cities which comply with the AQG.

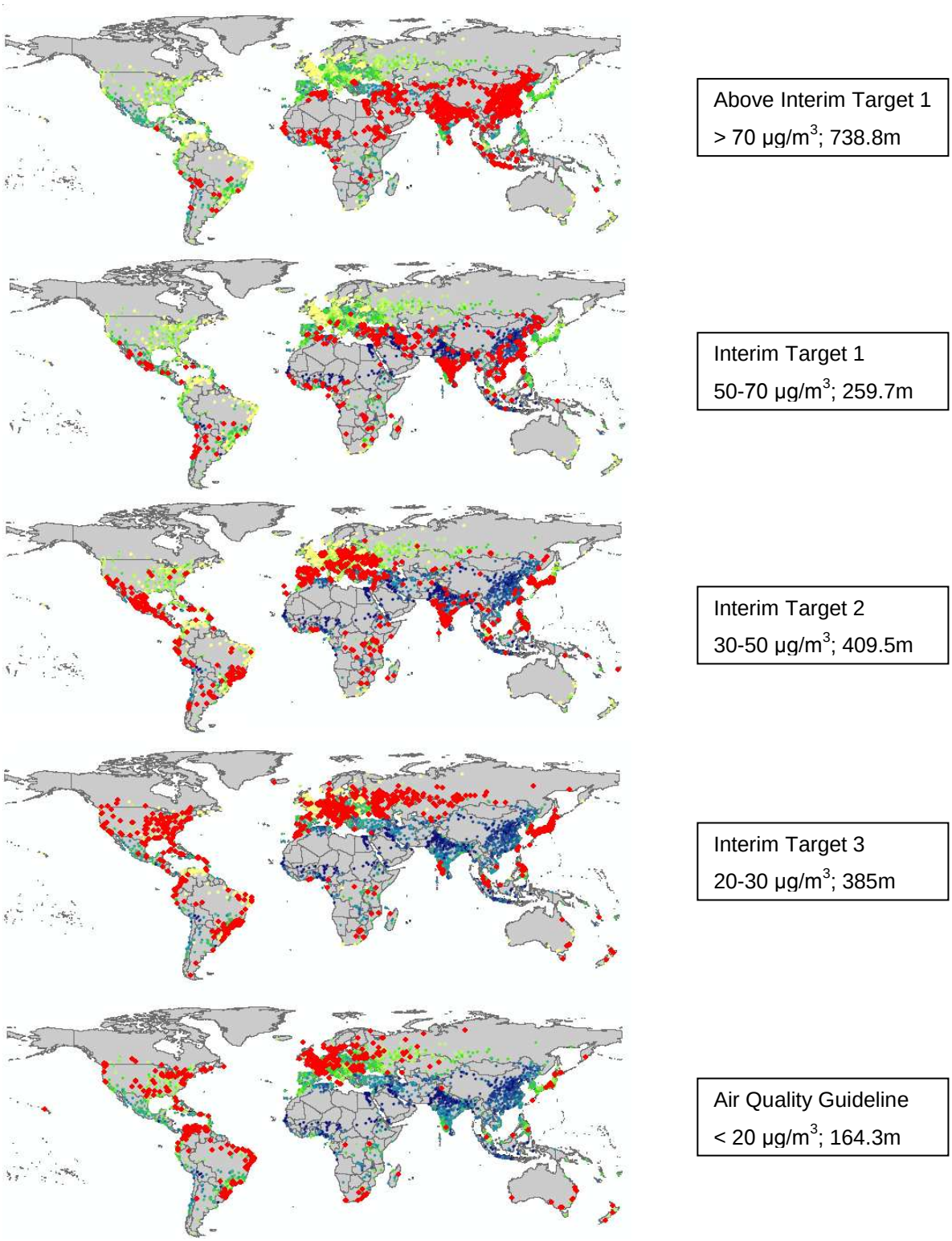

Figure 7. Locations (in red) and total population of cities with respect to the WHO interim guidelines on average annual PM10 concentrations. Background legend same as Figure 3.

An alternative way to analyze these data is to consider exposures, which here is defined as the product of concentration and population. There is little discernable shift in the location of the lowest 10% of exposures from concentration. Figure 8 shows the

spatial distribution of the top 10% of values by absolute concentration is spread throughout Northern India, Pakistan, China and Indonesia and Sahelian Africa along with the Middle East. However, the top 10% of exposures comprises many locations outside of these initial areas. Many more locations in China are included as are locations in main cities in the US, Japan and Europe. This is because such a representation will also include cities of low concentration but high population. The dataset gives figures for Paris of 12 $\mu g/m^3$ with 9.8 million inhabitants. Comparison to data from the European Apheis programme gives the equivalent figures for the year 2000 of 22 $\mu g/m^3$ but a smaller population of 6.2 million people (Medina *et al.*, 2005). Both exposures are would be within the top 10%. In order to qualify for this top 10% group, the population required for any city which meets the WHO AQG guideline of 20 ug/m^3 is 3.9 million people.

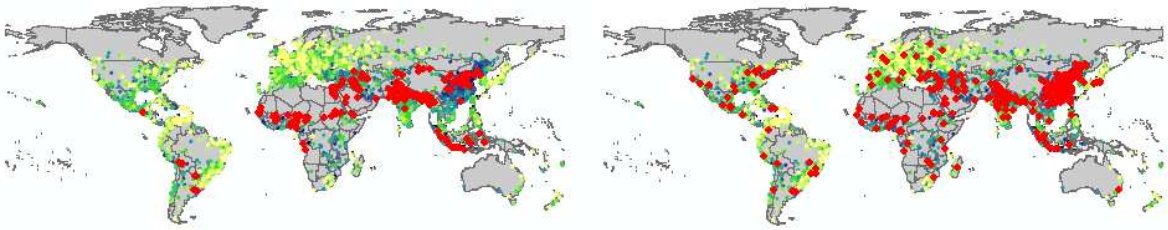

Figure 8. Location of the top 10% PM_{10} concentrations by absolute value on the left-hand panel and exposure concentration weighted by population (product) on the right-hand panel. Red points mark pertinent locations with the background legend same as Figure 3.

An integrated map of concentration and exposure is shown in Figure 9, where exposures are plotted according the underlying PM_{10} concentration. In this representation, large exposure circles of low concentration can be thought to have a large underlying population and vice-versa. We observe that some locations of low concentration actually have large exposures because of the high population associated with that point (e.g. Chicago). Cities will vary in population by more than over 2 orders of magnitude but concentrations less so. As stated in the introduction, it is noted here that data presented here deals with ambient outdoor air pollution and therefore exposure measures are also based on this parameter. As a point of comparison Smith (1993) considers the Global Exposure Equivalent (GEE) for the 8 human micro-environments (indoor and outdoor in urban and rural for developed and developing countries). Rural levels of PM_{10} concentration maybe as high or even higher due to indoor combustion of traditional fuels. Smith (1993) estimated PM_{10} concentrations for the 8 major human micro environments Indoor concentrations in developing countries area estimated to be 551$\mu g/m^3$.

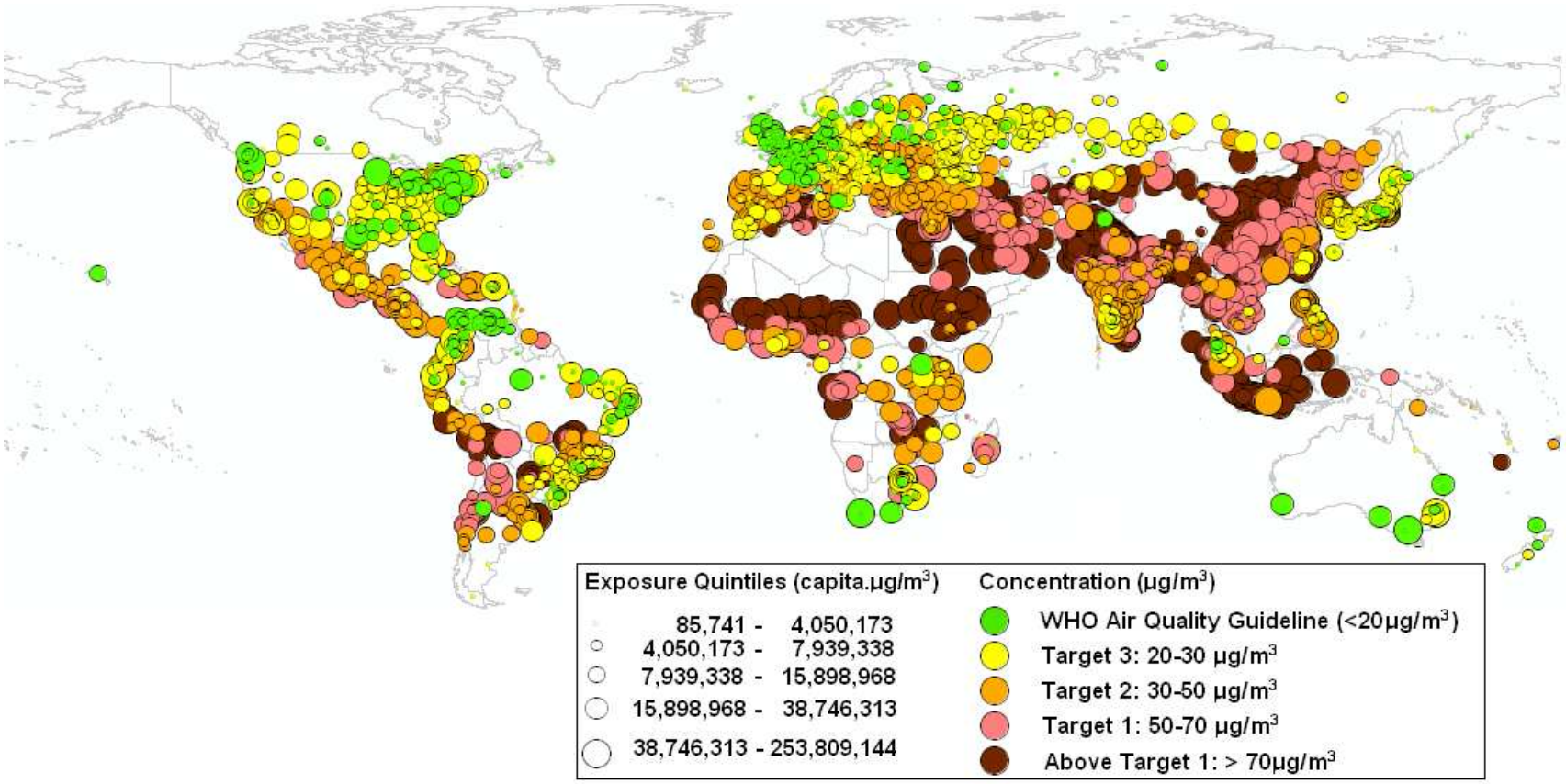

Figure 9. Outdoor air pollution exposure quintiles (size of circle) classed by WHO Guidelines (color) outlined in Table 3.

The GEE is the "equivalent annual concentration to which the entire global population would have to be exposed to the equal the population in that particular micro-environment" (Note from Table 4 in Smith, 1993). This is determined not only by the concentrations and populations of these environments but also by how much time is spent in them (indoors and outdoors). Figures show firstly that concentrations are universally higher indoors than outdoors and secondly that urban outdoor pollution only captures some 7% of the total GEE at that time, 95% of which is in the developing world. While there have been undoubted changes in both urbanization and concentrations since the time of the survey (some 20 years ago), the magnitude of these changes are likely to be lowest where concentrations and exposures are highest (i.e. rural developing areas) and it is therefore nonetheless insightful to give pause to the consideration that much global monitoring is directed in areas which actually contribute little to exposures. In particular rural indoor monitoring seems to be the pose the greatest challenge since the areas where this is highest are also the areas with lowest access to health care resources in terms of both space and affordability.

7 PM_{10} Comparison with EDGAR Emissions Data

PM_{10} concentrations were compared to spatially explicit datasets of two other pollutants connected to anthropogenic activities: Sulphur dioxide (SO_2) and oxides of nitrogen (NO_x). These two datasets of *total annual* SO_2 and NO_x were obtained from the EDGAR Fast track 2000 database (Olivier *et al.*, 2005; NEAA, 2005) and republished as grids at 1ºx1º resolution. To make the datasets comparable, the city the PM_{10} concentrations were averaged onto a 1ºx1º grid and converted to points as were the grids. Emission locations not coincident with PM_{10} locations were masked out thereby creating a subset of the emissions to compare with the PM_{10} concentrations. The reduced resolution of this dataset reduced the number of PM_{10} observations by over one-third to 2000 due to multiple points falling in a single 1-degree grid cell. Comparing SO_2 and NO_x emissions to PM_{10} concentrations reveals the same pattern for both gases shown in Figure 10 for SO_2 emissions. Emissions are uncorrelated to PM10 concentrations but higher emissions appear to extend along two limbs of high and low PM_{10} concentrations (~30 and 80 $\mu g/m^3$). The top 20% of SO_2 emissions are highlighted in green and can be seen to extend over a far larger range of emissions compared to the width of other quintiles.

Generally, the spatial patterns of SO_2 and NO_x emissions are similar to each other. High emissions occur in the US, Europe, Mediterranean Middle East, China and Southern India. They are lower throughout Africa and much of South America. However the locations of the very highest emissions are not entirely co-located. Figure 11 shows the location for the top 20% of each pollutant overlaid on each other. Areas which have co-located red and green points have both very high NO_x and SO_2 emissions. We note green areas (high SO_2) in East and South Eastern Europe and in parts of China along with parts locations in Russia as being places which have very high SO_2 emissions but lower NO_x emissions. The UK, Belgium, Holland and northern Germany contain elements of both sets.

Generally, we see a shift of peak location emissions for high NO_x in Western Europe to high SO_2 in Eastern Europe. We also note locations in California and Northern Texas having very high NO_x emissions for levels of SO_2 and relatively similar PM_{10} concentrations between 20-30 $\mu g/m^3$. Interestingly, with the exception of Northeastern

China and the Mediterranean Middle East, the top 20% of PM_{10} concentrations have very little spatial coincidence with the equivalent part of the distribution for NO_x or SO_2. The highest PM_{10} concentrations extend across northern India, Egypt and a band between the Sahelian and central Africa.

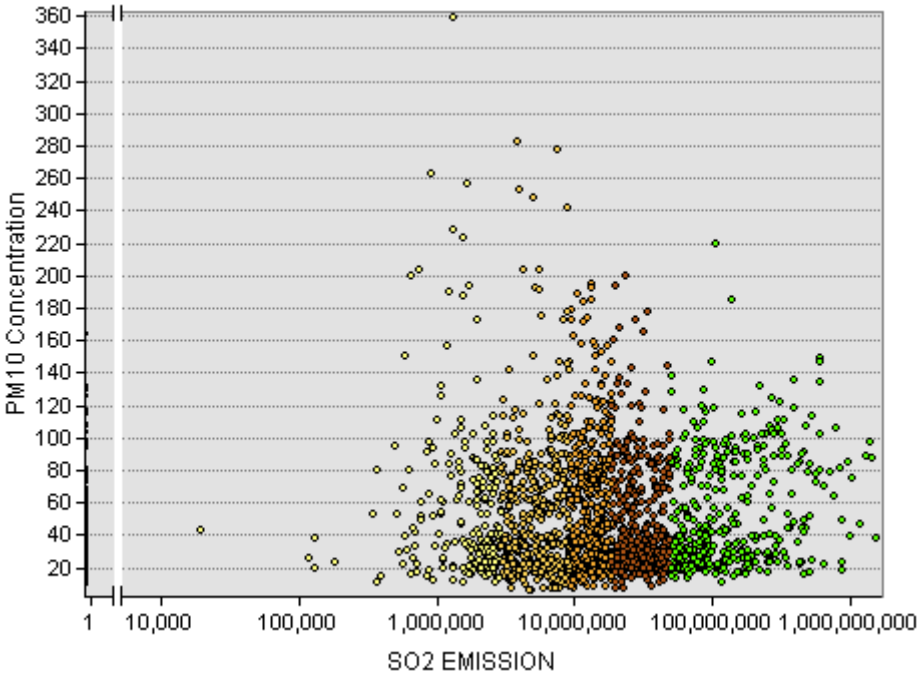

Figure 10. Scatterplot of Total SO_2 emissions (kg) with PM_{10} concentrations (µg/m3) shaded by quintile. The top 20% of PM_{10} concentrations are highlighted in green and located on the map in Figure 11.

The difference in location of peak emissions between gases could be indicative of the imposition of abatement policies for one gas over another (such as SO_2 in Europe). SO_2 and NO_x are two of the major inorganic pollutants which can combine with other gases and aerosols in the atmosphere to form secondary particles. In certain areas these can dominate the PM_{10} concentration but are more often found to exist in the finer fraction. In Europe 85% of fine particles are comprised of secondary emissions (formed in the atmosphere from precursor gases) rather than primary sources (EEA, 2008) which is unsurprising given the high amounts of precursor gases in the region. Fine *et al.*, (2008) support this finding with their figures of between 50-70% for the United States. Secondary particles have both organic and inorganic source contributions, the latter of which can be ameliorated through emission control policies. However, declining levels from inorganic sources over time will make secondary organic aerosols an increasingly important component of PM, the health implications of which are less well understood (Fine *et al.*, 2008).

There are numerous factors which help to explain why there is a mismatch between gaseous emissions and particulate concentrations. These are related to differences in the emissions structure and sources between locations but more obviously due to PM_{10} concentrations being strongly influenced by both natural and anthropogenic local contributions, and larger particles being less likely to be transported long distances. NO_x has a relatively short residence time in the atmosphere (around 1 day

due to photo dissociation in daylight) compared to that of a few days for SO_2 (Seinfeld & Pandis, 1998). The difference in lifetimes along with considerations of their chemical transformation and interdependences all affect their relative ability to form secondary particles.

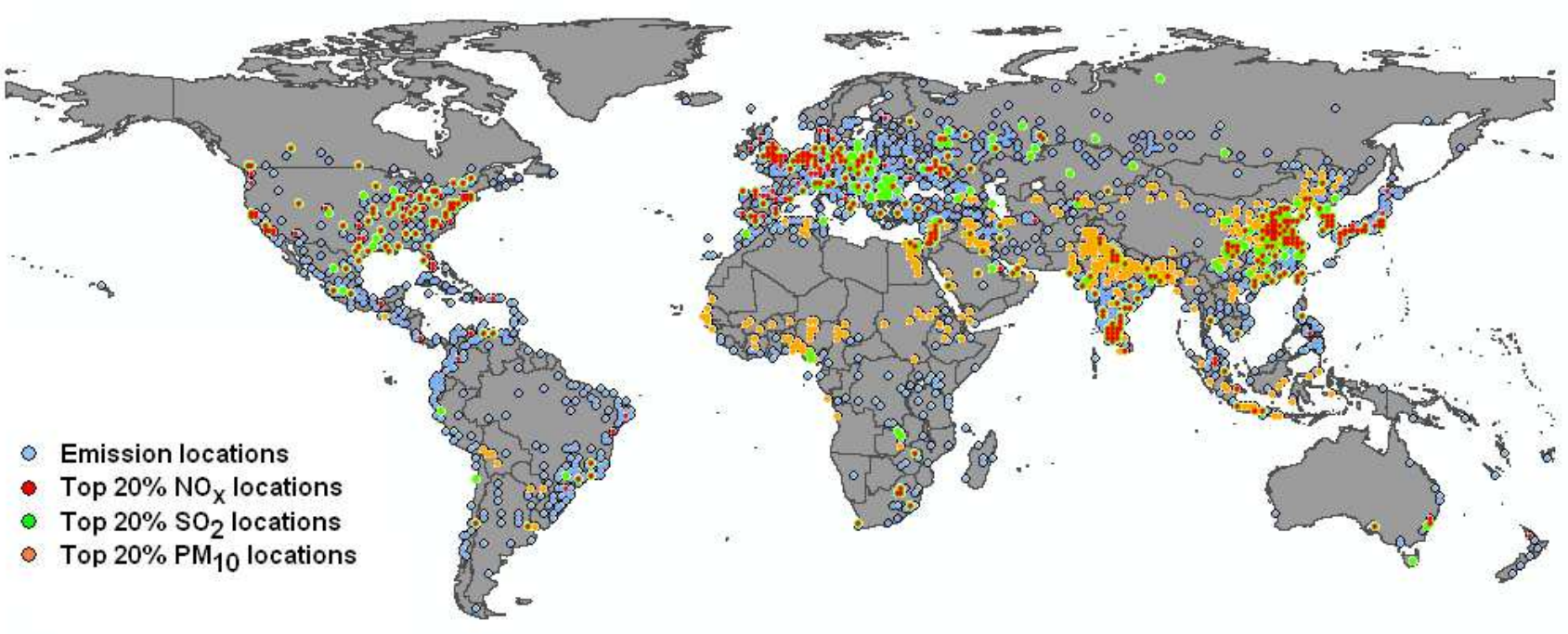

Figure 11. Spatial overlay of the top 20% SO_2, NO_x emissions with PM_{10} concentrations. Note the broadly coincident location of SO_2, and NO_x (red on green) is at odds with PM_{10} (orange) except in NE China.

However, despite these issues and the coarse nature of a global-scale analysis which precludes any detailed conclusions, we note even though PM_{10} concentrations are relatively low in Europe and North America, the emission of other pollutants in the finer fraction of particles are high and carry with them their own effects. China merits a special mention for the high exposure to very high co-incident levels of all three pollutants.

8 Conclusions

Spatial locations have been provided for 3,197 cities of over 100,000 inhabitants plus capital cities around the world. These have been mapped onto population density and summarized by the IIASA GGI regional classification scheme to reveal the global distribution of PM_{10} concentrations in urban areas. Spatially explicit representations show that there can be large variations between cities within countries. With few exceptions, aggregated country level PM_{10} concentrations decline with increasing per-capita GDP. Using the World Health Organization Air Quality Guidelines on PM_{10}, nearly 40% of the global urban population covered in the dataset lives outside of the first air quality guideline target. Many of these cities are in countries which will experience the greatest increases in urbanization. The explicit consideration of urban areas results in variable coverage of total population in the less urbanized parts of the world, especially Africa. As such only part the story of global PM_{10} concentrations can be told. The spatial distribution of concentrations are likely to be highly heterogeneous not just between urban and rural but between indoor and outdoor sources. In order to gain an understanding of these dynamics, detailed information on emission structures, their relative contributions and utilized technologies need to be more widely collected and synthesized to help complete the picture of air pollution in its many forms. Consideration of these issues will advise on what kinds of data are needed and how to best to use them. Basic exposure measures are presented here; more refined methods would consider the amount of time spent exposed to outdoor air pollution, which may in fact be low.

The complex combination of sources and factors which contribute to PM_{10} concentrations along with the aggregated nature of the parameters used in the model means there are significant uncertainties in the estimates of PM_{10} concentrations. However, comparisons of selected cities with other published values and show broad agreement but also the high variability in PM_{10} as a measure depending on the time period of the measurement and prevailing conditions climatic conditions.

Given that the PM_{10} data are modelled according to a range of proxies rather than an atmospheric model or observed in situ, and the demonstrated error in some estimates, it would be imprudent to use this dataset to draw far-reaching conclusions from the comparison of precursor gases at least in terms of policy decisions. However the dataset does provide a broad overview of the magnitude of PM_{10} across the world's urban centers and comparative analysis of the PM_{10} data to other pollutants does offer insights into the magnitude and composition of the finer fraction. Given the connected nature of air pollution to climate change and human health, emission of gases from combustion, such integrated analysis reveals differences in spatial patterns of emissions which are suggestive of further investigation.

This global dataset of PM_{10} concentrations adds to the toolchest of spatial datasets available to IIASA researchers. While principally a dataset on PM_{10} concentrations,

the purely spatial component of verified coordinates for capitals and cities along with their population should find wide and useful application across many projects. The list of omitted locations is in Appendix A and the location of the dataset and associated files on the IIASA network is given in Appendix B.

9 References

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; and Centro Internacional de Agricultura Tropical (CIAT). 2004a. Global Rural-Urban Mapping Project (GRUMP), Alpha Version: Settlement Points. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw. (Downloaded 23.5.2008).

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; and Centro Internacional de Agricultura Tropical (CIAT). 2004b. Global Rural-Urban Mapping Project (GRUMP), Alpha Version: Coastlines. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw. (Dowloaded 25.9.2007).

Chan, C.K. and Yao, X. Air pollution in mega cities in China – A review, *Atmospheric Environment* (2007), 42(1) 1-42.

Cohen A.J., Ross Anderson, H., Ostro, B., Pandey, K.D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Pope, A., Romieu, I., Samet, J.M., and Smith K. 2005. The global burden of disease due to outdoor air pollution. *J Toxicol Environ Health A.* 68(13-14):1301-7.

European Commission, 1999. Council Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air. Available from: *http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexplus!prod!CELEXnumdoc&lg=EN&numdoc=31999L0030* (accessed 28.7.2009)

EEA, 2008. Emissions of primary particles and secondary particulate matter precursors (version 1). CSI 003. European Environment Agency. Available from: http://themes.eea.europa.eu/IMS/IMS/ISpecs/ISpecification20041001123025/IAssessment1190630330756/view_content (accessed 6.8.2009)

EPA, 2009. Nattional Ambient Air Quality Standards. Available from: http://www.epa.gov/air/criteria.html (accessed 28.7.2009)

Fine P.M., Sioutas, C., and Solomon, P.A. 2008.Secondary particulate matter in the United States: insights from the Particulate Matter Supersites Program and related studies. *J Air Waste Manag Assoc.* 58(2):234-53.

Grubler, A., O'Neil, B., Riahi, K., Chirkov, V., Goujon, A., Kolp, P., Prommer, I, Scherbov, and Slentoe, E. 2007. Regional, national and spatially explicit scenarios of demographic and economic change based on SRES. *Technological Forecasting and Social Change Special Issue.* 74(7) 980-1029

Kim Oanh N. T., Upadhyay N., Zhuang Y-H., Hao Z-P., Murthy D.V.S., Lestari P., Villarin J.T., Chengchua K., H. X. Co., Dung N. T., Lindgren E. S. (2006). Particulate Air Pollution in Six Asian Cities: Spatial and Temporal

Distributions, and Associated Sources. *Atmospheric Environment, 40,* 3367-3380.

Larsen, B., Hutton, G. and Khanna, N. 2008. Copenhagen Consensus 2008 Challenge Paper: Air Pollution. Copenhagen Consensus Center.

Medina S., Boldo E., Saklad M., Niciu E.M., Krzyzanowski M., Frank F., Cambra K., Muecke H.G., Zorilla B., Atkinson R., Le Tertre A., Forsberg B. and the contribution memberss of the APHEIS group**.** 2005. *APHEIS Health Impact Assement of Air Pollution and Communications Strategy. Third year report.* Institut de Veille Sanitaire, Saint-Maurice; 232 pages.

NEAA, 2005. EDGAR Fast Track 2000 (32FT2000) Gridded 1ºx1º SO_2 emissions. Netherlands Environmental Assessment Agency. Dataset Available from: http://www.mnp.nl/edgar/model/v32ft2000edgar/edgarv32ft_acid/edgv32ft-so2.jsp
(accessed 12.09.2008).

NOAA-NGDC. 2008. DMSP-OLS Night-time lights Time Series v.2. Avaialble from: http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html (Dowloaded 25.9.2007)

Olivier, J.G.J., Van Aardenne, J.A., Dentener, F., Ganzeveld, L. and J.A.H.W. Peters (2005). Recent trends in global greenhouse gas emissions: regional trends and spatial distribution of key sources. In: "Non-CO_2 Greenhouse Gases (NCGG-4)", A. van Amstel (coord.), page 325-330. Millpress, Rotterdam, ISBN 90 5966 043 9.

Pandey, K.D., Bolt, K., Deichmann, U., and Hamilton, K. Urban Outdoor Particulate Air Pollution: New Estimates. (2002 est.). Powerpoint presentation from The World Bank Development Economics Research Group and the Environment Department Working Paper. *The World Bank*, Washington DC.

Seinfeld, J.H. and Pandis, S.N. 1998. Atmospheric Chemistry and Physics, from Air Pollution to Climate Change. Wiley & sons. New York.

Smith, K.R. 1993. Fuel combustion, air pollution exposure, and health: The situation in developing countries. *Annual Review of Energy and the Environment.* 18: 529-566.

United Nations, 2006. *World Urbanization Prospects: The 2005 Revision.* Urban and Rural Areas Dataset (POP/DB/WUP/Rev.2005/1/Table A.3), dataset in digital form. Available from: http://esa.un.org/unup/. New York: United Nations (accessed 6.8.2009)

U.S. Census Bureau. 2008. U.S. Gazetteer. Available from: http://www.census.gov/cgi-bin/gazetteer (accessed 23.7.2008).

World Bank. 2007. Word Development Indicators. *The World Bank.* Washington D.C.

World Bank. 2008. Air Pollution in World Cities database. Available from: http://go.worldbank.org/3RDFO7T6M0 (accessed 11.08.2008)

World Gazetteer, 2008. Search for a geographical entity. Available from: http://world-gazetteer.com (accessed 23.7.2008)

World Health Organization, 2005. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide - Global update 2005. Available from;
http://www.who.int/phe/health_topics/outdoorair_aqg/en/index.html (accessed 11.7.08)

Xie, S., Yu, T., Zhang, Y., Zeng, L., Qi, L. and Tang, X. 2005. Characteristics of PM_{10}, $S0_2$, NO_x and O_3 in ambient air during the dust storm period in Beijing. *Science of the Total Environment.* (345) 153-164.

Appendix A: Cities excluded from the original dataset.

iso3	Country	Citycode	City	Population	Concentration	Latitude	Longitude	Notes
CHN	China	1560172	Anyang	779,264	112.0			Henan/Hebei
CHN	China	1560353	Bodong	147,147	41.3			Cannot find
CHN	China	1560242	Da'an	510,280	76.4			Ambiguous
CHN	China	1560241	Dayuan	516,053	46.1	25.0000	121.3000	Taoyuan/Taoyuan shih (already in dataset)
CHN	China	1560060	Dengzhou	1,757,449	99.8			duplicate
CHN	China	1560174	Donglin	767,866	106.1	41.8200	123.5600	Dongling (district of shenang)?
CHN	China	1560345	Dongsen	185,751	77.8			Cannot find
CHN	China	1560055	Dongshan	1,785,218	64.3			duplicate
CHN	China	1560040	Dongwan	2,194346	72.3			Cannot find
CHN	China	1560084	Haozhou	1,513,916	37.5			Cannot find
CHN	China	1560314	Houzhou	313,527	53.2			Ambiguous
CHN	China	1560273	Hua	428,759	81.3			No firm data
CHN	China	1560333	Jining(Shanxi Sheng)	243,942	61.9			Cannot find
CHN	China	1560199	Shangzhou	646,005	86.9			No firm data
CHN	China	1560292	Shaown	377,368	62.7			Cannot find
CHN	China	1560318	Wuchuang	307,182	103.5			Cannot find
CHN	China	1560070	Yulin (Hunan)	1,671,985	71.9			Cannot find
CHN	China	1560252	Zicheng	486,031	73.3			Cannot find
CHN	China	1560110	Fuyu	1,193,819	91.8	47.6413889	124.6794444	Jilin/Heiliongjiang both counties
CHN	China	1560128	Beidong	1,007,469	83.3	35.6122222	111.5425	ambiguous
CHN	China	1560207	Zuozhou	614,753	117.1	22.6886111	107.4969444	Cannot verify
CHN	China	1560304	Dong chuan	353,062	68.8	26.1666667	103.0333333	Cannot verify
IRQ	Iraq	3680013	Mamoon	366,500	104.1			Cannot find
MEX	Mexico	4840057	Cuatlas	194,109	40.3			Cannot find

Suspicious match

iso3	Country	Citycode	City	Population	Concentration	Latitude	Longitude	Notes
CHN	China	1560105	Sifen	1,247,348	58.4	27.528900	113.487800	Appears mismatched with light
CHN	China	1560264	Zixing	455,848	42.9	25.97	113.4	Appears mismatched with light
CHN	China	1560063	Guikong	1740,184	52.1	22.34	111.715	Appears mismatched with light
CHN	China	1560304	Dong chuan	353,062	68.8	26.1666667	103.0333333	Appears mismatched with light
CHN	China	1560206	Baihua	616,969	82.1	29.1	104.6	Appears mismatched with light

Appendix B

At the time of writing, the location of the dataset and this report is placed on the TNT network drive at the following location: \\tigris\/p6tigris/tnt /Databases/PM10

Within this directory ArcGIS point shapefile:

PM10_Locations.shp

With associated files in the same location.

*.sbn, *.shx, *.dbf, *.sbx

Shapefile fields

ISO3	- ISO3 Country Code
CITYCODE	- Citycode from World Bank
UNCODE	- UN City code
CITY	- City Name
POPULATION	- Population (people)
CONCENTRATION	- PM_{10} Concentration ($\mu g/m^3$)
LATITUDE	- Latitide (Decimal Degrees)
LONGITUDE	- Longitude (Decimal Degrees)
Pop_Conc	- Exposure (Population x Concentration: people.$\mu g/m^3$)
Norm_Exp	- % contribution of Total Exposure

MS-Excel files of this list and the unmatched locations:

PM10_locations.xls

unmatched_locations.xls