Air Pollution and Human Health Burden in Urban India
Amruta Nori-Sarma / Study Co-Investigators: R.K. Thimulappa¹, G.V. Venkataramana², A.K. Fauzie², S.K. Dey¹, Jesse Berman³, Michelle L. Bell⁴
1 – Jagadguru Sri Shivarathreeswara University, Mysore; 2 – University of Mysore, Department of Environmental Science; 3 – University of Minnesota School of Public Health; 4 – Yale University School of Forestry and Environmental Studies

Abstract
Introduction: In India, rapid urbanization has led to a decrease in urban air quality. Few ongoing studies apply spatially heterogeneous sampling to assess ambient air pollution levels in urban India. Understanding air pollution exposure differentials is particularly important in cities, where socio-economic disparity by neighborhood may increase health impacts of varying exposures.

Methods: Seasonal sampling campaigns have been conducted at 150 sampling sites throughout urban Mysore, Karnataka, India. Sampling sites were selected using a 3-step combination of purposeful site selection, systematic random sampling, and purposeful site selection to fill in remaining gaps of exposure measurement. Nitrogen dioxide (NO₂) levels were assessed using passive Palmes tube and Ogawa badge technologies at these sites, to develop a spatial interpolation of ambient air pollution exposure in these cities. Seven hundred subjects from 17 city wards were selected for participation in a cohort for lung function testing as part of the Burden of Obstructive Lung Diseases (BOLD) study in 2012-2013; those subjects who have not moved were re-sampled for lung function in 2017-2018 and administered a concurrent qualitative survey, to assess long-term lung function decline and potential confounders of the relationship between air pollution and health.

Results: Mysore, considered a “clean city” of India, has minimum air pollution levels during the post-monsoon season in excess of 20 ppb averaged over a 2-week period. Maximum air pollution levels measured in the city around 50-60ppb over a 2-week period, indicating that pollution levels may exceed the health-protective thresholds established by the World Health Organization. Strong spatial and temporal patterns indicate a seasonal trend in air pollution levels, as well as the importance of point sources of pollution. Of the original 700 cohort subjects, 320 have been resampled to assess lung-function; subjects living in lower socioeconomic status (SES) neighborhoods appear to be less concerned with long-term health decline associated with air pollution exposure than their counterparts in higher income communities.

Conclusions: While traffic pollution has been indicated in the past as a major contributor to ambient air pollution levels in urbanizing centers of Asia, our results indicate that air pollution levels do not follow major roadway patterns. Other pollution factors (e.g., building density), as well as highly localized environmental factors (such as waste burning and small trash fires), might be critically influencing air pollution exposure patterns in urban India. Study participants who live in the lowest income neighborhoods seem to be the least concerned about the long-term effects of air pollution within their neighborhoods. Future interventions targeted to improve the health of people in these impoverished neighborhoods will depend on generating local will to address these health effects.

Introduction / Related Works
Air pollution exposure assessment studies and assessments of the corresponding health outcomes are needed in developing countries worldwide, but especially in heavily populated and urban countries such as India. Air pollution is likely to differ in Indian cities due to differences in traffic pollution as a result of fuel characteristics, vehicle technology, contributions of heavy polluters (e.g.,
more motorcycles), driving habits, traffic patterns (e.g., frequent congestion) and roadways (e.g., graded vs. ungraded). Further, the presence of heavy industry in urban areas of India is likely to impact air quality and the heterogeneity of air pollution within urban environments in developing countries. Compared to Western cities, cities in India likely have different built structures, topography, weather and land use patterns influencing the spatial distribution of air pollution, which can differ even between cities in northern versus southern India. Additionally, actual exposure may vary through differences in lifestyle and culture that affect indoor/outdoor activity patterns and housing characteristics as well as through different population characteristics that may be generalizable to other countries.

Novel approaches for estimation of air pollution exposures in sensitive communities in developing countries are needed. Traditionally, urban air pollution studies estimate air pollution exposure by assigning values from the nearest available monitor or averaging values from nearby monitors, an approach that uses readily available data from regulatory monitors and provides large spatial and temporal coverage. This approach either requires a large network of monitors, or relies on the assumption that the pollutant of interest is homogeneous within urban areas. However, my pilot study (described in the Preliminary Results, below) indicates spatial heterogeneity within small areas in urban India, leading to high potential for exposure misclassification that may notably alter the significance, direction and magnitude of health effect estimates in epidemiological studies using traditional methods. Newer approaches have emerged that make use of field samples of air pollution levels, as well as spatial analyst tools such as Geographical Information Systems (GIS), to estimate individual level exposures from more spatially dense air pollution monitoring sites.

The spatial gradient of air pollution exposure may be particularly varied due to point sources of emissions, which may be a larger factor in India than in more developed countries. Therefore, pollution measurements taken at a small number of fixed monitoring stations distributed across a city in India may not be a good characterization of air pollution exposures, and a more thorough understanding of the gradient of air pollution exposures is necessary. The need for understanding intraurban exposure variation relates to increasing contribution from factory and traffic sources, potential for intraurban gradients of traffic pollution near major roads, larger variation within cities than between cities for some traffic pollutants, and significant health effects due to exposure within cities.

Interest in assessing the exposure to ambient air pollution at the within-city scale has increased for several reasons. The contribution of traffic pollution has grown, and several studies agree that the demand for transportation will exceed improvements to emissions reduction technologies. This is especially important as a factor for within-city variation of air pollution levels in developing country settings such as India, where rapid economic growth has led to an increase in the number of a variety of vehicles on major roads. Regardless of regulatory interventions, higher exposure to traffic pollution with distinct within-city gradients have been shown in studies of urban areas around major roads and highways, suggesting higher exposure differentials within cities than between cities. Studies conducted in primarily Western and developed countries suggest that the exposure within cities may exert significant health effects, which differ by neighborhood. For example, a recent study of a cohort of 5,000 people in the Netherlands reported increased cardiopulmonary mortality near major roads. Urban background pollution interpolated from government monitoring sites was also shown to exert an independent effect on mortality. This study used a basic type of exposure measurement (buffers), and there is a need to test similar relationships with more robust exposure metrics in regions where the within-city heterogeneity of air pollution exposure may be higher due to urban characteristics.
Interpolation models of air pollution exposure rely on deterministic and stochastic geostatistical techniques. Several selected recent studies in developed areas of Europe and North America have tested a variety of empirical modeling methodologies using real data inputs, to assess the feasibility of using these analytical methods to characterize the spatial distribution of air pollution levels. The development of models to assess air pollution exposures within cities for assignment to subjects in health studies has been identified as a priority area for future research.\textsuperscript{9,10} While surrogate measures, such as distance to roads, have been related to large health effects, these may misclassify exposures and lead to issues of systematic bias or residual confounding.\textsuperscript{3} Various methods exist to address intraurban variation of air pollutants, including the use of indicator variables (e.g. distance to major road or point source\textsuperscript{11,12}), multiple regression models\textsuperscript{13}, interpolation (e.g. kriging\textsuperscript{14–20}, inverse distance weighting\textsuperscript{12,21}), dispersion modeling\textsuperscript{22}, and Land Use Regression (LUR)\textsuperscript{12,18–20} models.

For all of these different types of models, field measurements of the target pollutant are obtained at a set of monitoring stations distributed throughout the study area. On the basis of this information, the objective is to generate estimates of pollutant levels at sites other than the locations of monitoring stations. To date, there have been no published studies using interpolation models to assess intraurban air pollution exposure in India.

Socioeconomic status at both the individual and neighborhood level may be one of the most impactful determinants of health outcomes associated with air pollution exposure in urban India.\textsuperscript{23} The socioeconomic status of a neighborhood in India may be a driving factor in the relative health impacts of air pollution among different communities occupying the same urban space. Neighborhoods of low SES in India are composed of understudied populations. Residents are also underserved from a healthcare access perspective, and therefore may be particularly vulnerable to adverse health outcomes attributable to environmental exposures.

Prevalence of respiratory illness among the sensitive subpopulations living in low-income areas is likely to be relatively high, according to previous studies in other similar settings.\textsuperscript{24,25} Previous qualitative studies of the perceptions of people living in slum communities regarding their relative health risks illustrate an apparent disconnect between knowledge and opportunities, with individuals engaging in practices that placed them at high risk of exposure to air pollution. For example, residents living in a slum in Nairobi appeared from a survey of their perceptions of their living conditions to be resigned to poor air quality and low health access.\textsuperscript{26} Overall, it appears that urban poor communities have historically experienced a lack of agency in addressing prevalent air pollution.

\textbf{Objectives}

As shown in Figure 1, below, this project aims to develop a spatially-resolved estimate of long-term exposure to air pollution throughout the city of Mysore – a small but fast-growing city in the southern state of Karnataka. Modifications will be made to previously conducted air pollution sampling and models to account for local characteristics. A goal of this project is to develop a flexible methodology for exposure assessment suitable for Asian cities in developing countries. This research will answer critical questions: What are the neighborhood-level differences in air pollution exposures within urban centers of India? Where are the hotspots of high air pollution exposure in the cities being studied? What is the influence of seasonality on pollution exposure (e.g. spatial patterns of air pollution)? In what ways are exposure models developed in Western cities applicable, or not, for a rapidly growing city in India?
The output from the air pollution model will subsequently be used to estimate exposure to air pollution within sensitive low-income communities in Mysore. An additional aim of the project is to understand the relationship between a neighborhood’s socioeconomic indicators and air pollution exposure estimates, in order to assess the environmental inequities in air pollution by socioeconomic status in the poorest neighborhoods of India. Finally, this project will use qualitative survey data to assess the self-perception of neighborhood residents of their own environment, and the potential for detrimental health effects associated with air pollution levels in their neighborhood.

Figure 1: Project Proposal, Datasets, Analysis Models, and Outputs

Methods

1. Air Pollution - Selection of Sampling Sites
For the purpose of this sampling campaign, the city of Mysore was divided up into a 300m grid, shown in Figure 2 below. This grid size was selected upon examination of similar previous studies conducted in the US and Europe. This yielded approximately 1,420 potential air pollution sampling sites, at the centroid of each 300m x 300m grid cell. To further correspond to prior studies of air pollution distribution, a sample size of 150 points was selected, or approximately 10% coverage of the available grid points.

Air pollution sampling sites at the centroids of the 300m grid were selected using a 3-tier systematic approach. In tier 1, Purposeful Site Selection, 15 sampling sites for a pilot air pollution sampling campaign were selected in Mysore through assessment of areas likely to be low, medium, and high pollution, including collocation with existing government monitors as well as conversations with local collaborators to ascertain areas of interest.

In tier 2, Randomized Site Selection, a preliminary spatial interpolation of air pollution levels throughout the city across the same 300m grid described above was developed using the pilot data from the initial 15 air pollution sampling sites. This spatial interpolation was developed using an Inverse Distance Weighting (IDW) procedure, in which the assigned values to unknown points are calculated with a weighted average of the values available at the known points. Each of the grid cells was also assigned a value for the presence of a major road, the presence of different types of buildings, and Land Use Classifications from publicly available data sources. These data points were combined to randomly select 130 additional sites for air pollution sampling, with areas of high air
pollution and high road/building density preferentially selected over areas of low pollution and low road/building density.

Figure 2: 300m x 300m grid surface of Mysore (each intersection shown here represents the centroid of a grid cell)

Finally, tier 3, *Further Purposeful Site Selection / Filling in the Gaps*, was conducted in the field. Of the 130 sites selected randomly, 11 were discarded during the first round of field sampling due to infeasibility of the sampling site (for most of the sites, inaccessibility from the nearest road or location in the middle of a government compound). Of the final 16 sampling locations, 8 were systematically selected from the center of the city, where pollution levels are high. The remaining 8 were selected at the center of areas where at least 5x5 grid cells had not been selected for sampling, and where major roads or residential / commercial buildings were located.

2. *Air Pollution Field Measurements*

Nitrogen Dioxide (NO\(_2\)) was selected as the pollutant of interest given the feasibility and reliability of field measurements and relatively low cost of equipment and sampling for NO\(_2\), as well as the potential for use of NO\(_2\) samples as a proxy for levels of other pollutants. Further, current studies link short-term exposure to NO\(_2\) with adverse health outcomes, including airway inflammation in healthy people and increased respiratory symptoms in people with asthma. Exposure to NO\(_2\) at high concentrations is of particular concern among sensitive populations, including children, the elderly, and people with chronic respiratory ailments. Palmes tubes were primarily used for NO\(_2\) sampling at an average height of 2.5m (breathing level). Additionally, 10% of the data points (15 locations total) were collocated with Ogawa badges to increase the reliability of the estimates. Data were collected in the field and subsequently returned to the US to be analyzed using laboratory facilities provided by Johns Hopkins University collaborators, as well as in laboratories at Emory University, by the primary investigator, as well as a laboratory manager at Emory University who has prior experience for air pollution exposure data analysis using Palmes tubes for exposure data collection.

Mysore experiences a subtropical, temperate climate with four distinct seasons a year: winter (December – Feb), summer / pre-monsoon (March – June), monsoon / rainy season (July – August), and post-monsoon or autumn (September – November). Four sampling campaigns were performed, one for each season (September, January, April, and July), for two weeks each. The campaigns were planned such that they were not scheduled during known unusual events (e.g. major holidays), as
far as possible given the prevalence of religious holidays on the Indian calendar. All samplers were installed within a 3-5 day window, with time recorded. Samplers were collected in the same order as they were installed. The data were collected continuously as an aggregate measure of the air pollution during the whole 2-week sampling period.

3. Cohort Study – Statistical Sampling and Health Data Collection

a. Sampling
The sampling strategy for inclusion into the health data collection cohort was based on a multi-stage random sampling procedure, stratified by wards:

- In the first stage, sampling units were wards. There are 17 Wards in the constituency and all seventeen wards had been selected.

- In the second stage, houses were sampled in clusters of 10. A random start was made in each ward and 10 consecutive houses were included following this house number. That is, if we want to sample k clusters in a ward with 10 houses in each cluster, then k random numbers were selected with re-sampling if any subsequent house number leads to inclusion of a house (or houses) already included. All houses in a ward were consecutively numbered from 1 to N.

- In the third stage, all individuals aged 40 years or over from each selected household were invited to participate.

Unfortunately, it was difficult to reconstruct which participants came from the same household. For analysis purposes we treated the sample within each ward as if it was a random sample (in fact it will have been biased towards larger households, because all individuals in each randomly selected household took part). Census data in the Mysore is not broken down by age and sex, so it was not possible to use post-stratification by age and sex to try to reduce bias (this is a method of adjusting sampling weights so that they add up within subgroups to the known population totals in those subgroups).

b. Health Survey Data Collection
Data from this cohort had previously been collected in 2012/2013, for inclusion in the Burden of Obstructive Lung Diseases (BOLD) international cohort study. All of the members of the cohort who have not moved since their previous inclusion in the BOLD study were selected for resampling in the current project.

Subjects participating in the health data collection portion of the study were administered a spirometry test in their home, using the NDD EasyOne field spirometer. Additionally, participants were administered a qualitative questionnaire verbally, with answers recorded in a Qualtrics survey by field assistants or by the PI.

Results
Air Pollution Data - Pilot Sampling Campaign, Results
Figure 3, below, shows the locations that were chosen for the pilot air pollution sampling campaign. From the 15 locations, a spatial interpolation was developed using Inverse Distance Weighting (IDW) of the data from the Ogawa badges at the 15 locations, to generate the demand surface for air
pollution exposures. The demand surface, shown in Figure 4 below, was incorporated into the random sampling for subsequent site selection.

Figure 3: Pilot Air Pollution Sampling Locations

Figure 4: Demand Surface, Pilot Sampling (Inverse Distance Weighted; PPB NO₂)

Air Pollution Data – Seasonal Sampling Campaigns, Preliminary Results
Using the information from the pilot air pollution sampling and spatial interpolation of the air pollution exposures, as well as purposeful site elimination and site selection during the first field sampling campaign, 150 sites were chosen for each of the 4 seasonal air pollution sampling campaigns. These 150 sites are shown in Figure 5, below.

Figure 5: 150 Sites Selected for Seasonal Air Pollution Sampling, Mysore (2016-2017)

Data were collected at each of these 150 points, with some sampler losses (on average, 12% loss over the entire year of data collection; the highest losses were experienced at the end of the
summer / beginning of the monsoon season). Preliminary results from the sample data collection are provided in Figures 6 and 7, below.

Figure 6: Preliminary Results, Data from Seasonal Air Pollution Sampling Campaign

While it appears from Figure 6 that the highest air pollution levels occurred during the summer air pollution sampling campaign, it is important to point out that this summer campaign also experienced the greatest losses of samplers. It is possible that samplers which were prone to losses due to meteorological variables may have been located in lower pollution areas, which would lead to a bias in the estimates of exposures during the summer season.

Figure 7: Preliminary Results, Ordinary Kriging of NO₂ Levels, by Season
Amruta Nori-Sarma  
Hixon Center for Urban Ecology – Final Report

Figure 7 shows the preliminary ordinary kringing spatial interpolation of the air pollution exposures throughout the city of Mysore, by season. The center of the city stands out as the area of highest air pollution exposure in the city, which is a consistent finding throughout all of the seasons for which the air pollution field sampling was conducted. This city center area also represents communities of lowest income, and densest population. While traffic has been postulated as the major driver of air pollution levels in developing cities, these results indicate that there are likely characteristics of the built environment which are influential on long-term air pollution exposures.

Additionally, the World Health Organization (WHO) has developed health-protective guidelines for annual and daily NO₂ levels. The annual mean recommendation for NO₂ exposure is 40 µg/m³; at an annual average temperature of 25℃, this translates to 21.3 ppb annual mean air pollution exposure. The areas at the center of the city are consistently in excess of this WHO recommendation, indicating that despite the reputation of Mysore as being one of the “clean cities” of India, residents are likely experiencing the long-term health effects of air pollution exposures.

Cohort Data – Lung Function, Qualitative Survey, Preliminary Results
Of the original 700 participants in the BOLD study, through November 2017, 320 subjects have been resampled for inclusion in the current study. Because of the method for the current round of data collection, no statistically significant results have been prepared for report as yet. However, one interesting preliminary qualitative result is that the people who are residents in the lowest-income areas of Mysore, appear to be the least concerned about air pollution levels in their neighborhood. Those study subjects living in relatively higher SES communities nearby appear to be more concerned about the potential for long-term health impacts from air pollution.

Conclusions, Study Strengths / Limitations
One of the main strengths of this study is that it represents the development of a sustainable method for site selection and field sampling in developing country settings. The methods for field site selection combine the knowledge of local co-investigators, as well as a randomization process to distribute the sites appropriately to reduce measurement bias. The preliminary results indicate that the air pollution exposure estimates derived using these methods capture the spatial variation in air pollution exposures well.

The study does have some significant limitations. Related to the sampling method, local conditions require very intensive sampling campaigns that are difficult to replicate. Each sampling campaign required several days of field work to install samplers, as well as several days to retrieve the samplers. Stationary passive sampling technologies are cost effective and relatively less resource intensive, which is an important factor in developing country air pollution sampling campaigns. However, deploying samplers for longer periods makes them much more prone to losses, particularly in places such as India where passers-by tend to disturb samplers frequently.

Methodologically, this study provides a single pollutant model in an area where there are likely to be complex mixes of multiple pollutants that interact in a non-linear manner to impact human health. While we selected NO₂ for sampling due to the relative ease and low costs associated with spatially heterogeneous sampling for NO₂, it may also be the case that other pollutants such as particulate matter (PM) or the chemical components of PM are more impactful on health. Additionally, environmental data in the region is sparse, particularly the spatial data that would be relevant to Kriging, Land-Use Regression (LUR) and other sophisticated statistical models of air pollution exposure. Developing relationships and cooperation with government agencies is difficult but vital to efforts to assess air pollution exposures throughout cities in India.
Until now, Mysore has been seen as a model city for low air pollution levels and overall good health. However, our results indicate that air pollution levels, especially in the center of the city where the population density is highest, may exceed the health protective levels of NO₂ recommended by the WHO. Additionally, those persons who experience the highest exposure to air pollution may be least interested in or able to cope with the long-term effects of air pollution.

**Next Steps**
The air pollution modeling effort is well under-way. The next stage of the air pollution exposure assessment involves building a universal kriging or land-use regression model, depending on which model provides the best estimates of air pollution exposure given the availability of environmental co-variates. Additionally, it will be important to assess the desired output from the model – this project is most interested in assessing which environmental factors are MOST influential on the distribution of air pollution throughout the city.

The human health survey is currently ongoing among the members of the BOLD cohort in Mysore. The co-investigators aim to complete the data collection process by March of 2018. Statistical analysis of collected data will begin at that point, including assessing long-term lung function decline and the relationship to air pollution levels in the city.

One of the main limitations of the study, described above in the *Limitations* section, is the use of a single pollutant (NO₂) to assess levels of air pollution exposure in Mysore. Co-investigators have developed and are currently implementing a campaign to sample the levels of PM10 (particulate matter less than 10 microns in diameter) as well as PM2.5 (particulate matter less than 2.5 microns in diameter) in Mysore. This air pollution sampling will be collocated at 5-10 sites where NO₂ measurements were taken during the current study, to assess the relationship between different types of pollution. The co-investigators will also seek to assess the oxidative potential of the chemical components of PM, to understand the potential pathways for health effects.

In the long-term, this group of co-investigators seeks to expand the current study to neighborhoods of Bangalore. The proposed work would use similar methods for site selection, field air pollution data collection, and recruitment of a long-term cohort for health data collection, to assess the health effects of air pollution in a larger and more complex city in India.
Acknowledgements

This work was funded by the Yale Tropical Resources Institute, the Yale MacMillan Center, the Air & Waste Management Association, and the Yale Hixon Center for Urban Studies.

This research has been supported by a grant from the U.S. Environmental Protection Agency’s Science to Achieve Results (STAR) program. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.

Special thanks to Dr. Ana Rule @ Johns Hopkins Bloomberg School of Public Health for providing Ogawa samplers and laboratory space for sample analysis, and Prof. P. Barry Ryan @ Emory University Rawlins School of Public Health for sample analysis and support.

Many thanks to Prof. Michelle Bell, the Bell research group @ Yale, and Drs. Rajesh and Raman and their teams in Mysore.

Signatures:

_______________________________________________
Amruta Nori-Sarma
Hixon Student Research Fellow

_______________________________________________
Michelle L. Bell
Faculty Advisor
Bibliography


