Land Use Regression Modeling of Intra-Urban Residential Variability in Multiple Traffic-Related Air Pollutants

Jane E. Clougherty[1], Rosalind J. Wright[2,3], Lisa K. Baxter[1], Jonathan I. Levy[1]

Institutional Affiliations:

[1] Harvard School of Public Health, Department of Environmental Health, Landmark Center Room 404, P.O. Box 15677, Boston, MA, 02215, USA

[2] Channing Laboratory, Brigham & Women's Hospital, Harvard Medical School, Landmark Center, 401 Park Drive, Boston, MA 02215, USA

[3] Harvard School of Public Health; Department of Society, Human Development, and Health, 677 Huntington Avenue, Boston, MA 02215, USA

Corresponding Author: Jane E. Clougherty, ScD; Harvard School of Public Health, Department of Environmental Health, Landmark Center Room 404, P.O. Box 15677, Boston, MA, 02215, USA. jcloughe@hsph.harvard.edu, phone 617-384-8811, fax 617-384-8859

Running title: Intra-Urban Variability in Multiple Air Pollutants

Key words: geographic information systems (GIS), intra-urban variability, traffic, fine particles ($PM_{2.5}$), elemental carbon (EC), nitrogen dioxide (NO_2)

Word count: Abstract: 315; Main text: 4,912; References: 1,436

ABSTRACT

Background

There is a growing body of literature linking GIS-based measures of traffic density to asthma and other respiratory outcomes. However, no consensus exists on which traffic indicators best capture variability in different pollutants or within different settings. As part of a study on childhood asthma etiology, we examined variability in multiple traffic-related air pollutants within urban communities, using a range of GIS-based predictors and land use regression techniques.

Methods

We measured fine particles ($PM_{2.5}$), nitrogen dioxide (NO_2), and elemental carbon (EC) outside 44 homes representing a range of traffic densities and neighborhoods across Boston, Massachusetts. Three to four-day average samples were collected during winters and summers from 2003 to 2005. Traffic indicators were derived using Massachusetts Highway Department data and direct traffic counts. Multivariate regression analyses were performed separately for each pollutant, using traffic indicators, land use, meteorology, site characteristics, and central site concentrations.

Results

$PM_{2.5}$ was strongly associated with the central site monitor (R^2 = 0.68). Additional variability was explained by total roadway length within 100 m of the home, smoking or grilling near the monitor, and block-group population density (R^2 = 0.76). EC showed greater spatial variability, especially during colder months, and was predicted by roadway length within 200 m. The influence of traffic was greater under low wind speed conditions, and concentrations were lower during warmer months (R^2 = 0.52). NO_2 showed significant spatial variability, predicted by population density and roadway length within 50 m, modified by site characteristics (obstruction), and with higher concentrations during warmer months (R^2 = 0.56).

Conclusions

Each pollutant examined displayed somewhat different spatial patterns within urban neighborhoods, and were differently related to local traffic and meteorology. Our results indicate a need for separate exposure modeling to disentangle causal agents in epidemiological studies, as well as further investigation of site-specific and meteorological modification of the traffic-pollutant concentration relationship to better assess exposure residential variability across urban neighborhoods.

BACKGROUND

There is a growing body of literature linking geographic information system (GIS)-based measures of proximity to roadway to asthma and other respiratory outcomes. In the U.S. and Europe, children living or attending school near truck routes and highways show greater asthma symptoms (Brauer et al. 2002; Zmirou et al. 2004; Gordian et al. 2006), asthma hospitalizations (Edwards et al. 1994; Lin et al. 2002), respiratory illness (Brauer et al. 2002), allergic rhinitis (Duhme et al. 1996), and reduced lung function (Brunekreef et al. 1997). However, proximity measures can represent a variety of pollutants or other near-roadway exposures (i.e., noise, poverty). As there is no consensus on which traffic indicators may best capture variability in different pollutants within different settings, there is a need to distinguish the relative spatial patterns of multiple traffic-related air pollutants, and to estimate concentrations using different GIS-based traffic indicators applicable across larger epidemiological studies.

Pollutants of interest include nitrogen dioxide (NO_2), fine particulate matter ($PM_{2.5}$), and elemental carbon (EC); each has been linked to both respiratory health and vehicular emissions. One recent study distinguished their relative spatial distributions within urban settings using GIS; this study found greater intra-urban variability and stronger traffic influences for NO_2 and EC than for $PM_{2.5}$ in European cities (Hochadel et al. 2006). Comparable multi-pollutant analyses in the United States or in other settings have been limited.

Addressing multiple pollutants in large cohort studies is valuable but imposes constraints on the exposure assessment. For example, it generally limits the number of sites that can be sampled simultaneously, reinforcing the need for models with spatial and temporal components, which can calibrate spatial models over time. For outcomes like asthma etiology, models estimating long-term exposures are needed, implying that measurements need to be taken at a number of points in time and that models need to separate spatial from temporal factors to the extent possible.

Issues regarding choice of traffic indicators and spatial-temporal separation may be exacerbated within dense urban neighborhoods, as predictors shown to be significant elsewhere (i.e., population density, land use type) lack adequate variability for predicting concentrations in this setting. Moreover, unlike measurements collected in open spaces near major roadways, residential measures are affected by near-home sources and site characteristics altering the traffic-concentration relationship. Modification of the traffic-pollutant concentration relationship by meteorology and site characteristics has received little attention in land use regression models to date. Traffic data quality can be poor in residential areas, where measurements are limited and lesser variability in traffic counts is observed. Monitoring at residences can also impose additional logistical constraints related to site configuration and limiting the number of samples that can be simultaneously collected. Finally, in North America, the diesel fraction of total traffic in residential neighborhoods is generally smaller than in Europe, such that traffic measures may be less predictive of EC concentrations.

Land-use regression (LUR), a standard approach for predicting pollutant concentrations using GIS-derived spatial parameters and site characteristics, has been shown to better capture small-scale intra-urban variability than does kriging, integrated meteorological-emission (IME) models, or dispersion models (Jerrett et al. 2004). In this study, we used LUR techniques and GIS-derived variables to investigate the varying associations between multiple traffic indicators and outdoor residential concentrations of multiple air pollutants within the dense urban neighborhoods of Boston, Massachusetts. We evaluated a suite of GIS-based traffic indicators, and explored meteorology and residential site characteristics as potential modifiers of the traffic-concentration relationship.

METHODS

Site selection

This exposure modeling effort is nested within the Asthma Coalition on Community, Environment and Social Stress (ACCESS) birth cohort study. Sample homes were selected to represent variability in traffic densities across Boston neighborhoods. Candidate homes were geocoded using U.S. Census TIGRE files and City of Boston street parcel data, and initial traffic scores for each home were assigned using Massachusetts Highway Department (MHD) traffic volume data. As we anticipated first-order (Gaussian) decay of fine particles in the first 100-300 meters near major roadways (Zhu et al. 2002), we opted to create initial traffic scores for site selection by applying a kernel weighting function to total traffic counts for all road segments within 100 meters of the home. The kernel function approximates concentration gradients expected under Gaussian decay, assigning higher weights to road segments nearer to the home. Resultant traffic scores were divided into tertiles, and sampling homes were selected to represent the observed range of traffic scores and neighborhoods. Due to unbalanced cohort recruitment in the study's early stages, additional non-cohort participants were recruited to capture a wider range of traffic scores, and neighborhoods where further recruitment was anticipated were over-sampled. The spatial distribution of our final sampling cohort is shown in Figure 1, where homes are shaded by 100-meter kernel-weighted traffic score, against a surface of the same measure for each 50-meter cell across urban Boston.

Sampling methods

We measured indoor and outdoor concentrations of $PM_{2.5}$, NO_2, and EC in two seasons (summer: May through early October, winter: December through March) at 44 homes across urban Boston, though only outdoor measures are included in this analysis. $PM_{2.5}$ was measured using the Harvard Personal Environmental Monitor (PEM) (Marple et al. 1987), EC using reflectance analysis of $PM_{2.5}$ filters, and NO_2 using Yanagisawa passive filter badges (Yanagisawa and Nishimura 1982). Integrated measurements for each pollutant were collected for one week per season per home whenever feasible, in two sessions of 3 to 4 days duration. Traffic counts were collected using the Trax I Plus (JAMAR Technologies, Horsham, PA), on the highest-density road within 100 m of the home. Questionnaires were administered to identify nearby sources and sampling week activities that may influence concentrations, as detailed elsewhere (Baxter et al. 2006).

Additional data sources

Traffic Data

Road networks and traffic data were obtained from MHD. Because different aspects of traffic including density, roadway configuration, and average vehicle speed may affect emission rates, pollutant mix, and dispersal, we opted to create a suite of 25 traffic indicators (Table 1) capturing varying aspects of traffic. We built raster-based cumulative density scores for average daily traffic (ADT) counts within radii of 50 to 500 meters around each home. Because roadway segments nearer to the home may have greater influence on concentrations, we also explored inverse-distance quadratic functions (kernel-weighted buffers) for the same radii. As traffic counts on smaller residential roads were sparse, we created cumulative density scores including only larger roads (above 8,500 cars/day), summary measures of total roadway length within radii of 50 to 500 meters, and the product of roadway length and average daily traffic counts within 200 meters. We considered distance to various roadway types, including the nearest larger road (greater than 8500 cars/day), major road (13,000 cars/day), highway, and designated truck route. Lastly, to explore the influence of major roads on nearby neighborhoods, we created indicators of its average daily traffic, diesel traffic (using axle length from Trax I Plus measures as an indicator), and weighted each by the home's distance to the road.

We considered other GIS covariates that may be associated with traffic, represent other pollutant sources, or modify the observed traffic-concentration relationship. Block group-level population and area measures were used to estimate population density. NCLD-50 land use categories and elevation data were downloaded from the USGS National Land Cover Dataset (NLCD) and National Elevation Dataset (NED) (http://gisdata.usgs.gov), respectively.

Temporal variability: Background concentrations and meteorology

With a residential multi-pollutant approach, we were able to sample at a maximum of three homes per week, creating the need to account for temporal variability in background concentrations and meteorology. We estimated the influence of temporal heterogeneity in our data by regressing measured concentrations against mean central site concentrations for specific hours that each sample was collected. This temporal correction method is similar to that used elsewhere (Brauer et al., 2003), though annual averages were not calculated.

Thus, we regressed outdoor concentrations of $PM_{2.5}$, EC, and NO_2 against mean concentrations reported at a central site monitor (Massachusetts Department of Environmental Protection (DEP) monitor in the central Roxbury neighborhood) for the specific hours that each residential sample was collected. Temporally-corrected residuals were then used for selection of spatial covariates.

Meteorological data were collected from the same central site. Mean windspeed and direction were calculated for daytime hours (6am-9pm) within each sampling period, when we anticipate significant traffic on roadways, our main source of interest. It is noteworthy that several wind parameters were created in relation to traffic sources (i.e., percent of sampling hours during which the home is downwind from the nearest road), such that significance of the wind term implies significance of the source. Lastly, although meteorological texts generally define 'still winds' as below 1 m/s (Hanna et al. 1982), we used 2.0 m/s to better dichotomize our high-windspeed dataset (median = 4.9 m/s). Meteorological factors and other covariates considered as effect modifiers of the traffic-pollution relationship are summarized in Table 2.

Analytic methods and model-building

We built models separately by pollutant, allowing different aspects of traffic, meteorology, and site-specific factors to predict concentrations of different pollutants. We selected candidate traffic indicators and modifiers against the temporally-corrected residuals, using nonparametric univariate correlations (Spearman correlations, $p < 0.3$) of concentrations against traffic indicators as our primary selection method. Because traffic indicators are highly correlated, however, we considered cluster analysis as a secondary selection method; the tree command in R groups observed concentrations by applying an impurity criterion to minimize within-group variances while maximizing between-group differences. The command compared concentration groups created using the 25 examined traffic indicators as predictors, and returned the indicators which best distinguished, as a group, high and low pollution locations, and the most effective binary cut-point for each indicator. Multivariate models were built using those traffic indicators selected by both correlation and clustering methods.

Using a stepwise forward process, we first included central site data, then traffic indicators, meteorological and site-specific modifiers as interaction terms with traffic indicators. Finally, we added indicators of other site-specific sources (e.g., grilling or smoking noted near outdoor monitor, block group population density, land use type, proximity to industry). We note that several of these indicators may be associated with traffic, capturing some traffic effect. We used the general form of Equation 1, and a maximum p-value of 0.1 to retain variables at each stage.

Equation 1: $Concentration_{ijt} = ß_{0ijt} + ß_{1jt}*DEP_{jt} + ß_{2ji}*Traffic_{i} + ß_{3ijt}*Traffic_{i}*Modifier_{it} + ß_{4ijt}*Other\ sources_{it} + e_{ijt}$

Where $Concentration_{ijt}$ is the measured concentration of pollutant *j* at location (home) *i* during sampling period (time) *t*. DEP_{j} is the mean concentration of pollutant *j* at the central site during sampling period *t*. $Traffic_{i}$ is the value of each traffic indicator listed in Table 1, tested separately in prediction models, at location *i*. $Modifier_{ijt}$ is the value of meteorological or site characteristics altering the association between traffic indicators and $Concentration_{ijt}$. $PM_{2.5}$, EC, and DEP_{ij} values for $PM_{2.5}$ and EC were log-distributed, and thus transformed prior to covariate selection and model building. NO_2 values were normally distributed, and not transformed.

Finally, we anticipated that residential EC may display a different relationship with central site EC by season, and allowed for season-specific slopes in the model. Key sources (i.e., traffic, wood smoke, home heating fuel) may increase spatial variability during winter, when lower atmospheric mixing height may increase their influence. EC from different sources may also be detected differently by the reflectance-based method we used and the optical method (aethelometer) at the central site.

Sensitivity Analyses

Extensive sensitivity tests were performed on the final model for each pollutant. Models were examined for sensitivity to the selection of traffic indicator by individually substituting each traffic indicator from Table 1. Likewise, we examined the selection of meteorological and site-specific modifiers by individually substituting other candidates. In each case, the final model was retained based upon overall model fit (R^2).

The resolution of raster-based traffic indicators was examined by considering a range of base cell sizes (the smallest spatial unit employed in variable creation), varying in width from 10 to 50 meters. We also examined the quality of MHD traffic data by creating and comparing with comparable traffic indicators built using three other data sources. We tested the effect of log-transformation of $PM_{2.5}$ and EC data, and evaluated the use of the central site monitor for temporal variability by substituting data from other DEP monitors for the same period. To assess residual seasonality not captured by the central site monitor, we examined the influence of categorical season variables and seasonally-varying slopes on the central site term. Finally, we examined the robustness of each model to within-site autocorrelation owing to multiple measures at each site, using random effects by household.

All traffic and land use variables were created in ArcGIS 9, clustering analyses were performed using the tree command in R version 2.2.0, and model-building in SAS version 9.1.

RESULTS

We conducted 66 sampling sessions in total, consisting of 86 three-to-four day measurements in 44 homes. Fifty-one measurements were taken in 36 homes during summer months, and 35 measurements were taken in 25 homes during winter months. Table 3 summarizes the within-season average concentrations by sampling session for each pollutant. EC values are indicated by filter absorbance (units $m^{-1}*10^{-5}$), and converted to mass measures using a conversion factors used in prior urban northeastern U.S. studies of 0.83 $\mu g/m^3$ per reflectance unit (Kinney et al. 2000). $PM_{2.5}$ and EC were significantly correlated during winter and summer ($p < 0.05$), while EC and NO_2 were marginally correlated in both seasons, and $PM_{2.5}$ and NO_2 were not.

Pollutant-specific modeling results

Outdoor $PM_{2.5}$ was highly correlated with central-site $PM_{2.5}$ ($R^2 = 0.68$), as suggested in Figure 2a, indicating a predominance of temporal variability and relative spatial homogeneity in $PM_{2.5}$ across the urban area. In multivariate regressions including central site data, the best traffic indicator was total roadway length within 100 meters of the home (Table 4). Final multivariate model results indicate that the traffic-$PM_{2.5}$ relationship was not significantly altered by any of our candidate modifiers. Other combustion sources (smoking or grilling) and population density significantly contributed to concentrations (overall $R^2 = 0.76$).

EC shows relatively poor associations with central site data overall ($R^2 = 0.08$), though this is partly attributable to seasonal differences in the relationship (Figure 2b), with varying slopes and stronger correlations during summer (Spearman r = 0.66) than winter (r = 0.37). In the final multivariate model ($R^2 = 0.52$), EC was best predicted by total roadway length within 200 meters, and the association between EC and traffic was increased under low wind speed conditions. During summer months, residential EC concentrations were somewhat lower and displayed stronger associations with central site data. Approximately 30% of the variability in EC was explained by temporal terms, and 14% by the traffic term (spatial component). The interaction of traffic with hours of low wind speed, incorporating both spatial and temporal variance, accounted for an additional 8%.

NO_2 was weakly associated with central site concentrations (R^2 = 0.21), suggesting significant spatial heterogeneity within urban residential areas (Figure 2c). The final multivariate model (R^2 = 0.56) includes total roadway length within 50 meters of the home, significantly attenuated by an obstruction (i.e., building) between the monitor and nearest major road. NO_2 concentrations were higher during summer months, and positively associated with population density (Table 4). Spatial terms (traffic, obstruction between the monitor and nearest major road, population density) together account for approximately 23% of NO_2 variability. Temporal terms (central site, warmer season) account for about 34%.

Sensitivity analyses

Selection of traffic indicator

Sensitivity analyses indicate that other traffic indicators could not be substituted to create a comparable model for $PM_{2.5}$ (Table 5) For EC (Table 6), diesel-based measures can explain more variability, with R^2 values of approximately 0.54, but were available for only a subset of locations (n = 35). For the full cohort, no indicator was exchangeable with roadway length within 200 meters. In addition, the interaction term of traffic modified by low wind speeds remained significant in several cases where the main effect of traffic did not maintain significance. For NO_2, sensitivity tests (Table 7) support the finding that shorter buffer lengths were most effective. Larger buffer lengths did not produce a comparable model, but kernel-weighted traffic density within 50 meters of the home could be substituted effectively, as could unweighted cumulative density within 100 meters.

Accuracy of traffic indicators

To validate raster-based traffic indicators, we considered a range of base cell sizes from 10 to 50 meters square, bearing little difference on traffic indicator values compared to our default 25 meter cell size. Given concerns about data quality, where possible we verified MHD counts against traffic data obtained from the Massachusetts Executive Office of Transportation, ESRI Business Analyst, and our traffic counts collected outside cohort homes using the Jamar Trax I device. Correlations across traffic sources were generally above 0.7.

Use of the central site monitor for temporal variability

We considered several alternatives to the use of central site monitor concentrations for temporal correction, including average concentrations from all urban monitors available during each sampling period, and the mean concentration at a background monitor south of Boston (available for summer months only). No alternative to the central site sampling period mean explained greater variability in concentrations or significantly altered traffic-pollution relationships in multivariate models.

Selection of meteorological and site-specific modifiers for EC, NO_2

All EC models showed a significant, positive effect of low wind speeds on the traffic-pollution relationship. Sensitivity analyses indicated that other wind variables (mean daytime windspeed, percent of day downwind from road) were significant and may be substituted for percent of low wind speed hours, losing only marginal explanatory power ($R^2 = 0.52$ and 0.49, respectively).

In sensitivity analyses for NO_2, no other modifier could replace obstruction between the monitor and nearest major road with significance in the final model. Because presence of an obstruction could theoretically proxy for distance to nearest major road, we replaced the obstruction term with a continuous distance to major road term, and found highly non-significant results, indicating that this was not likely the case.

Log-transformation of $PM_{2.5}$ and EC data

The selection of the 100-meter roadway length term and other predictors for the $PM_{2.5}$ model was not dependent on log transformation. Using un-transformed $PM_{2.5}$, we achieve an R^2 of 0.73, and retain significance in all predictors. For un-transformed EC, the same traffic term and all other predictors retained significance, with an R^2 of 0.51.

Inclusion of a categorical variable for season

Because the season term may be extraneous in models including temporal data from a central site, we explored the effect of removing this term from the final EC and NO_2 models. For EC, removing the season term caused the central site monitor estimate to drop by half and fall out of significance, while the effect of low wind speed increased by almost 50%, and overall model fit declined. Because of this decline in overall explanatory power when removed, we opted to maintain the season term and season-specific slopes on the central site monitor term in the final model.

For NO_2, dropping the season term decreased the effect of the central site monitor by approximately 50%, but did not affect overall model fit or other parameters. Thus, although NO_2 is higher during summer months, the effect is largely captured by the central site monitor; because the term did not significantly alter other parameters, we opted to leave it in the final model. Finally, we tested the addition of a season term to the $PM_{2.5}$ model, found no effect on the central site term or overall fit, and opted to leave it out. The effect of other combustion sources (i.e., smoking or grilling near outdoor monitor) would be increased by approximately 35%, however, potentially indicating seasonal differences in these sources.

Robustness to within-site autocorrelation

For all pollutants, because the majority of homes were monitored in two seasons, we examined the effect of within-site autocorrelation using random effects by household. Autocorrelation by site was not found to influence either model fit or parameter estimates for any of the three final models.

Finally, a one-at-a-time exclusion cross-validation was performed to assess the internal consistency of model results. The Pearson correlation between estimated and measured log $PM_{2.5}$ was 0.84, 0.63 for log EC, and 0.66 for NO_2 ($p < .0001$ in all cases). Each correlation is in keeping with model R^2's, and indicates acceptable internal validity.

DISCUSSION

Working strictly within dense urban neighborhoods and employing a multi-pollutant approach, our study offers several findings useful to future research exploring and modeling air pollution exposures for epidemiological purposes. These observations broadly apply to four areas: (1) urban residential variability in traffic densities and pollution concentrations, (2) variability in urban residential pollution that is attributable to traffic, (3) selection of traffic indicators for residential exposure estimation, and (4) modification of traffic-concentration relationships by site characteristics and meteorology.

(1) Urban residential variability in traffic densities and pollution concentrations

We found significantly greater variability and stronger relationships with local traffic for EC and NO_2 than for $PM_{2.5}$, consistent with prior literature (Hochadel et al, 2006), and which corroborates evidence that $PM_{2.5}$ patterns are largely regional in nature for the Eastern U.S. (Burton et al. 1996; Suh et al. 1997). We found somewhat weaker correlations across the three pollutants than have been shown in prior European studies (Brauer et al. 2002; Hochadel et al, 2006), potentially because of our focus within dense urban neighborhoods, while most prior intra-urban studies have actually examined metropolitan regions.

Though we found significant variability in traffic across cohort homes, traffic varied somewhat less within residential neighborhoods than across the entire urban core, as shown in Figure 1. Across the entire urban core, 100-meter kernel-weighted traffic scores ranged from 0 to 3,305 vehicle miles traveled (VMT) per m^2 per day; at cohort homes, this measure ranged from 5.8 to 168 VMT/ m^2-day. This difference is driven largely by major highways, alongside which relatively few homes are located, but this observation may be important for exposure estimation; many models are derived from near-roadway concentration data, which may inaccurately reflect traffic-concentration associations at the lower end.

(2) Variability in urban residential pollution that is attributable to traffic

Using the traffic-pollution relationship observed across our sampling sites, we can estimate the portion of residential concentrations that are attributable to traffic. For $PM_{2.5}$, the mean 100-meter summary roadway length of 1,110 meters accounted for a marginal contribution of 1.2 $\mu g/m^3$, or 9.7% of predicted $PM_{2.5}$. Applying our predictive models with mean values for all terms, mean predicted concentration is 13.2 $\mu g/m^3$; a increase of one standard deviation in roadway length (371 meters) increases concentrations to 13.9 $\mu g/m^3$. Using non-transformed $PM_{2.5}$, the predicted traffic contribution is somewhat larger (2.6 $\mu g/m^3$). Overall, traffic explains relatively little variability in

$PM_{2.5}$, though population density, likely capturing some traffic-related influence, adds 1.1 μg/m^3 on average.

For EC, the mean contribution of modeled traffic terms is approximately 0.17 μg/m^3, using a mean 200-meter buffer roadway length of 3,560 meters, accounting for 36% of measured EC. A mean roadway length of 3,560 meters, predicts 0.47 μg/m^3, and increasing roadway length by one standard deviation (1,156 meters) increases predictions to 0.54 μg/m^3. Using non-transformed EC concentrations, predicted traffic contributions are somewhat larger (0.39 μg/m^3).

Modeled traffic terms accounted for approximately 2.8 ppb, or 21% of modeled NO_2. A mean 50-meter buffer roadway length of 441 meters, with mean values for other terms, predicts a concentration of 17.8 ppb. A one standard deviation increase (179 meters) increases concentrations to 18.9 ppb. The range of 50-meter roadway lengths observed predicts a NO_2 range of 15.4 to 20.9 ppb. Population density, again, likely captures some traffic effect, and accounts for 4.4 ppb on average.

(3) Selection of traffic indicators for residential exposure estimation

Total roadway length within varying buffer radii provided useful traffic indicators for each pollutant. Because actual traffic counts for residential roads are generally sparse, length measures can provide more stable traffic indicators in these areas than do ADT-based estimates. There is also differential bias in traffic count accuracy by roadway size in the traffic data available for many cities; actual traffic counts are generally collected on a regular basis for highways and major roads, and rough estimates are created for smaller residential roads.

While correlated, total roadway lengths within various buffers were not interchangeable in predictive models. The 100-meter buffer length effective for $PM_{2.5}$ coincides with our original buffer radius created for site selection. This original buffer was selected because maximum declines in particle concentrations occur in the first 100-300 meters alongside major roadways, dependent on particle size distributions and wind characteristics (Zhu et al. 2002). A slightly larger buffer length of 200 m was effective for EC, as expected, as EC generally lies in the smaller range of $PM_{2.5}$ and thus may deposit further from the source under similar conditions. For NO_2, shorter buffer lengths of 50 meters were most effective. The precise distance-concentration decay relationship for NO_2 varies by setting, as dispersion, dilution, and chemical transformation are affected by the local pollution mix, wind characteristics, and meteorology. Secondary formation of NO_2 from NO might suggest longer buffer lengths, with one LUR study showing 300 m buffers to be effective (Jerrett et al. 2003). Other data, however, suggests that NOx concentrations can decline up to 60 percent, and NO_2 by 30 percent, within the first 50 meters from a roadway (Kanaroglou et al. 2005).

Beyond total roadway length, sensitivity tests suggest that diesel-based indicators might provide stronger predictors for EC and possibly for NO_2. Diesel influence raises a critical difference between traffic-based exposure modeling in North America and Europe, as diesel vehicles constitute a relatively small and poorly measured fraction of North American traffic, especially in residential areas. In this study, we rely on imperfect axle-length estimates from our traffic counter on only one major road near the home, to approximate vehicle type (i.e., CNG-powered buses, for example, would be categorized as "diesel" under our approach). In Europe, the higher prevalence of diesel passenger vehicles implies that total traffic may itself act as a diesel marker, providing more stable predictors for EC.

(4) Modification of traffic-concentration relationships by site characteristics and meteorology

We found that accurate exposure modeling near urban residences required some consideration of site characteristics, such as population density and obstructions, which explained significant variability in concentrations and altered traffic-concentration associations. Population density, significant for $PM_{2.5}$ and NO_2, may proxy either for other residential sources or for traffic, and may indicate higher per-mile emission rates from 'stop-and-go' traffic in denser residential neighborhoods. Our finding of significant obstruction effects on NO_2 supports recent evidence that residential NO_2 concentrations differed significantly depending on whether the home faced onto the courtyard or street, after accounting for distance to road (Reungoat et al. 2005).

Limitations

Although our models are physically interpretable and explain significant variability, there are some limitations to our analysis. Our small sample size may be a limitation in model-building, though sensitivity analyses show our findings to be robust. Similarly, our multi-pollutant approach limited our sampling design to sample at only a few homes simultaneously and distributing sampling sessions over the course of a season, such that samples incorporate both temporal and spatial variability. A temporally-staggered sampling design, however, is necessary for long-term residential exposure estimation of multiple pollutants; simultaneously-collected data lacks within-season temporal variability, and temporal covariates are needed for long-term exposure estimation models. The central site monitor used to account for temporal heterogeneity may capture some local-source component as well, and thus we opted to maintain the central site data as one temporal predictor in our models, rather than predicting temporally-corrected concentration residuals, and included other temporal terms such as season and meteorological parameters.

Ultimately, most GIS-based residential exposure models are intended to allow exposure estimation across large cohorts, and thus we rely on readily-created GIS-based traffic indicators generally available across urban neighborhoods, such as total roadway length measures. Several predictors in our models, however, such as obstruction between the home and nearest major road, are effectively correction factors for the restrictions associated with residential monitoring – i.e. samplers often need be set up behind the buildings, wherever power sources are available, on a porch where smoking or grilling also occurred, or some such non-ideal location. These parameters may not be appropriate for extrapolation, as they may not reflect mean concentrations near the home, but are important to correctly interpreting residential data.

5. CONCLUSIONS

Our analysis explored a range of GIS-based traffic indicators to capture small-scale variability in multiple air pollutants within dense urban neighborhoods. Because our measures were collected outside residential homes rather than at roadside locations, our measures likely reflect more realistic exposures to urban residents. The resultant sampling design, however, raised a number of methodological challenges, including the need to account for spatial and temporal variability in measures collected during different weeks, and the need to account for home site characteristics and other residential sources which may obscure the true traffic-concentration relationship.

We found that traffic indicators not reliant on ADT estimates (i.e. roadway lengths) provided more stable predictors in residential settings. As shown elsewhere, greater spatial variability was observed in NO_2 and EC than in $PM_{2.5}$, and LUR techniques worked well within the urban setting to capture pollutant variability, although parameters useful at the metropolitan scale (e.g., land use type) displayed low variability and limited predictive power. Finally, the relationship between traffic and pollution concentrations was significantly modified by meteorological factors and site characteristics, indicating the importance of incorporating small-scale spatial and temporal predictors to accurately capture exposure variability in urban residential settings.

List of Abbreviations

ACCESS – Asthma Coalition on Community, Environment, and Social Stress study

ADT – average daily traffic

CNG – compressed natural gas

DEP – Massachusetts Department of Environmental Protection

EC – elemental carbon

GIS – geographic information systems

IME – integrated meteorological-emission models

LUR – land use regression

MHD – Massachusetts Highway Department

NED – National Elevation Dataset

NLCD – National Land Cover Dataset

NO_2 – nitrogen dioxide

PEM – Harvard Personal Environmental Monitor

$PM_{2.5}$ – fine particulate matter, smaller than 2.5 microns mean diameter

USGS – United States Geological Survey

VMT – vehicle miles travelled

Acknowledgements: JEC gratefully acknowledges funding from the American Lung Association, which supported this work. In addition, the ACCESS project is supported by NIH R01 ES10932, U01 HL072494, and R01 HL080674 (RJ Wright, PI), and the analyses were supported by NIH R03 ES013988 and by the Health Effects Institute (HEI 4727-RFA04-5/05-1), J Levy, PI. Many thanks to our participating families, and to Brent Coull, Chris Paciorek, and Jack Spengler for valuable suggestions and comments on the manuscript.

Declaration of competing interests: The authors declare that they have no competing interests.

Authors' Contributions: JEC participated in exposure study design, sampling design, and data collection, created all spatial and meteorological variables, performed statistical analyses, and drafted the manuscript. RJW oversaw cohort recruitment and overall ACCESS study design. LKB participated in sampling design, data collection, and analysis. JIL oversaw sampling design, data collection, data analysis, and manuscript development. All authors read and approved the final manuscript.

REFERENCES

Baxter, L.K., Clougherty, J.E., Laden, F., Levy, J.I, 2006. Predictors of concentrations of nitrogen dioxide, fine particulate matter, and particle constituents inside of lower socioeconomic status urban homes. Journal of Exposure Science and Environmental Epidemiology doi: 10.1038/sj.jes.7500532.

Brauer, M., Hoek, G., Van Vliet, P., Meliefste, K., Fischer, P., Gehring U., et al. 2003. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and Geographic Information Systems. Epidemiology. 14:228-239.

Brauer, M., Hoek, G., Van Vliet, P., Meliefste, K., Fischer, P., Wijga, A., Koopman, L., Neijiens, H., Gerritsen, J., Kerkhof, M., Heinrich, J., Bellander, T., Brunekreef, B., 2002. Air Pollution from Traffic and the Development of Respiratory Infections and Asthmatic and Allergic Symptoms in Children. American Journal of Respiratory and Critical Care Medicine 166, 1092-98.

Brunekreef, B., Janssen, N.A., Hartog, J., Harssema, H., Knape, M., van Vliet, P., 1997. Air pollution from truck traffic and lung function in children living near motorways. Epidemiology 8 (3), 298-303.

Burton, R.M., Suh, H.H., Koutrakis, P., 1996. Spatial Variation in Particulate Concentrations within Metropolitan Philadelphia. Environmental Science and Technology 30 (2), 400-407.

Chauhan, A., Inskip, H., Linaker, C., Smith, S., Schrieber, J., Johnston, S., Holgate, S., 2003. Personal exposure to nitrogen dioxide (NO_2) and the severity of virus-induced asthma in children. Lancet 361, 1939-44.

de Hoogh, K., Gulliver, J., Briggs, D., Mitchell, J., 2002. A comparison of techniques for modelling exposure to traffic-related air pollution in contrasting urban environments. 12th Conference of the International Society of Exposure Analysis (ISEA), Vancouver BC Canada.

Delfino, R., 2004. Association of FEV-1 in asthmatic children with personal and microenvironmental exposure to airborne particulate matter. Environmental Health Perspectives 112, 932-41.

Diaz-Sanchez, D., Penichet-Garcia, M., Saxon, A., 2000. Diesel exhaust particles directly induce activated mast cells to degranulate and increase histamine levels and symptom severity. Journal of Allergy and Clinical Immunology 106, 1140-46.

Duhme, H., Weiland, S.K., Keil, U., Kraemer, B., Schmid, M., Stender, M., Chambless, L., 1996. The association between self-reported symptoms of asthma and allergic rhinitis and self-reported traffic density on street of residence in adolescents. Epidemiology 7(6), 578-82.

Edwards, J., Walters, S., Griffiths, R.K., 1994. Hospital admissions for asthma in preschool children: relationship to major roads in Birmingham, United Kingdom. Archives of Environmental Health 49 (4), 223-7.

English, P., Neutra, R., Scalf, R., Sullivan, M., Waller, L., Zhu, L., 1999. Examining associations between childhood asthma and traffic flow using a geographic information system. Environmental Health Perspectives 107, 761-67.

Gauderman, W., McConnell, R., Gilliand, F., London, S., Thomas, D., Avol, E., Vora, H., Berhane, K., Rappaport, E.B., Lurmann, F., Margolis, H.G., Peters, J., 2002. Association between air pollution and lung function growth in Southern California children: results from a second cohort. American Journal of Respiratory and Critical Care Medicine 166 (4), 76-84.

Gordian, M.E., Haneuse, S., Wakefield, J., 2006. An investigation of the association between traffic exposure and the diagnosis of asthma in children. Journal of Exposure Science and Environmental Epidemiology 16, 49-55.

Hanna, S., Briggs, G., Hosker Jr., R., 1982. Handbook on Atmospheric Diffusion, Technical Information Center, U.S. Department of Energy.

Hasselblad, V., Eddy, D., Kotchmar, D., 1992. Synthesis of environmental evidence: Nitrogen dioxide epidemiology studies. Journal of the Air Waste Management Association 42 (5), 662-71.

Hochadel, M., Heinrich, J., Gehring, U., Morgenstern, V., Kuhlbusch, T., Link, E., Wichmann, H.-E., Kramer, U. 2006. Predicting long-term average concentrations of traffic-related air pollutants using GIS-based information. Atmospheric Environment 40, 542-53.
Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., van den Brandt, P., 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. Lancet 360, 1203-9.
Hussain, I., Jain, V.V., O'Shaughnessy, P., Businga, T.R., Kline, J., 2004. Effect of nitrogen dioxide exposure on allergic asthma in a murine model. Chest 126, 198.
Jenkins, H., Devalia, J., Mister, R., Bevan, A., Rusnak, C., Davies, R., 1999. The effect of exposure to ozone and nitrogen dioxide on the airway response of atopic asthmatics to inhaled allergen: Dose- and time-dependent effects. American Journal of Respiratory and Critical Care Medicine 160, 33-9.
Jerrett, M., Arain, A., Kanaraglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2004. A review and evaluation of intraurban air pollution models. Journal of Exposure Analysis and Environmental Epidemiology 15 (2), 185-204.
Jerrett, M., Arain, M.A., Kanaraglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., Finkelstein, N., Finkelstein, M.M., 2003. Modelling the intra-urban variability of ambient traffic pollution in Toronto, Canada. Strategies for Clean Air and Health, Rome, NERAM: Network for Environmental Risk Assessment and Management.
Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Cilbert, N.L., Brook, J.R., 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment. Atmospheric Environment 39, 2399-409.
Kinney, P.L., Aggarwal, M., Northridge, M.E., Janssen, N.A., Shepard, P., 2000. Airborne concentrations of PM2.5 and Diesel Exhaust Particles on Harlem Sidewalks: A Community-Based Pilot Study. Environmental Health Perspectives 108, 213-8.
Lin, S., Munsie, P., Hwang, S.-A., Fitzgerald, E., Cayo, M.R., 2002. Childhood Asthma Hospitalization and Residential Exposure to State Route Traffic. Environmental Research 88 (A), 73-81.
Marple, V., Rukow, K., Turner, W., Spengler, J., 1987. Low flow rate sharp cut impactors for indoor air sampling: Design and calibration. Journal of the Air Pollution Control Association 37, 1303-7.
Martuzevicius, D., Grinshpun, S.A., Reponen, T., Gorny, R.L., Shukla, R., Lockey, J., Hu, S., McDonald, R., Biswas, P., Kliucininkas, L., LeMasters, G., 2004. Spatial and temporal variations of PM2.5 concentration and composition throughout an urban area with high freeway density - the Greater Cincinnati study. Atmospheric Environment 38, 1091-1105.
Mastrangelo, G., Clonfero, E., Pavanello, S., Fedeli, U., Fadda, E., Turato, A., Piccinni, S., Montagnani, R., Marcer, G., 2003. Exposure to diesel exhaust enhances total IgE in non-atopic dockers. Archives of Occupational and Environmental Health 76: 63-8.
Pandya, R., Solomon, G., Kinner, A., Balmes, J., 2002. Diesel exhaust and asthma: hypotheses and molecular mechanisms of action. Environmental Health Perspectives 110 (Suppl 1), 103-12.
Parnia, S., Frew, A., 2002. Role of pollutants in allergic sensitization and the development of asthma. Annals of Allergy, Asthma and Immunology 57, 1111.
Pope III, C., Thun, M., Namboordiri, M., Dockery, D., Evans, J., Speizer, F., 1995. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. American Journal of Respiratory and Critical Care Medicine 151 (3), 669-74.
Reungoat, P., Chiron, M., Gauvin, S., Zmirou-Navier, D., Momas, I., 2005. Retrospective assessment of exposure to traffic air pollution using the ExTra index in the VESTA French epidemiological study. Journal of Exposure Analysis and Environmental Epidemiology 15, 524-33.

Rusznak, C., Devalia, J., Davies, R., 1996. Airway response of asthmatic subjects to inhaled allergen after exposure to pollutants. Thorax 51 (11), 1105-8.
Sahsuvaroglu, T., Jerrett, M., 2004. Modeling Intra-urban Variability of Traffic Pollution in Hamilton, Canada. International Soceity for Exposure Analysis (ISEA), Philadelphia, PA.
Schauer, J.J., 2003. Evaluation of elemental carbon as a marker for diesel particulate matter. Journal of Exposure Analysis and Environmental Epidemiology 13, 443-453.
Suh, H.H., Nishioka, Y., Allen, G.A., Koutrakis, P., Burton, R.M., 1997. The Metropolitan Acid Aerosol Characterization Study: Results from the Summer 1994 Washington D.C. Field Study. Environmental Health Perspectives 105, 826-34.
Van Strien, R., Gent, J.F., Belanger, K., Triche, E., Bracken, M.B., Leaderer, B.P., 2004. Personal exposure to nitrogen dioxide (NO_2) and nitrous acid: respiratory symptoms in the first year. Epidemiology 15, 471-8.
Venn, A.J., Lewis, S.A., Cooper, M., Hubbard, R., Britton, J., 2001. Living Near A Main Road and the Risk of Wheezing Illness in Children. American Journal of Respiratory and Critical Care Medicine 164, 2177-2180.
Wu, J., Lurmann, F., Winer, A., Lu, R., Turco, R., Funk, T., 2005. Development of an individual exposure model for application to the Southern California Children's Health Study. Atmospheric Environment 39, 259-273.
Yanagisawa, Y., Nishimura, H., 1982. A badge-type personal sampler for measurement of personal exposure to NO_2 and NO in ambient air. Environment International 8: 235-42.
Zhu, Y., Hinds, W.C., Kim, S., Sioutas, C., 2002. Concentration and Size Distribution of Ultrafine Particles Near a Major Highway. Journal of the Air and Waste Management Association 52, 1032-41.
Zmirou, D., Gauvin, S., Pin, I., Sahraoui, F., Just, J., Moullec, Y.L., Bremont, F., Cassadou, S., Reungoat, P., Albertini, M., Lauvergne, N., Chiron, M., Labbe, A., Investigators, V., 2004. Traffic related air pollution and incidence of childhood asthma: results of the Vesta case-control study. Journal of Environment and Community Health 58, 18-23.

FIGURE LEGENDS

Figure 1: 100-meter kernel-weighted traffic scores for Boston urban area and sampling homes (Vehicle-miles per day/km^2). Darker shading indicates higher traffic density.

Figure 2: Scatter plots of $PM_{2.5}$, EC, and NO_2 measured outside of homes, vs. average concentrations during the sampling hours at the central site DEP monitor.

> **Figure 2a:** $PM_{2.5}$ at homes vs. central site ($\mu g/m^3$)
>
> **Figure 2b:** EC at sampled homes ($m^{-1}*10^{-5}$) vs. central site ($\mu g/m^3$), after removing one influential point per season
>
> **Figure 2c:** NO_2 at homes vs. central site (ppb)

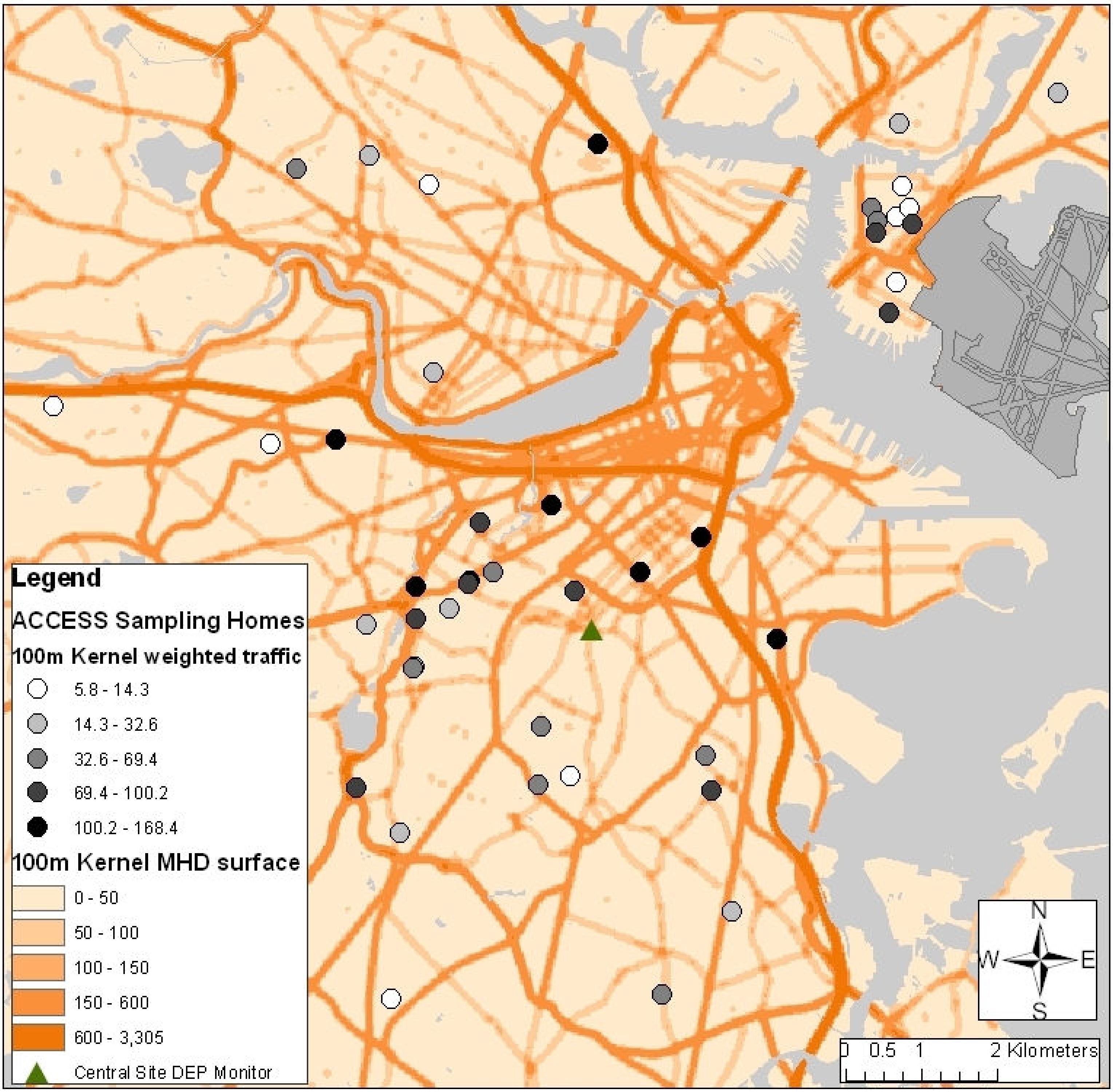
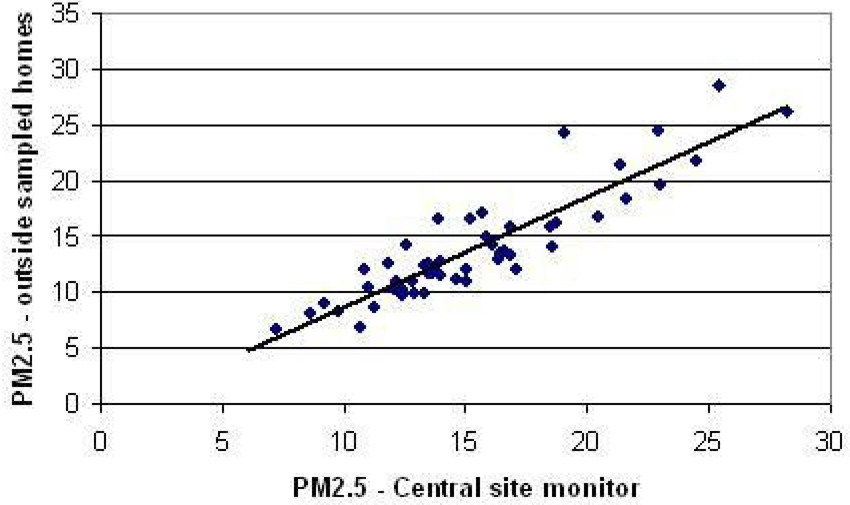
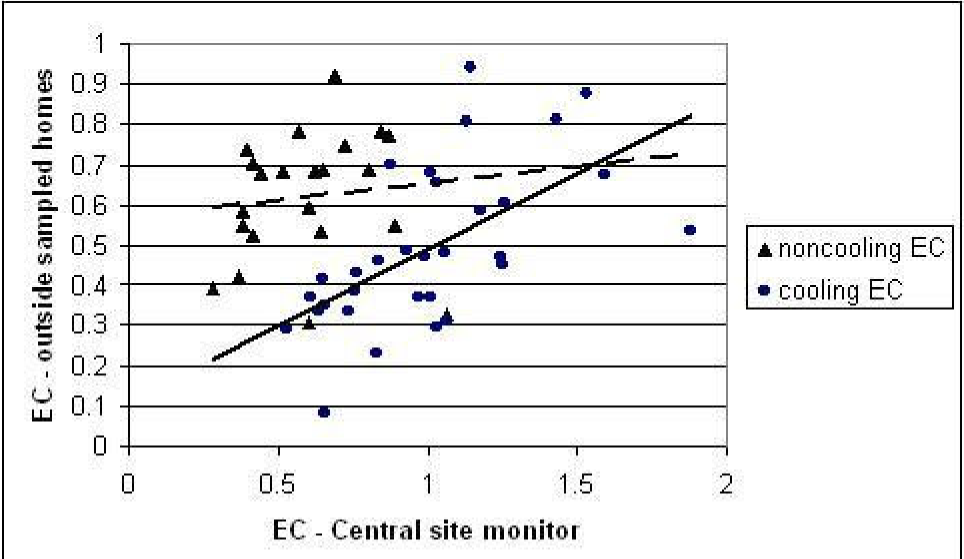

Legend
ACCESS Sampling Homes
100m Kernel weighted traffic
5.8 - 14.3
14.3 - 32.6
32.6 - 69.4
69.4 - 100.2
100.2 - 168.4
100m Kernel MHD surface
0 - 50
50 - 100
100 - 150
150 - 600
600 - 3,305
Central Site DEP Monitor
N
W
E
S
0 0.5 1 2 Kilometers

Figure 1



PM2.5 - outside sampled homes
PM2.5 - Central site monitor
0
5
10
15
20
25
30
35

Figure 2



EC - outside sampled homes
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
EC - Central site monitor
noncooling EC
cooling EC

Figure 3

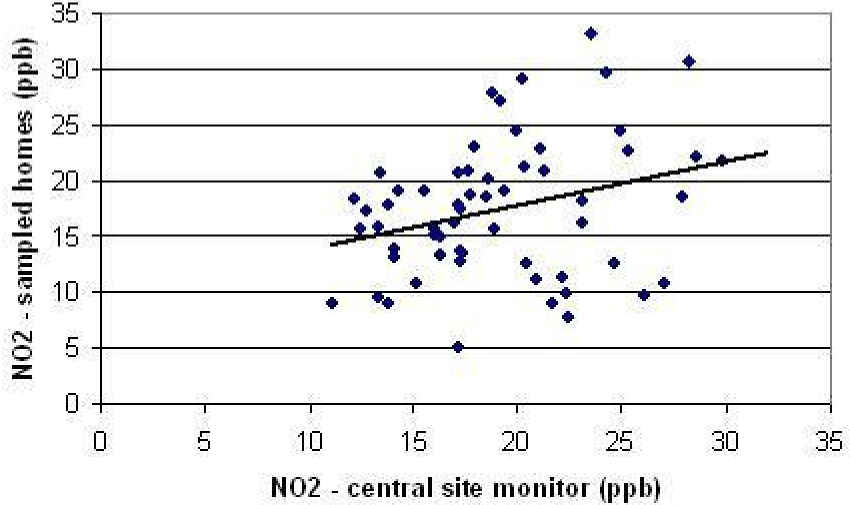

35
30
25
20
15
10
5
0
NO2 - sampled homes (ppb)
0
5
10
15
20
25
30
35
NO2 - central site monitor (ppb)

Figure 4

Additional files provided with this submission:

Additional file 1: axs table 1.pdf, 19K
http://www.ehjournal.net/imedia/5256679161522375/supp1.pdf
Additional file 2: axs table 2.pdf, 14K
http://www.ehjournal.net/imedia/2487890491522375/supp2.pdf
Additional file 3: axs table 3.pdf, 21K
http://www.ehjournal.net/imedia/8063383401522375/supp3.pdf
Additional file 4: axs table 4.pdf, 93K
http://www.ehjournal.net/imedia/1515532205152237/supp4.pdf
Additional file 5: axs table 5.pdf, 20K
http://www.ehjournal.net/imedia/1071522359152237/supp5.pdf
Additional file 6: axs table 6.pdf, 24K
http://www.ehjournal.net/imedia/2042892008152237/supp6.pdf
Additional file 7: axs table 7.pdf, 24K
http://www.ehjournal.net/imedia/1258416631152237/supp7.pdf