

**A Comparison of Exposure Metrics for Traffic-Related Air Pollutants:
Application to Epidemiology Studies in Detroit, Michigan**

Stuart Batterman^{1*}, Janet Burke², Vlad Isakov², Toby Lewis³, Bhramar Mukherjee⁴, Thomas Robins¹

¹University of Michigan, Department of Environmental Health Sciences, School of Public Health, Room 6075
SPH2, 1420 Washington Heights, Ann Arbor, MI 48109-2029 USA
tel: 734 763 2417 fax: 734 763-8095 email: stuartb@umich.edu

²US Environmental Protection Agency, National Exposure Research Laboratory, 109 T.W. Alexander Drive,
Research Triangle Park, NC 27711 USA

³University of Michigan, Medical School, Pediatric Pulmonary, 1500 East Medical Center Drive, Room L2221,
Ann Arbor, MI 48109-5212, USA

⁴University of Michigan, Department of Biostatistics, School of Public Health, 1420 Washington Heights, Ann
Arbor, MI 48109-2029 USA

*Corresponding author
4/29/14

Abstract

Vehicles are major sources of air pollutant emissions, and individuals living near large roads endure high exposures and health risks associated with traffic-related air pollutants. Air pollution epidemiology, health risk, environmental justice, and transportation planning studies would all benefit from an improved understanding of the key information and metrics needed to assess exposures, as well as the strengths and limitations of the various alternatives. Several metrics for characterizing exposure to traffic-related air pollutants were developed for the 218 residential locations of participants in the NEXUS epidemiology study conducted in Detroit, Michigan, U.S. Exposure metrics included proximity to major roads, traffic volume, vehicle mix, traffic density, vehicle exhaust emissions density, and pollutant concentrations predicted by dispersion models. Results presented for each metric include comparisons of exposure distributions, spatial variability, intraclass correlation, concordance and discordance rates, and overall strengths and limitations. While showing some agreement, the simple categorical and proximity classifications (e.g., high diesel/low diesel traffic roads and distance from major roads) do not reflect the range and overlap of exposures seen in the other metrics. Information provided by the traffic density metric, defined as the number of kilometers traveled (VKT) per day within a 300 m buffer around each home, was reasonably consistent with the more sophisticated metrics. Dispersion modeling provided spatially- and temporally-resolved concentrations, along with apportionments that separated concentrations due to traffic emissions and other sources. While several of the exposure metrics showed broad agreement, including traffic density, emissions density and modeled concentrations, these alternatives still produced exposure classifications that differed for a substantial fraction of study participants, suggesting the potential for exposure misclassification and the need for refined and validated exposure metrics. While data and computational demands for dispersion modeling of traffic emissions are non-trivial concerns, once established, dispersion modeling systems can provide exposure information for both on- and near-road environments that would benefit future traffic-related assessments.

Keywords

Air pollution, dispersion modeling, epidemiology, exhaust, exposure misclassification, PM_{2.5}, traffic, vehicle

Introduction

Residential location and proximity to major roads have been widely used in analyses of exposures and adverse health effects that can result from traffic-related air pollutants, reflecting the elevated concentrations found near large roads (Zhu et al. 2006; Baldauf et al. 2008; Barzyk et al. 2009; Hagler et al. 2009; Hu et al. 2009; Health Effects Institute 2010; Karner et al. 2010). As an indicator or exposure surrogate, residential distance to roads, or more generally, residence location, is intended to reflect the portion of exposure received at home, which is generally the dominant share since most individuals spend the majority of their time at home (Huang et al. 2000). Residence location or proximity to roads can be used as a surrogate exposure metric itself, or as an input to land use regression, dispersion, space-time (geostatistical), and hybrid models, which are designed to estimate ambient air concentrations and sometimes personal exposures in the case of the hybrid models (Brauer 2010; Rioux et al. 2010; Batterman 2013; Beevers et al. 2013; Özkaynak et al. 2013).

Actual exposure for any particular individual will be the result of many factors, e.g., the amount of time spent indoors and outdoors, building and vehicle cabin air exchange rates, and breathing rates (Özkaynak et al. 2013). Unfortunately, direct measurement of traffic-related air pollutant exposure is rarely practicable due to cost and logistical issues (Rioux et al. 2010). Ambient air quality monitoring networks do not provide the spatial coverage needed to estimate near-road exposures (Batterman 2013), and personal, home or biomarker measurements rarely are feasible in large scale studies. Still, it remains important to improve exposure estimates that are used in epidemiology, health impact, environmental justice and other studies (Jerrett et al. 2005; Brauer 2010; Institute 2010; Sheppard et al. 2012). Improved estimates will minimize exposure misclassification that can bias results of epidemiology studies towards the null (Kioumourtzoglou et al. 2014), incorrectly predict risks in health impact studies, and misidentify affected populations in environmental justice studies.

This paper explores alternate metrics for characterizing exposure to traffic related air pollutants, including metrics based on proximity to major roads, traffic volume and density, traffic type, traffic emissions density, and pollutant concentrations from dispersion modeling. These metrics are being utilized in an ongoing epidemiology study investigating effects of diesel exhaust emissions on the respiratory health of asthmatic children in Detroit, Michigan, USA (Vette et al. 2013). The evaluation of the exposure metrics presented in this paper includes a comparison of exposure distributions, spatial variability using maps coded by exposure group, intraclass correlations, and concordance rates.

Methods

Study population

The Near-road EXposures and effects of Urban air pollutants Study (NEXUS) was designed to examine near-roadway exposures to air pollutants and respiratory health in children with asthma living close to major roads in Detroit, MI. This community-based participatory research (CBPR) study was designed and conducted with community input and consent. Children with asthma or symptoms of asthma from 6 to 14 years of age were recruited on the basis of the proximity of their home to major roads in three traffic categories: high diesel/high traffic (HDHT), defined as homes within 175 m of roads with >6,000 commercial vehicles/day (commercial annual average daily traffic; CAADT) and >90,000 total vehicles/day (annual average daily traffic; AADT); low diesel/high traffic (LDHT), defined as homes within 175 m of roads with >90,000 AADT and <4,500 commercial vehicles/day; and low diesel/low traffic (LDLT) homes located >300 m from roads with >25,000

AADT and greater than 500 m from roads with >90,000 AADT. In the initial groups, the designation of commercial vehicles was used as a surrogate for diesel vehicles. Children in the LDLT group were drawn from the same neighborhoods and school catchment areas as the other groups, but lived further from high-traffic corridors, thus minimizing possible confounding from unmeasured neighborhood-associated covariates.

Ultimately, 139 children were recruited and participated in the study from September 2010 to December 2012. Because a number of children moved during the study, a total of 218 residence locations were considered (Figure 1). The study population had approximately equal distribution across the three traffic categories. The population was primarily minority (non-Hispanic Blacks constituted 82% of the participants, Hispanics 8%, Non-Hispanic Whites 4%, and other/multiracial 6%). Many households were poor (about one-third of families reported annual household incomes below \$15,000).

Residential proximity to major roadways

Because concentrations of traffic-related air pollutants rapidly decrease with distance from roads, considerable effort was spent to obtain accurate estimates of home locations and the distance to major road. Initially, to guide field staff in their recruitment efforts, candidate homes were identified using GIS-produced maps, which identified buffers within 200 m of selected highways and the street and house numbers of residences within these buffers. For the children recruited and enrolled into the study, the resulting 218 home locations were geocoded using three approaches. The first used a hand-held GPS device (60CS, Garmin International Inc., Olathe, KS, USA) operated by our technician who stood as close as possible to the residence's front door. When the indicated accuracy was <10 m, the location was recorded on a data entry form and as a waypoint in the device's memory. The calibration of the device was confirmed using several other GPS units. Second, online automated geocoding software provided by 'Bing Maps' (<http://www.bing.com/maps/>) was used by manually entering the number, street, city and ZIP code of each residence into this application, which returned latitude and longitude. The Bing Map estimates used the European Petroleum Survey Group (EPSG) code, a Mercator projection, and a spherical model of the earth (Aitchison 2011). The third geocoding estimate used the address locator in the online U.S. Streets Geocode Service in ArcGIS, ESRI ArcMap 10.0 (Redlands, CA, USA). This system uses a cascading sequence of geolocators starting with the Tele Atlas Address Points database, followed by the Tele Atlas Street Address Range database, 9-digit ZIP code, and then the 5-digit ZIP code locators (ESRI 2010).

All coordinates were converted to Universal Traverse Mercator (UTM) coordinates for use in subsequent analyses. In the event of discrepancies exceeding 100 m between the GPS and automated geocoding coordinates, data were checked, plotted, and if needed, our technician was sent out to confirm GPS coordinates a second (and sometimes a third) time, at which point all GPS measurements agreed. Final home locations are plotted in Figure 1. On average, the automated geocoding estimates diverged from GPS measurements at the NEXUS homes by an average 30 ± 23 m, although much larger errors were not infrequent, e.g., 75th and 95th percentile errors were ~50 and 75 m (Ganguly et al. 2014). The bigger errors can represent a large fraction of the buffer width (200 m) used in the proximity metrics, suggesting that a fraction of homes geocoded using automated systems would be misclassified by such surrogate exposure metrics, especially since automated geocoding programs typically give larger errors than those found for Detroit (Rioux et al. 2010; Ganguly and Batterman 2014).

Distances from each residence to the nearest roads and highways were determined using ESRI ArcMap (version 10.0) "NEAR" function within the Proximity toolset, the 2012 Topologically Integrated Geographic Encoding

and Referencing (TIGER) 2012 road shape files, and the North American Datum (NAD) for 1983. A second measure of distance to the nearest highway was calculated similarly using the link-based road network for the emissions inventory described next. This distance averaged 20 m greater than the TIGER estimate, representing the distance to the road centerline rather than to the road edge.

Road network, traffic data and emissions inventory

The modeled road network used 9,701 links (linear segments) to represent 3,109 km of roads, which included all but the smaller and numerous local roads in the ~800 km² study area (Figure 1). Major roads (e.g., freeways) were represented using separate links for each direction, large service roads, if any, and ramps. Road network data obtained from the Southeast Michigan Council of Governments (SEMCOG) included the locations, number of lanes, and roadway type (e.g., freeway, arterial). The road network extended at least 5 km beyond the locations of the NEXUS homes, and fully encompassed the city of Detroit (area of ~355 km²).

Hourly vehicle traffic volume and speed on each link were estimated using the SEMCOG Travel Demand Forecast Model, which uses a TransCAD-based suite of applications, and temporal profiles for monthly, weekly and daily flows by vehicle class (e.g., heavy-duty diesel, light-duty gasoline). Where possible (e.g., for interstates), estimated vehicle flows were checked against observed traffic counts.

Estimates of total and diesel vehicle volumes on major roads formed the basis of the initial participant recruitment and the exposure classification. The initial grouping of roads (as HDHT, LDHT and LDLT) used 2006 and 2007 maps with AADT and CADT flows prepared by the Michigan Department of Transportation (http://mdotcf.state.mi.us/public/maps_adtmaparchive/). As described earlier, subsequent analyses used state and regional data and a traffic demand model to derive AADT for each road link for the year 2010. Traffic on the major road closest to each HDHT and LDHT residence was estimated by summing AADT values on the corresponding road-links, which included at least two links (one for each direction) and occasionally additional links if the road split or if ramps added to the road's traffic in the vicinity of the residence. At three homes, two major roads had similar proximity and both were counted in the AADT estimate. The corresponding number of lanes on these roads was summed as another measure of road size. Finally, the number of diesel vehicles on the closest major road was calculated as the product of the AADT (vehicles/day) and the fraction of diesel vehicles for the road type, which was 5.23% for "other freeways" (National Functional Class or NFC 12) and 9.18% for interstates (NFC 11). These fractions represent the sum of light-duty diesel trucks, heavy-duty diesel trucks, and heavy-duty diesel vehicles (LDDV, LDDT, HDDV) as derived using state-level data from the US Federal Highway Administration (Table VM-4 from the FHWA Highway Statistics Series www.fhwa.dot.gov/policyinformation/statistics/2010/vm4.cfm), and the US EPA Emission Inventory Improvement Program (www.epa.gov/ttn/chief/eiip/).

Primary mobile source emissions of particulate matter below 2.5 µm dia (PM_{2.5}), oxides of nitrogen (NO_x), carbon monoxide (CO) and other pollutants were estimated for each link on an hourly basis, thus producing a spatially and temporally resolved mobile source emissions inventory. The link-based inventory used emission factors representative of each vehicle class in the study area for the year 2010 calculated using the MOVES2010a model with inputs including the average speed and flow on each link, local vehicle mix and age distribution, ambient temperature, season, and road type (Cook et al. 2008; Isakov et al. 2009). Due to large uncertainties, particulate emissions for brake and pavement wear, and resuspension of dust were not included in the PM emissions estimates.

Emissions-based exposure metrics incorporate the quantity of traffic-related pollutant emissions released,

providing an exposure metric that may be particularly relevant to policies addressing emission controls, transportation control measures, and other actions that directly affect emissions. Emissions density was estimated as the daily vehicle exhaust emissions (g/day) within a 300 m buffer around each home. PM_{2.5}, NO_x and CO emissions were considered.

Concentration estimates and dispersion modeling

We evaluated whether simple “box” models using the emissions density information discussed above could provide useful estimates of near-road pollutant concentrations. We assumed a wind speed $u = 3.66$ m/s (the long term Detroit average), a mixing height $h = 100$ m, and the average traffic-related PM_{2.5} emission rate Q in the 300 m buffer ($r = 300$ m) around high traffic homes. Concentrations were calculated as $C = Q / (2 u r h)$.

Hourly pollutant concentrations for the year 2010 were predicted at each home for three cases: annual average concentrations due to on-road exhaust emissions; the maximum 24-hour concentration also due to on-road exhaust emissions, and the “total” annual average concentration due to on-road, non-road and background sources. Each case used the road-link emissions inventory for the Detroit area described above, the new RLINE dispersion model specifically designed for roadway emissions (Snyder et al. 2013a; Venkatram et al. 2013), and hourly meteorological data from the Detroit City airport processed by AERMET. In addition, the third case used a hybrid model system that integrated RLINE, the AERMOD model for area and point sources in the region using source locations (Cimorelli et al, 2005), emission rates and other parameters from the 2008 National Emissions Inventory (NEI), and estimated regional (background) concentrations determined using the Community Multiscale Air Quality (CMAQ) model, observations from air quality monitoring networks in the region, and a space/time kriging model. As described in Isakov et al. (2014), this system is highly flexible, and model outputs can provide spatial and temporal patterns of air pollutants by source category.

Detailed descriptions and evaluations of RLINE and the other dispersion models have been presented elsewhere (Heist et al. 2013; Snyder et al. 2013a; Snyder et al. 2013b; Venkatram et al. 2013; Isakov et al. 2014). In Detroit, model results have been compared to ambient observations collected in both routine monitoring networks (AQS) and during the NEXUS intensive campaign. Compared to AQS data, 24-hr average PM_{2.5} concentrations correlated well ($0.78 < r < 0.94$) with 2010 data collected at four PM_{2.5} monitoring sites in Detroit, and most predictions were within a factor of two of observations. NO_x concentrations predicted at the sole AQS monitoring site in Detroit reproduced morning and afternoon peaks but overpredicted the concentrations, likely due to contributions from regional sources since this site was several km from major highways. Compared to black carbon measured outside of 25 of the NEXUS homes and NO_x measured at 9 homes in (September-November) 2010, the model generally captured the magnitude and dynamics of observed concentrations, although concentrations were overpredicted or missed at some sites and some specific hours, likely due to uncertainty in hourly traffic activity and emissions at the road link level. Further description of the evaluation of the modeling system in the Detroit application is presented elsewhere in this issue (Isakov et al. 2014).

Data analysis

Descriptive analyses included graphs of distributions stratified by the original HDHT, LDHT and LDLT groups. Differences in means between the HDHT and LDHT groups were evaluated using t tests, and difference in distributions for the same groups were evaluated using the non-parametric Mann-Whitney (MW) tests. (Sample size $n=96$ in all cases for both tests). Comparisons between exposure metrics used Spearman's and Kendall's τ -b correlation coefficients to measure interclass agreement. The latter correlation coefficient

considers the number of concordant pairs of observations minus the number of discordant pairs, expressed as the fraction of total pairs, and accounts for ties. Both are non-parametric measures that range from -1 to 1. Several additional measures of concordance/discordance rates were derived for exposures divided into “high”, “medium” and “low” categories to provide estimates potentially relevant to exposure misclassification, as discussed in the text.

Results

Grouping of homes by distance to roads

By design, many of the children participating in NEXUS lived close to major roads. Most children in the high traffic groups (HDHT and LDHT) lived near two interstate highways and two State of Michigan highways: I75 (n=30 residence locations), I94 (n=20), M10 (n=30), and M39 (n=16). A few children lived near I96 (n=3), which was classified as a “medium” diesel high traffic road and not considered further in the present analysis.

The distributions of distances from high traffic roads for the HDHT, LDHT and LDLT groups are shown in [Figure 2](#). Means and distributions in the two high traffic groups did not differ (t-test: $p=0.641$; MW test: $p=0.376$), and the mean distance was 96 ± 45 m from the road edge. The distributions were rectangular in nature, and one residence was as close as 1 m from the road edge (along I75 in southwest Detroit). In contrast, distances in the LDLT group were much longer, averaging $1,562 \pm 1,133$ m (\pm standard deviation).

Among the HDHT and LDHT groups, distances from roads did not vary among the four major roads (Kruskal-Wallis test, $p=0.628$), although the average distance for M39 was 10 m less than those for the three other major roads. No general trends were noted for residence-road distances with respect to region of the city or distance from downtown.

Total and diesel traffic metrics

The distributions of total and diesel traffic volume on major roads near the NEXUS homes are shown in [Figures 3A](#) and [3B](#) and mapped in [Figures 4A](#) and [4B](#); descriptive statistics are in [Table 1](#). As anticipated, diesel vehicle traffic was highly correlated to the total traffic ($r = 0.93$), however, there are important differences, as discussed below. The maps show the clustering of high traffic homes along five major highways: HDHT homes fall mostly in the south and east along I75 and I94 (red circles, [Figure 4A](#)); LDHT homes are mostly in the north and west along M10 and M39 (blue circles). The LDLT homes (green circles) are distributed throughout the region.

Considering the total traffic (AADT) near the high traffic homes, the average volume was $134,000 \pm 30,000$ vehicles/day and the average number of lanes was 7.2 ± 1.6 . As noted above, three homes were equally near two major roads, including the home with the highest AADT, which was located at the I94-M10 intersection (AADT = 212,000 vehicles/day, counting both roads); the two other homes were at the I94-I96 and I75-I96 junctions (178,000 and 134,000 vehicles/day, respectively). Traffic volumes at such junctions can be difficult to estimate given the numbers of ramps and highway segments involved. Otherwise, the traffic volume estimates were due to a single road. The AADT distributions for HDHT and LDHT homes were similar (t-test: $p=0.091$; MW test: $p=0.262$; [Figure 3A](#)), although the HDHT group had a greater range and its lowest tertile had 25,000 fewer vehicles/day. AADT had negligible correlation with the home’s distance to the road ($r = -0.006$). The maps suggest that each road had a wide range of traffic volumes, i.e., no apparent geographic bias.

Considering diesel traffic near the high traffic homes, the volume averaged $8,640 \pm 2,240$ vehicles/day, and eight homes had over 12,500 diesel vehicles/day. HDHT homes had an average of 2,100 diesel vehicles/day more than LDHT homes, a statistically significant difference (t-test: $p < 0.001$; MW test, $p < 0.001$). Still, the distributions of HDHT and LDHT homes had considerable overlap, and roughly half of the two sets of homes had a comparable range of diesel volume (7,000-9,000 vehicles/day; Figure 3B). The highest diesel volume occurred for homes along I75, while homes along the eastern portion of I94 had lower diesel traffic volumes than other HDHT homes (Figure 4B). This resulted in part from reclassifying this section of I94 to an “other highway” (NFC 12) designation rather than “interstate” in an attempt to better match the observed traffic classification. Diesel volume had negligible correlation with distance to the road ($r = 0.001$). Overall, estimates of diesel traffic are less certain than those for total traffic due to the variation in fleet mix and the limited classification data available.

By design, the LDHT homes were intended to have commercial annual average daily traffic (CADT) below 4,500 vehicles/day. There are several reasons for differences between the initial groupings and those computed using the road-link network. First, not all diesel traffic is “commercial” (and *vice-versa*). Based on the fleet composition data used, the bulk of diesel traffic on major roads are heavy duty diesel vehicles (HDDV), which represent 84% of the diesel volume on interstates, and 71% on “other highways.” The balance of diesel vehicles is contributed by light duty diesel vehicles (LDDV) and light duty diesel trucks (LDDT), which respectively represent 11 and 6% of the diesel vehicles on interstates, and 20 and 9% on other highways. With the exception of the 6 to 9% of the diesel vehicles classified as LDDVs, most LDDTs and all HDDVs form part of CADT. With the LDDV fraction removed, differences between the HDHT and LDHT categories slightly increase, e.g., the average volume of diesel traffic is $8,700 \pm 2,500$ vehicles/day near the HDHT homes, and $6,300 \pm 950$ vehicles/day near the LDHT homes. Second, CADT will include some gasoline-powered vehicles. However, the emissions inventory classifications do not correspond well to the AADT/CADT classifications. Third, the initial and road-link estimates of CADT used different data sources: the former depended on local measurements, while the latter used a mix of local, Michigan and national-level data.

Traffic density metric

A measure of traffic density, the number of vehicle-km traveled (VKT) per day within a 300 m radius of each home, was calculated using the road-link network. The average VKT/day was $94,200 \pm 24,900$ for HDHT homes, $102,200 \pm 17,230$ for LDHT homes, and $12,700 \pm 12,400$ for LDLT homes. Differences in traffic density between HDHT and LDHT homes approached statistical significance (t-test: $p = 0.071$; MW test, $p = 0.117$; Figure 3C). This metric varied considerably within each home group, e.g., the range spanned a factor of 2.8 for HDHT homes, and variation for homes along the same highway could be considerable, e.g., the southwest section of I75 shows several HDHT homes with low traffic density (Figure 4C). Of the LDLT homes, 4% ($n=4$) had moderately high traffic density (41,000 to 74,000 VKT/day), overlapping or nearly overlapping the high traffic homes. Based on this exposure metric, the LDHT homes had slightly greater exposure than HDHT homes.

For most high traffic homes, the largest share by far of the traffic density metric was contributed by interstates and other highways. This exposure metric can be very sensitive to the distance criterion since roads will drop out as the distance criterion (buffer size) decreases. For example, a 150 or 200 m buffer excluded the nearby highway for a few HDHT homes, dropping the KMT/day to near zero, despite large values obtained with the 300 m buffer.

Emissions density metrics

Figures 3D and 4D show the PM_{2.5} emissions density distributions and levels at each home. For PM_{2.5}, the emissions density averaged $2,226 \pm 568$ g/day at HDHT homes, $2,029 \pm 379$ at LDHT homes, and only 315 ± 320 at LDLT homes. The distributions of PM_{2.5} emissions near HDHT and LDHT homes appeared similar, although outliers caused results of the statistical tests to vary (t-test: $p=0.052$; MW test: $p=0.153$), e.g., a few HDHT homes near intersections of major roads (e.g., I95 and I75) had very high emissions. For the high traffic homes, nearby highways contributed the bulk of emissions. Emissions rates in the 300 m buffers around LDLT homes were much lower, although a few homes had emissions densities that overlapped those in the high traffic group.

Distributions and maps for NO_x and CO emissions density are shown in Supplemental Figures S1 and S2. Results for NO_x parallel the findings for PM_{2.5}. However, results for CO diverged as HDHT homes tended to have lower CO emissions within the 300 m buffer, although statistical significance was not reached (t-test: $p=0.072$; MW test: $p=0.093$). Some of the higher levels were found for homes along M39 (in the western portion of the study region), due to the lower CO contributions from diesel vehicles and the dominance of CO emissions from much more numerous (and gasoline-powered) automobiles. Overall, the pattern for CO closely resembled the traffic density metric (Figures 3C and 4C).

Emission-based metrics may be especially relevant for the traffic analysis zones (TAZ) commonly used by metropolitan planning organizations for various purposes, although TAZs may be relatively large relative to the spatial variability of traffic-related air pollutants. Emission-based metrics can be developed for the major road nearest point of interest (e.g., a home), for roads within a local area, or for buffers around points of interest, as presented here. This metric is pollutant-specific and thus can account for differences in fleet mix, technology and transportation control measures (TCMs) that affect emissions.

Concentration metrics

The “box” model using the average emissions density around high traffic homes for PM_{2.5} (2,226 g/day) gave a predicted PM_{2.5} concentration of $0.12 \mu\text{g}/\text{m}^3$. This low estimate does not account for temporal variation in emission rates (e.g., rush-hour emissions are approximately twice the average rate) or meteorology (e.g., wind speeds near 1 m/s are not uncommon and mixing heights also vary). In addition, some homes had twice the average emission rate, and results will be sensitive to the buffer size and source configuration. The most significant limitation, however, is the validity of the fully mixed assumption near major roads where vertical dispersion of roadway emissions may not reach 100 m. As shown below, dispersion models designed for near-road applications yielded concentrations one to two orders of magnitude greater. Thus, while box models have been used at a city-wide or regional level (Apte et al. 2012), they do not appear useful for near-road applications.

Dispersion modeling results for PM_{2.5} are summarized in Table 2, and concentration distributions and maps are shown in the bottom panels of Figure 3 and 4. Considering emissions from local traffic and annual average concentrations, PM_{2.5} levels at high traffic homes were $1.6 \mu\text{g}/\text{m}^3$ higher than those at LDLT homes, or about twice that at the LDLT homes. Considering 24-hour peak concentrations, PM_{2.5} levels at several HDHT homes were estimated to exceed $40 \mu\text{g}/\text{m}^3$, which is high relative to PM_{2.5} standards. (The U.S. National Ambient Air Quality Standards are currently 12 and $35 \mu\text{g}/\text{m}^3$ for annual and 24-hr averages, respectively). The average 24-hr peak across the high traffic homes was $11 \mu\text{g}/\text{m}^3$ more than levels at the LDLT homes, again about twice the level. The peak metric tended to decrease the relative difference between high and low traffic homes.

Considering the “total” PM_{2.5} estimated by the hybrid model, regional sources were dominant, contributing about 12 µg/m³ on an annual average basis across the homes. Given the relatively high contribution of PM_{2.5} by regional sources, the spatial variability due to local sources was significantly reduced, e.g., PM_{2.5} levels at low and high traffic homes differed by only 10 to 20%. (Distributions and maps of total PM_{2.5} are shown in [Supplemental Figure S3](#)).

The distributions of annual average PM_{2.5} concentrations at HDHT and LDHT homes due to local traffic were very similar (t-test: p=0.637; MW test: p=0.523, [Figures 3E, S3](#)); peak 24-hr concentrations tended to be slightly but not statistically higher at the LDHT homes (t-test: p=0.251; MW test: p=0.235; [Figures 3E, 3F](#)). The highest prediction, considered an outlier since it considerably exceeded any other prediction, occurred at a home very near I75. Excluding this point, concentrations due to local traffic varied by about 3-fold in each exposure group.

Unlike the metrics discussed earlier, dispersion modeling results are expressed as concentrations that permit direct and meaningful comparisons to air quality standards, monitoring data and other studies. For example, a recent application of a hybrid dispersion model in London, England showed annual average PM_{2.5} contributions from 1 to 5 µg/m³ due to local vehicle exhaust (depending on site), and about 11 µg/m³ due to other local and regional sources (Beevers et al. 2013). Both the traffic and regional PM_{2.5} contributions were very similar to those predicted for Detroit. While an intercity comparison incorporating meteorology and spatial factors is beyond the present scope, the volume of traffic and the number of large and high emitting diesel trucks on Detroit’s highways may produce PM_{2.5} concentrations that are comparable to those in London, despite Detroit’s smaller size and lower fraction of diesel vehicles.

Comparison of exposure metrics

[Table 3](#) compares the exposure metrics using Spearman’s and Kendall’s τ -b correlation coefficients. To include the original categorical proximity classifications in this analysis, LDLT, LDHT and HDHT groups were coded two ways: as 1, 2 and 3 respectively in the “Group1” variable, with the assumption that exposures were ranked as HDHT > LDHT > LDLT; and as 1, 2 and 2 in the “Group2” variable with the assumption that exposures were HDHT \approx LDHT > LDLT, as suggested by most metrics. Results for these variables are only indicative of agreement. [Table 3](#) colors the higher (>0.6 and >0.8, absolute value) correlation coefficients. For the Spearman coefficients, absolute values above 0.17 (n=96 for comparisons involving only high traffic groups) or 0.12 (n=198 to 218 for other comparisons) are statistically significant (p<0.05, 1-sided test). While the measures have some differences, e.g., the Spearman coefficients are 3 to 23% higher than the Kendall τ -b coefficients and contrasts in Kendall’s τ -b coefficients tend to be larger, similar patterns emerge. The original proximity classifications (“Group1” and “Group2”) are related to distance to the nearest highway (defining variables), to traffic density (VKT/day in the 300 m buffer around each home), and to PM_{2.5} and CO emissions density metrics using the same buffers. While maximum 24-hr and annual average PM_{2.5} concentrations from local traffic were closely correlated, predicted PM_{2.5} concentrations had only modest agreement with most of the other exposure metrics, although correlations with the traffic density and distance metrics might be viewed as reasonable.

Differences among the metrics can occur for a number of reasons. Metrics that depended on a single road, i.e., the original groups, distance to, traffic volume on, and number of lanes on the nearest major road, fared poorly in comparisons since these metrics excluded the influence of other roads, among other reasons. Still, these simple metrics have some value. Second, traffic density and emissions density are very strongly related, and

may yield equivalent spatial patterns. Third, predicted concentrations (e.g., annual average and 24-hr maxima) were not as strongly related to the two density metrics, most likely a result of the dispersion modeling accounting for meteorological influences, e.g., directional effects. Fourth, while concentrations of different pollutants and concentrations using different averaging times and statistics (e.g., average versus peak) show differences, the long term spatial patterns are relatively similar and a single metric may be sufficient for many applications. (In contrast, short-term patterns, e.g., daily levels, will vary considerably due to both the influence of meteorology and traffic patterns.) In health effects studies such as NEXUS, it is also useful to understand the relative ranking of exposures, discussed next.

Table 4 presents a concordance analysis related to exposure misclassification if one of the metrics in the pairwise comparison reflects the true exposure. This measure is defined as the percentage of homes identically classified after grouping each metric into tertiles (high, medium and low categories). Thus, 100% agreement indicates that each home is placed in the same tertile for the pair of metrics considered. Random assignment would be expected to yield 33% agreement. For distance, tertiles were reversed, i.e., the shortest distances (presumably the highest exposure) were placed in the third tertile. The skewness of several metrics produced non-uniform tertile ranges, i.e., for the number of lanes (0-6, 6-8 and 8-14), PM_{2.5} emissions (0-306, 306-1837, 1837-4027 g/day), 24-hr peak (6-14, 14-18, 18-47 µg/m³) and annual average (0.8-1.6, 1.6-2.7, 2.7-9.4 µg/m³) PM_{2.5} concentrations. Thus, another concordance measure was used to examine agreement with “thirds,” defined using three evenly spaced bins between the metric’s minimum and a nominal maximum value, either the actual maximum or an adjusted value that provided 15 observations in the top third. This gave ranges for the distance metric of 3067-5025, 1534-3067 and 1-1534 m (reversed); 0-1300, 1300-2600 and >2600 g/day for PM_{2.5} emissions; 0.8-1.6, 1.6-2.7 and >2.7 µg/m³ for annual average PM_{2.5} concentrations; and 6-17, 17-27 and >27 µg/m³ for peak 24-hr PM_{2.5} concentrations. (Figure 3D illustrates differences between tertiles and thirds.) Again, 100% agreement denotes that the pair of metrics placed each home at the same (low, medium or high) level. Table 4 highlights metrics with higher (>60 and >80%) agreement. These concordance measures give an estimate of misclassification intended for comparison to the three original groups (HDHT, LDHT, LDLT), as discussed next, but they have several disadvantages, e.g., potential sensitivity to the metric’s distributions and the bin cutoffs.

The original proximity classification for NEXUS homes (“Group1” in Table 4) matched the tertiles and thirds groupings for PM_{2.5} concentrations predicted by the dispersion model for 54 to 65% of homes. Considering only the low and high traffic categories (“Group2”), the percentage of homes that matched PM_{2.5} concentration tertiles was lower (45 to 52%), but higher for thirds (60 to 80%). The same five metrics that were highly correlated (traffic density, emissions density for CO and PM_{2.5}, annual average and peak 24-hr average PM_{2.5} concentrations from traffic, Table 3), and sometimes the total PM_{2.5} (including all sources) had higher agreement, e.g., 50 to 90% using tertiles, and 59 to 80% using thirds. This excludes the very high concordance found between (and expected for) traffic density and CO emissions density. Overall, the concordance measures in Table 4 and the correlations in Table 3 produced similar patterns among the metrics. While over half of the homes were similarly placed into “low,” “medium” or “high” traffic exposure classes using either the original groups or one of the other metrics discussed, 20 to 50% of homes, depending on the metric, were categorized differently.

We also examined rates of “severe” discordance using both tertiles and third, but considering homes placed in the high exposure group by one metric, but in the low exposure group by the second metric. Paralleling the analysis above, severe discordance rates among five more comprehensive metrics (traffic density, CO and PM_{2.5}

emissions densities, and annual average and 24-hr peak concentration) were low (0-6%). However, rates using total or diesel traffic volume on the nearest major road with these five metrics were 11 to 22% using tertiles, and 43 to 52% using thirds. Severe discordance rates using the simple distance-to-highway metric compared to the five metrics were also low (1-3% using tertiles), meaning that few LDLT homes were placed in the upper tertiles of the five metrics and few HDHT or LDHT homes were placed in low tertiles of the same metrics. In addition to reinforcing earlier conclusions, this analysis indicates severe discordance is rare among exposure metrics using traffic density, emissions density, or concentrations.

Table 5 summarizes the exposure metrics, describing their strengths, limitations, data requirements, and key results in NEXUS. Each metric provides different information (noted in the table and not repeated here). Several additional points are noteworthy. First, traffic-related exposures should be viewed as a continuum, and exposure groups are not “homogeneous.” For example, most of exposure metrics for the NEXUS homes in each of the three initial groups (LDLT, LDHT and HDHT) typically spanned a 3-fold range, with considerable overlap between groups. Second, exposures result from multiple emission sources, and concentrations depend on source characteristics (e.g., emission rate) and dispersion (e.g., wind direction and stability), factors that vary in time. While an emphasis on nearby sources (e.g., major roads) makes analyses more tractable, such simplifying assumptions can be inaccurate, and the choice of a distance cut-off is arbitrary. As examples, concentrations of traffic-related pollutants at NEXUS homes near major highways varied greatly from day-to-day, and homes that were distant (>500 m) from highways sometimes still received considerable levels of traffic-related pollutants. Third, the minimum information needed to evaluate spatial variability (typically using the long term average) of traffic-related exposure includes the distance-to-road and traffic flows for the larger and nearby roads. This can be expressed in the relatively simple traffic density metric, which correlated reasonably well with predicted PM_{2.5} concentrations ($r=0.78$ for annual averages, $r=0.74$ for 24-hr peaks). Traffic density has been identified as one of the strongest predictor variables in recent land-use regression (LUR) models of traffic related air pollutants (Hoek et al. 2008). However, traffic density metrics do not provide concentrations and, even if incorporated into LUR models, results may be limited in terms of comparability across cities and time (Allen et al. 2011; Wang et al. 2013). The dispersion models provided spatial and temporal concentration predictions as well as source apportionment information, e.g., the PM_{2.5} share due to traffic. While data and computationally intensive, such models have a strong physical basis and also are amenable to forecasting and scenario analysis (Brauer 2010).

Conclusions

Surrogate exposure metrics for traffic-related air pollution exposures developed and compared in this study include proximity to highways, traffic volume, traffic density, number of lanes, emissions density, and concentration predictions from dispersion models. These metrics were critiqued individually and collectively, focusing on results obtained at the 218 home locations of participants in the Detroit-area NEXUS epidemiology study. Comparisons included examination of exposure distributions, spatial variability using maps coded by exposure group, Spearman’s and Kendal’s τ rank correlation coefficients, and concordance rates. While showing some agreement, simple categorical and proximity classifications (high diesel/low diesel traffic roads and distance from major roads) did not reflect the range and overlap of exposures seen in the other metrics. Information provided by the traffic density metric, defined as the number of kilometers traveled per day within a 300 m buffer around the home, was reasonably consistent with the more sophisticated metrics, although this metric does not provide information related to concentrations or temporal variability. Dispersion modeling

provided this information, along with source apportionment results that separated concentrations from traffic emissions and other sources. At the NEXUS homes, the annual average and 24-hour peak concentrations showed a high degree of spatial agreement. While there is broad agreement between several of the surrogate exposure metrics, including traffic density, emissions density and dispersion modeling, most of these alternatives still produced significantly different exposure classifications, suggesting the potential for exposure misclassification and the need for refined and validated exposure metrics.

The analyses also indicated the need for accurate geocoding of homes and roads given the spatial variability of pollutant levels near roads. Positional errors in the range of 30 to 50 m, and sometimes much more, should be anticipated using automated geocoding software.

Dispersion modeling systems can provide exposure information relevant to on- and near-road environments, not only at homes, as demonstrated in the Detroit, but also at schools, parks, workplaces, commuting routes and other locations where people are exposed. While assembling the data and the computational demands for dispersion modeling of traffic emissions in large urban areas are non-trivial issues, future traffic-related health assessments, including epidemiological, risk and environmental justice studies, would benefit from such information. Further, this information can be used in hybrid models that simulate indoor exposures and time-activity behaviors, thus providing a refined estimate of air pollution exposure. At the community level, exposure assessments used in conjunction with transportation planning tools would advance policy initiatives aimed at mitigating traffic-related air pollutant exposures and health effects.

Acknowledgements

This research was conducted with input from several partners, including the Community Action Against Asthma (CAAA) Steering Committee and our EPA partners. We thank Rich Cook of U.S. EPA's Office of Transportation and Air Quality and Michelle Snyder, Sarav Arunachalam and others at the University of North Carolina at Chapel Hill for help in developing locally-resolved traffic emissions and dispersion modeling. The NEXUS study involves a community-based participatory research partnership, and we thank CAAA and its member organizations. We also appreciate assistance from Ricardo Mayo, Laprisha Berry-Vaughn, Sonya Grant, Graciela Menz, Rajiv Ganguly, Christopher Godwin, Leonard Brakefield, and Feng-Chiao Su at the University of Michigan.

The United States Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research described here under Cooperative Agreement R834117 (University of Michigan). It has been subjected to Agency review and approved for publication. Funding sources that also supported this research included the National Institute of Environmental Health Sciences (NIEHS) through grants 5R01 ES014677-02 and R01 ES016769-01 for NEXUS. NIEHS provided additional support in grant P30ES017885 "Lifestage exposures and adult disease" as did the Health Effects Institute. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

- Aitchison, A. (2011) "The Google Maps / Bing Maps Spherical Mercator Projection." <http://alastaira.wordpress.com/2011/01/23/the-google-maps-bing-maps-spherical-mercator-projection/>.
- Allen, R. W., O. Amram, A. J. Wheeler and M. Brauer (2011). "The transferability of NO and NO₂ land use regression models between cities and pollutants." *Atmospheric Environment* **45**(2): 369.

Apte, J. S., E. Bombrun, J. D. Marshall and W. W. Nazaroff (2012). "Global intraurban intake fractions for primary air pollutants from vehicles and other distributed sources." Environmental science & technology **46**(6): 3415.

Baldauf, R., E. Thoma, M. Hays, R. Shores, J. Kinsey, B. Gullett, S. Kimbrough, V. Isakov, T. Long, R. Snow, A. Khlystov, J. Weinstein, F. L. Chen, R. Seila, D. Olson, I. Gilmour, S. H. Cho, N. Watkins, P. Rowley and J. Bang (2008). "Traffic and meteorological impacts on near-road air quality: summary of methods and trends from the Raleigh Near-Road Study." J Air Waste Manag Assoc **58**(7): 865-878.

Barzyk, T. M., B. J. George, A. F. Vette, R. W. Williams, C. W. Croghan and C. D. Stevens (2009). "Development of a distance-to-roadway proximity metric to compare near-road pollutant levels to a central site monitor." Atmospheric Environment **43**(4): 787-797.

Batterman, S. (2013). "The near-road ambient monitoring network and exposure estimates for health studies." EM Journal(July): 24-30.

Beevers, S. D., N. Kitwiroon, M. L. Williams, F. J. Kelly, H. Ross Anderson and D. C. Carslaw (2013). "Air pollution dispersion models for human exposure predictions in London." Journal of exposure science & environmental epidemiology **23**(6): 647.

Brauer, M. (2010). "How Much, How Long, What, and Where." Proceedings of the American Thoracic Society **7**(2): 111-115.

Brauer, M. (2010). "How much, how long, what, and where: air pollution exposure assessment for epidemiologic studies of respiratory disease." Proceedings of the American Thoracic Society **7**(2): 111.

Cimorelli A.J., Perry S.G., Venkatram A., Weil J.C., Paine R.J., Wilson R.B., et al. (2005) AERMOD: A Dispersion Model for Industrial Source Applications. Part I: General Model Formulation and Boundary Layer Characterization. Journal of Applied Meteorology and Climatology **44**(5): 682-693.

Cook, R., V. Isakov, J. S. Touma, W. Benjey, J. Thurman, E. Kinnee and D. Ensley (2008). "Resolving local-scale emissions for modeling air quality near roadways." Journal of the Air & Waste Management Association **58**(3): 451-461.

ESRI (2010). ESRI Geocoder Information.

Ganguly, R. and S. Batterman (2014). "Effect of Geocoding Errors on Traffic-Related Air Pollutant Concentration Estimates." submitted.

Hagler, G. S. W., R. W. Baldauf, E. D. Thoma, T. R. Long, R. F. Snow, J. S. Kinsey, L. Oudejans and B. K. Gullett (2009). "Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants." Atmospheric Environment **43**(6): 1229-1234.

Health Effects Institute (2010). Traffic-related air pollution: A Critical review of the literature on emissions, exposure, and health effect. Boston, MA, HEI.

Heist, D., V. Isakov, S. Perry, M. Snyder, A. Venkatram, C. Hood, J. Stocker, D. Carruthers, S. Arunachalam and R. C. Owen (2013). "Estimating near-road pollutant dispersion: A model inter-comparison." Transportation Research Part D-Transport and Environment **25**: 93.

Hoek, G., R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer and D. Briggs (2008). "A review of land-use regression models to assess spatial variation of outdoor air pollution." Atmospheric Environment **42**(33): 7561.

Hu, S. S., S. Fruin, K. Kozawa, S. Mara, S. E. Paulson and A. M. Winer (2009). "A wide area of air pollutant impact downwind of a freeway during pre-sunrise hours." Atmospheric Environment **43**(16): 2541-2549.

Huang, Y. L. and S. Batterman (2000). "Residence location as a measure of environmental exposure: a review of air pollution epidemiology studies." J Expo Anal Environ Epidemiol **10**(1): 66-85.

Institute, H. E. (2010). Traffic-related air pollution: A critical review of the literature on emissions, exposure, and health effects. Boston, Mass., Health Effects Institute.

Isakov, V., S. Arunachalam, S. Batterman, S. Bereznicki, J. Burke, K. Dionisio, V. Garcia, D. Heist, S. Perry, M. G. Snyder and A. Vette (2014). "Air Quality Modeling in Support of the Near-Road Exposures and Effects of Urban Air Pollutants Study (NEXUS)." International Journal of Environmental Research and Public Health submitted.

- Isakov, V., J. S. Touma, J. Burke, D. T. Lobdell, T. Palma, A. Rosenbaum and H. Ozkaynak (2009). "Combining Regional- and Local-Scale Air Quality Models with Exposure Models for Use in Environmental Health Studies." Journal of the Air & Waste Management Association **59**(4): 461-472.
- Jerrett, M., A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuvaroglu, J. Morrison and C. Giovis (2005). "A review and evaluation of intraurban air pollution exposure models." Journal of exposure analysis and environmental epidemiology **15**(2): 185.
- Karner, A. A., D. S. Eisinger and D. A. Niemeier (2010). "Near-roadway air quality: synthesizing the findings from real-world data." Environ Sci Technol **44**(14): 5334-5344.
- Kioumourtzoglou, M.-A., D. Spiegelman, A. A. Szpiro, L. Sheppard, J. D. Kaufman, J. D. Yanosky, R. Williams, F. Laden, B. Hong and H. Suh (2014). "Exposure measurement error in PM_{2.5} health effects studies: A pooled analysis of eight personal exposure validation studies." Environmental Health: A Global Access Science Source **13**: 2.
- Özkaynak, H., L. K. Baxter, K. L. Dionisio and J. Burke (2013). "Air pollution exposure prediction approaches used in air pollution epidemiology studies." Journal of exposure science & environmental epidemiology **23**(6): 566.
- Rioux, C. L., D. M. Gute, D. Brugge, S. Peterson and B. Parmenter (2010). "Characterizing urban traffic exposures using transportation planning tools: an illustrated methodology for health researchers." Journal of urban health : bulletin of the New York Academy of Medicine **87**(2): 167.
- Sheppard, L., R. T. Burnett, A. A. Szpiro, S.-Y. Kim, M. Jerrett, C. A. Pope III and B. Brunekreef (2012). "Confounding and exposure measurement error in air pollution epidemiology." Air Quality, Atmosphere & Health **5**(2): 203.
- Snyder, M., S. Arunachalam, V. Isakov, D. Heist, S. Batterman, K. Talgo, R. Ganguly and P. Harbin (2013a). Sensitivity Analysis of Dispersion Model Results in the NEXUS Health Study due to Uncertainties in Traffic-related Emissions Inputs. Air & Waste Management Association Annual Conference & Exhibition, June 25-28th, 2013.
- Snyder, M. G., A. Venkatram, D. K. Heist, S. G. Perry, W. B. Petersen and V. Isakov (2013b). "RLINE: A line source dispersion model for near-surface releases." Atmospheric Environment **77**(0): 748-756.
- Venkatram, A., M. G. Snyder, D. K. Heist, S. G. Perry, W. B. Petersen and V. Isakov (2013). "Re-formulation of plume spread for near-surface dispersion." Atmospheric Environment **77**(0): 846-855.
- Vette, A., J. Burke, G. Norris, M. Landis, S. Batterman, M. Breen, V. Isakov, T. Lewis, M. I. Gilmour, A. Kamal, D. Hammond, R. Vedantham, S. Bereznicki, N. Tian and C. Croghan (2013). "The Near-Road Exposures and Effects of Urban Air Pollutants Study (NEXUS): study design and methods." The Science of the total environment **448**: 38.
- Wang, R., S. B. Henderson, H. Sbihi, R. W. Allen and M. Brauer (2013). "Temporal stability of land use regression models for traffic-related air pollution." Atmospheric Environment **64**: 312.
- Zhu, Y., T. Kuhn, P. Mayo and W. C. Hinds (2006). "Comparison of daytime and nighttime concentration profiles and size distributions of ultrafine particles near a major highway." Environ Sci Technol **40**(8): 2531-2536.

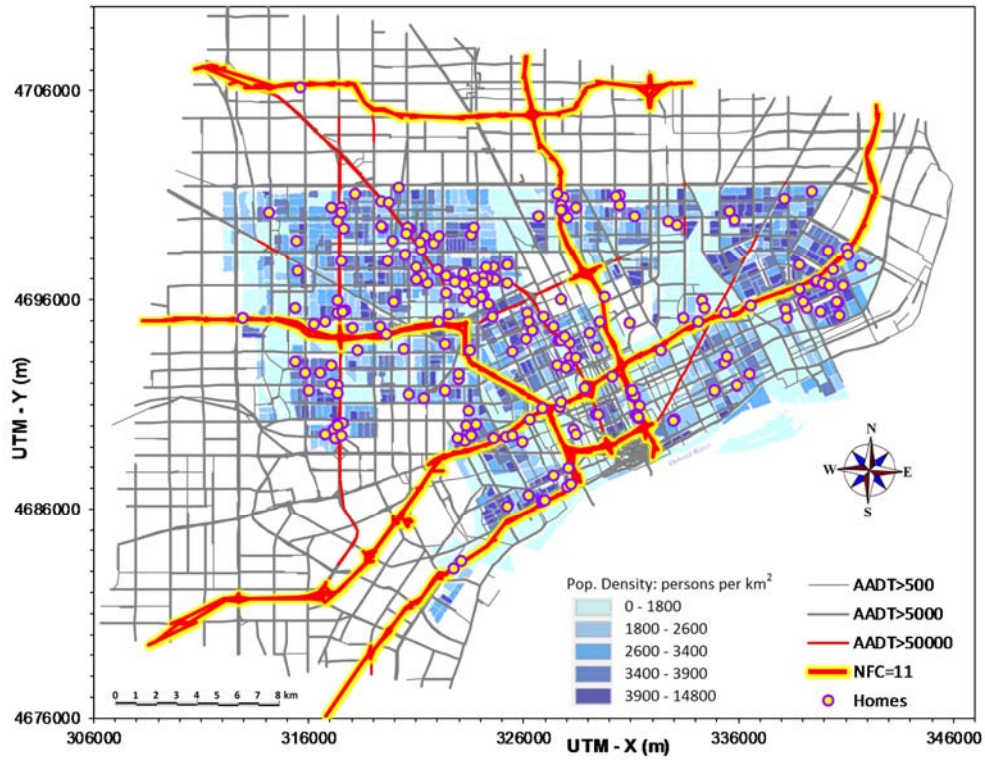


Figure 1. Map of modeled road network in study area, and locations of 218 homes of participants in NEXUS. Shaded area defines city of Detroit and population by Census Block group. Axis scales are Universal Transverse Mercator coordinates (m). Highlighted roads are NFC=11 (HDHT).

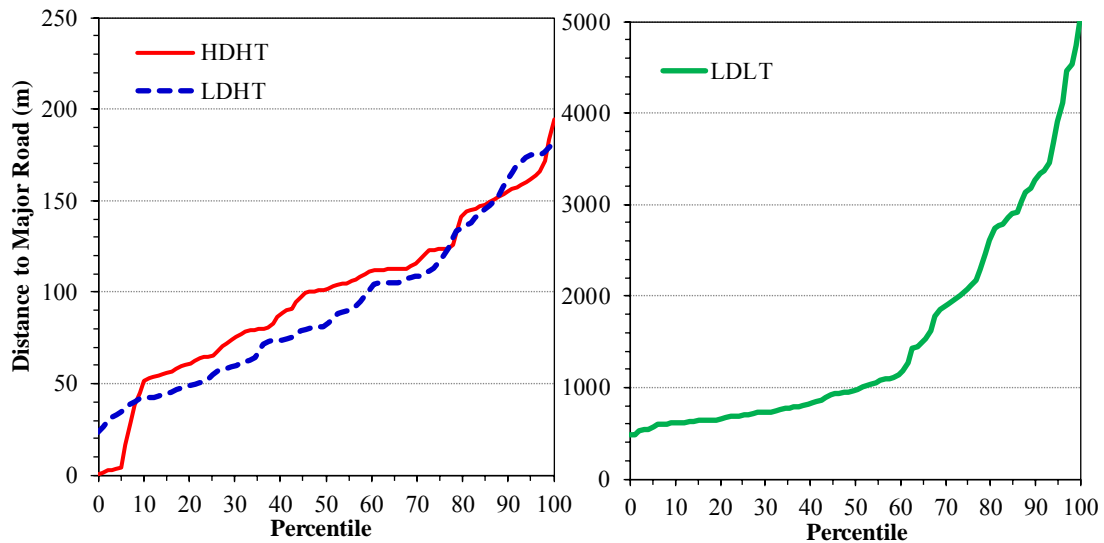


Figure 2. Distribution of distances of homes to major roads for the three traffic exposure groups (HDHT=high diesel/high traffic; LDHT=low diesel/high traffic; LDLT=low diesel/low traffic). Based on GPS home location and road edge.

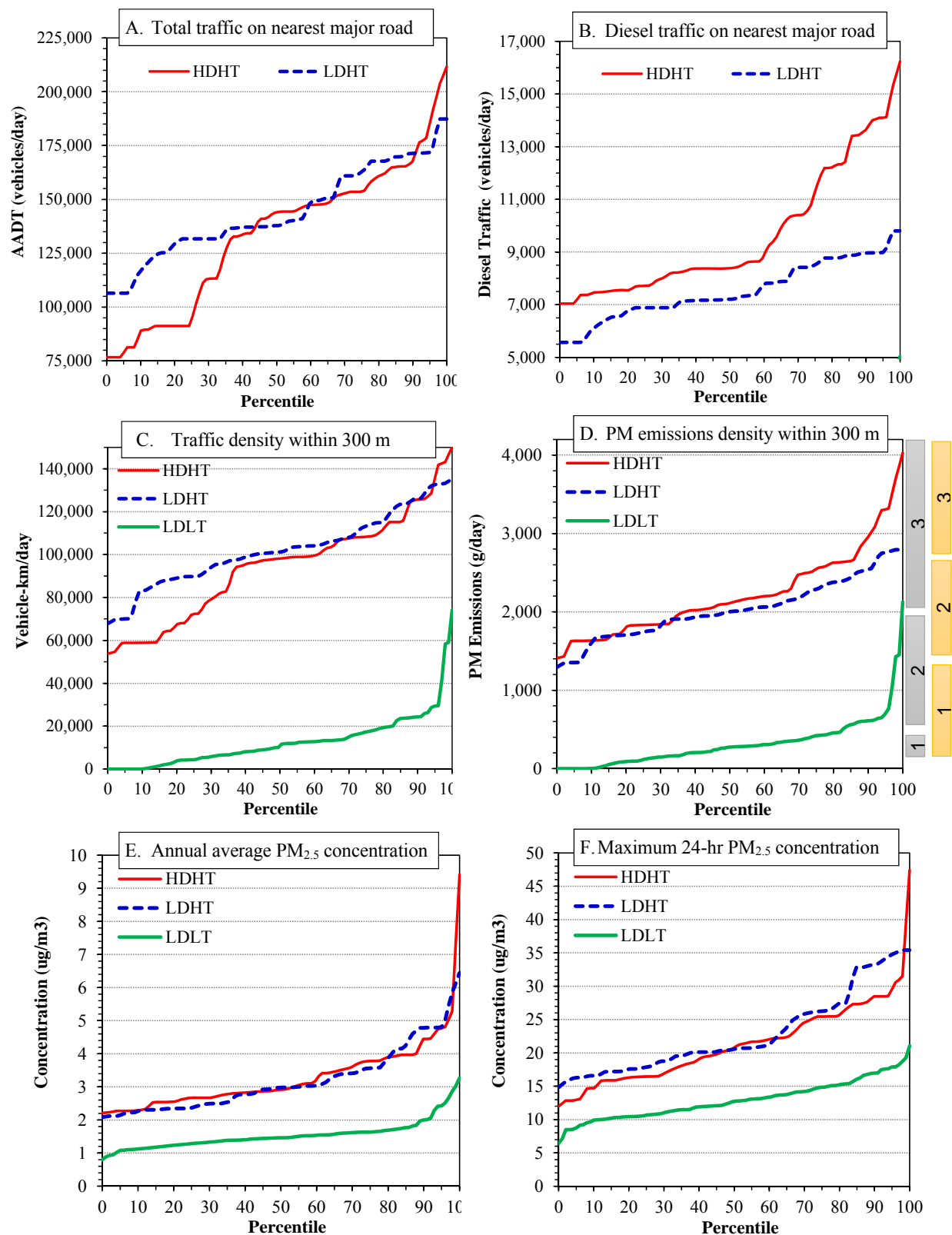


Figure 3. Distributions of selected exposure metrics coded by home group (HDHT=high diesel/high traffic; LDHT=low diesel/high traffic; LDLT=low diesel/low traffic). A. Total traffic volume on major roads nearest to children's homes. B. Diesel traffic volume on major roads nearest children's home. C. Traffic density for roads within 300 m of the home. D. PM_{2.5} emissions density for roads within 300 m of the home. E. Annual average PM_{2.5} concentrations due to on-road emissions. F. Maximum 24-hour average PM_{2.5} concentrations due to on-road emissions. Panel D also depicts differences between tertiles (in grey) and thirds (in orange).

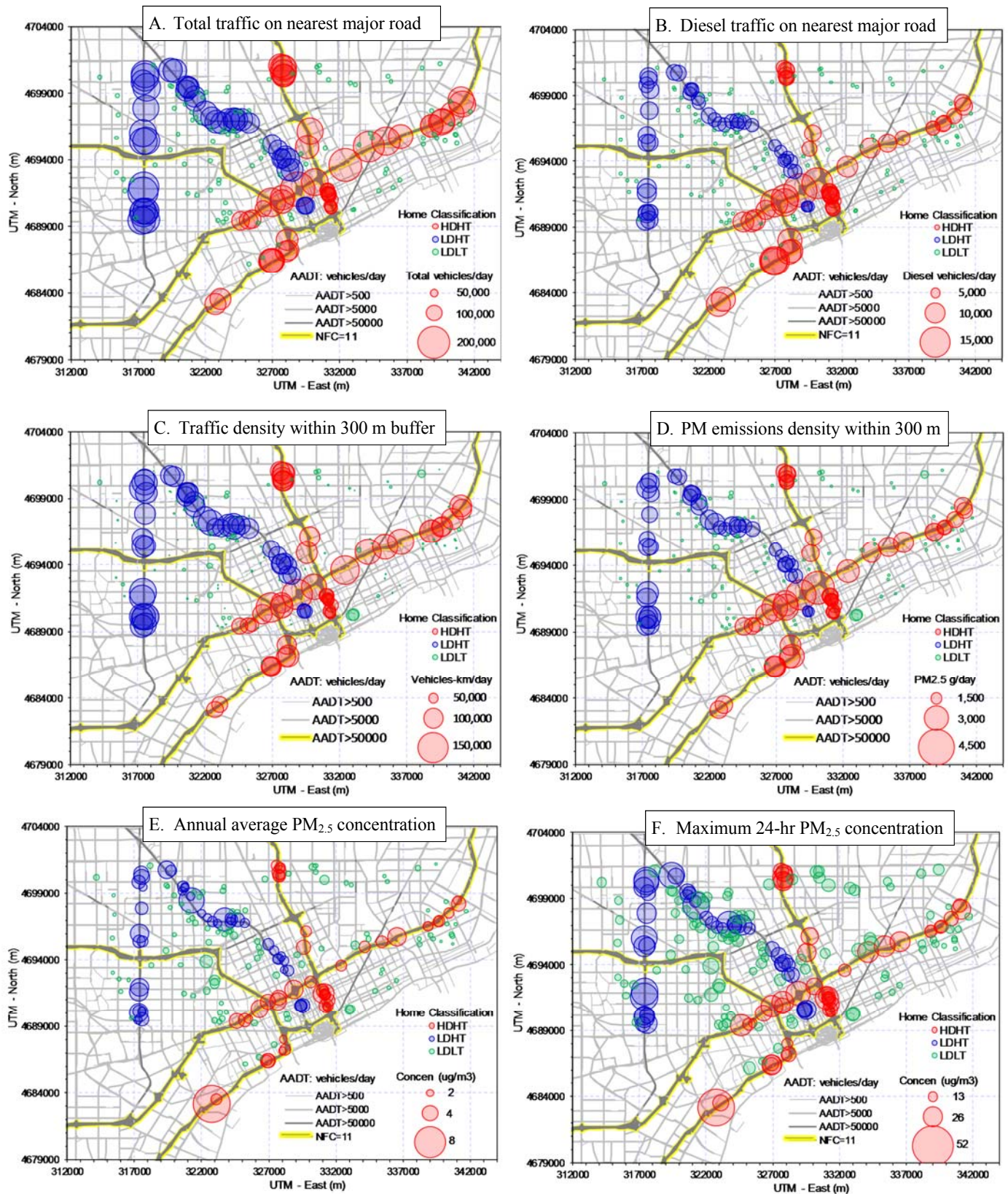


Figure 4. Depiction of various exposure metrics mapped and coded by home group (HDHT=high diesel/high traffic; LDHT=low diesel/high traffic; LDLT=low diesel/low traffic). High diesel roads are highlighted in yellow. A. Total traffic volume on major roads nearest to children's homes. B. Diesel traffic volume on major roads nearest children's home. C. Traffic density for roads within 300 m of the home. D. PM_{2.5} emissions density for roads within 300 m of the home. E. Annual average PM_{2.5} concentrations due to on-road emissions. F. Maximum 24-hour average PM_{2.5} concentrations due to on-road emissions. In A and B, LDLT homes are shown as green circles without coding for traffic volume.

	All Traffic (vehicles/day)			Diesel Traffic (vehicles/day)			Number of Lanes		
	HDHT	LDHT	All	HDHT	LDHT	All	HDHT	LDHT	All
Average	133,737	143,965	138,638	9,663	7,529	8,640	7.8	6.6	7.2
St. Dev.	34,962	21,614	29,634	2,510	1,130	2,237	1.8	1.1	1.6
Minimum	76,723	106,508	76,723	7,043	5,570	5,570	6	6	6
25th Percentile	94,202	131,718	124,586	7,716	6,889	7,182	6	6	6
Median	144,013	137,845	140,722	8,386	7,209	8,218	8	6	6
75th Percentile	153,576	162,808	160,968	11,297	8,515	8,974	9	7	8
95th Percentile	185,442	171,849	180,417	14,098	8,988	13,711	11	9	10
Maximum	211,750	187,373	211,750	16,235	9,800	16,235	14	10	14
Number	50	46	96	50	46	96	50	46	96

Table 1. Statistics of AADT volume, diesel traffic volume, and number of lanes for the nearest highways at the high traffic (HDHT and LDHT) homes.

	Onroad Sources, Annual Ave.			Onroad Sources, 24-hr Peak			Hybrid, Annual Average		
	HDHT	LDHT	LDLT	HDHT	LDHT	LDLT	HDHT	LDHT	LDLT
Average	3.3	3.2	1.5	21.4	22.8	12.9	15.6	15.6	13.9
St. Dev.	1.2	1.0	0.4	6.3	6.2	2.9	1.4	1.5	1.3
Minimum	2.2	2.1	0.8	12.0	14.9	6.4	13.3	13.7	12.1
25th Percentile	2.7	2.4	1.3	16.4	17.9	10.7	14.6	14.4	13.0
Median	2.9	3.0	1.5	20.8	20.6	12.7	15.4	15.3	13.6
75th Percentile	3.8	3.6	1.6	25.4	26.2	14.8	16.2	16.3	14.4
95th Percentile	4.8	4.8	2.4	29.6	34.7	17.9	18.2	18.6	16.5
Maximum	9.4	6.4	3.3	47.4	35.4	21.1	20.7	19.2	19.7
Number	50	46	102	50	46	102	50	46	102

Table 2. Statistics of PM_{2.5} concentrations ($\mu\text{g}/\text{m}^3$) predicted at home of NEXUS participants, classified by exposure group (HDHT=high diesel/high traffic; LDHT=low diesel/high traffic; LDLT=low diesel/low traffic). The hybrid annual average includes on-road, area, point and regional sources.

	Group1 (group)	Group2 (group)	Distance (m)	AADT (veh/day)	Lanes (no)	Diesel (veh/day)	VKT (km/day)	PMemis (g/day)	COemis (g/day)	PMave (ug/m3)	PMmax (ug/m3)	PMtot (ug/m3)
Spearman correlation coefficients												
Group1	1.00											
Group2	0.95	1.00										
Distance	-0.81	-0.87	1.00									
AADT	-0.12	(a)	0.06	1.00								
Lanes	0.39	(a)	0.31	-0.08	1.00							
Diesel	0.47	(a)	-0.02	0.66	0.12	1.00						
VKT	0.80	0.86	-0.66	0.50	-0.43	0.26	1.00					
PMemis	0.83	0.85	-0.66	0.36	-0.33	0.47	0.98	1.00				
COemis	0.79	0.86	-0.66	0.48	-0.43	0.25	1.00	0.98	1.00			
PMave	0.78	0.82	-0.75	0.15	-0.24	0.23	0.78	0.78	0.78	1.00		
PMmax	0.70	0.76	-0.71	0.35	-0.23	0.27	0.74	0.73	0.75	0.90	1.00	
PMtot	0.56	0.59	-0.57	0.13	-0.08	0.17	0.55	0.54	0.55	0.74	0.71	1.00
Kendall Tau-B coefficients matrix												
Group1	1.00											
Group2	0.90	1.00										
Distance	-0.62	-0.71	1.00									
AADT	-0.10	(a)	0.04	1.00								
Lanes	0.36	(a)	0.23	-0.07	1.00							
Diesel	0.39	(a)	-0.02	0.60	0.09	1.00						
VKT	0.61	0.71	-0.42	0.42	-0.34	0.19	1.00					
PMemis	0.66	0.70	-0.42	0.28	-0.26	0.36	0.90	1.00				
COemis	0.60	0.70	-0.41	0.40	-0.34	0.18	0.98	0.91	1.00			
PMave	0.62	0.67	-0.58	0.12	-0.18	0.16	0.55	0.56	0.55	1.00		
PMmax	0.54	0.62	-0.52	0.27	-0.18	0.19	0.53	0.52	0.53	0.74	1.00	
PMtot	0.44	0.49	-0.40	0.09	-0.06	0.12	0.38	0.37	0.38	0.56	0.52	1.00

Table 3. Comparison of exposure metrics at the NEXUS homes using Spearman (top) and Kendall τ -b correlation coefficients. Shaded numbers show absolute value of correlations above 0.6 and 0.8. Variables: Group1 = LDLT, LDHT and HDHT assigned 1, 2, 3, respectively; Group2 = LDLT, LDHT and HDHT assigned 1, 2, 2, respectively; Distance = distance to nearest major highway (LDHT and HDHT only); AADT = traffic volume on nearest major road (LDHT and HDHT only); Lanes = number of lanes on nearest major road (LDHT and HDHT only); Diesel = diesel vehicles volume on nearest major road (LDHT and HDHT only); VKT = traffic density as vehicles-km-traveled/day in 300 m buffer around home; PMemis = PM_{2.5} emissions in 300 m buffer; COemis = CO emissions in 300 buffer; PMave = 2010 annual average PM_{2.5} concentration due to local traffic; PMmax = 2010 maximum 24-hr average PM_{2.5} concentration from local traffic; PMtot = 2010 annual average PM_{2.5} concentration from all sources. Sample size is n=218, except for comparisons involving AADT, Lanes, and Diesel metrics where n=116 since only high traffic homes are considered. (a) not calculated due to variable definition.

	Group1 (group)	Group2 (group)	Distance (m)	AADT (veh/day)	Lanes (no)	Diesel (veh/day)	VKT (km/day)	PMemis (g/day)	COemis (g/day)	PMave (ug/m3)	PMmax (ug/m3)	PMtot (ug/m3)
Agreement among tertiles (percent)												
Group1	100											
Group2	75	100										
Distance	58	44	100									
AADT	28	36	33	100								
Lanes	24	26	18	38	100							
Diesel	33	39	27	56	34	100						
VKT	52	45	49	34	19	37	100					
PMemis	56	45	49	37	17	43	89	100				
COemis	53	45	48	33	21	35	96	90	100			
PMave	60	51	66	32	20	34	58	61	59	100		
PMmax	54	45	62	41	18	35	58	56	57	75	100	
PMtot	55	52	58	42	22	36	51	50	50	64	64	100
Agreement with thirds (percent)												
Group1	100											
Group2	75	100										
Distance	31	7	100									
AADT	45	71	16	100								
Lanes	15	33	15	34	100							
Diesel	53	76	21	67	40	100						
VKT	70	76	28	35	28	42	100					
PMemis	76	90	14	60	31	74	80	100				
COemis	70	73	30	33	28	38	97	78	100			
PMave	64	80	12	47	33	50	70	77	69	100		
PMmax	65	76	15	60	30	48	68	76	66	77	100	
PMtot	56	68	13	44	39	43	61	62	59	72	67	100

Table 4. Comparison of exposure metrics at the NEXUS homes showing concordance with classifications using tertiles and “thirds”. Shaded numbers show percentage agreement above 60 and 80%. Variables and sample size are defined in Table 3.

Exposure Metric (as defined for NEXUS)	Strengths	Limitations	Results in Detroit
1 Distance to major road (distance from home to road edge, and distance from home to road centerline, using GPS home location)	Simple to construct. Low data needs. Can potentially distinguish roads with varying traffic volume, vehicle mix, or other characteristics.	Distance limit used as cutoffs for classifying homes/receptors is arbitrary. May not consider traffic volume, vehicle mix, and other factors. Sensitivity to distance calculation, e.g., using road edge or centerline.	HDHT and LDHT roads had comparable distances to homes. LDLT distances considerably exceeded HDHT and LDHT groups, by design and recruitment approach.
2 Total traffic volume on nearby roads (AADT roads within 200 m of homes, using nearest road edge and GPS home location)	Relatively simple to construct. Reasonably good traffic volume estimates on major roads. Could potentially select period of day, e.g., rush-hour.	Traffic volume estimates needed. Distance criterion used to determine road is arbitrary. Does not provide metric for low traffic groups.	HDHT and LDHT groups largely indistinguishable. HDHT group had considerable range.
3 Diesel traffic volume on nearby roads (roads within 200 m of homes using road edge and GPS home location)	Relatively simple to construct. May relate to PM emissions from diesel traffic. Could potentially select period of day.	Difficult to estimate diesel traffic volume accurately. Does not account for type of diesel vehicles and emissions. Otherwise as 2 above.	HDHT and LDHT groups were largely indistinguishable. HDHT group had roughly 10-20% higher diesel volumes than LDHT group, but about 2/3 of the values overlapped.
4 Local traffic density (AADT on road segments with 300 m distance (buffer) around each home, based on distance to road centerline, GPS home location, and traffic-demand model estimates of AADT)	Includes local traffic emissions that might affect receptor. Result (VKT/day) is easily interpretable and possibly generalizable. Large range across sites. Can be applied to irregular shapes, including sources and receptors. Can select period of day. Relevant to traffic analysis zones used by planners.	Moderately high data needs and computationally intensive. Sensitive to distance criterion, which is somewhat arbitrary. Uncertainty of traffic estimates on all but major roads. Excludes smallest roads.	LDHT group had slightly greater exposure than the HDHT group All but a few LDLT homes had low values
5 Emissions on local roads (as 4 above with addition of annual average road-link emissions estimates for PM _{2.5} , NO _x and CO)	Incorporates vehicle emissions of pollutants of interest. Reflects vehicle mix on roads. Also as 4 above	Results depend on pollutant, to an extent. High data needs and computationally intensive. Difficult to estimate emissions accurately.	For PM _{2.5} and NO _x , HDHT had slightly higher exposure than LDLT. For CO, results are reversed but very similar All but a few LDLT homes had much lower values.
6 Pollutant concentration predictions (PM _{2.5} predictions at homes used road-link emissions inventory for traffic emissions with RLINE dispersion model; area and point sources using AERMOD and regional sources handled using CMAQ and kriging interpolations of monitoring data)	Incorporates effects of emissions, meteorology, and location in a physically-based approach. Quantification and apportionment of ambient concentrations due to different sources of interest, e.g., traffic. Can be derived for specific periods of day, season or year, e.g., daily predictions at rush hour periods. Inter-study comparisons are possible and meaningful.	Results depend on pollutant, averaging time, and statistic. High data needs and computationally intensive. Uncertainty not well characterized. Results potentially sensitive to many factors, including home placement.	For PM _{2.5} , HDHT and LDHT distributions were very similar although some dependence on averaging time and statistic. PM _{2.5} contributions from local traffic at HDHT and LDHT homes were about twice that at the LDLT homes. Regional sources provide large fraction (80%) of total PM _{2.5} concentration, but smaller contributions of NO _x and CO.

Table 5. Summary of metrics used for exposure to traffic-related air pollutants in NEXUS.

Supplemental Tables and Figures

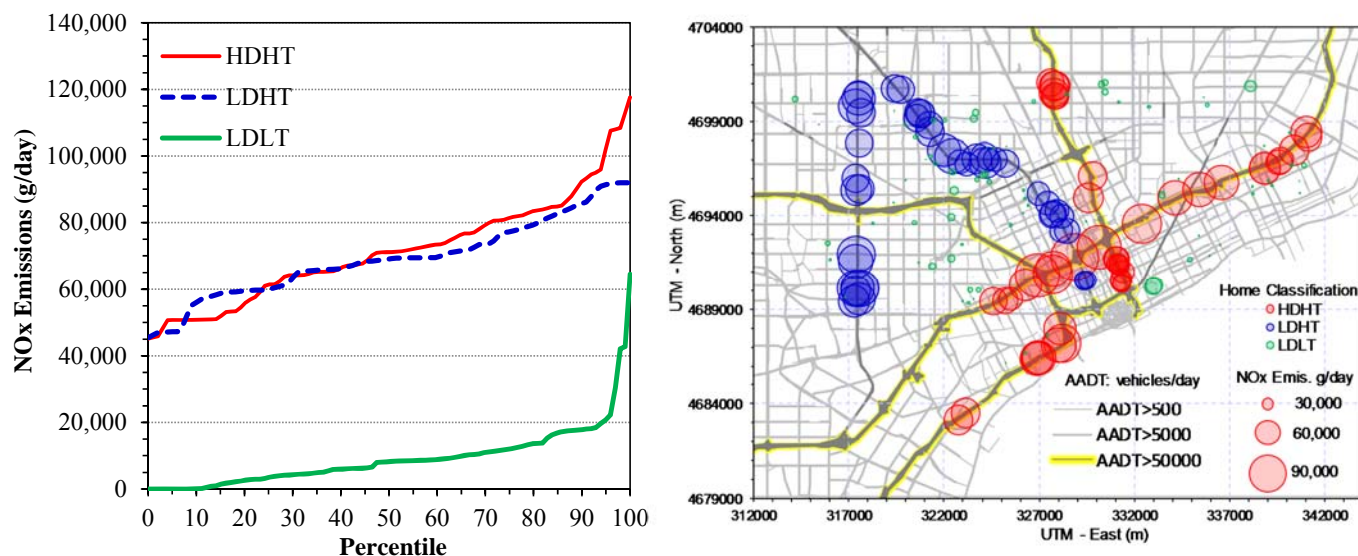


Figure S1. Distribution and maps of NO_x emissions, defined as g/day for roads within 300 m of the home, coded by home group. High diesel roads are highlighted in yellow.

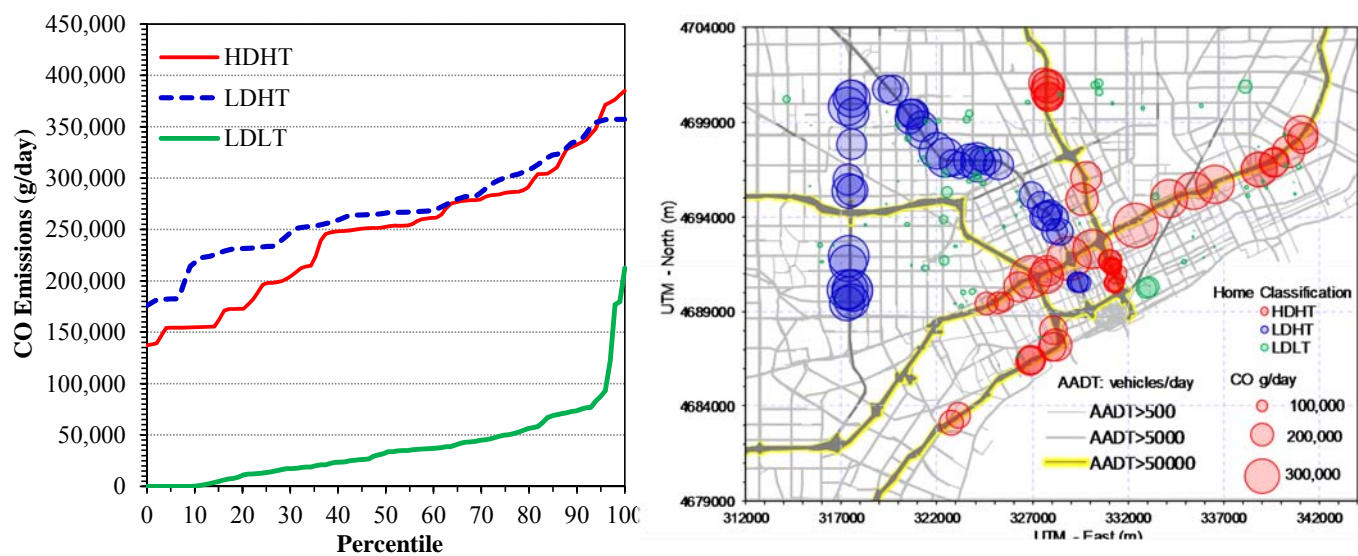


Figure S2. Distribution and maps of CO emissions, defined as g/day for roads within 300 m of the home, coded by home group. High diesel roads are highlighted in yellow.

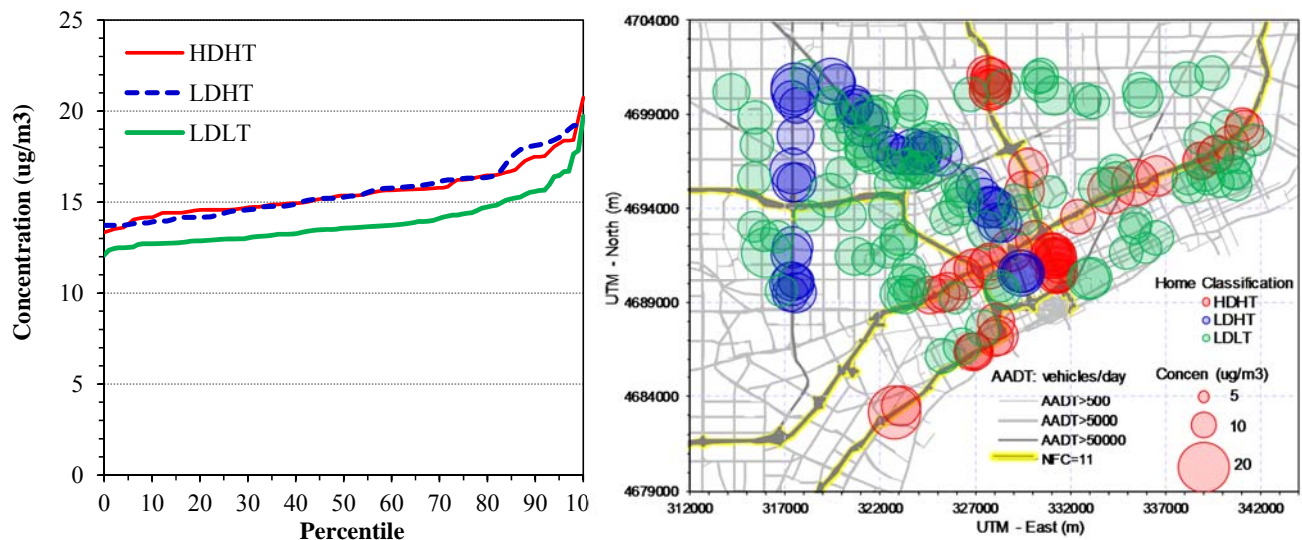


Figure S3. Distribution and maps of total annual average $PM_{2.5}$ predicted for 2010 from hybrid model. Each figure is coded by home group. High diesel roads are highlighted in yellow.

	Group	Distance	AADT	Lanes	Diesel	VKT	PM Emis	CO Emis	PM Max
	(group)	(m)	(veh/day)	(no)	(veh/day)	(km/day)	(g/day)	(g/day)	(ug/m3)
Severe disagreement among tertiles (percent)									
Distance	51								
AADT	11	20							
Lanes	31	7	19						
Diesel	20	20	4	16					
VKT	47	1	11	9	22				
PM Emis	48	1	15	9	19	0			
CO Emis	47	1	15	10	19	0			
PM Max	46	2	15	8	16	1	1	1	
PM Ave	53	3	16	3	18	1	1	1	0
Severe disagreement with thirds (percent)									
Distance	33								
AADT	66	34							
Lanes	82	26	1						
Diesel	57	25	0	3					
VKT	1	33	51	65	50				
PM Emis	0	33	46	61	44	0			
CO Emis	1	33	52	65	51	0	0		
PM Max	8	42	44	56	43	3	1	3	
PM Ave	4	46	50	62	48	6	2	7	0

Table S1. Comparison of exposure metrics for NEXUS homes using rates of “severe” exposure misclassification in percent for tertiles and “thirds” groups. Shaded numbers show scores above 60 and 80%. Variable definitions as follows: Group = LDLT, LDHT and HDHT (assigned 1, 2, 3) categorical classification; Distance = m to nearest major highway; AADT = total traffic on nearest major highway; Lanes = number of lanes on nearest major highway; Diesel = diesel vehicles on nearest major highway; VKT = traffic density defined as vehicles-kilometers-traveled/day in 300 m buffer around home; PM Emis = $PM_{2.5}$ emissions in same buffer; CO Emis = CO emissions in same buffer; PM Max = maximum 24-hr $PM_{2.5}$ concentration; PM Ave = 2010 annual average $PM_{2.5}$ concentration. Sample size is n=218, except for comparisons involving AADT, Lanes, Diesel metrics where n=116 due to use of only high traffic homes.