9 Supplemental Materials

9.1 Roadway links and traffic activity

A detailed, link-based NO_x emission inventory was compiled for roads in Detroit and surrounding Wayne County for the year 2010. NO_x was selected as an air pollutant representative of emissions from traffic. A major part of this link-based emission inventory was the road network, including the locations of individual road links, link type based on the National Functional Classification (NFC), annual average daily traffic (AADT), and average speed information. These data were obtained for 9,701 road links that represent the study area (Figure 1). For the larger roads, e.g., major arterials and interstate highways, each traffic direction and/or group of lanes used separate links. All but local neighborhood streets and alleys were included in this network. AADT and speed data were derived using road counts and travel demand modeling (TDM) with link-specific inputs, e.g., AADT, number of lanes, roadway type and location, data provided by the Southeast Michigan Council of Governments (SEMCOG), the Michigan Department of Transportation (MDOT), and the US EPA Office of Transportation and Air Quality (OTAQ). The average speed for each link was estimated for four time periods: morning rush hour peak (7-9 AM), mid-day (9 AM – 3 PM), afternoon rush hour peak (3 PM – 6 PM), and off-peak (6 PM – 7 AM).

Hourly traffic volume, fleet mix, and vehicle speed were estimated for each link as:

$$V_{i,k,t} = FM_{NFC(i),k}\ MAF_{MON(t)}\ DAF_{k,DAY(t)}\ HAF_{NFC(i),t}\ AADT_i \qquad (1)$$

where $V_{i,k,t}$ (counts h^{-1}) = number of vehicles on link i (i = 1...9701) for vehicle class k (k = 1…8) and hour of the year t (t = 1...8760); $FM_{NFC(i),k}$ (dimensionless) = fleet mix allocation factor; $MAF_{MON(t)}$ $DAF_{k,DAY(t)}$, and $HAF_{NFC(i),t}$ are monthly, daily and hourly temporal allocations; and $AADT_i$ = annual average daily flow. The 8 vehicle classes are aggregations from the MOVES emissions model that represent motorcycles, light-duty gasoline vehicles, light-duty diesel vehicles, light-duty gasoline trucks with gross vehicle weight (GVW) less than 6001 pounds, light-duty gasoline trucks with GWV>6001 pounds, light-duty diesel trucks, heavy-duty diesel trucks, heavy-duty gas vehicles, and heavy-duty diesel vehicles (MC, LDGV, LDDV, LDGT1, LDGT2, LDDT, HDGV, HDDV, respectively). These classes were derived using state-level data from the US Federal Highway Administration, and information from the US EPA Emission Inventory Improvement Program. NFC designations included classes 11, 12, 14, 16, 17, 19 and 90, which represent interstates, other freeways, other principal arterials, minor arterials, major collectors, minor collectors and bridges, respectively.

The fleet mix allocation factor, $FM_{NFC(i),k}$, in eq. (1) gives the fraction of vehicles in vehicle class k for link i. This depends on the link's NFC designation, using allocation factors from Table VM-4 from the FHWA Highway Statistics Series (http://www.fhwa.dot.gov/policyinformation/ statistics/2010/vm4.cfm) and information from the US EPA Emission Inventory Improvement Program (http://www.epa.gov/ttn/chief/eiip/). Summed across the 8 vehicle classes, $\Sigma_{k=1...8}\ FM_{NFC(i),k} = 1$ for each road link. Only 3 short (<245 m) links were designated as NFC=90, for which NFC=11 was substituted, which had the highest fraction of diesel vehicles.

The temporal allocation factors account for variation in traffic flow by month, day of the week, and hour of the day. The monthly factor $MAF_{MON(t)}$ ranged from 0.86 (December) to 1.10 (August), reflecting higher summer traffic. (Summed across the 12 months, $\Sigma_{t=1..12}\ MAF_t = 12$). The day-of-week factor $DAF_{k,DAY(t)}$ portrayed slightly increased daily total volume for most vehicle classes on Friday (by 8%), and decreased volumes on Saturday (by 9%) and Sunday (21%), compared to other weekdays. Patterns for HDGV and HDDV classes differed: volumes were slightly lower flows on Friday (by 3%) and significantly lower on Saturday and Sunday (61 and 71%, respectively). (Summed across the 7 days in a week, $\Sigma_{t=1..7}\ DAF_{k,DAY(t)} = 7$ for each vehicle class k.) The hour-of-day factor $HAF_{NFC(i),HR(t),DT(t)}$ gave the diurnal pattern (HR(t) = 1...24) for three day types (DTs, weekdays, Saturday, and Sunday), and was obtained from the SMOKE modeling system (http://www.cmascenter.org/smoke/). Generally, weekday patterns were bimodal with morning and afternoon rush hour peaks; weekends patterns were unimodal with a broad afternoon peak, but patterns vary by road type as given by NFC. (Summed across the 24 hours in a day, $\Sigma_{t=1\ldots24}\ HAF_{NFC(i),DT(t),HR(t)} = 1$ for each NFC and DT.) For holidays, a Sunday schedule was assumed ($DAF_{k,DAY(t)}$ and $HAF_{NFC(i),HR(t),DT(t)}$ were set to Sunday values). Holidays in year 2010 were New Year's Day (Jan. 1), Memorial Day (May 31), Independence Day (July 5), Thanksgiving (Nov. 25), and Christmas (Dec. 25). We confirmed that eq. (1) obtained the correct AADT by summing link specific-flows over vehicle classes and hours of the year:

$$AADT_i \approx 365^{-1}\ \Sigma_{k=1..8}\ \Sigma_{t=1,8760}\ V_{i,k,t} \qquad (2)$$

Because AADT does not account for holidays, eq. (2) is not an equality, although differences are small.

9.2 Emissions

Hourly emissions were estimated for each of the 9,701 links as follows. First, emission factors for primary exhaust emissions of each pollutant were calculated using MOVES2010a (http://www.epa.gov/otaq/models/moves/), which uses a power-based approach that varies by vehicle class, vehicle speed, ambient temperature, and fuel properties. Emission factors $EF_{k,\ SPEED,\ TEMP,\ MON}$ (g mile^{-1} vehicle^{-1}) were calculated for 8 vehicle class (k=1…8), 16 vehicle speeds (2.5, 5, 10, 15 … 75 mph), 11 ambient temperatures (0, 10, 20 ... 90, 100 ^{o}F), and 12 months (Jan. through Dec.) using monthly average properties for fuels in the modeling domain, based on survey information from SEMCOG. MOVES inputs were adjusted for the 2010 Detroit vehicle age distribution, based on an analysis of vehicle registration information by the Lake Michigan Air Directors' Consortium (LADCO). Emission factors from MOVES were applied to each road link to generate an hourly and link-based emissions inventory that accounted for traffic activity, i.e., the volume of each vehicle class and the average speed for each link and hour. Link-specific emission rates $E_{i,t}$ (g m^{-1} s^{-1}) for link i and hour t were calculated as:

$$E_{i,t} = 1.72604E\text{-}07\ \Sigma_{k=1\ldots8}\ EF_{k,\ SPEED(i,\ t),\ TEMP(t),\ MON(t)}\ V_{i,k,t} \qquad (3)$$

where the first constant converts units of distance (1 mile/1609 m) and time (1 h/3600 s) to match the vehicle counts and the MOVES emission factor units; $EF_{k,SP0EED(i,\ t),\ TEMP(t),Month(h)}$ is the emission factor (g vehicle^{-1} mile^{-1}) from MOVES for link i, vehicle class k, link speed SPEED(i, t), hourly average ambient temperature TEMP(t), and month MON(t); and $V_{i,k,t}$ is the number of vehicles per hour for link i, vehicle class k, and hour t given in eq. (1). Temperature and vehicle speed were placed into 11 and 16 bins, described earlier, and lookup tables were used to select values. Temperatures were taken as the average across five airport weather stations in the Detroit area, which yielded a complete and robust dataset.

9.3 Dispersion Modeling

Pollutant concentrations from vehicle emissions were predicted using RLINE (http://www.cmascenter.org/r-line/), a research grade dispersion model for near-roadway assessments under development by US EPA. This steady-state plume-dispersion model incorporates newly developed algorithms for predicting concentrations from road sources, including at receptors very near roads and at 'upwind' receptors due to plume meandering. (Snyder, Venkatram et al. 2013, Venkatram, Snyder et al. 2013) RLINE utilizes numerical methods to integrate multiple point sources along a line source, or an analytical approximation that provides similar results. Dispersion parameters were derived from field data and recent wind tunnel experiments for near road sources. Hourly meteorological data were taken from Detroit City airport, which was determined to be representative of the study area after examining land use in the vicinity of the meteorological site, as well as correlations among surface measurements obtained at this and four other area meteorological monitoring sites. These data were processed by AERMET, which completed quality assessment checks, merged surface, upper air and on-site data, and estimated boundary layer parameters.

Several sets of receptors were used. (Each receptor represents a discrete point or location.) The first used extremely fine (10 m) resolution to model a high impact area, the intersection of I75 and I94. In this case, 12,221 receptors were laid out over a 1.0 x 1.2 km rectangular grid on 10 m centers (SW Universal

Transverse Mercator (UTM) coordinate of 330,000, 4,692,400). The second set used the same resolution to estimate concentrations in the vicinity of each of three NO_x monitoring sites in the Detroit area (121 receptors, 100 x 100 m area). The third modeled the entire Detroit area using 27,622 receptors over a 34.5 x 23.0 km grid on 150 m centers (SW UTM coordinate of 311,500, 4,680,500). The SE corner of this region covering portions of the Detroit River, Lake St. Clair and Canada was excluded. In all cases, the modeled road network extended well beyond the receptor network.

9.4 Computational Considerations

Estimating concentrations with high spatial resolution at the urban scale is computationally intensive. As an example, for the 9,701 road links, the 150 m grid spanning Detroit, and 8,760 hours per year, 2.34 trillion source-receptor calculations are needed to determine annual average concentrations. This problem would require literally centuries for a standard workstation. Calculations were speeded up by using the analytical approximation in RLINE; estimating annual averages based on a subset of meteorology, specifically, every 6^{th} day (found to provide representative results); revising the RLINE code to allow hourly emissions without post-processing; using look-up tables with precomputed emission profiles for each NFC and speed class combination; and breaking down receptors into several subsets that were run simultaneously on multiple computers. In addition, we considered only receptor-link distances less 25 km (roads at longer distances provide negligible impacts), and used an adaptive algorithm to further select link-receptor pairs based on distance and road link emissions. This algorithm was tailored to Detroit but is easily generalized. First, 1-hr concentrations were estimated for several "model" line sources (unitary emission rate, 20 and 2000 m long sources) using a radial receptor network (16 equally spaced “arms” with a logarithmic progression of distances and receptors from 10 m to 100 km from the source), RLINE, and 2010 Detroit meteorology. Then, the maximum annual average concentration at each distance was found and fitted to a simple model:

$$C_x = 10^{[2.68346\ -\ 1.45929\ \log10(x)\ ]} \qquad (4)$$

where C_x = fitted concentration at distance x (m), and the two coefficients are best fit parameters (calculated using linear regression on transformed variables). This model fitted the concentration envelope reasonably well (r=0.96) and reflects a roughly logarithmic decay of concentrations with distance. For example, $C_{x=20000m}$ is five orders of magnitude lower than $C_{x=0m}$. Next, the maximum possible near-field concentration for a high emitting source was calculated as the product of C_x with x=25 m, the highest hourly emission rate (over the year), and a relatively high volume link (AADT=100,000, NFC=11). Then, the potential impact of each link-receptor pair was estimated as the product of the link’s emission rate for that hour and C_x, using the link’s closest distance to the receptor (precomputed to save time). Only if the potential impact exceeded 1.5×10^{-5} of the maximum possible concentration was that link-receptor pair and hour calculated. Thus, only those link-receptor pairs and hours that can contribute at least 1.5×10^{-5} of the maximum concentration that can result from a high volume link are selected; low traffic and/or distant links are screened out. This algorithm slightly decreases concentrations at locations distant from large roads, but no significant changes are seen for receptors near larger roads. The 1.5×10^{-5} threshold, AADT=100,000, and x=25 m parameters are empirically determined values; smaller values increase computation time but lead to more accurate results.

The above steps dramatically increased computational speed and obtained results that were nearly identical to those using all source-receptor pairs and the numerical algorithm. Still, computation of annual averages for the largest receptor network required several days on a workstation.

9.5 Interpolation concentrations and estimating errors

As noted in the text, concentrations were estimated from concentrations predicted at receptors using two interpolation methods. The first uses the “nearest neighbor” (NN) technique and estimates the unknown as the concentration at the nearest receptor. Errors were calculated as the absolute difference in concentrations between all pairs of receptor separated by distance Δx:

$$\Delta C_{NN, i, \Delta x} = | C_i - C_{i+\Delta x} | \quad (5)$$

where C_i and $C_{i+\Delta x}$ = known concentrations at receptor i and a second receptor displaced in either E or N direction by Δx. The second interpolation method is equivalent to an inverse-distance-weighted (IDW) average, and the unknown concentration is estimated as average of concentrations at the four nearest receptors:

$$\Delta C_{ID, i, \Delta x} = | C_i - (C_{i+\Delta x} + C_{i-\Delta x} + C_{i+\Delta y} + C_{i-\Delta y})/4 | \quad (6)$$

where Δx and Δy=equal separation distances in E and W directions, respectively.

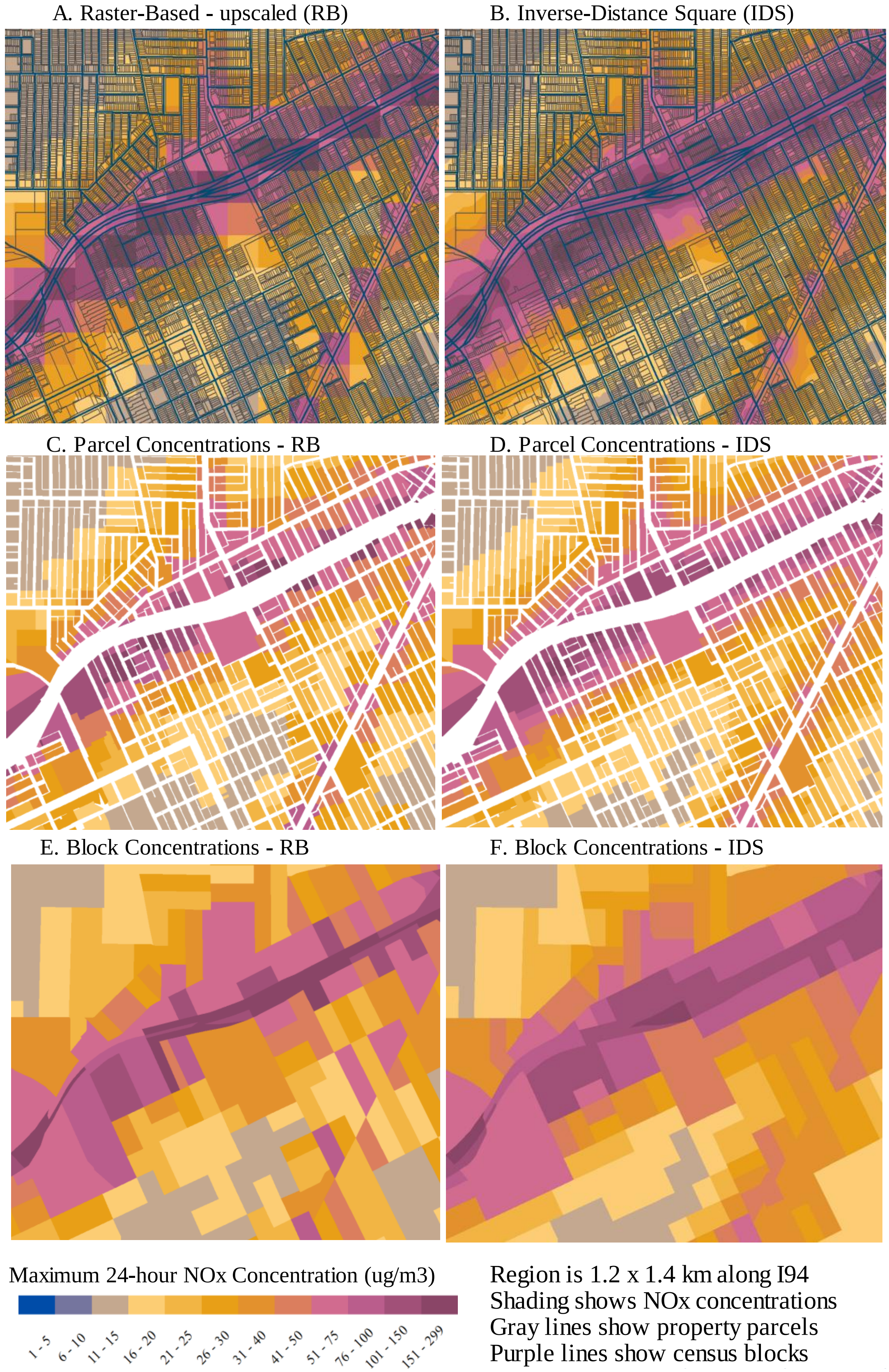
Concentration errors were expressed as absolute concentration differences, as given by eqs. (5) and (6), and by the relative absolute concentration differences (RΔC, %) by dividing by the average concentration:

$$R\Delta C_{NN,i,\Delta x} = 100\% \ \Delta C_{NN, i, \Delta x} / [(C_i + C_{i+\Delta x}) / 2] \quad (7)$$

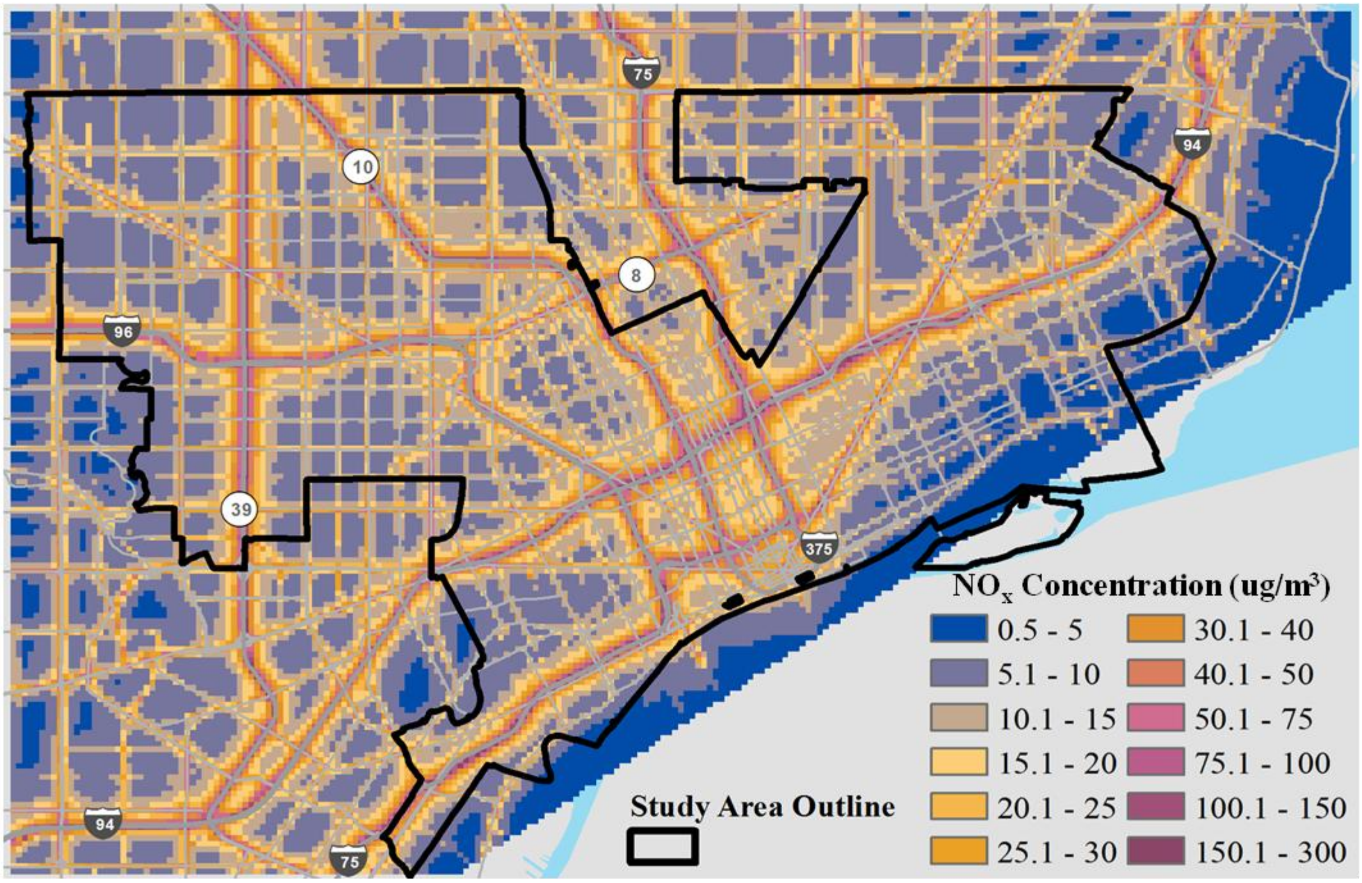
$$R\Delta C_{ID,i,\Delta x} = 100\% \ \Delta C_{ID, i, \Delta x} / [C_i /2+ (C_{i+\Delta x} + C_{i-\Delta x} + C_{i+\Delta y} + C_{i-\Delta y})/8] \quad (8)$$

The RΔC is equivalent to the absolute fractional bias. At finer spatial resolutions, i.e., as $\Delta x \rightarrow 0$, both error metrics approach 0.

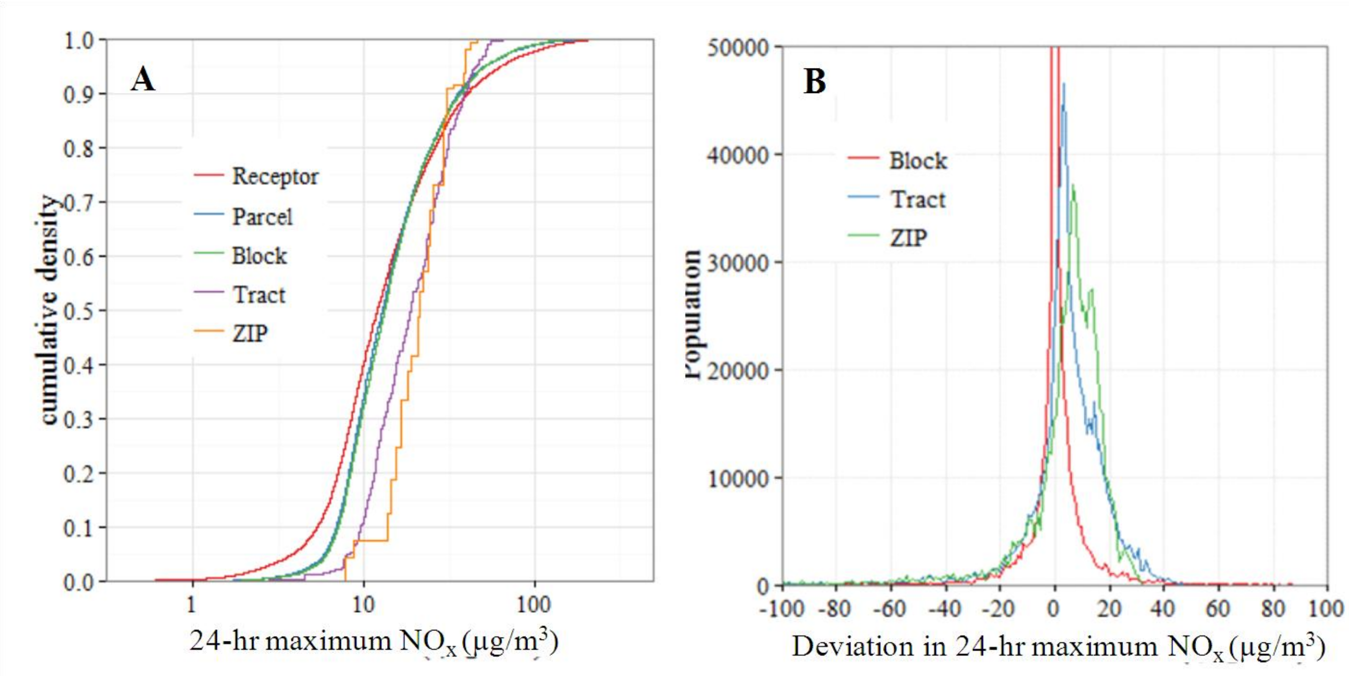
Supplemental Figure 1. Comparison of up-scaled raster-based (RB) and inverse-distance squared weighting (IDS) joining schemes for parcel- and block-level concentrations. Top figure shows underlying raster, roads, and parcel and block boundaries. Middle figures show estimated parcel-level concentrations. Bottom figures show estimated block-level concentrations.



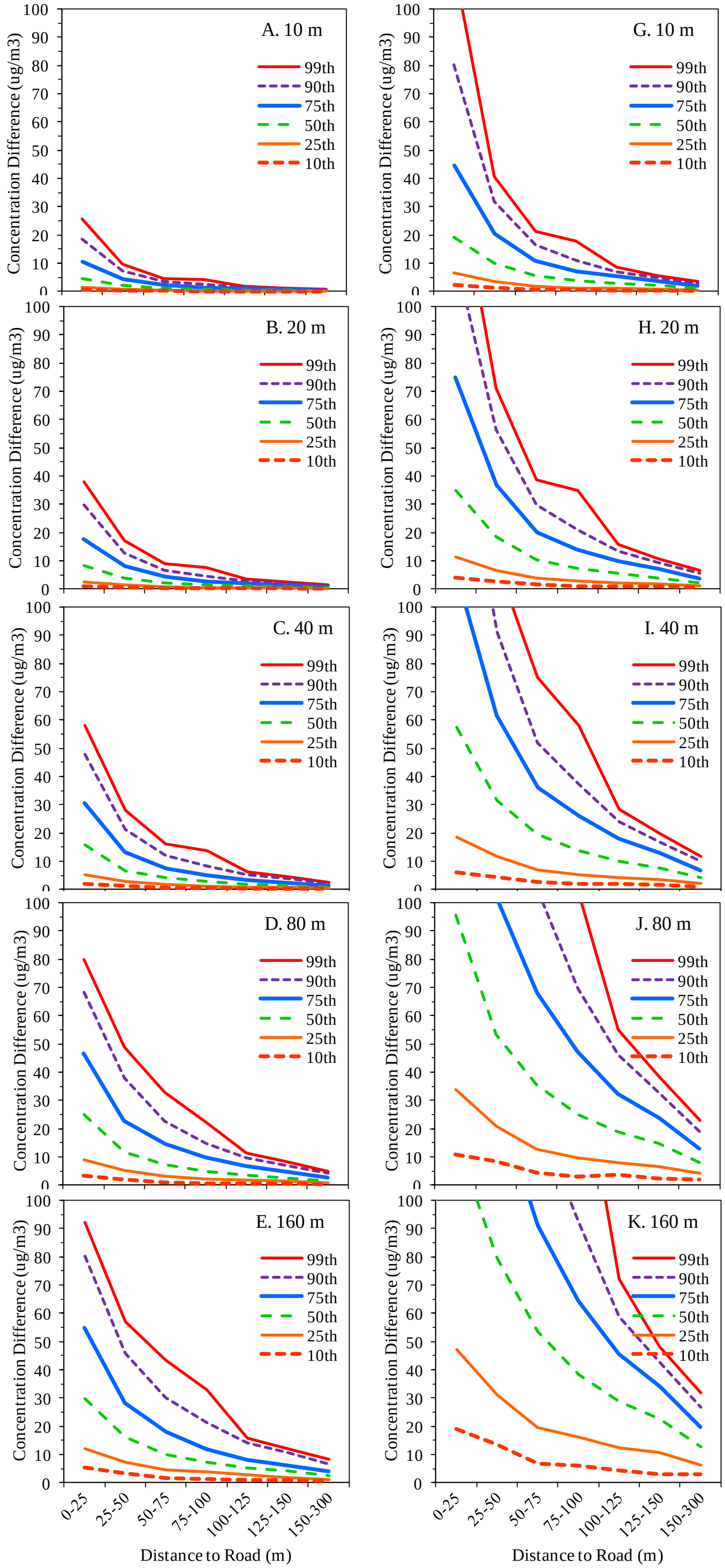
Supplemental Figure 2. Modeled maximum 24-hr NO_X concentrations ($\mu g/m^3$) over study region using receptors on 150 m centers. The area for which data were analyzed in shown with a black outline.



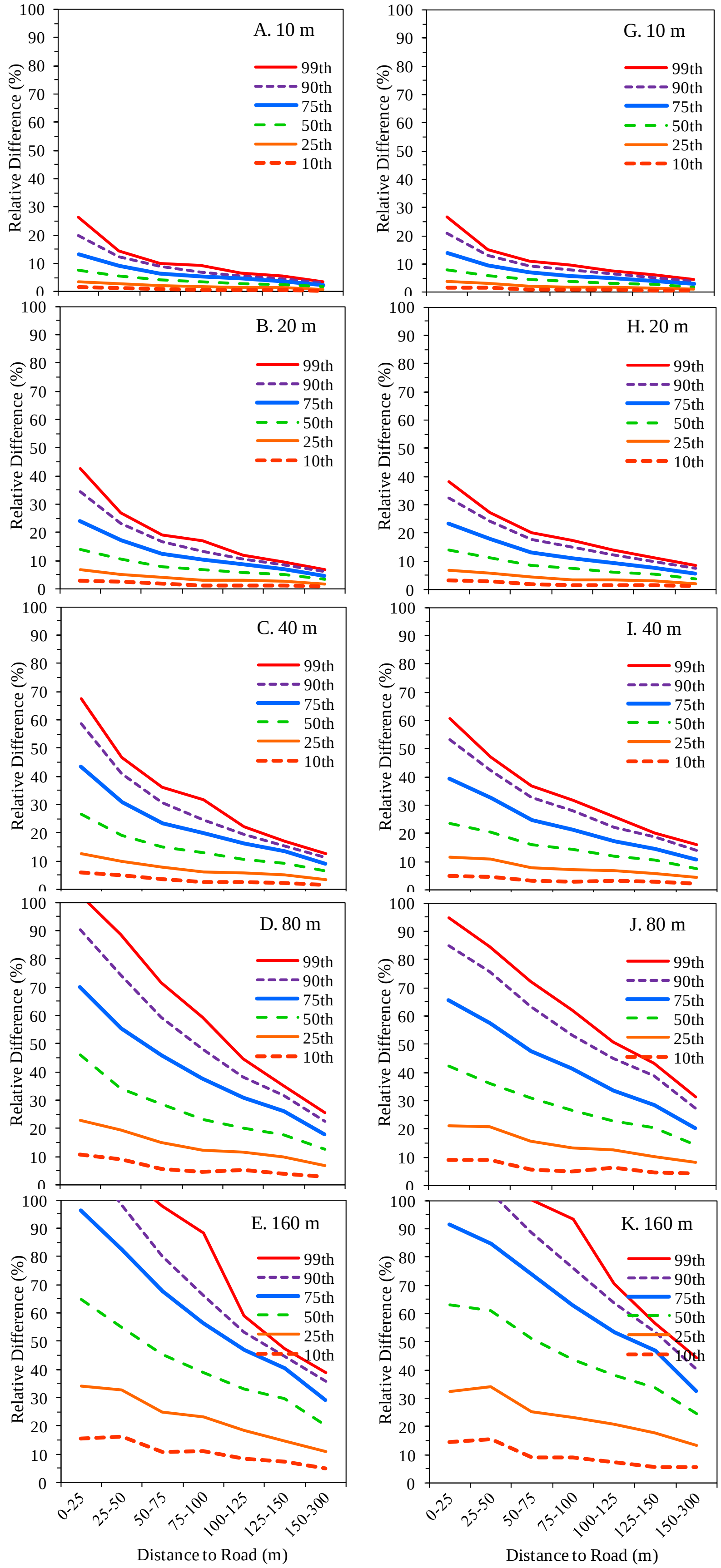
Supplemental Figure 3. Comparison of NO_x concentration estimates for several geographic zones. A: Distribution of maximum 24-hour NO_x concentrations in Detroit at receptors and four geographic zones (not weighted by population). B: Distribution of concentration deviations for three zones (from parcel concentrations). Figure B is truncated at 100 $\mu g/m^3$, thus the most elevated concentrations are not shown.


A
Receptor
Parcel
Block
Tract
ZIP
cumulative density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1
10
100
24-hr maximum NOx (μg/m3)
B
Block
Tract
ZIP
0
10000
20000
30000
40000
50000
-100 -80 -60 -40 -20 0 20 40 60 80 100
Deviation in 24-hr maximum NOx (μg/m3)

Supplemental Figure 4. Absolute concentration differences grouped by distance to roads for estimate of 10, 20, 40, 80 and 160 m distance. Left: monthly average; Right: 98th percentile 1-hr average.



Supplemental Figure 5. Relative concentration differences grouped by distance to roads for estimates of 10, 20, 40, 80 and 160 m from known values. Left: monthly average; Right: 98th percentile 1-hr average.



Supplemental Table 1. Population-weighted concentration deviations (1) and absolute concentration deviations (2) for maximum 24-hr NO_x concentration (µg/m^3). R denotes linear correlation with parcel level data; F2 is factor of two agreement; n=sample size. Deviations are defined as the difference between concentrations estimated for individuals within the zone compared to that at the parcel level. Note that at tract and ZIP-code levels, the 75th percentile concentration deviation exceeds the median concentration (shown in Table 1), indicating large exposure errors for at least 25% of the population.

Statistic	Concentration Deviations (1)			Absolute Concen.Deviations (2)		
	Block	Tract	ZIP	Block	Tract	ZIP
Mean	0.1	3.4	3.7	4.2	11.1	12.0
Std. Dev	11.3	17.7	18.2	10.5	14.2	14.2
Min	-260.3	-263.5	-278.3	0.0	0.0	0.0
10th	-4.8	-9.9	-11.1	0.0	1.3	2.1
25th	-0.4	-0.3	0.6	0.0	3.2	4.9
50th	0.0	4.2	6.9	0.6	7.1	9.2
75th	1.0	11.5	12.8	4.1	14.6	14.8
90th	6.0	19.2	17.4	11.2	23.5	20.9
99th	30.5	34.5	26.9	46.6	69.6	72.8
Max	238.7	51.8	35.2	260.3	263.5	278.3
R	-	-	-	-	-	-
F2	0.006	0.168	0.182	-	-	-
N	357,962	357,962	357,962	357,962	357,962	357,962