Spatial Modeling of Diesel Exhaust Markers in South Seattle

Jill Katherine Schulte

A thesis
submitted in partial fulfillment
of the requirements for the degree of

Master of Public Health

University of Washington

2013

Committee:
Joel Kaufman
Timothy Larson

Program Authorized to Offer Degree:
School of Public Health

University of Washington

ABSTRACT

Spatial Modeling of Diesel Exhaust Markers in South Seattle

Jill Katherine Schulte

Chair of the Supervisory Committee:
Professor Joel Kaufman, MD, MPH
Department of Environmental and Occupational Health Sciences

Background: South Park and Georgetown, two of Seattle's most diverse and affordable neighborhoods, contain the primary commercial traffic corridors from the Port of Seattle to interstates and state highways. Residents of these communities have expressed concern about exposure to diesel exhaust emitted by the large number of commercial trucks that pass through their neighborhoods. The aim of this project was to model the spatial distribution of diesel exhaust markers at a fine scale across these neighborhoods using measurements from a high-density air sampling campaign.

Methods: Two-week average concentrations of two markers of diesel exhaust, 1-nitropyrene (1-NP) and light-absorbing carbon (LAC), were measured in summer and winter at 24 sites. Land-use regression models were built using spatial characteristics of sampling sites, including land use and road density. Mobile source emissions predictions from the CAL3QHCR dispersion model were included in spatial models. Light-scattering particle concentrations measured by a mobile monitoring platform that drove through the neighborhoods were also included as model covariates. Model predictions were generated using land-use regression equations for a grid of points 50m

apart across the study area. Universal kriging was applied to these grid points to generate a raster surface of the gradient of predictions.

Results: 1-NP concentrations ranged from 0.263 pg/m^3 to 2.51 pg/m^3 in summer and 1.11 pg/m^3 to 5.71 pg/m^3 in winter. LAC concentrations, measured as the absorption coefficient of collected fine particles, ranged from 4.31E-06 m^{-1} to 7.84E-06 m^{-1} in summer and 6.30E-06 m^{-1} to 9.42E-06 m^{-1} in winter. The summer 1-NP model had an R^2 of 0.87 and a leave-one-out cross-validated R^2 of 0.73. No prediction model of winter 1-NP was identified. The LAC models had R^2 values of 0.78 and 0.79 and leave-one-out-cross-validated R^2 values of 0.66 and 0.70 for August and December, respectively.

Conclusions: Spatial modeling was successfully used to identify a clear gradient in concentrations of diesel exhaust markers at a fine scale within the neighborhoods of South Park and Georgetown. Spatial features that predicted diesel exhaust concentrations included dispersion model predictions, mobile monitoring results, land use, and distance to railroad tracks, roads and intersections. The existence of this gradient suggests that particularly in stagnant periods, the health and environmental impacts of diesel traffic are not evenly distributed across these neighborhoods.

TABLE OF CONTENTS

INDEX OF FIGURES

INDEX OF EQUATIONS

INDEX OF TABLES

ACKNOWLEDGEMENTS

I owe an enormous debt of gratitude to the dedicated community of faculty, staff, students and fellows who have supported this project. I would like to express my appreciation to Dr. Joel Kaufman for his ever-insightful guidance throughout this study. Thank you to the DEEDS dream team of Dr. Julie Richman Fox, Dr. Sheryl Magzamen and Nancy Beaudet for steering this project forward with such skill. I would especially like to thank statistical mentor Dr. Assaf Oron for his patience, humor and creative mathematical analogies. Thank you also to Dr. Tim Larson and Dr. Chris Simpson for providing so many ideas to improve our study; Mark Davey for his superb organizational skills; Mike Paulsen, Tim Gould and Miyoko Sasakura for their mastery of lab analysis; and Anne Ho and Amanda Gassett for their statistical expertise. Thank you as well to the large cadre of students and staff who generously volunteered their time on our long days of ladder-climbing.

Our research team is indebted to the many public agencies that have supported this project, including the Puget Sound Clean Air Agency, the Washington Department of Ecology, Seattle City Light, Seattle Parks and Recreation, and the South Seattle Community College.

Thank you as well to the staff of Puget Sound Sage for their extraordinary outreach efforts and their tireless commitment to environmental justice. Most importantly, thank you to the South Park and Georgetown communities. So many community members have given their trust, time, ideas, yards, and electrical outlets to support this study, all in the hope of making this city a better place.

BACKGROUND

INTRODUCTION

Seattle's South Park and Georgetown neighborhoods are located a short distance from the Port of Seattle and contain a large network of commercial traffic corridors. The high volume of diesel trucks in these neighborhoods prompted concern among community members about exposure to ambient diesel exhaust. In a 2009 survey by Puget Sound Sage, a Seattle-based community organizing coalition, 60% of residents surveyed in a convenience sample responded that they believed pollution from commercial trucks affected the health of their family (Puget Sound Sage, 2009). In response to these concerns, the University of Washington partnered with Puget Sound Sage to conduct the Diesel Exhaust Exposure in the Duwamish Study (DEEDS). This study sought to characterize the gradient of diesel exhaust across the South Park and Georgetown neighborhoods and identify fine-scale spatial variations in diesel exhaust concentrations. This thesis represents a portion of the work of the larger DEEDS study, which was funded by the Kresge Foundation.

SPECIFIC AIMS

The specific aims of this thesis are:

1. Implement a high-density air sampling campaign in South Park and Georgetown to measure 1-nitropyrene (1-NP) and light-absorbing carbon (LAC), both markers of diesel exhaust.

2. Develop hybrid land-use regression/dispersion models of 1-NP and LAC in order to characterize and map the gradient of diesel exhaust at the neighborhood scale.

DUWAMISH VALLEY NEIGHBORHOODS

South Park and Georgetown are located south of downtown and the Port of Seattle along the Duwamish River (see Figure 1).

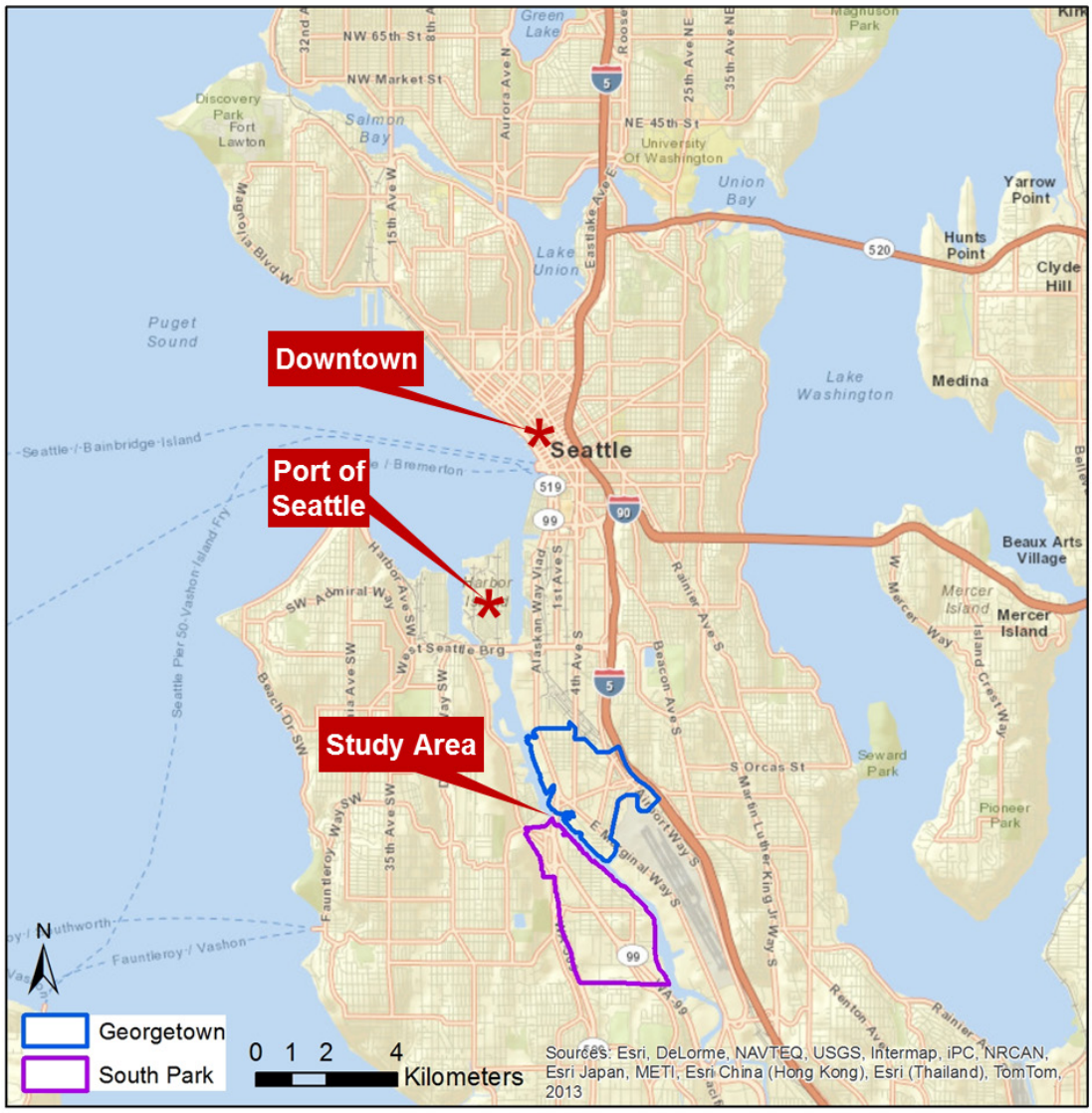


Figure 1. Overview map of study area

Commercial traffic routes serving these neighborhoods include Interstate 5, Washington State Routes 99 and 509, and a number of arterial roads. Because of this dense network of commercial traffic corridors, the Duwamish Valley neighborhoods are a key thoroughfare for commercial traffic between the Port of Seattle and points across the Pacific Northwest. Over 10 trucking companies are headquartered in South Park and Georgetown, and associated trucks typically access these facilities several times a day. Several other sources of ambient air pollution are located within and around South Park and Georgetown. These include an 11-acre Waste Management transfer station, a major First Student school bus depot that houses approximately 200 buses, commercial and passenger rail lines, Seattle's primary industrial zone, and the King County Airport (see Figure 2).

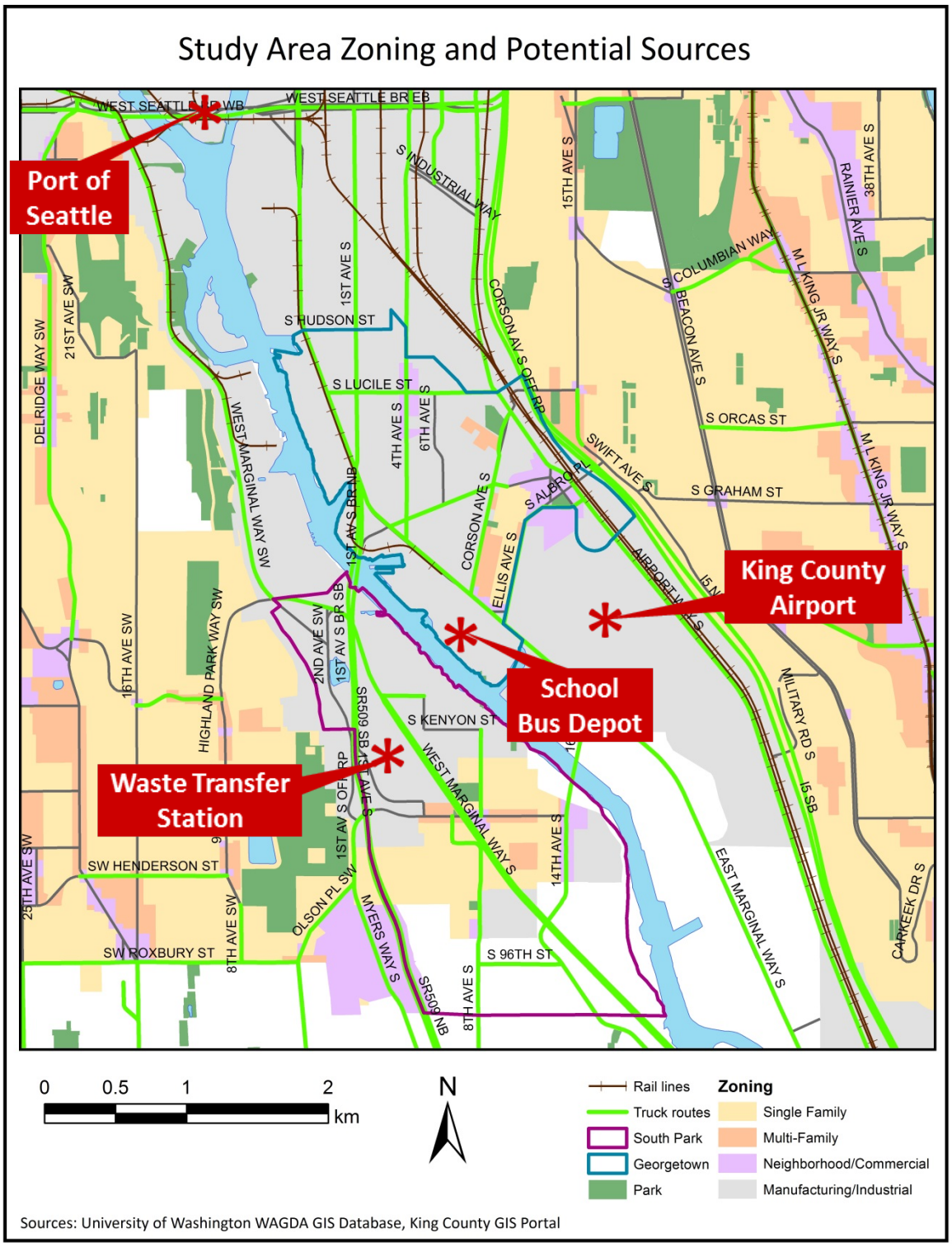


Figure 2. Duwamish Valley zoning and potential sources

The total area of the South Park and Georgetown neighborhoods is 2.79 square miles or 7.23 square kilometers. Approximately 73% of this area is zoned for industrial use, 6% for commercial use and 20% for residential use. As of the 2010 census, the combined population in both neighborhoods was approximately 5,803 people (King County, 2012).

PREVIOUS RESEARCH

Diesel exhaust exposure is a health concern in these neighborhoods because of its association with a number of adverse health outcomes. Diesel engine exhaust was reclassified by the International Agency for Research on Cancer (IARC, 2012) as carcinogenic to humans (Group 1) in 2012. Epidemiological studies have shown an elevated risk of lung cancer associated with long-term diesel exhaust exposure in occupationally exposed populations (Silverman et al., 2012; Attfield et al., 2012; Olsson et al., 2011; Garschick et al., 2008). Acute effects of diesel exhaust exposure include respiratory irritation and inflammation (Hesterberg et al., 2009; Sydbom et al., 2001; Nordenhäll et al., 2000; Nightingale et al., 2000).

Two recent studies in south Seattle have identified diesel exhaust exposure as a predominant environmental health risk for residents in the area. In response to a petition by south Seattle residents, the Washington State Department of Health conducted an air toxics health assessment in the Duwamish Valley in 2008. This assessment found that the increased cancer risk attributed to diesel exhaust exposure in these neighborhoods was three times that attributed to gasoline exhaust (Palcisko,

2008). The emissions estimates used in the health assessment were modeled from estimated traffic counts on highways and interstates. They did not involve any direct measurement of pollutants or account for truck traffic on non-highway arterial roads.

In 2010, the Puget Sound Clean Air Agency conducted an air toxics evaluation in partnership with the University of Washington to characterize the health risks of exposure to air toxics in the Duwamish Valley and other area neighborhoods. This risk assessment was based on direct measurement of over 100 pollutants, including several markers of diesel exhaust. Two of the sampling sites were located in the Duwamish Valley. The study estimated that diesel exhaust exposure accounted for over 70% of the elevated cancer risk of the air toxics measured for residents of the Duwamish Valley (Gilroy, Strange, & Yost, 2010).

These previous studies highlight the health concerns associated with diesel exhaust exposure in South Park and Georgetown and the need for further information about the gradient of diesel exhaust within these individual neighborhoods. The DEEDS project sought to expand upon the existing body knowledge by modeling the spatial distribution of diesel exhaust markers at a fine scale using measurements from a high-density air sampling campaign.

POLLUTANTS MEASURED

The two pollutants selected for modeling, 1-NP and LAC, vary in their specificity to diesel exhaust. 1-nitropyrene, a nitrated polycyclic aromatic hydrocarbon (nitro-PAH), is a byproduct of combustion and the most prevalent nitro-PAH found in diesel

engine exhaust. It is classified by the International Agency for Research on Cancer (IARC, 1985) as possibly carcinogenic to humans (Group 2B). In addition to diesel exhaust, 1-NP has also been detected in airplane exhaust, kerosene stove and heater emissions, coal combustion fly ash, and cigarette smoke (Chan, 1996). As a central goal of the DEEDS project was to characterize diesel exhaust as distinct from gasoline engine exhaust, 1-NP was a suitable pollutant to measure because it is only detectable in trace amounts in gasoline exhaust. In a study of 1-nitropyrene concentrations in lab-generated diesel and gasoline engine exhaust, Hayakawa *et al*. (1992) determined that 1-nitropyrene concentrations were over 200 times as high in diesel engine particles relative to gasoline engine particles. Neither extensive 1-NP field sampling campaigns nor spatial models of 1-NP have been described in the literature to date.

Light-absorbing carbon is a general term for carbonaceous aerosols that absorb visible light. It is measured as the absorption coefficient of particles in units of m^{-1} (Bond, Anderson, & Campbell, 1999). A previous study of diesel exhaust markers in Harlem by Kinney *et al.* (2000) found that particle absorption coefficient and elemental carbon mass concentration were highly correlated, with r=0.95. Elemental carbon is frequently used as a surrogate for diesel exhaust, though other sources include wood combustion and, to a lesser extent, gasoline exhaust. The contribution of diesel exhaust to fine particle elemental carbon as estimated by 11 source apportionment studies reviewed by Schauer (2003) using National Institute for Occupational Safety and Health (NIOSH) methods ranged from 57% - 96%, with a median of 86%. In addition, results from the two channels of Aethalometer® readings collected in the study area, described

in detail on page 22, were nearly identical in both seasons. The presence of wood smoke is generally indicated by large differences between the ultraviolet and black carbon channel readings, and such differences were not observed in the DEEDS study. Based on previous research and these measurement results, it is likely that diesel exhaust was the predominant source of the LAC measured in the DEEDS study.

METHODS

SITE SELECTION

Sampling was conducted simultaneously at 23 outdoor sites in both August and December. One additional site was sampled in each season, for a total of 24 data points per season. Of these 24 sites, 19 were located in the study area and 5 were located in other neighborhoods to serve as comparison sites. Four of these comparison sites (the Queen Anne, Beacon Hill, and downtown agency monitors, and the King County Airport) were excluded from the spatial models due to their distance from the study neighborhoods. The remaining 20 sites will be referred to as the "core study area." The density of sites within the South Park and Georgetown neighborhoods was approximately 7 sites per square mile.

Table 1. Number of sampling sites by neighborhood

	South Park	Georgetown	Other Neighborhoods	Total	Included in Modeling
August	12	7	5	24	20
December	11	8	5	24	20

The majority of sampling sites were at homes, though sampling was also conducted at 4 businesses and 8-9 public sites. Public sites included 4 monitoring sites operated by public agencies (the Puget Sound Clean Air Agency and the Washington Department of Ecology) in the Queen Anne, Beacon Hill, downtown Seattle and SoDo (south of downtown) neighborhoods. Other public sites sampled were the South Park Neighborhood Center (August only), the South Park Community Center playground, the Georgetown campus of South Seattle Community College, and the King County Airport (see Figure 3).

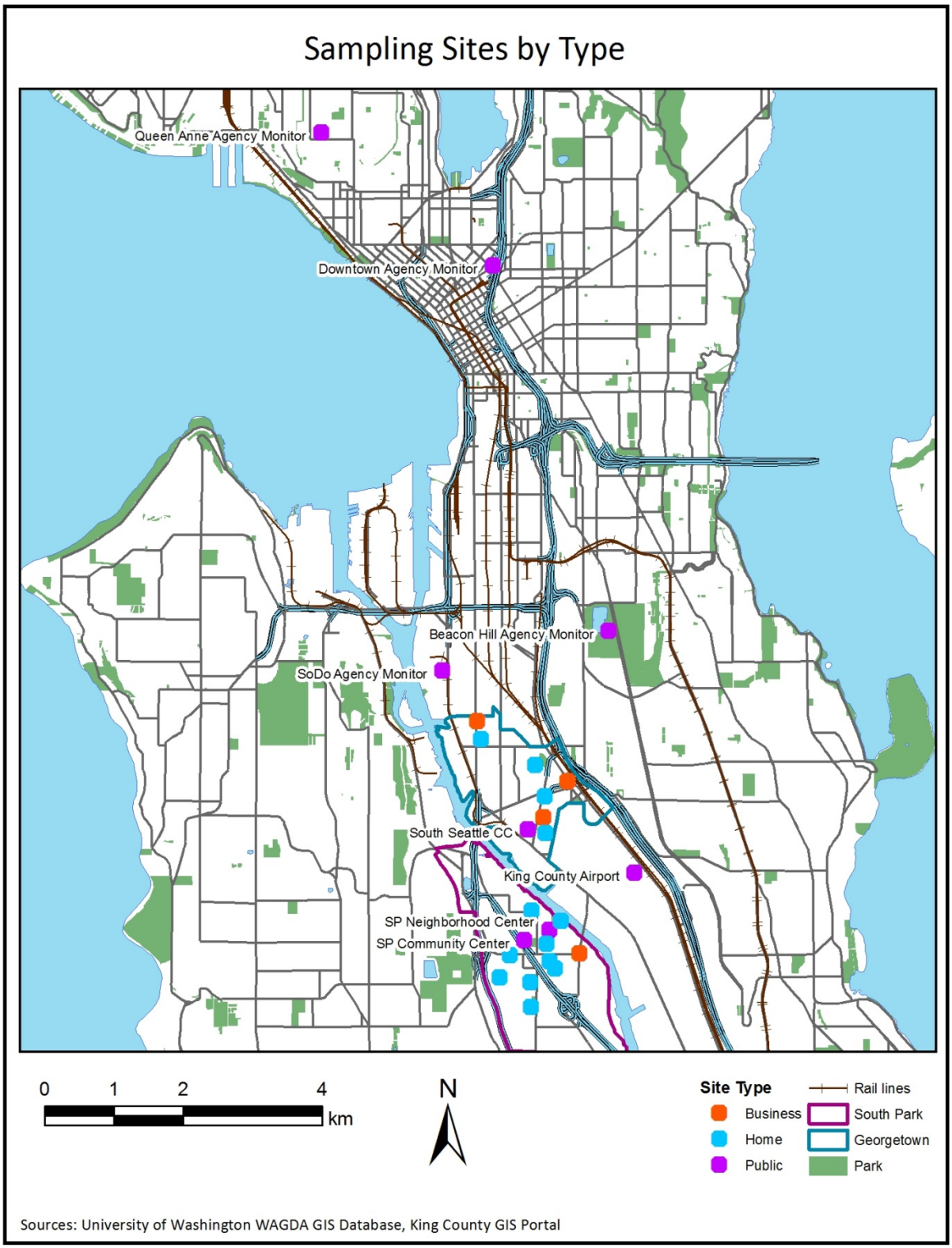


Figure 3. Map of all sampling sites by type

Feedback from community members guided the selection of sampling sites. In spring 2012, Puget Sound Sage surveyed approximately 550 members of the South Park and Georgetown communities who either lived or worked in the study area at the time. Respondents identified commercial trucks as their top diesel exhaust source of concern and residential areas as their priority locations to monitor. Sampling sites were selected to maximize coverage in residential areas in accordance with these responses. Of the 19 sites within the Duwamish Valley, 12 were located in residential zones in August and 11 in December (see Figure 4).

In addition, sites were selected at a wide range of distances from commercial truck routes to increase the likelihood that statistical models would be able to discern the impacts of commercial trucking. Before sites were selected, the study area was divided into 18 geographic zones. Zone boundaries were drawn to maximize variation in distance to truck route between zones and minimize this variation within zones. Puget Sound Sage recruited a minimum two volunteers per zone who were interested in hosting a monitor at their homes or businesses. The research team visited each home or business to screen for sampling site locations. The most suitable site in each zone was selected based on the security of the location, the availability of canopy-free space away from walls or fences, and the proximity of outdoor electrical outlets.

Table 2. Zoning designation of sampling sites within core study area

	Residential	Commercial	Industrial
August	13	2	4
December	12	2	5

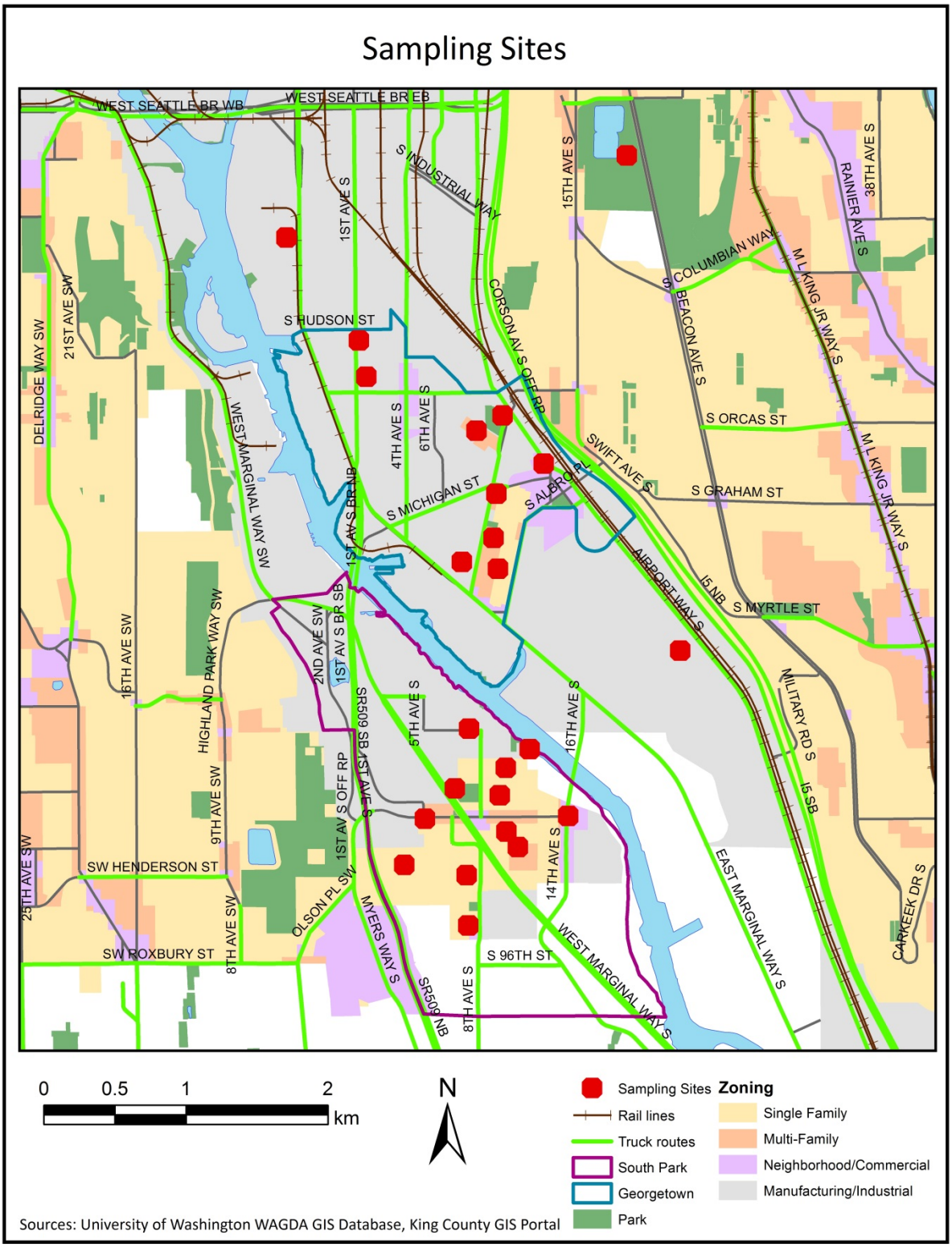


Figure 4. Map of Duwamish Valley sampling sites and zoning

Monitors were located either in yards or on rooftops. Sites in back or front yards were located at least 6 feet away from walls or fences and outside of canopy cover. Residents were asked to record any barbecuing, lawn mowing or idling of personal vehicles that took place in the vicinity of the monitor and to avoid these activities if possible. Rooftop sites were located at least 6 feet away from roof edges (see Figure 5). Two monitors (one business and one public site) were flagged as atypical set-ups because they were suspended from the edge of a roof; this public site was only sampled in August.

Figure 5. Typical 2nd-story rooftop sampler

SAMPLING PROTOCOL

Sampling was conducted August 18-30 (12 days) and December 1-14 (13 days), 2012. Particles were collected on 37mm Teflon filters (Pall Life Sciences, Port Washington, NY) loaded in Harvard Personal Exposure Monitors (HPEMs, Harvard School of Public Health, Boston, MA) with a 50% cut point of 2.5µm. At each sampling site, two HPEMs were suspended beneath waterproof rain caps at a height of approximately 5 feet from the ground or rooftop. Sampling pumps (Medo USA Inc., Hanover Park, IL) operated at approximately 1.8 L/min and were equipped with dual valve timers to direct airflow through each HPEM 50% of the time in alternating 5-minute periods.

Continuous sampling flow rates were measured in the field using rotameters (Cole Parmer, Inc., Vernon Hills, IL) that were calibrated against a primary flow meter during the August sampling period. Flow rates were adjusted to within 3% of the 1.8L/min target flow at the onset of sampling and at a mid-sampling check roughly halfway (5-7 days) into the sampling period. The total air volume per sample (V) was calculated using Equation 1:

$$V(m^3) = \frac{\left(\frac{flow_a + flow_b}{2} * time_1\right) + \left(\frac{flow_c + flow_d}{2} * time_2\right)}{2}$$

Equation 1. Calculation of total airflow per sample

Where $flow_a$ = initial flow (m^3/minute)

$flow_b$ = mid-sampling check flow (m^3/minute)

$flow_c$ = mid-sampling check adjusted flow (m^3/minute)

$flow_d$ = final flow (m^3/minute)

$time_1$ = length of first half of sampling period (minutes)

$time_2$ = length of second half of sampling period (minutes)

The lapsed times between sampling onset and the mid-sampling check ($time_1$) and between the mid-sampling check and the end of the sampling period ($time_2$) were calculated from log sheets completed by field staff. The total airflow was divided by 2 to account for the 50% timers.

LAB ANALYSIS

1-Nitropyrene

Collected particles were analyzed for 1-nitropyrene content according to a previously reported method (Miller-Schulze 2010). In brief, samples were spiked with an isotopically labeled internal standard, extracted in solvent, evaporated, and then re-suspended in a solution of ethanol, sodium acetate and acetic acid. The suspension was filtered prior to analysis via high-performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS). 1-NP results are reported as a mass concentration in units pg/m^3. The 1-NP sample loading was determined using the fine particulate matter ($PM_{2.5}$) mass concentration of the filter samples, which was assessed by gravimetric analysis.

Light-Absorbing Carbon

Filter reflectance was assessed both before and after sampling using a Smokestain Reflectometer (Diffusion Systems Ltd Model 43D, London, United Kingdom) as described in Dotse, Asane and Ofosu (2012). Absorption coefficients (b_{ap}) of collected particles in units m^{-1} were calculated from Equation 2:

$$b_{ap} = \left(\frac{A}{2V}\right) * \ln\left(\frac{\overline{R_b}}{R_s}\right)$$

Equation 2. Absorption coefficient equation

Where A = area of the exposed filter (0.000730246 m^2)

V = sample volume (m^3) (see Equation 1)

$\overline{R_b}$ = average change in reflectance of field blanks

R_s = change in reflectance of sample filter

QUALITY ASSURANCE AND CONTROL

In each season, at least 5 lab blank and 6 field blank HPEMs (one of each for every 10 sample filters) were assembled to identify any contamination issues in the lab or field. Lab blanks were immediately placed in sealed plastic bags upon assembly and stored at room temperature throughout the sampling period. Field blank HPEMs were also stored in plastic bags, and each field technician carried at least one field blank with them on the first day of each sampling campaign. The field blanks were removed from

plastic bags for approximately 5 seconds at a sampling site of the technician's choosing, then resealed and stored. This process was repeated on the final day of sampling only during the August sampling period. Field blanks were stored in the laboratory during the remainder of the sampling period.

Both lab and field blank filters were analyzed for 1-NP and LAC. The average 1-NP mass on blank filters was less than 3% of the average mass on sample filters in both seasons, indicating that significant lab or field contamination of filters was unlikely. The change in reflectance values of the field blank filters, which was within the margin of error of the reflectometer, was used to adjust the absorption coefficient results as shown in Equation 2.

Equivalent volumes of air were sampled by duplicate HPEMs operated by a single pump at each site. Duplicate setups were also installed at two sampling sites. Every filter was analyzed for LAC. The average percent error in absorption coefficient between duplicate pairs on the same pump was 16% in August and 7.4% in December, calculated as the ratio of the difference in b_{ap} to the average of the two b_{ap} values. The average percent error in b_{ap} between filters on separate duplicate setups was 10% in August and 8.6% in December. Final LAC results by site were calculated as the mean of duplicate/replicate filters collected at each site.

Only one filter per site was analyzed for 1-nitropyrene, with the exception of 2 duplicates in August and 5 duplicates/replicates in August. The average percent error in 1-NP concentration between duplicates/replicates was 11% in August and 14% in December. Six analytical replicates from the August sampling period were selected for

reanalysis with the December filters in order to compare the measurement precision between batches. The average percent error among the analytical replicates was 40%. Average percent error was calculated as the ratio of the largest difference among any pair of replicate filters to the average concentration of all replicate filters. Based on this high level of error in replicates between batches, pollutants were modeled separately in each season and were first standardized within each season when combined into multi-season models. Final 1-NP concentrations by site were calculated as the average of only the duplicates/replicates analyzed in the same batch. August results were not adjusted for the analytical replicates analyzed in December because of the poor between-batch precision.

STATISTICAL METHODS

Land-use regression (LUR) was used to develop spatial models of the two pollutants. LUR is a multiple linear regression technique that uses measured concentrations of a pollutant as the dependent variable and spatial covariates such as land use and road density as independent variables. Spatial covariates can then be used to predict pollutant concentrations in locations without measurements (Jerrett et al., 2005). LUR has been used to predict concentrations of traffic-related pollutants in urban environments in several previous studies (Mercer et al., 2011; Novotny et al., 2011; Larson, Henderson & Brauer, 2009; Hoek et al., 2008; Henderson et al., 2007; Briggs et al., 1997), though usually at a broader geographic scale than the DEEDS effort.

Mobile-source emissions estimates from both dispersion modeling and direct-reading instruments were included as covariates to create hybrid land-use regression/dispersion models. The hybrid model was first developed by Wilton *et al*. (2010), who found that inclusion of line source dispersion model predictions enhanced the predictive ability of a LUR model of oxides of nitrogen (NO_x) in Seattle and Los Angeles. The additional inclusion of on-road emissions estimates from direct-reading instruments as model covariates has not been described in the literature to date.

DISPERSION MODELING

Emissions estimates from mobile sources were generated using the CAL3QHCR dispersion model (U.S. EPA, 1995). This model predicts concentrations of a non-reactive pollutant at receptor sites using a Gaussian dispersion equation. Specific emissions estimates were generated for the 12-13 days of each sampling period based on roadway locations, estimates of traffic volumes and meteorological data. Traffic inputs were derived from the Travel Demand Model Version 1C developed by the Puget Sound Regional Council (PSRC, 2008). This model contains car and truck volume forecasts on each segment of highway and major arterial road in the study area, based on commuter surveys and a sample of empirical traffic counts. Meteorological data were obtained from the National Weather Service station at the King County Airport with the exception of mixing height, which was obtained from the Sea-Tac International Airport.

MOBILE MONITORING

Additional on-road emissions data were collected using a dual-channel Aethalometer® (microAeth® Model AE52, Magee Scientific, Berkeley, CA). The instrument's inlet was affixed to a hybrid vehicle that drove a fixed route throughout the study area 5-6 times per day (see Figure 6). This route included 4 loops per trip of 4 consecutive right turns around 11 of our sampling sites. Mobile monitoring was conducted on 6 days in September 2012 and 6 days in December 2012. The December mobile monitoring campaign was concurrent with the December air sampling campaign, but September was the earliest the mobile platform was available following the August sampling period. In each season, 5 monitoring days were weekdays and 1 was a weekend day. All mobile monitoring took place during the hours of 2pm to 7pm to capture peak traffic during the evening rush hour. A total of 11,561 data points were collected within the South Park and Georgetown neighborhoods.

As aerosols collected on a filter, the instrument continuously measured the change in rate of absorption of transmitted light at the 880 nm and 370 nm wavelengths. The 880 nm channel indicates the presence of black carbon (BC), and the 370 nm channel (UV) indicates the presence of both black carbon and additional aromatic organic compounds, both in units of $\mu g/m^3$ (Magee Scientific, 2010). The log-transformed concentrations from the BC and UV channels on all 12 sampling days were averaged within buffers of 300 meters and 500 meters around each sampling site and included as model covariates.

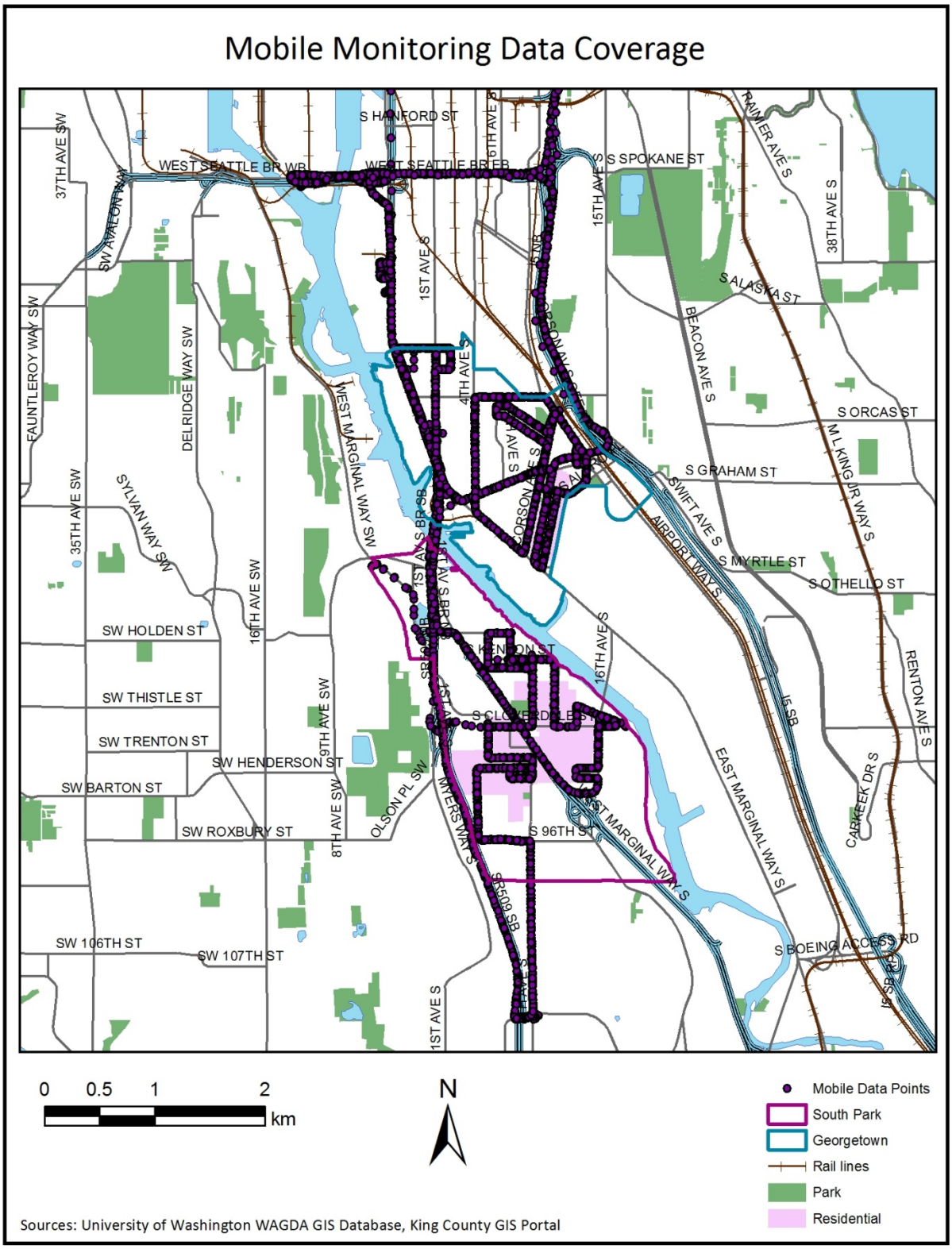


Figure 6. Mobile monitoring data coverage

SPATIAL COVARIATES

Geographic attributes of the sampling sites, including land use, road density, and population, were extracted using Tele Atlas (TomTom, Amsterdam, Netherlands) and ArcGIS 10.1 (ESRI, Redlands, CA). Spatial data were obtained from the following sources: National Emissions Inventory (U.S. Environmental Protection Agency), Tele Atlas, Google Maps, U.S. Census Bureau, Multi-Resolution Land Characteristics Consortium 2006 National Landcover Dataset (U.S. Geological Survey), National Geospatial Intelligence Agency, and the Bureau of Transportation Statistics. The 107 spatial variables initially considered for the models are summarized in Table 3.

Table 3. Spatial covariates considered for modeling of 1-NP and LAC

Variable	Type	Buffer Radius	Description
log10.m.to.<type>	a1, airp, coast, port, ry, rr, truck, road, intersect, interchange12, interchange3	N/A	Log_{10} meters to a1 road, airport, coast, port, rail yard, railroad, truck route, road, intersection, a1 and a2 road interchange, a3 road interchange
ll.a1.<buffer>	N/A	100m, 150m, 300m, 500m, 750m, 1500m	Length of a1 roads in various buffer distances
ll.a23.<buffer>	N/A	100m, 150m, 300m, 500m, 750m, 1500m	Length of a2 and a3 roads in various buffer distances
log10.pop.<buffer>	N/A	500m, 1500m, 2500m	Log_{10} population in various buffer distances
intersect.<buffer>	N/A	500m, 1000m	Intersections in various buffer distances
pop.s01000	N/A	1000m	Population in 1000m
log10.pop.<buffer>	N/A	500m, 1000m, 1500m, 2500m	Log_{10} population in various buffer distances
interchange <type>.<buffer>	a12, a3	500m, 1000m	Interchanges with A1 roads by road type in various buffer distances
imp.<buffer>	N/A	50m, 150m, 300m	Impervious surface in various buffer distances
elev.elevation	N/A	N/A	Elevation
rlu.dev.<type>.<buffer>	open, openlow, openbasic, openplus, medhi, hi, anyflat, anyforest	50m, 150m, 300m, 750m, 1000m	Land use in various buffer distances
log10_trucking	N/A	N/A	Log_{10} meters to trucking company
log10.caline <type><buffer>	Cars.sm, Cars.wn, Trux.sm, Trux.wn	1500m, 3000m, 4500m	CAL3QHCR emissions estimates for cars and trucks in various buffer distances in summer and winter
<type>LogBuff <buffer>	bc, uv	300m, 500m	Mean of log_{10} mobile monitoring values in black carbon and ultraviolet channels in various buffer distances

Variables were excluded from consideration if they were correlated by greater than 95% with another variable that appeared previously in the dataset. Variables were also excluded if their total change across the study area was less than 10% or if their coefficient of variation was less than 0.1. The least absolute shrinkage and selection operator (lasso) was then used to select from among the remaining 77 variables. Lasso is a method described by Tibshirani (1996) that selects a model to minimize the sum of squared residuals and the absolute value of the model coefficients. The lasso estimate $(\hat{\alpha}, \hat{\beta})$ for a set of standardized x_{ij} is defined by Equation 3:

$$(\hat{\alpha}, \hat{\beta}) = argmin\left\{\sum_{i=1}^{N}\left(y_1 - \alpha - \sum_j \beta_j x_{ij} \right)^2\right\} subject\ to\ \sum_j |\beta_j| \leq t$$

Equation 3. The lasso equation

The constraint on $\sum_j |\beta_j|$ is defined by the value of the lasso penalty *t*. The optimal value of *t* to minimize the mean-square error for each model was selected by 5-fold cross-validation. Cross-validation groups were randomly selected, with any sites closer than 250 meters kept in the same group.

Reverse stepwise regression was performed on the model terms selected by the lasso method to minimize the Akaike Information Criterion (AIC). Due to the small number of sampling sites (n=20), the lasso and stepwise procedures were repeated 500 times for each model to ensure their stability as cross-validation groups were rearranged. The performance of the final models was assessed using leave-one-out cross-validation. These statistical analyses were conducted in R version 2.15.2.

Predictions were calculated from land-use regression equations for a grid of points 50 meters apart across the study area. Universal kriging was applied to these points to generate a smoothed raster surface of the gradient of predictions for each of the three models using ArcGIS 10.1. Kriging is a minimum mean-squared error technique for interpolating a prediction surface between points with fixed values. The universal kriging tool in ArcGIS assumes spatial correlation based on a linear semivariogram with linear drift (Mercer et al., 2011).

RESULTS

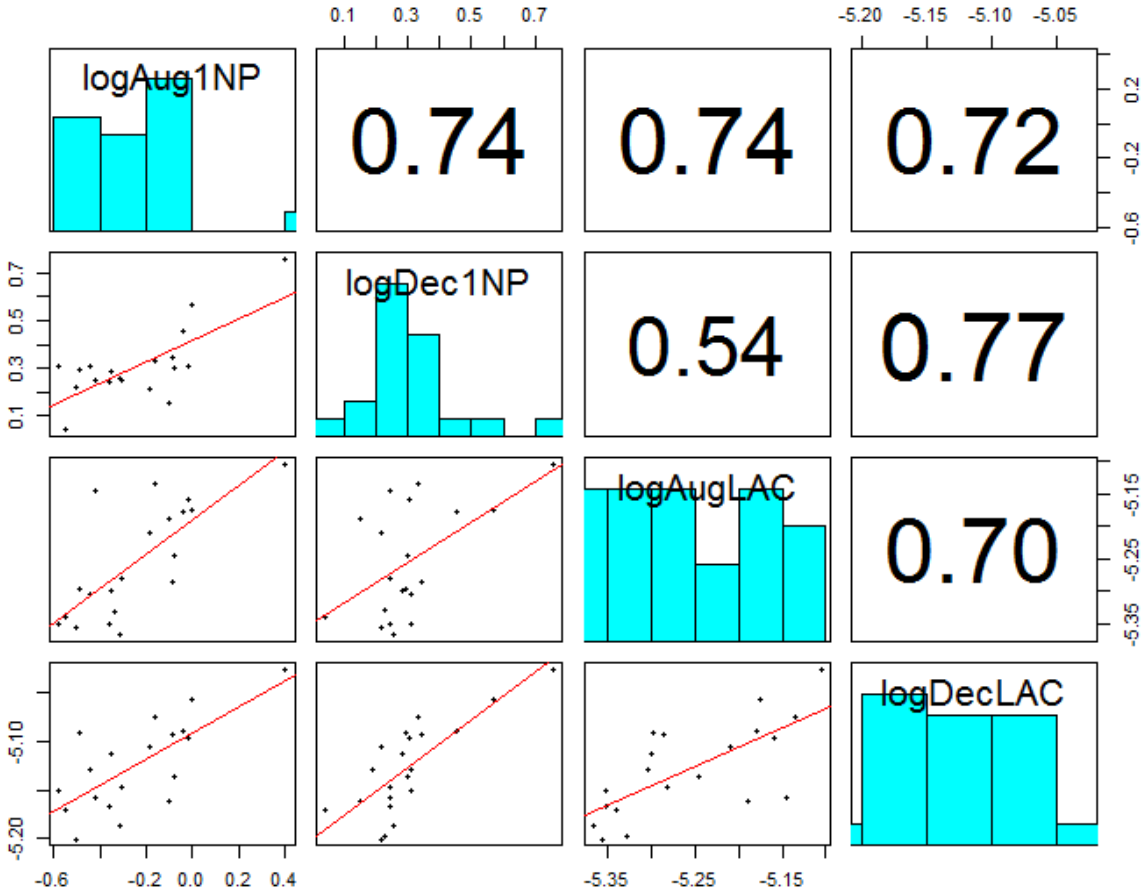
MONITORING RESULTS

Concentrations of both pollutants were higher on average in December than August, though August measurements showed a greater degree of variability. The measurement results from the August and December sampling campaigns in the core study area are summarized in Table 4.

Table 4. Summary of 1-NP and LAC concentrations (core study area)

	Mean	Median	Standard Deviation	Minimum	Maximum
August 18-30					
1-Nitropyrene (pg/m^3)	0.658	0.485	0.505	0.263	2.51
LAC (m^{-1})	5.57E-06	5.11E-06	1.13E-06	4.31E-06	7.84E-06
December 1-14					
1-Nitropyrene (pg/m^3)	2.10	1.86	0.985	1.11	5.71
LAC (m^{-1})	7.47E-06	7.38E-06	8.32E-07	6.30E-06	9.42E-06

In each season, LAC and 1-NP were highly correlated ($r > 0.7$). The matrix in Figure 7 shows Pearson correlation coefficients and scatterplots between log-transformed pollutants and between seasons.



Where *logAug1NP* = Log_{10} August 1-NP measurements
logDec1NP = Log_{10} December 1-NP measurements
logAugLAC = Log_{10} August LAC measurements
logDecLAC = Log_{10} December LAC measurements

Figure 7. Correlation matrix of 1-NP and LAC results in August and December

Figures 8-11 below display the measured concentrations of 1-NP and LAC in both seasons. Symbol colors were classified by quintiles of data.

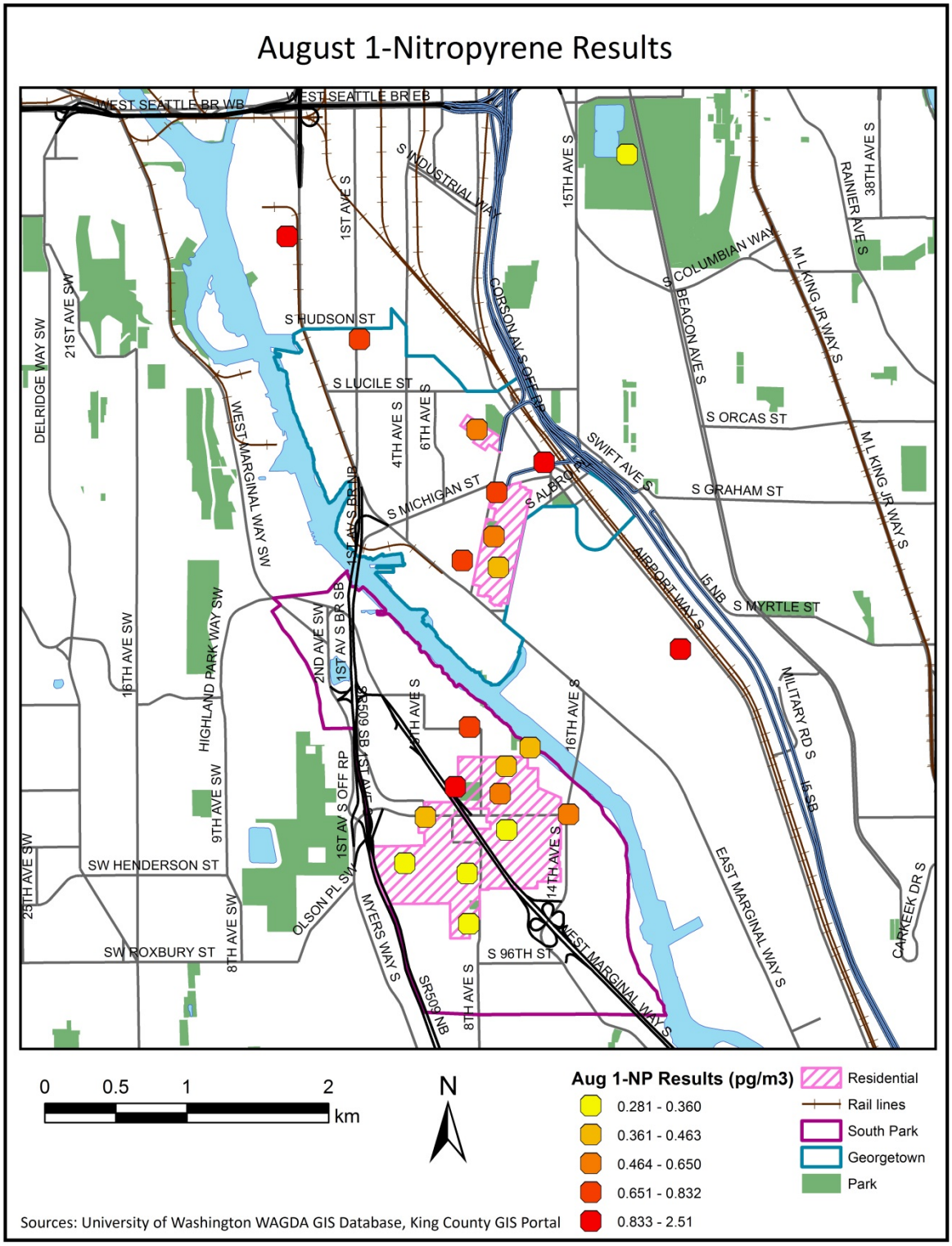


Figure 8. August 1-NP sampling results

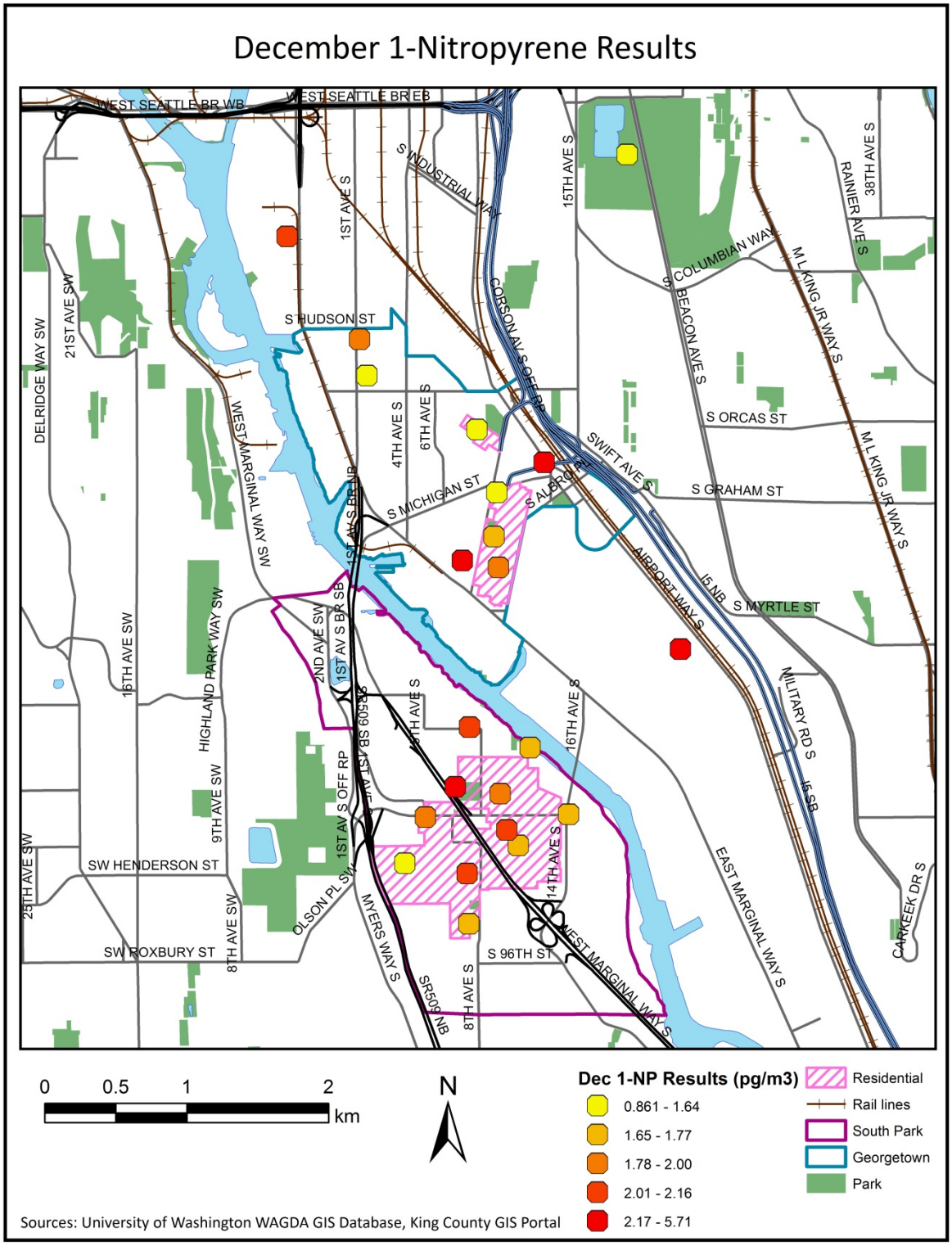


Figure 9. December 1-NP sampling results

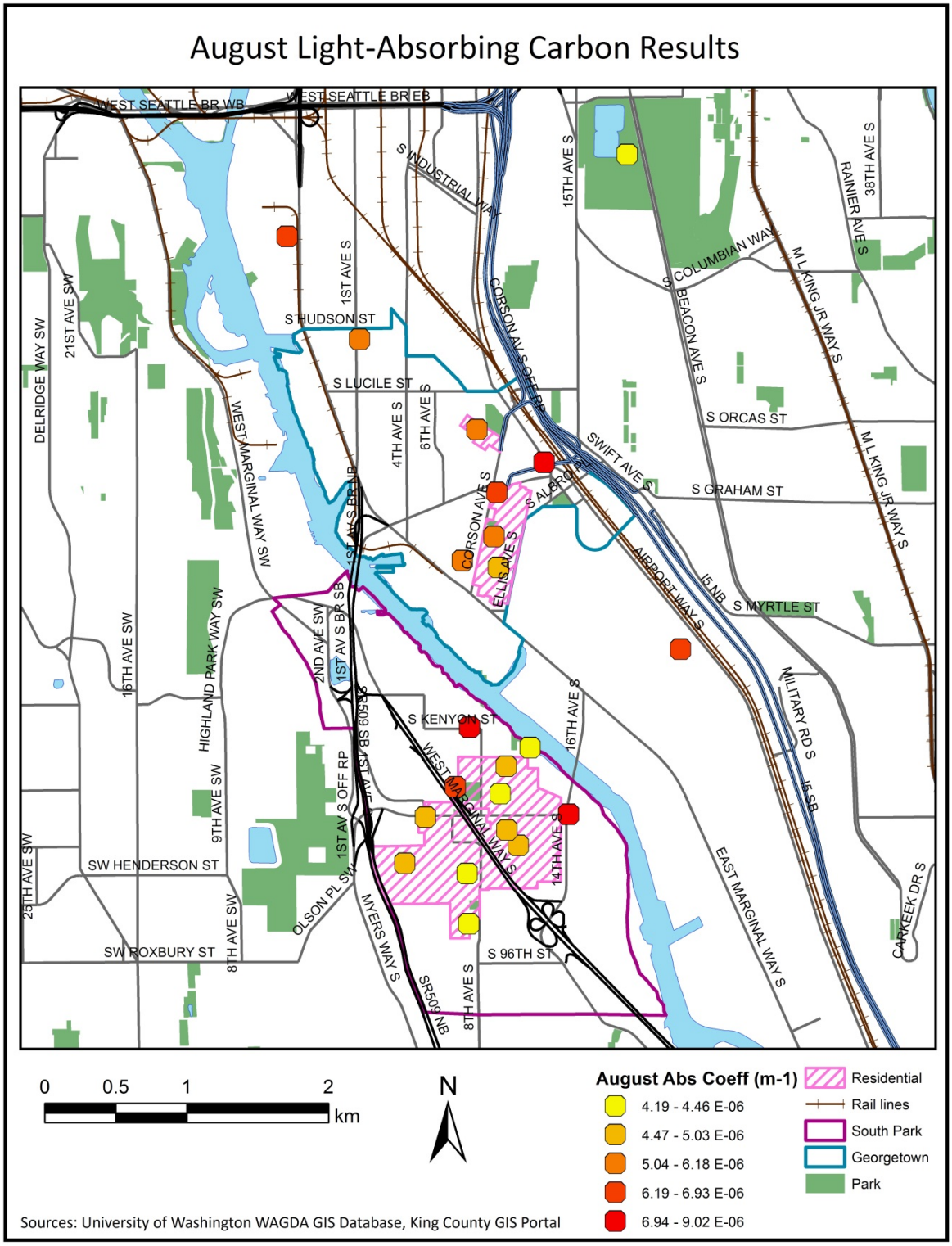


Figure 10. August LAC sampling results

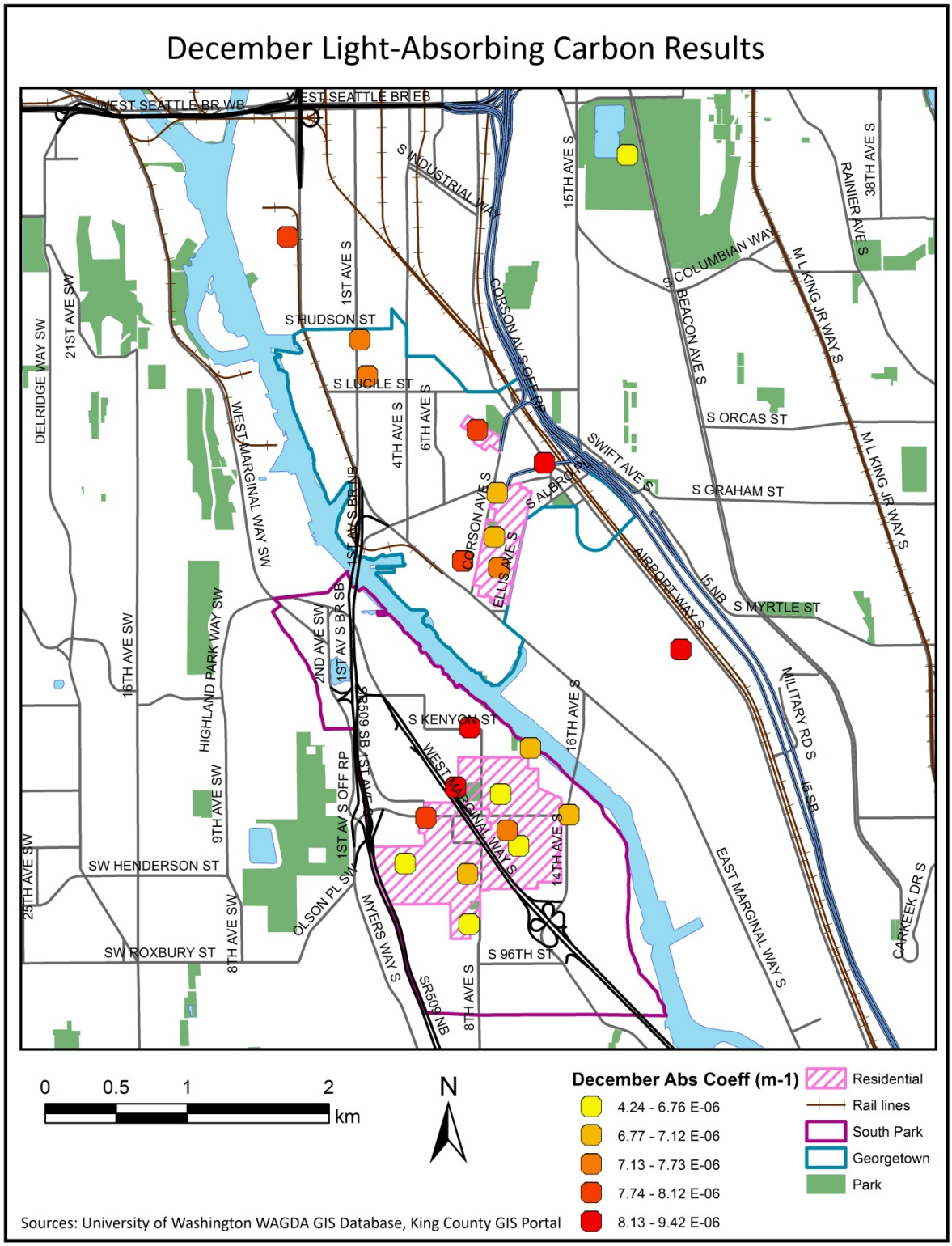


Figure 11. December LAC sampling results

MODELING RESULTS

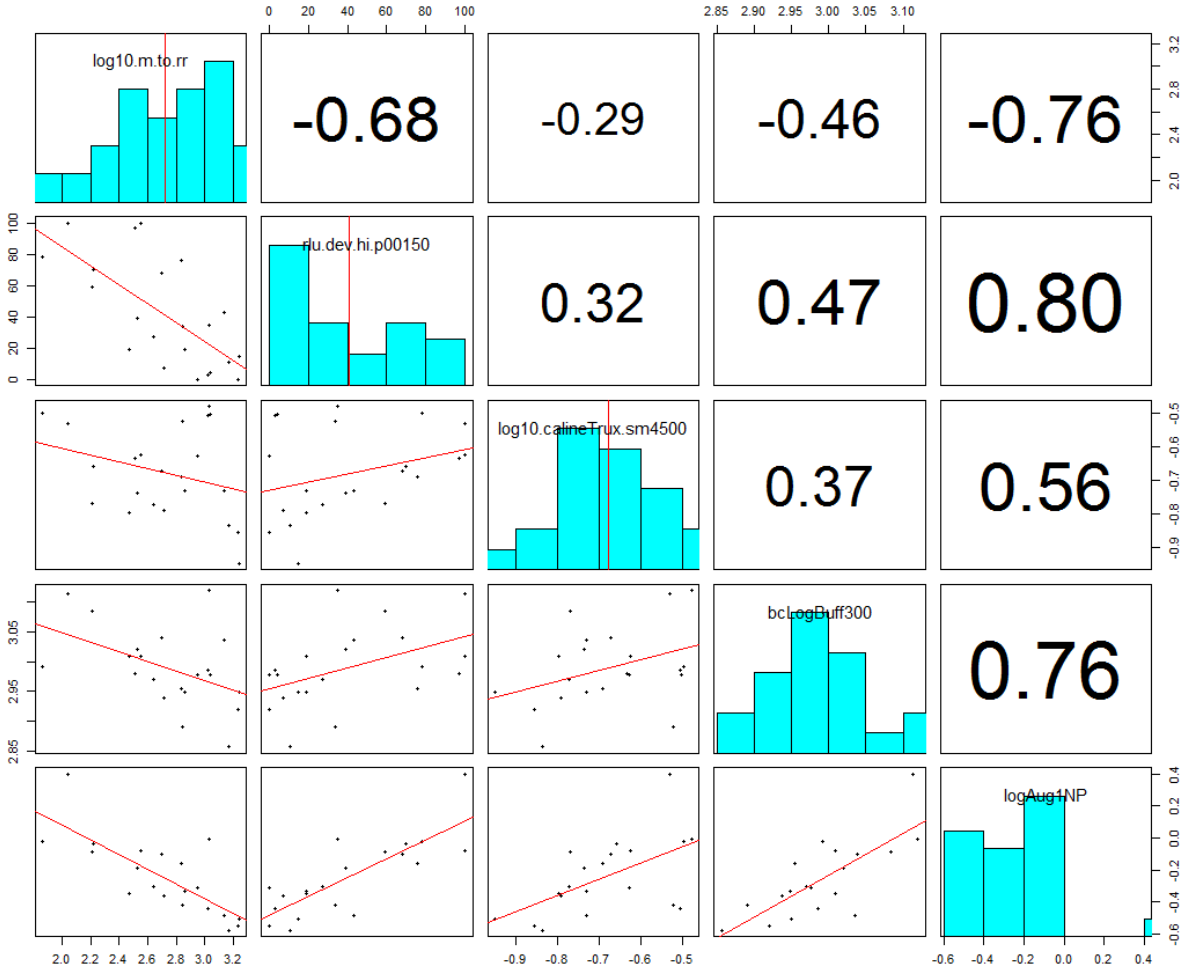
August 1-Nitropyrene

The spatial covariates included in the model of log_{10} August 1-NP are summarized in Table 5 below. The correlation matrix in Figure 12 shows the correlations and scatterplots for each pair of terms in the model. This model has an R^2 of 0.868, a cross-validated R^2 of 0.730, and a cross-validated RMSE of 0.123 log_{10} pg/m^3. The final model covariates are log_{10} meters to railroad, developed high intensity area[1] in 150m, CAL3QHCR truck emission predictions in 4500m, and mean log_{10} black carbon-channel Aethalometer® readings in 300m. The model residuals are mapped in Figure 13, and Figure 14 shows the gradient of August 1-NP predictions across the core study area.

Table 5. Log_{10} August 1-NP model terms

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
log10.m.to.rr	-0.180	0.0813	-2.72	0.04	(-0.355, -0.00608)
rlu.dev.hi.p00150	0.00249	0.00102	-2.22	0.03	(0.000296, 0.00468)
log10.calineTrux.sm4500	0.291	0.196	2.43	0.16	(-0.129, 0.711)
bcLogBuff300	1.31	0.412	1.49	0.01	(0.427, 2.19)
(Intercept)	-3.57	1.32	3.18	0.02	(-6.39, -0.755)

[1] "Developed high intensity" areas are defined as "highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80% to 100% of the total cover" (MRLC, 2006).



Where *log10.m.to.rr* = Log_{10} meters to railroad
rlu.dev.hi.p00150 = High intensity development in 150 meters
log10.calineTrux.sm4500 = Log_{10} CAL3QHCR truck emission predictions in 4500m
bcLogBuff300 = Mean of log_{10} black carbon-channel Aethalometer® readings in 300m
logAug1NP = Log_{10} August 1NP measurements

Figure 12. Correlation matrix of log_{10} August 1-NP model terms

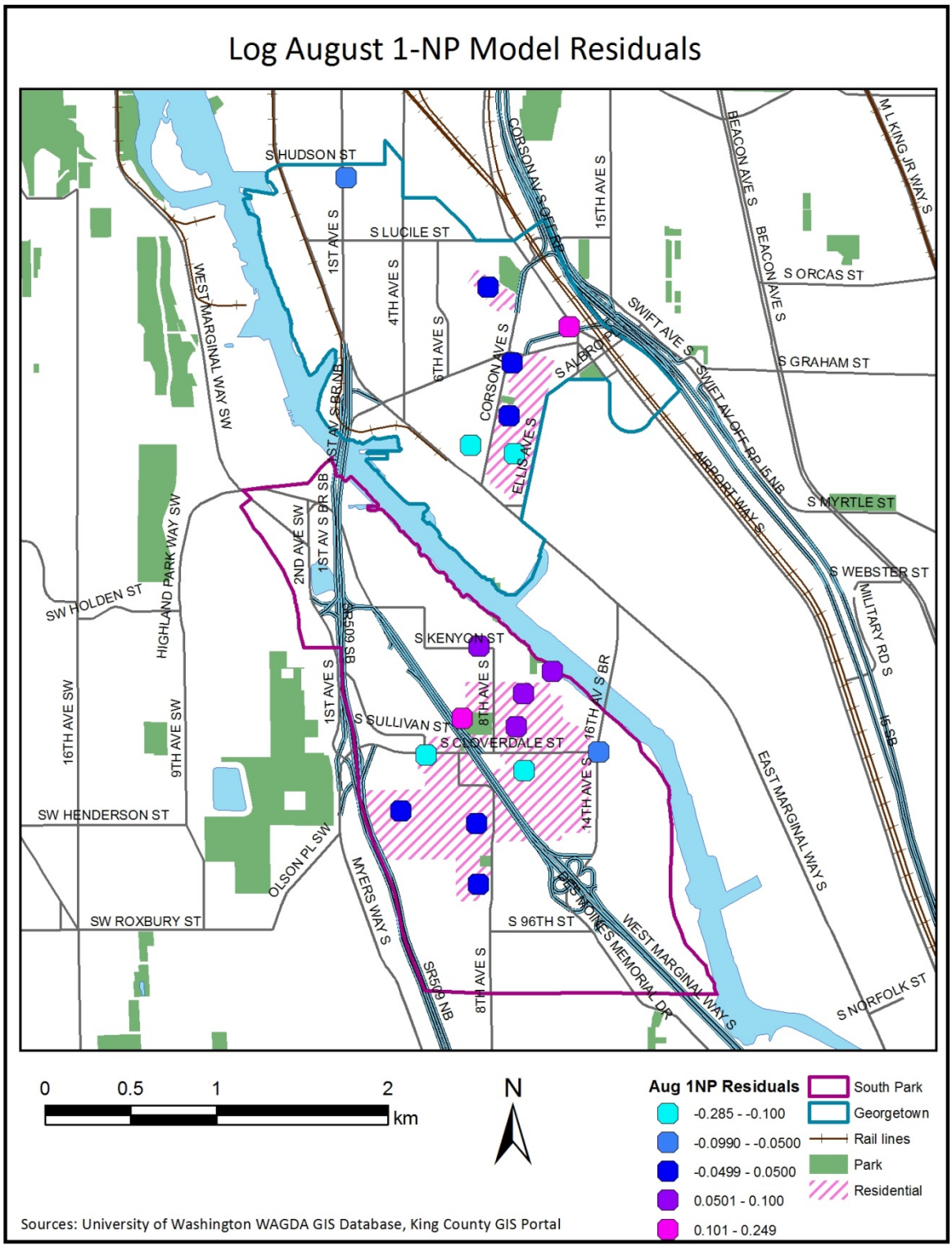


Figure 13. Map of $\log_{10}$ August 1-NP model residuals

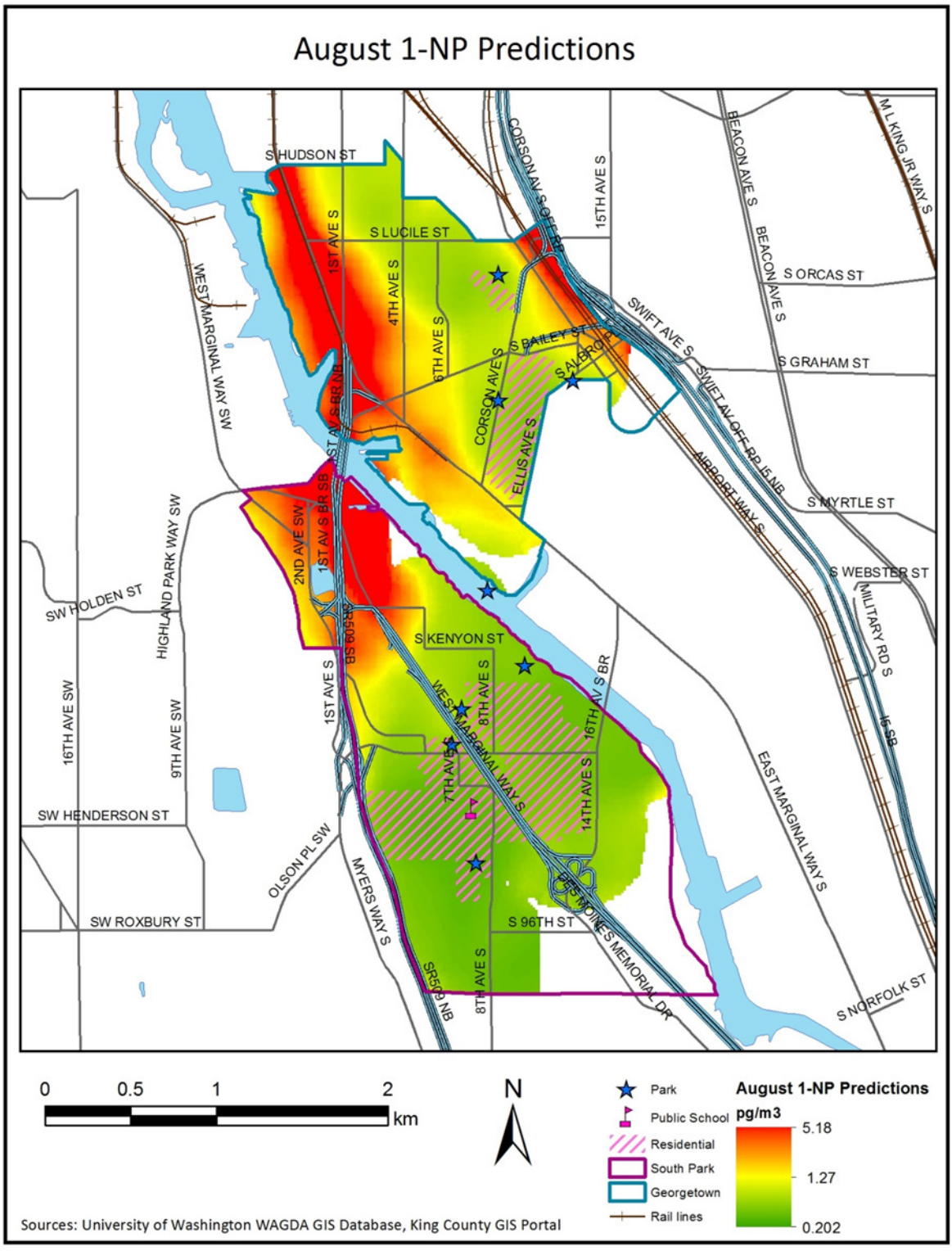


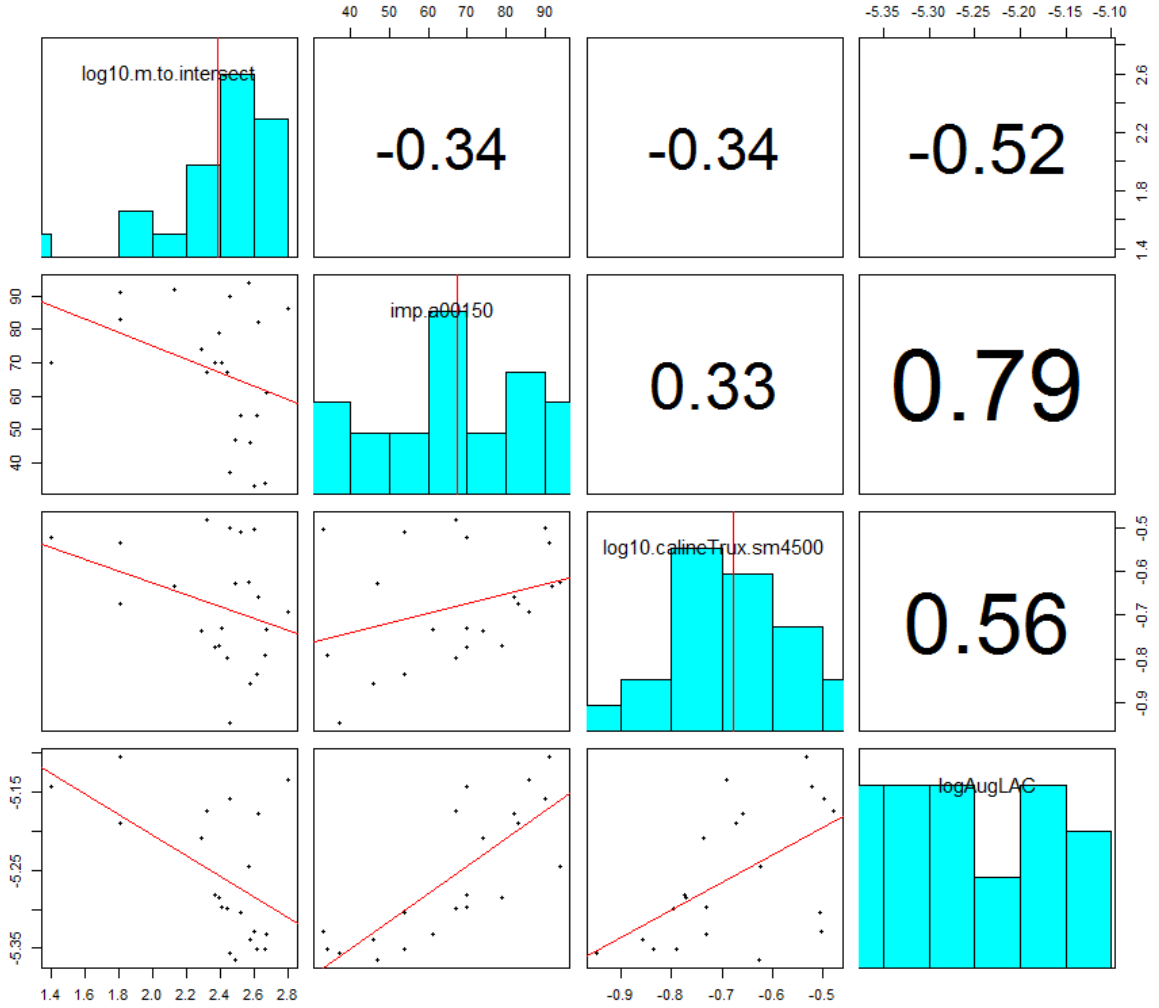
Figure 14. Map of August 1-NP prediction gradient

August Light-Absorbing Carbon

The spatial covariates included in the model of log_{10} August LAC are summarized in Table 6 below. The correlation matrix in Figure 15 shows the correlations and scatterplots for each pair of terms in the model. This model has an R^2 of 0.784, a cross-validated R^2 of 0.658, and a cross-validated RMSE of 0.0489 log_{10} m^{-1}. The final model covariates are log_{10} meters to intersection, area of impervious surface in 150m, and CAL3QHCR truck emission predictions in 4500m. The model residuals are mapped in Figure 16, and Figure 17 shows the gradient of August LAC predictions across the study area.

Table 6. Log_{10} August LAC model terms

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
log10.m.to.intersect	-0.0666	0.0321	-2.07	0.06	(-0.135, 0.00152)
imp.a00150	0.00265	0.000564	4.71	<0.001	(0.00146, 0.00385)
log10.calineTrux.sm4500	0.173	0.0776	2.23	0.04	(0.00847, 0.337)
(Intercept)	-5.16	0.102	-50.6	<0.001	(-5.38, -4.94)



Where *log10.m.to.intersect* = Log_{10} meters to intersection
imp.a00150 = Area of impervious surface in 150m
log10.calineTrux.sm4500 = CAL3QHCR truck emission predictions in 4500m
logAugLAC = Log_{10} August LAC measurements

Figure 15. Correlation matrix of August LAC model terms

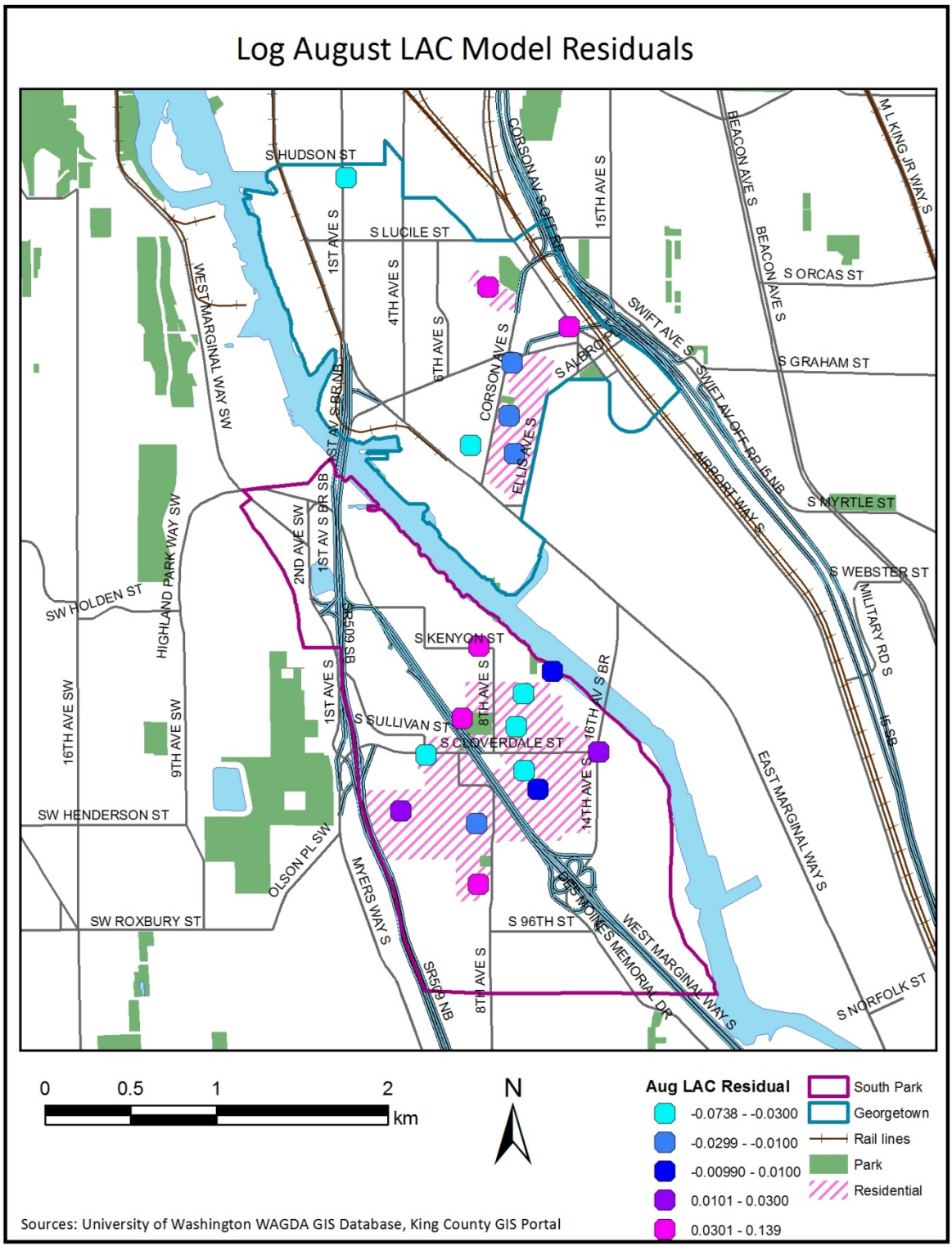


Figure 16. Map of $\log_{10}$ August LAC model residuals

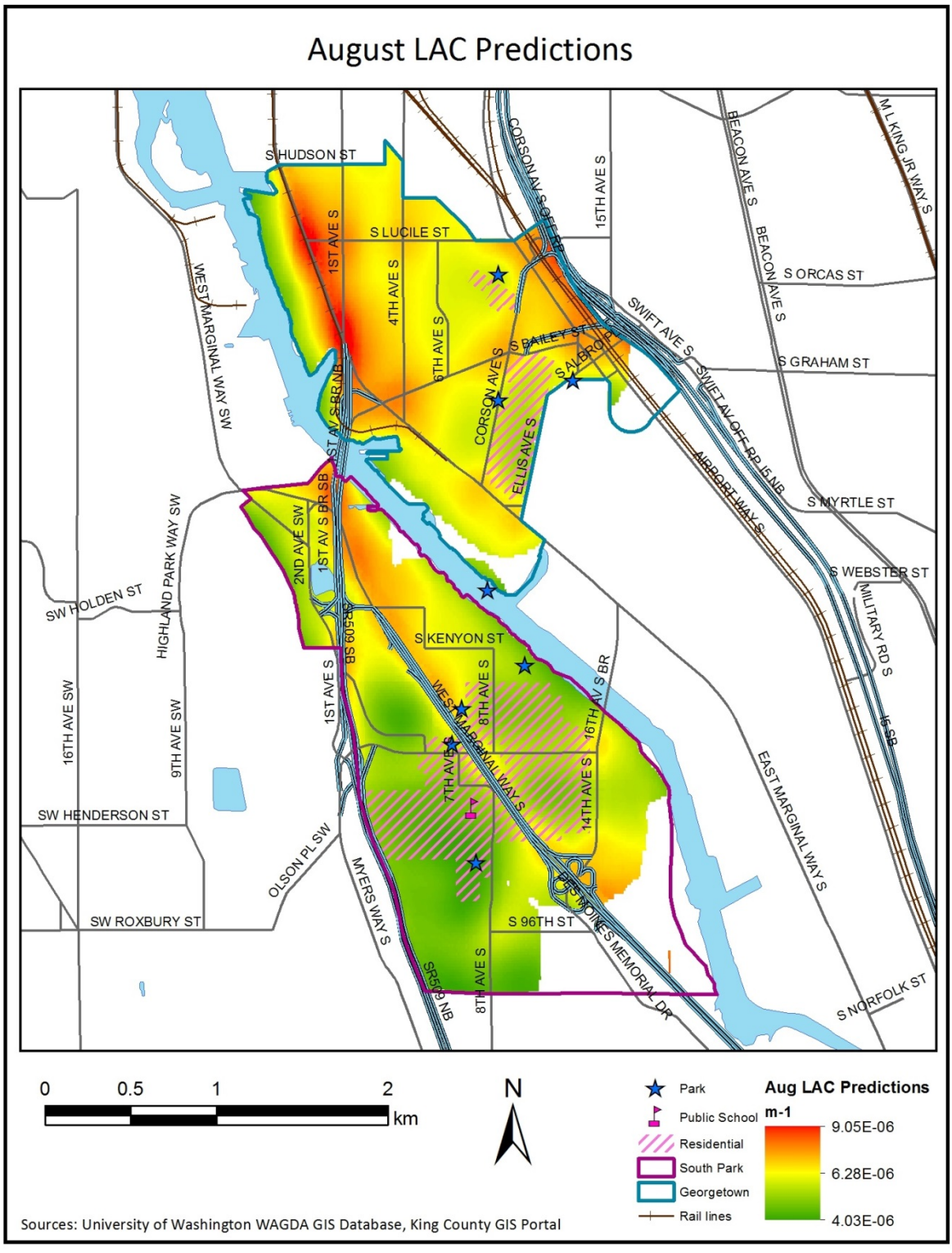


Figure 17. Map of August LAC prediction gradient

December Light-Absorbing Carbon

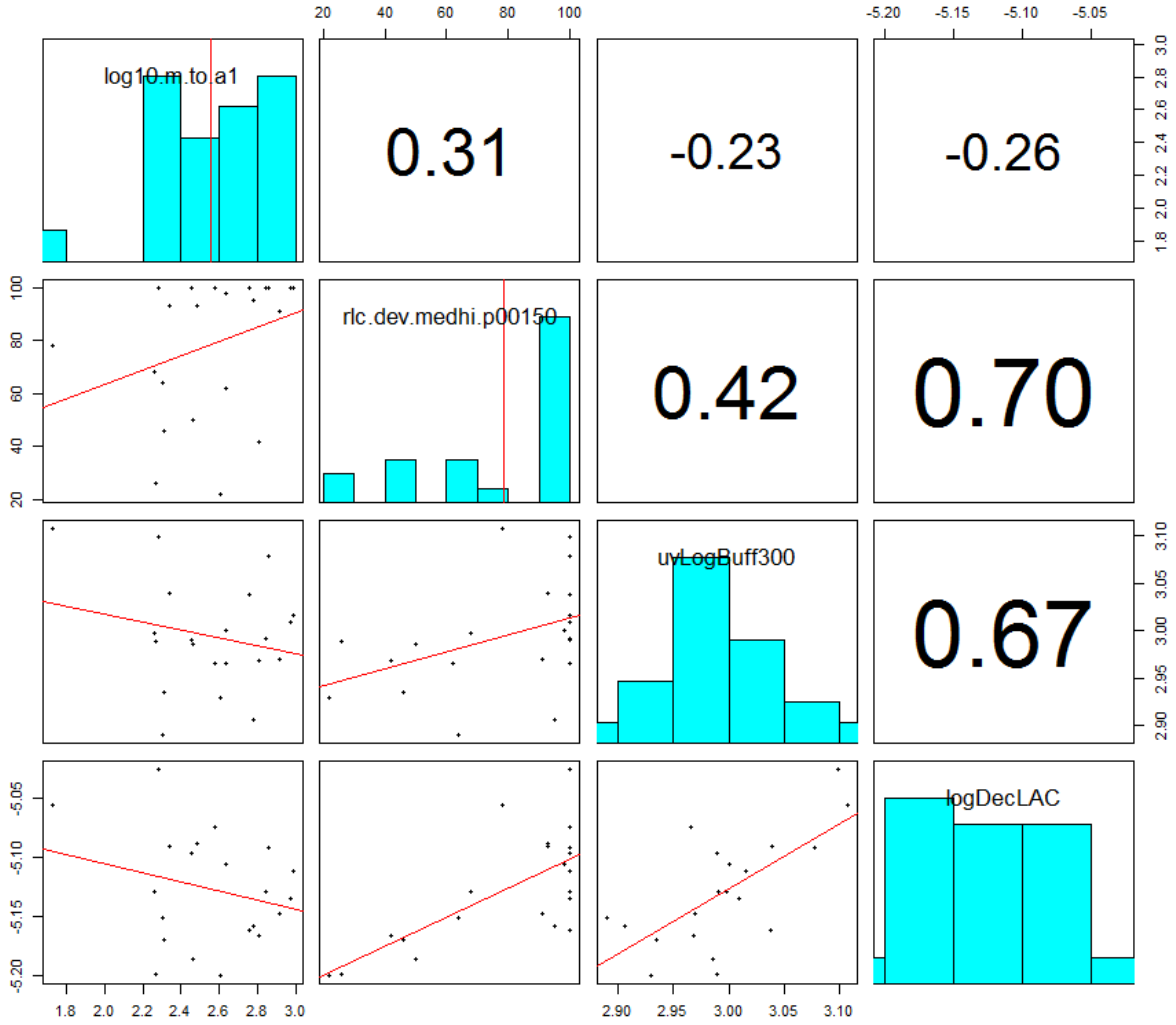
The spatial covariates included in the model of $\log_{10}$ December LAC are summarized in Table 7 below. The correlation matrix in Figure 18 shows the correlations and scatterplots for each pair of terms in the model. This model has an R^2 of 0.794, a cross-validated R^2 of 0.695, and a cross-validated RMSE of 0.0260 $\log_{10}$ m^{-1}. The final model covariates are $\log_{10}$ meters to a1 road[2], developed medium and high intensity areas[3] in 150m, and mean $\log_{10}$ ultraviolet-channel Aethalometer® readings in 300m. The model residuals are mapped in Figure 19, and Figure 20 shows the gradient of December LAC predictions across the core study area.

Table 7. $\log_{10}$ December LAC model terms

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
log10.m.to.a1	-0.0632	0.0196	-3.22	0.01	(-0.105, -0.0216)
rlc.dev.medhi.p00150	0.00124	0.000247	5.04	<0.001	(0.000721, 0.00177)
uvLogBuff300	0.230	0.112	2.05	0.06	(-0.00774, 0.468)
(Intercept)	-5.75	0.349	-16.5	<0.001	(-6.49, -5.01)

[2] A1 roads consist of interstates, state highways and the upper West Seattle bridge.

[3] In addition to the high intensity areas described in the August 1-NP model, the rlc.dev.medhi.p00150 land use category also includes medium intensity development, which is defined as "areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50% to 79% of the total cover. These areas most commonly include single-family housing units" (MRLC, 2006).



Where *log10.m.to.a1* = Log_{10} meters to a1 road
rlc.dev.medhi.p00150 = Developed medium and high intensity areas in 150m
uvLogBuff300 = Mean log_{10} ultraviolet-channel Aethalometer® readings in 300m
logDecLAC = Log_{10} December LAC measurements

Figure 18. Correlation matrix of log_{10} December LAC model terms

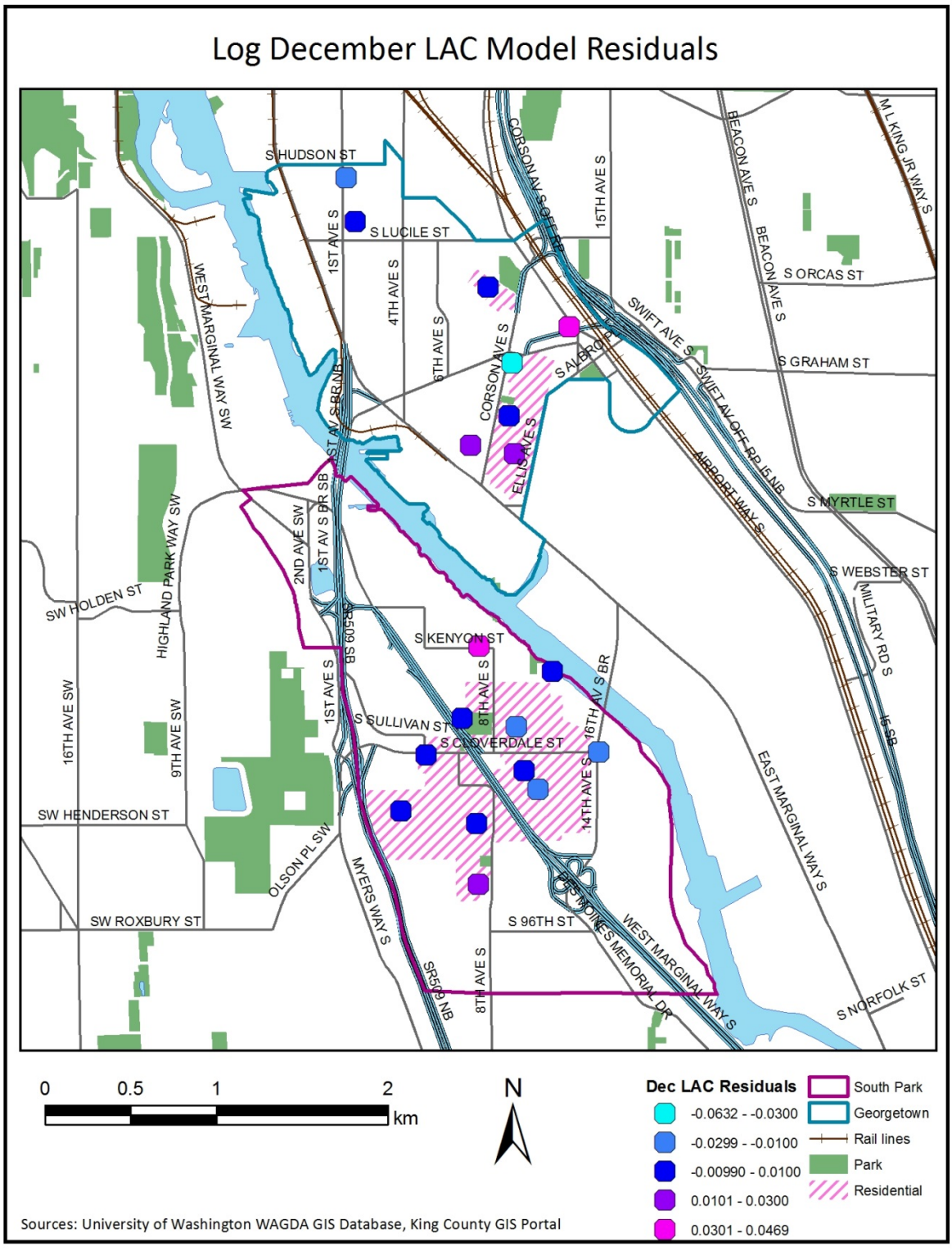


Figure 19. Map of $\log_{10}$ December LAC model residuals

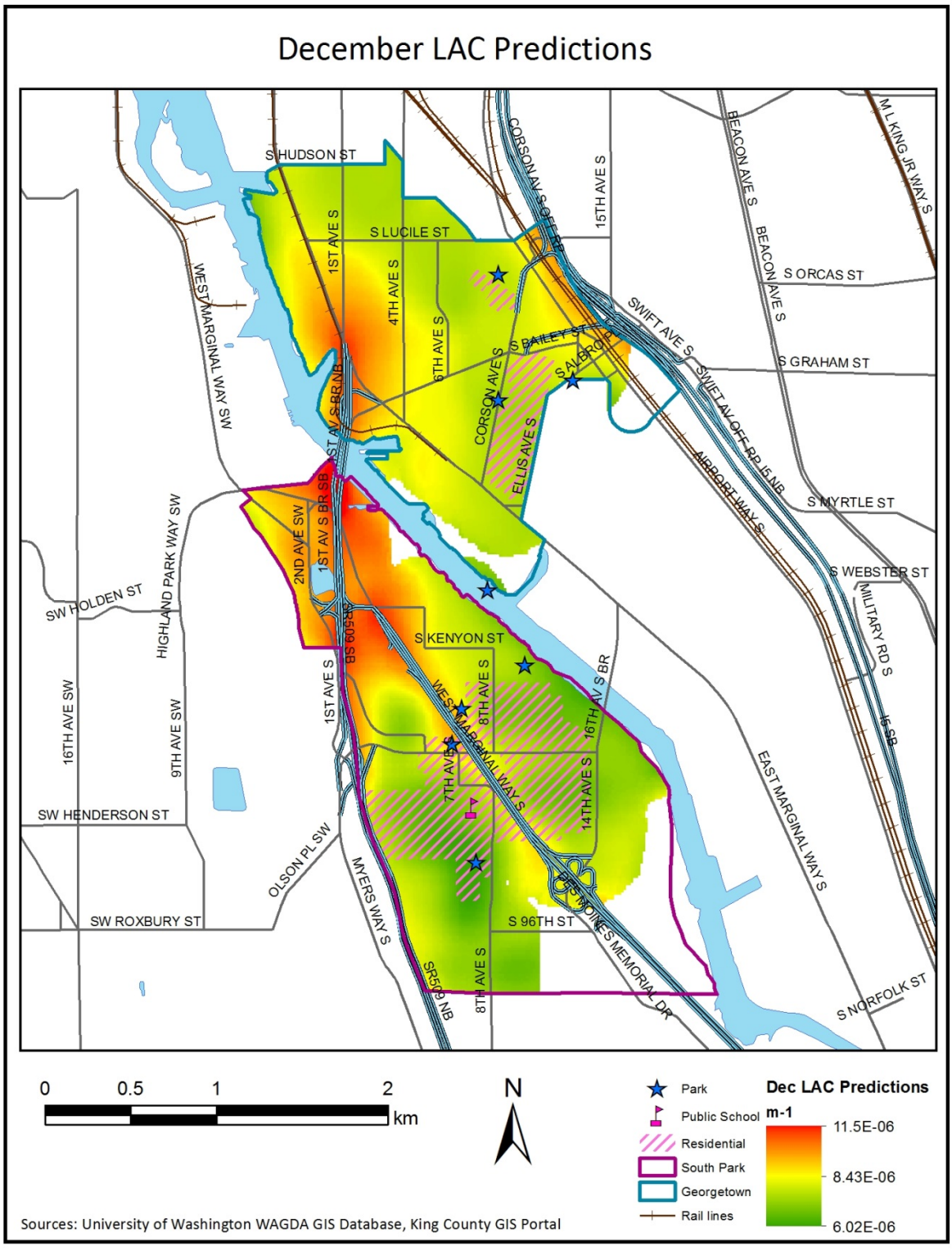


Figure 20. Map of December LAC prediction gradient

MODEL PERFORMANCE

The performance statistics for each model are displayed in Table 8 below, both in the $\log_{10}$ units used to develop the models and in the original units, where exponentiated predictions were compared to the original measurements.

Table 8. Summary of model performance statistics.

Model	Model R^2	Cross-Validated R^2 (Log_{10} Units)	Cross-validated RMSE (Log_{10} Units)	Cross-validated R^2 (Original Units)	Cross-validated RMSE (Original Units)
August 1-NP	0.868	0.730	0.123 $\log_{10}$ pg/ m^3	0.644	0.301 pg/m^3
August LAC	0.784	0.658	0.0489 $\log_{10}$ m^{-1}	0.646	6.59E-07 m^{-1}
December LAC	0.794	0.695	0.0260 $\log_{10}$ m^{-1}	0.698	4.50E-07 m^{-1}

In Figure 21, cross-validated model predictions at the 20 sampling sites are plotted against the observed values at those sites.

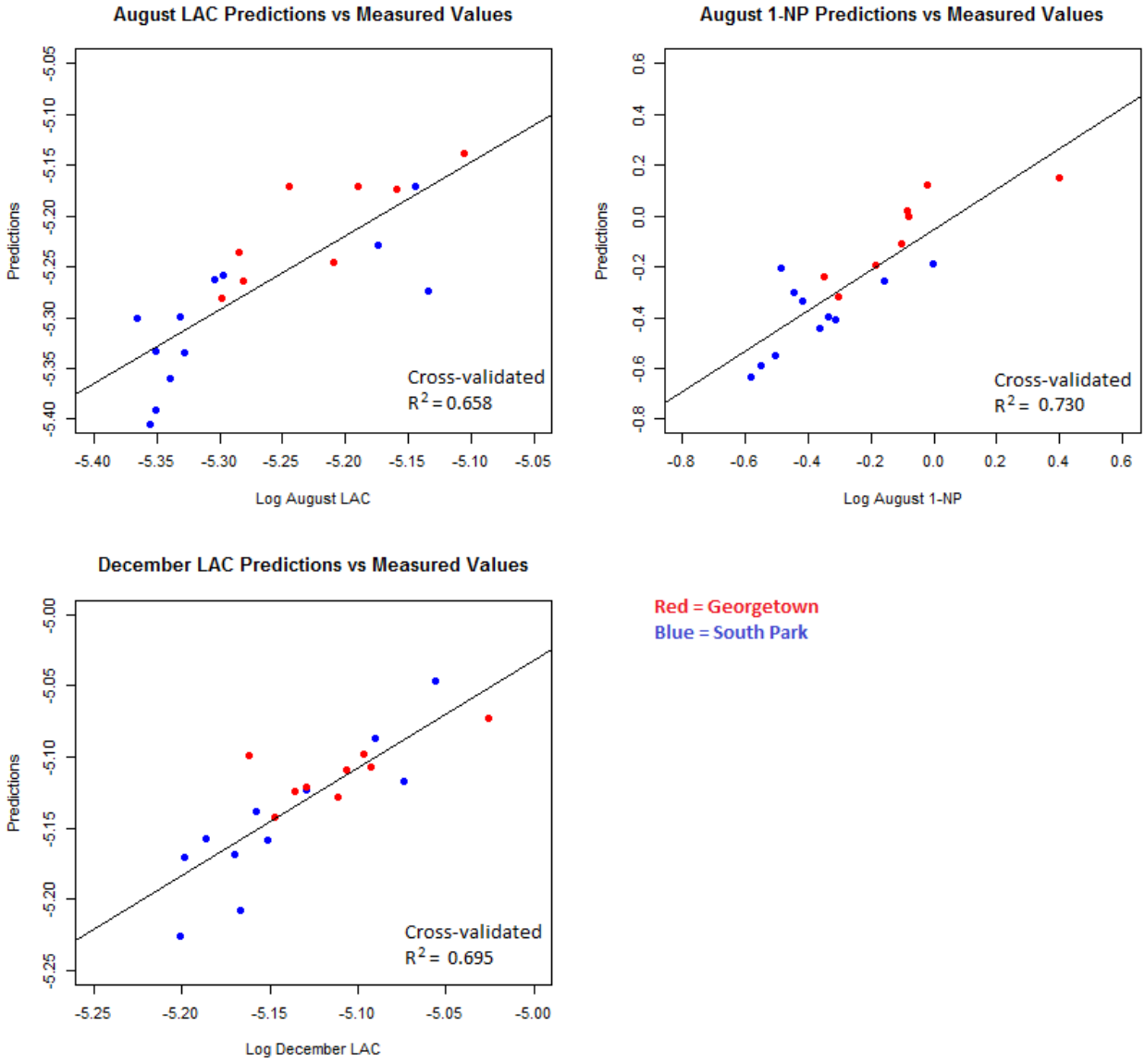


Figure 21. Scatterplots of predicted versus measured values for the 3 models.

SUMMARY OF PREDICTED VALUES

The mean predictions of the three pollutants are summarized in Table 9 by neighborhood, within both the full neighborhood boundaries and areas of residential zoning by neighborhood. The measurement results from the Queen Anne and Beacon Hill agency sites are included for comparison.

Table 9. Mean predictions by neighborhood with comparison site measurements

	Georgetown Mean*	South Park Mean*	Georgetown Residential Mean*	South Park Residential Mean*	Queen Anne Agency Site**	Beacon Hill Agency Site**
August 1-NP (pg/m^3)	1.58	0.910	0.631	0.379	0.211	0.341
August LAC (m^{-1})	6.50E-06	5.54E-06	5.69E-06	5.04E-06	1.74E-06	4.19E-06
December LAC (m^{-1})	8.19E-06	7.98E-06	7.40E-06	7.18E-06	4.22E-06	4.24E-06

* Mean of predictions
**Measured value

MODEL COMPARISONS

The modeling process was repeated excluding mobile monitoring data for August 1-NP and December LAC to identify the "next best" models in the absence of mobile data. As the inclusion of mobile monitoring data as a spatial covariate is a novel technique, an assessment of the degree to which these data enhanced model performance may help justify their use in future studies.

August 1-Nitropyrene

The $\log_{10}$ August 1-NP model without mobile data is summarized in Table 10 below. This model has an R^2 of 0.919, a cross-validated R^2 of 0.682 and a cross-validated RMSE of 0.136 $\log_{10}$ pg/m^3.

Table 10. Log_{10} August 1-NP model without mobile data

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
log10.m.to.rr	-0.269	0.0627	-4.29	<0.001	(-0.402, -0.135)
ll.a1.s00150	0.000776	0.000153	5.09	<0.001	(0.000451, 0.00110)
intersect.s00500	0.0116	0.00332	3.49	0.003	(0.00450, 0.0186)
rlu.dev.hi.p00150	0.00302	0.000755	4.01	0.001	(0.00142, 0.00463)
(Intercept)	0.278	0.194	1.43	0.17	(-0.136, 0.691)

December Light-Absorbing Carbon

The log_{10} December LAC model without mobile data is summarized in Table 11 below. This model has an R^2 of 0.805, a cross-validated R^2 of 0.652 and a cross-validated RMSE of 0.0281 log_{10} m^{-1}.

Table 11. Log_{10} December LAC Model without mobile data

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
log10.m.to.a1	-0.0984	0.0177	-5.55	3.57E-05	(-0.136, -0.609)
log10.m.to.rr	-0.0310	0.0174	-1.78	0.0935	(-0.0677, 0.00581)
rlc.dev.openlow.p00150	-0.00139	0.000279	-4.97	0.000117	(-0.00198, -0.000798)
(Intercept)	-4.77	0.0680	-70.1	2.00E-16	(-4.91, -4.62)

A comparison of performance statistics for both sets of models with and without mobile data is provided in Table 12 below.

Table 12. Comparison of log_{10} August 1-NP and log_{10} December LAC models with and without mobile data.

Model	Model R^2	Cross-Validated R^2 (Log_{10} units)	Cross-validated RMSE (Log_{10} units)	Cross-validated R^2 (Orig. units)	Cross-validated RMSE (Orig. units)
Log_{10} August 1-NP with mobile data	0.868	0.730	0.123 log_{10} pg/m^3	0.644	0.301 pg/m^3
Log_{10} August 1-NP without mobile data	0.919	0.682	0.136 log_{10} pg/m^3	0.704	0.293 pg/m^3
Log_{10} December LAC with mobile data	0.794	0.695	0.0260 log_{10} m^{-1}	0.698	4.50E-07 m^{-1}
Log_{10} Dec LAC without mobile data	0.805	0.652	0.0281 log_{10} m^{-1}	0.644	4.93E-07 m^{-1}

This comparison shows that the models with mobile data outperform the models without mobile data for each pollutant by a marginal amount in most but not all performance measures. Both models without mobile data have a higher R^2 than their corresponding models with mobile data before cross-validation. Both models with mobile data have a higher cross-validated R^2 and a lower cross-validated RMSE than their corresponding models without mobile data, but only in log_{10} units. In original units, the model with mobile data outperforms the model without mobile data for December LAC, but the model without mobile data outperforms the model with mobile data for August 1-NP. The inclusion of mobile monitoring data appears to enhance the performance of most cross-validated models to a small extent.

Two-Season Models

To assess the need for season-specific models, additional models were built using aggregated data from both seasons. The averages of $\log_{10}$ summer and winter pollutant concentrations, standardized to the mean and standard deviation within each season, were used as dependent variables in the lasso method. For each pollutant, stepwise regression was then performed on the variables selected by lasso using a full, non-standardized dataset with 2 data points per site and a binary "season" variable.

The 2-season $\log_{10}$ 1-NP model is summarized in Table 13.

Table 13. Two-season $\log_{10}$ 1-NP model terms

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
log10.m.to.a1	-0.205	0.0875	-2.34	0.03	(-0.383, -0.0265)
log10.m.to.rr	-0.170	0.0759	-2.24	0.03	(-0.325, -0.0149)
ll.a1.s00150	0.000368	0.000219	1.68	0.10	(-7.95E-05, 8.15E-04)
intersect.s00500	0.00889	0.00376	2.37	0.02	(0.00122, 0.0165)
rlu.dev.hi.p00150	0.00173	0.000757	2.28	0.03	(1.84E-04, 0.00327)
rlc.anyforest.p00750	-0.0111	0.00668	-1.66	0.11	(-0.0247, 0.00252)
season	0.542	0.0368	14.7	<0.001	(0.467, 0.617)
(Intercept)	0.0812	0.350	0.232	0.82	(-0.633, 0.795)

The 2-season $\log_{10}$ 1-nitropyrene model has an R^2 of 0.908, a cross-validated R^2 of 0.765 and a cross-validated RMSE of 0.164 $\log_{10}$ pg/m^3. Based on the R^2 statistics, this model appears to compare favorably to the August-specific $\log_{10}$ 1-NP model. However, when this 2-season model is assessed within each individual season, its predictive ability is diminished. When restricted to August data, the 2-season model has a cross-validated R^2 of 0.546 and a cross-validated RMSE of 0.161 $\log_{10}$ pg/m^3. When restricted

to December data, the 2-season model has a cross-validated R^2 of 0.120 and a cross-validated RMSE of 0.167 $\log_{10}$ pg/m^3.

The 2-season $\log_{10}$ LAC model is summarized in Table 14.

Table 14. Two-season $\log_{10}$ LAC model terms

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
(Intercept)	-6.17	0.376	-16.4	<0.001	(-6.93, -5.41)
season	0.129	0.0151	8.54	<0.001	(0.0986, 0.160)
imp.a00150	0.00215	0.000447	4.82	<0.001	(0.00125, 0.00306)
bcLogBuff300	0.214	0.130	1.65	0.11	(-0.0488, 0.477)

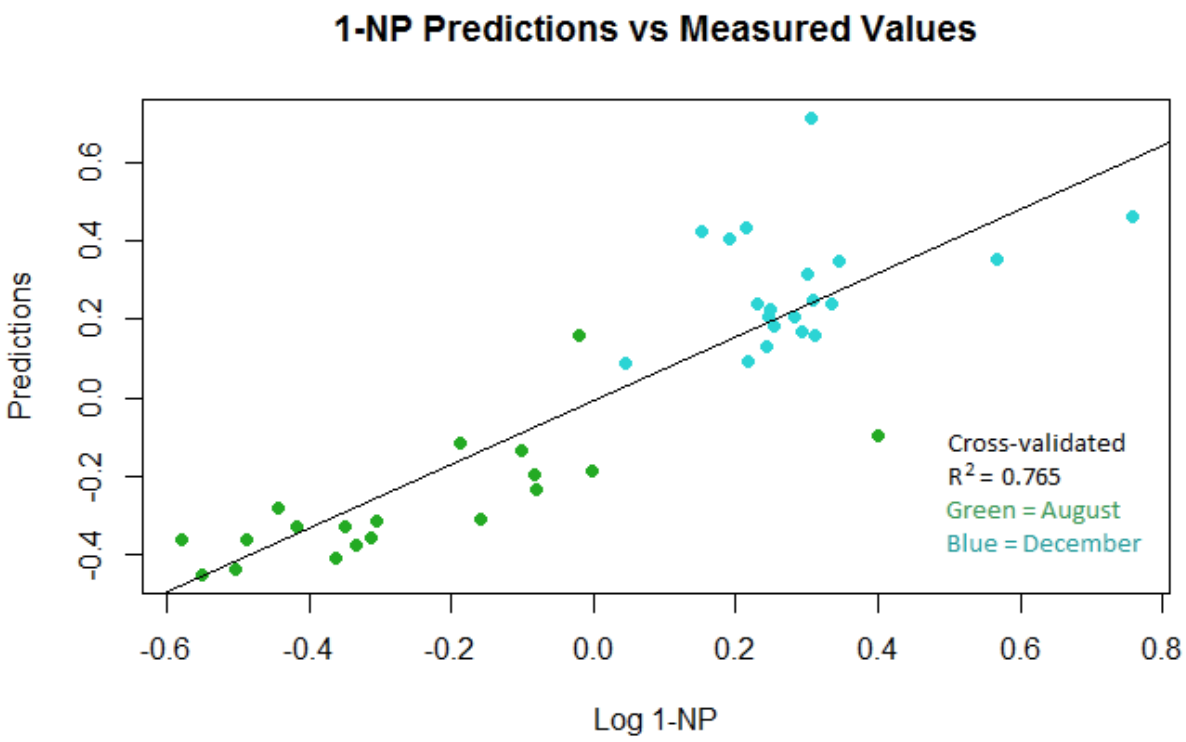
Similarly, the performance of the 2-season $\log_{10}$ LAC model differs when applied to the full dataset versus the individual seasons. For the full dataset, its R^2 is 0.769, cross-validated R^2 is 0.695 and cross-validated RMSE is 0.0523 $\log_{10}$ m^{-1}. Restricted to the August data, the cross-validated R^2 is 0.431 and cross-validated RMSE is 0.0627 $\log_{10}$ m^{-1}. Restricted to the December data, the cross-validated R^2 is 0.458 and cross-validated RMSE is 0.0392 $\log_{10}$ m^{-1}.

Table 15 summarizes the model performance of the original season-specific models and the 2-season models.

Table 15. Summary of model performance statistics for season-specific models and 2-season models

Model	Model R^2	Cross-Validated R^2 (Log_{10} units)	Cross-validated RMSE (Log_{10} units)
Two-season 1-NP	0.908	0.765	0.164 $\log_{10}$ pg/m^3
Two-season 1-NP (August-restricted)	N/A	0.546	0.161 $\log_{10}$ pg/m^3
August 1-NP (original model)	0.868	0.730	0.123 $\log_{10}$ pg/m^3
Two-season 1-NP (December-restricted)	N/A	0.120	0.167 $\log_{10}$ pg/m^3
Two-season LAC	0.769	0.695	0.0523 $\log_{10}$ m^{-1}
Two-season LAC (August restricted)	N/A	0.431	0.0627 $\log_{10}$ m^{-1}
August LAC (original model)	0.784	0.658	0.0489 $\log_{10}$ m^{-1}
Two-season LAC (December restricted)	N/A	0.458	0.0392 $\log_{10}$ m^{-1}
December LAC (original model)	0.794	0.695	0.0260 $\log_{10}$ m^{-1}

The scatterplots in Figure 22 show the predicted concentrations from the two-season models, plotted against the measured values at those sites. In particular, the poor predictive ability of the 1-NP model for December data is visible in the absence of a linear trend at the right side of the upper plot.



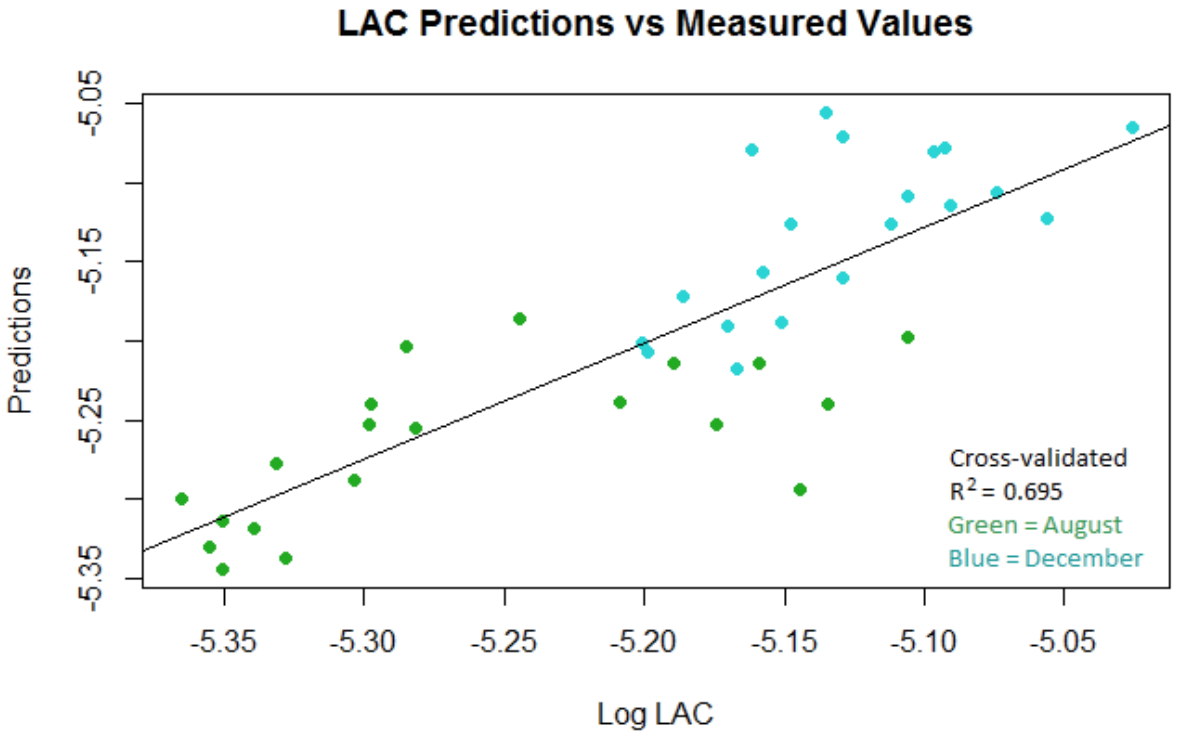


Figure 22. Scatterplots of predicted versus measured values for two-season models

DISCUSSION

Land-use regression techniques were successfully applied to measurements of diesel exhaust markers from a high-density sampling campaign, yielding finely-resolved models within individual neighborhoods. With the exception of 1-NP in December, clear gradients in concentrations of diesel exhaust markers were observed across small areas. Models of August 1-NP, August LAC and December LAC were stable throughout several hundred repetitions of the lasso and stepwise procedures. Spatial features that predicted concentrations of diesel exhaust markers in South Park and Georgetown varied between pollutants and between seasons, but in general they included road and railroad proximity, land use, and mobile source emissions predictions and measurements.

The maps of model residuals do not show obvious trends by zoning category, indicating that models performed comparably well both inside and outside of residential zones. The August 1-NP model under-predicted several sites in the northern half of the South Park residential area, though other residential sites were over-predicted. The August LAC model under-predicted some sites near residential arterials and over-predicted some inner residential sites. The map of December LAC model residuals shows some under-prediction at the outskirts of the South Park residential area, though residuals were near 0 at most residential sites. These trends suggest that while the spatial models do not capture all the fine-scale variations in measured concentrations, they are not systematically under- or over-predicting in residential areas.

Spatial modeling of 1-NP is novel to this study, but land-use regression models of LAC were previously developed by Larson *et al*. (2009). Significant predictors of log(mean) particle light absorption coefficient in that study were truck volumes within 300m, major roads within 300m and latitude. Larson *et al*. reported a cross-validated R^2 of 0.57 for this model and a RMSE of 0.40 $\log_{10}$ m^{-1}. Similarly, meters to A1 roads and intersections appear in spatial models of LAC for the DEEDS project, as road density is a consistent predictor of LAC concentrations. Dispersion model predictions appear in DEEDS models in lieu of truck volumes, as they are based in part on truck volumes on nearby roads. The improved performance of DEEDS models over this previously reported model indicates that CAL3QHCR predictions and the use of mobile monitoring data as spatial covariates can enhance the predictive ability of land-use regression models of LAC.

As expected, mean predicted pollutant concentrations in the South Park and Georgetown neighborhoods were higher than measured concentrations at the Queen Anne and Beacon Hill comparison sites, both of which were located atop hills in areas with less commercial truck traffic. Mean predicted concentrations in Georgetown were higher than those in South Park, both in residential zones and in the neighborhoods as a whole.

Predicted August 1-NP concentrations were highest along the 1st Avenue bridge that crosses the Duwamish River and near the Georgetown commercial district, which is located a short distance from both I-5 and a network of railroad tracks. Several multi-family apartment complexes are located in the Georgetown commercial district, and

these models suggest that residents of these buildings may face the highest levels of 1-NP pollution of Duwamish Valley residents. Outside of these areas, predictions of August 1-NP in the residential zones in both South Park and Georgetown are relatively homogenous and low.

In both August and December, LAC predictions show greater heterogeneity across the neighborhoods than 1-NP predictions. Predicted concentrations of LAC are also highest along the 1st Avenue bridge and in the Georgetown commercial district, similar to August 1-NP. Other areas of elevated LAC predictions include E Marginal Way S and 1st Avenue S in Georgetown, in the heart of Georgetown's industrial district. Predictions are also higher along W Marginal Way S (Washington State Route 99), which bisects South Park's residential zone and borders two of South Park's public park facilities. These models predict that South Park and Georgetown residents whose homes are near major arterial roads face higher concentrations of LAC in both summer and winter. While LAC concentrations may be driven by other sources in addition to diesel exhaust, these near-arterial areas of South Park and Georgetown may be the most appropriate areas to target for policy interventions to improve air quality.

December 1-NP was found to be unsuitable for spatial modeling because of a low coefficient of variation (0.49) and low signal-to-noise ratio within residential neighborhoods where the majority of sites were located. The lasso procedure consistently yielded a model with 0 variables, indicating that the mean measured concentration in December was a better predictor of winter 1-NP across the neighborhoods than any of the spatial covariates.

The most plausible explanation for the unsuitability of December 1-NP data for modeling is the difference in meteorological conditions between the August and December sampling periods. The August sampling period was relatively stagnant and dry, with an average wind speed of 4.0 mph and 0.0 inches of precipitation. In these stagnant conditions, measured concentrations would be driven largely by local sources, which were well represented by the spatial covariates considered for modeling. The average wind speed during December sampling was slightly higher at 7.0 mph (see Figure 23), and 3.2 inches of precipitation fell during the sampling period. Higher winds and precipitation would likely dilute the effects of local sources, causing greater regional mixing of pollutants and lower overall variability. That a suitable model was found for LAC during the same December sampling period suggests that there may be differences in dispersion patterns between the two pollutants, though little information on the dispersion of 1-NP exists in the literature.

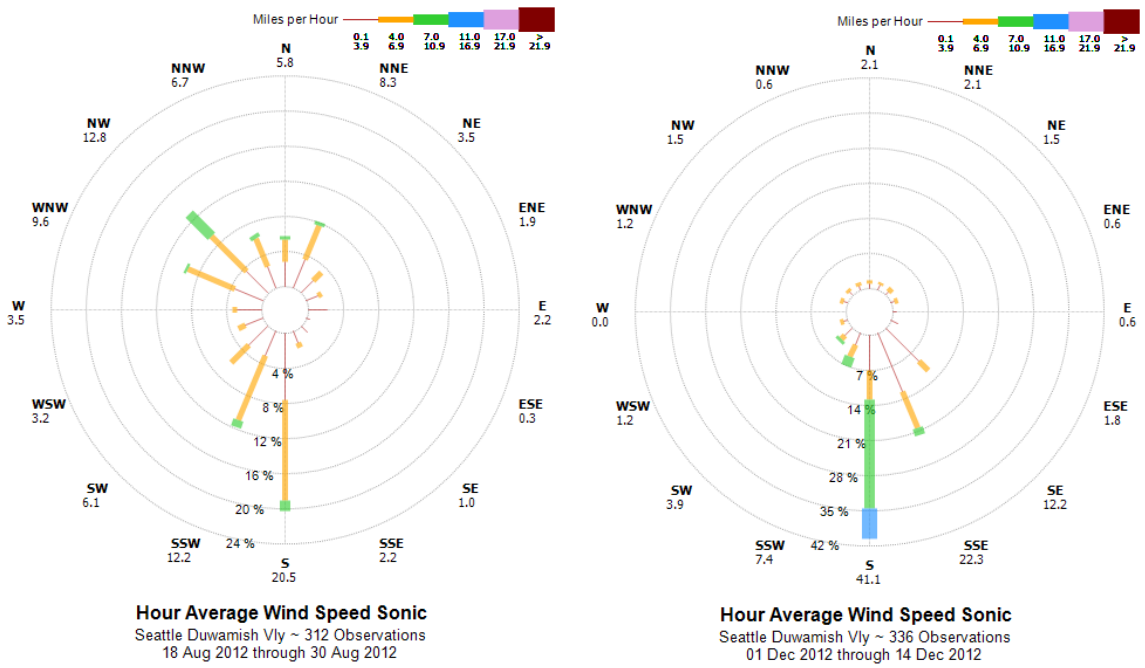


Figure 23. Wind roses from the August (left) and December (right) sampling periods. The length of each spoke represents the frequency of the wind direction, while the color represents the speed. Source: Puget Sound Clean Air Agency.

The comparison of season-specific and two-season models shows that two-season models are not able to predict concentrations of either pollutant within an individual season as well as season-specific models. The R^2 value of the two-season models is relatively high, but this is mainly due to the models' ability to predict higher values in winter than in summer, not to predict the spatial distribution of pollution. Though single-season models are commonly used in land-use regression models of air pollution, the modeling results from this study show that the spatial predictors of diesel exhaust concentrations vary between seasons. A key advantage of separate season-specific models is that they capture the specific spatial predictors of pollution relevant during the sampling period.

The results of the model comparison exercises also provide insight into the utility of mobile monitoring data in future spatial models of traffic-related pollutants. The comparison of models with and without mobile monitoring data suggests that these data can improve some measures of cross-validated model performance to a small degree. In future studies, the marginal increase in model performance may not justify the additional cost, time and effort involved in collecting and processing mobile monitoring data for use as an independent variable. However, in the DEEDS study the collection of mobile monitoring data was substantially less expensive and staff-intensive than the filter-based sampling campaigns. While the mobile monitoring data was of limited utility as a supplement to filter-based sampling data, in future studies it may be a cost-effective alternative to active sampling at fixed sites.

In addition, the design of the mobile monitoring route may have introduced bias by over-representing roads near sampling sites. Previous studies using mobile monitoring data, including Larson *et al.* (2009), have used aggregated mobile data as the dependent variable in land-use regression modeling rather than an independent variable. For that type of study design, driving repeated loops around sites provided the best approximation of concentrations at sites of interest and supplied enough data points per site to dilute the influence of any short-term spikes in pollution from traffic congestion.

By using mobile monitoring data as an independent variable, the DEEDS models instead considered these data a proxy for mobile source emissions. The bcLogBuff300 and uvLogBuff300 variables in the spatial models of 1-NP and LAC can be interpreted

as the mean intensity of on-road soot emissions in the vicinity of sampling sites. In spite of the different application of the mobile data, the mobile monitoring route was selected with guidance from previous studies and included repeated loops around sampling sites. Consequently, mobile monitoring data points were concentrated near sampling sites and were less abundant in other areas of the neighborhood grid, which likely introduced bias into the model predictions, particularly in outlying areas. The smallest buffer distance that offered sufficient coverage of the neighborhoods was 300m, and as a result the mean mobile monitoring value in a 300m buffer was driven in large part by which streets in a 300m area were selected for sampling.

One example of this bias was observed in the southwest corner of the South Park neighborhood on either side of Washington State Route 509 (SR509). The mobile monitor drove along SR509 and along several residential streets to the east. No residential streets to the west of SR509 were covered by the mobile monitor. The bcLogBuff300 variable is highest west of SR509, because the mobile monitoring points along SR509 are the only points that fall within the 300m buffer in these westernmost areas. Along SR509, the bcLogBuff300 values are diluted by the residential roads to the east that fall within the 300m buffer. In reality, the highest black carbon concentrations would be expected along SR509, and concentrations would be expected to drop with distance away from this major road on either side. Instead, concentrations appear to rise to the west because of the streets selected for mobile monitoring (see Figure 24).

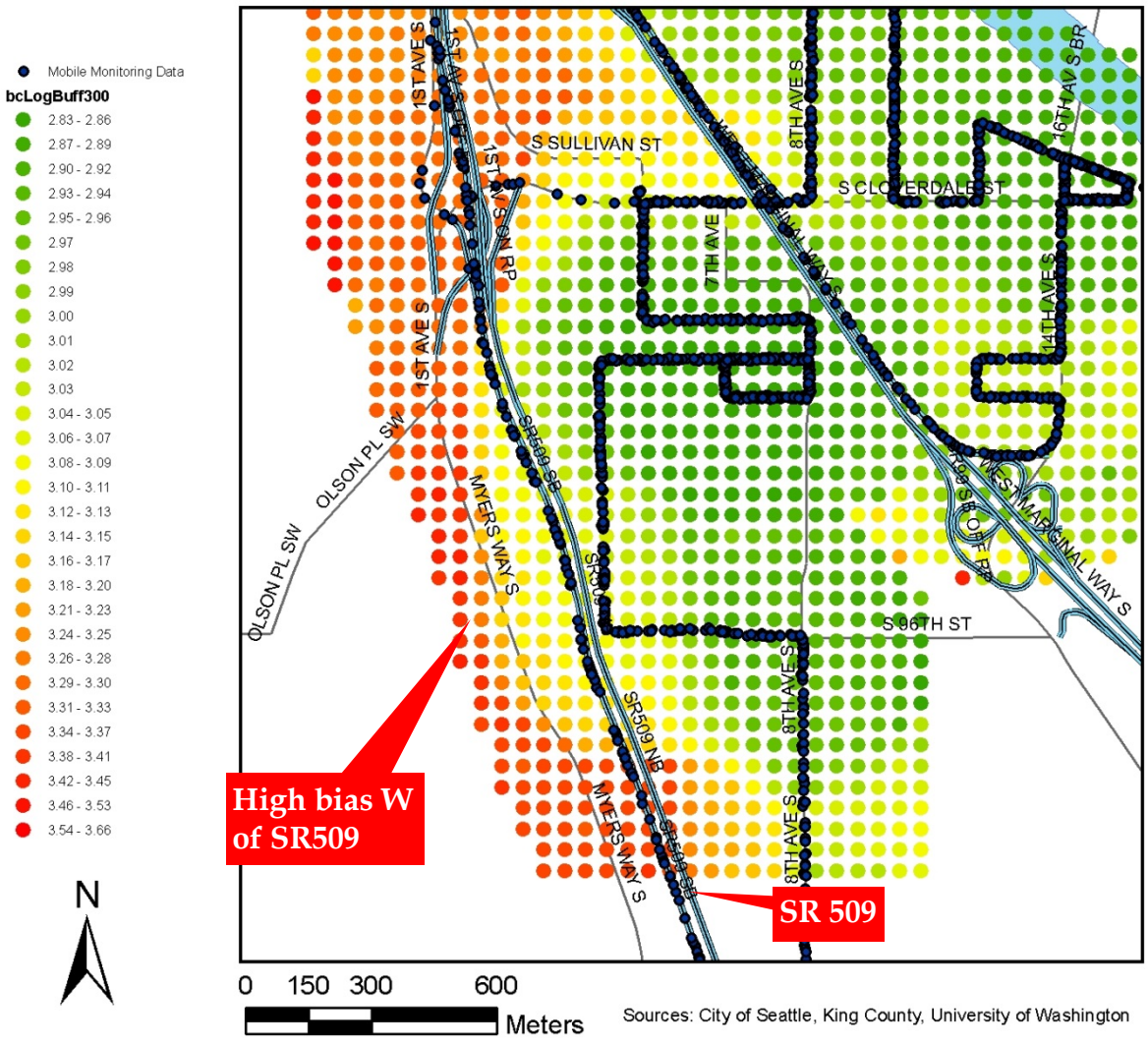


Figure 24. Close-up of bcLogBuff300 variable and mobile monitoring data points in southwest South Park

This issue suggests that the mobile monitoring route chosen for the DEEDS study was not the most appropriate route for inclusion of mobile monitoring data as model covariates. In future studies using mobile data as an independent variable, a better strategy would be to design a mobile route on a representative sample of neighborhood roads evenly distributed throughout the neighborhoods, regardless of the locations of

sampling sites. Such a route would be a more suitable proxy for mobile source emissions.

In addition to the difficulties with December 1-NP and mobile monitoring data, this study has a number of limitations. Models were based on the results of two snapshot campaigns of less than 2 weeks each, both of which were conducted during somewhat atypical weather patterns. The August sampling period was drier (0.0 inches of total precipitation) than the summer 2012 average (0.022 inches per day), though the average wind speed of 4.0 mph was the same during the sampling period and summer 2012 as a whole. The average of 0.25 inches of precipitation per day during the December sampling period was more than twice the daily average of 0.11 inches per day during all of winter 2012-2013. The average December sampling wind speed of 7.0 mph was also higher than the average of 5.0 mph from the entire winter season.

Since the weather patterns during the sampling periods were not entirely representative of typical weather for that season, the generalizability of the measurement results is limited. The Puget Sound Clean Air Agency collects continuous Aethalometer® black carbon readings at the SoDo agency monitoring site, which were used to compare pollutant concentrations during the sampling periods to seasonal averages. During the August sampling period, the mean black carbon concentration at this site was 0.92 $\mu g/m^3$, compared to the summer 2012 average of 1.1 $\mu g/m^3$. This comparison indicates that even though the average precipitation during the August sampling period was lower than the summer 2012 average, this did not result in unseasonably high concentrations of ambient black carbon. The average black carbon

concentration during the winter sampling period was 1.0 μg/m^3 and in winter 2012-2013 was 1.4 μg/m^3. The larger difference between these values suggests that the December sampling period was less representative of its season than the August sampling period. The particularly windy and wet weather in December likely resulted in unseasonably low pollutant concentrations during the sampling period.

Though the August sampling period was more representative of its season, the sunny summer weather may have contributed to atmospheric photodegradation of 1-NP prior to deposition, resulting in the relatively low concentrations of 1-NP measured in August. Previous studies show that 1-NP is readily decomposed in the presence of ultraviolet light (García-Berríos & Arce, 2012; van den Braken—van Leersum, Tintel, van 't Zelfde, Cornelisse, & Lugtenburg, 2010). This effect of sunlight may partially explain why August concentrations of 1-NP were lower than December 1-NP concentrations. Other factors that may have contributed to these seasonal differences include batch differences in lab analysis and differences in atmospheric conditions, such as mixing height.

Another limitation of the sampling time frame was that the study was conducted while the 16th Ave. S bridge connecting Georgetown and South Park was closed for reconstruction. The traffic patterns seen during the study period may not accurately reflect typical traffic patterns during normal bridge operation, which will limit the relevance of study results once the bridge reopens. Before the bridge closure, King County estimated that the bridge carried approximately 20,000 vehicle trips on an average weekday, roughly 2,800 of which were made by commercial trucks. A traffic

impacts analysis estimated that when the bridge closed, the majority of this traffic was diverted to the 1st Avenue S bridge and a smaller portion was diverted to the International Blvd S bridge south of the study area (King County, 2010). If traffic returns to its pre-closure pattern once the bridge reopens, concentrations of diesel exhaust markers can be expected to fall near the 1st Avenue S bridge and to rise near the 16th Avenue bridge in both South Park and Georgetown.

Another limitation of this study was the use of predictions from the CAL3QHCR model as a proxy for mobile-source emissions. These emissions predictions were based on modeled traffic counts rather than empirical traffic counts. In addition, these traffic counts were estimates of annual average daily traffic and were not specific to the sampling periods. They also were based on traffic patterns modeled before the closure of the 16th Avenue S bridge, which likely introduced additional error between the traffic model and the actual traffic patterns during the sampling periods.

There were also a number of limitations to site selection. The total number of sites included in the models (n=20) was relatively small for spatial modeling. The distribution of sites by land-use category overrepresented residential areas and underrepresented industrial areas at the request of community members. Because of the limited number of industrial sites, the high end of the range of pollutant predictions from all three models may be less precise. A substantial effort was made on the part of the research team to collaborate with the Port of Seattle and place a sampling site within port property, though ultimately permission was not granted to sample in any of the suitable locations identified. The nearest site to the Port of Seattle was several hundred

meters south, and thus the highest concentrations of diesel exhaust markers in the Duwamish Valley may not have been captured by the sampling campaigns.

CONCLUSIONS

This study demonstrated that hybrid dispersion/land-use regression models were able to identify a clear gradient in concentrations of diesel exhaust markers at a fine scale within individual neighborhoods. In addition, this study developed a spatial model with strong predictive ability for 1-NP, a novel marker in spatial modeling. Because 1-NP is more specific to diesel exhaust than other traffic-related pollutants previously modeled, this study was able to capture the specific sources and distribution of diesel exhaust with a higher degree of confidence.

The results indicate that residents of South Park and Georgetown are likely exposed to higher concentrations of diesel exhaust than residents of the Beacon Hill and Queen Anne comparison neighborhoods. The highest predictions of diesel exhaust markers were seen near the 1st Avenue bridge, the Georgetown commercial district, and major arterial roads in both neighborhoods. The existence of a gradient of diesel exhaust within these neighborhoods has implications for policy and community advocacy. Particularly in stagnant periods, the health and environmental impacts of diesel traffic do not appear to be evenly distributed across these neighborhoods. Residents in high intensity development areas near major roads and truck corridors, where a number of multifamily buildings are located, likely face disproportionate impacts of diesel traffic and higher exposure to diesel exhaust. Puget Sound Sage hopes to collaborate with

community members and use these results to advocate for improved air quality in South Park and Georgetown. Policy solutions to reduce idling, decrease individual truck emissions and divert commercial traffic away from residential areas may help to reduce levels of pollution in these neighborhoods, particularly in areas facing the highest impacts of truck activity.

REFERENCES

Andreae, M. O., & Gelencsér, A. (2006). Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. Atmospheric Chemistry and Physics Discussions, 6(3), 3131–3148.

Attfield, M. D., Schleiff, P. L., Lubin, J. H., Blair, A., Stewart, P. A., Vermeulen, R., Coble, J. B., et al. (2012). The diesel exhaust in miners study: a cohort mortality study with emphasis on lung cancer. Journal Of The National Cancer Institute, 104(11), 869–83.

Bond, T., Anderson, T., & Campbell, D. (1999). Calibration and intercomparison of filter-based measurements of visible light absorption by aerosols. Aerosol Science & Technology, (May 2013), 37–41.

Briggs, D., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., et al. (1997). Mapping urban air pollution using GIS: a regression-based approach. International Journal of Geographical Information Science, 11(7), 699-718. Taylor & Francis.

Chan, P. C. (1996). NTP Technical Report on Toxicity Study of 1-Nitropyrene (CAS No. 5522-43-0) Administered by Inhalation to F344/N Rats. Research Triangle Park, NC. Retrieved from http://ntp.niehs.nih.gov/ntp/htdocs/ST_rpts/tox034.pdf

Sam-Quarcoo Dotse, Joshua Kwame Asane and F.G. Ofosu, 2012. Particulate Matter and Black Carbon Concentration Levels in Ashaiman, a Semi-Urban Area of Ghana, 2008. Research Journal of Environmental and Earth Sciences, 4(01): 20-25.

García-Berríos, Z. I., & Arce, R. (January 01, 2012). Photodegradation mechanisms of 1-nitropyrene, an environmental pollutant: the effect of organic solvents, water, oxygen, phenols, and polycyclic aromatics on the destruction and product yields. The Journal of Physical Chemistry. A, 116, 14, 3652-64.

Garshick, E., Laden, F., Hart, J. E., Rosner, B., Davis, M. E., Eisen, E. A., & Smith, T. J. (October 01, 2008). Lung Cancer and Vehicle Exhaust in Trucking Industry Workers. Environmental Health Perspectives, 116, 10, 1327-1332.

Gilroy, M., Strange, K., & Yost, M. (2010). Tacoma and Seattle Area Air Toxics Evaluation. Puget Sound Clean Air Agency and University of Washington, Seattle, WA. Retrieved from http://www.pscleanair.org/news/library/reports/2010_Tacoma-Seattle_Air_Toxics_Report.pdf

Hayakawa, K., Butoh, M., & Miyazaki, M. (1992). Determination of dinitro- and nitropyrenes in emission particulates from diesel and gasoline engine vehicles by liquid chromatography with chemiluminescence detection after precolumn reduction. Analytica Chimica Acta, 266(2), 251–256.

Henderson SB, Beckerman B, Jerrett M and Brauer M (2007). Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. Environmental Science & Technology. 41(7); 2422-2428.

Hesterberg TW, Long CM, Bunn WB, Sax SN, Lapin CA, Valberg PA. Non-cancer health effects of diesel exhaust: a critical assessment of recent human and animal toxicological literature. Crit Rev Toxicol. 2009;39(3):195-227.

Hoek, G., Beelen, R., de, H. K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (October 01, 2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric Environment, 42, 33, 7561-7578.

IARC. (1985). 1-Nitropyrene (pp. 321–358). International Agency for Research on Cancer Monograph. Retrieved from: http://monographs.iarc.fr/ENG/Monographs/vol46/mono46-19.pdf

IARC. (2012). IARC: DIESEL ENGINE EXHAUST CARCINOGENIC. Press Release. Retrieved from http://www.iarc.fr/en/media-centre/pr/2012/pdfs/pr213_E.pdf

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., et al. (2005). A review and evaluation of intraurban air pollution exposure models. Journal of Exposure Analysis and Environmental Epidemiology, 15(2), 185–204.

Kakimoto, H., & Kitamura, M. (2000). Comparison of atmospheric polycyclic aromatic hydrocarbons and nitropolycyclic aromatic hydrocarbons in Kanazawa, Sapporo and Tokyo. Journal of health …, 39(32), 5817–5826.

King County (2010). South Park Bridge TIGER II Grant - Traffic analysis summary. Retrieved 5/28/2013, 2010, from http://kingcounty.gov/transportation/TIGERII/EconomicsAndJobs/TrafficAnalysis.aspx

King County (2012). KCGIS data download. Retrieved 02/01/2013, 2012, from http://www5.kingcounty.gov/gisdataportal/Default.aspx

Kinney PL, Aggarwal M, Northridge ME, Janssen NA, Shepard P. Airborne concentrations of PM(2.5) and diesel exhaust particles on Harlem sidewalks: a community-based pilot study. Environ Health Perspect. 2000 Mar;108(3):213-8.

Larson, T., Henderson, S. B., & Brauer, M. (January 01, 2009). Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. Environmental Science & Technology, 43, 13, 4672-8.

Magee Scientific. (2010). Introducing the new microAeth ® Model AE52 (pp. 1–2). Berkeley, CA.

Mercer, L. D., Szpiro, A. A., Sheppard, L., Lindström, J., Adar, S. D., Allen, R. W., Avol, E. L., et al. (2011). Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Atmospheric environment (Oxford, England : 1994), 45(26), 4412–4420.

Miller-Schulze JP, Paulsen M, Toriba A, et al. Exposures to particulate air pollution and nitro-polycyclic aromatic hydrocarbons among taxi drivers in Shenyang, China. Environ Sci Technol. 2010;44(1):216-221.

MRLC (2006). National Land Cover Database 2006 (NLCD2006) Product Legend. U.S. Geological Survey. Retrieved 5/9/2013 from http://www.mrlc.gov/nlcd06_leg.php.

Murakami, M., Yamada, J., Kumata, H., & Takada, H. (2008). Sorptive behavior of nitro-PAHs in street runoff and their potential as indicators of diesel vehicle exhaust particles. Environmental science & technology, 42(4), 1144–50.

Nightingale JA, Maggs R, Cullinan P, Donnelly LE, Rogers DF, Kinnersley R, Chung KF, Barnes PJ, Ashmore M, Newman-Taylor A. Airway inflammation after controlled exposure to diesel exhaust particulates. Am J Respir Crit Care Med. 2000 Jul;162(1):161-6.

Nordenhäll C, Pourazar J, Blomberg A, Levin JO, Sandström T, Adelroth E. Airway inflammation following exposure to diesel exhaust: a study of time kinetics using induced sputum. Eur Respir J. 2000 Jun;15(6):1046-51.

Novotny, E., & Bechle, M. (2011). National satellite-based land-use regression: NO2 in the United States. Environmental science & technology, (2), 4407–4414.

Olsson, A. C., Gustavsson, P., Kromhout, H., Peters, S., Vermeulen, R., Brüske, I., Pesch, B., et al. (2011). Exposure to Diesel Motor Exhaust and Lung Cancer Risk in a Pooled Analysis from Case-Control Studies in Europe and Canada. American Journal of Respiratory and Critical Care Medicine, 183(7), 941–948.

Palcisko, G. (2008). Health Consultation: Summary of Results of Duwamish Valley Regional Modeling and Health Risk Assessment. U.S. Department of Health and Human Services, Atlanta, GA.

PSRC. (2008). PSRC Travel Model Documentation (for Version 1.0). Puget Sound Regional Council, Seattle, WA.

Puget Sound Sage. (2009). Community Health Impact Survey Results: Port of Seattle Operations Hazardous to Health in Georgetown and South Park. Seattle, WA. Retrieved from http://pugetsoundsage.org/downloads/Community Health Impact Survey Overview.pdf.

Rao, V., & Somers, J. H. (2010). Black Carbon as a Short-Lived Climate Forcer : A Profile of Emission Sources and Co-Emitted Pollutants. U.S. Environmental Protection Agency.

Schauer JJ. Evaluation of elemental carbon as a marker for diesel particulate matter. J Expo Anal Environ Epidemiol. 2003 Nov;13(6):443-53.

Silverman, D. T., Samanic, C. M., Lubin, J. H., Blair, A. E., Stewart, P. a, Vermeulen, R., Coble, J. B., et al. (2012). The Diesel Exhaust in Miners study: a nested case-control study of lung cancer and diesel exhaust. Journal of the National Cancer Institute (Vol. 104, pp. 855–68).

Sydbom, a, Blomberg, a, Parnia, S., Stenfors, N., Sandström, T., & Dahlén, S. E. (2001). Health effects of diesel exhaust emissions. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology, 17(4), 733–46.

Tibshirani, R. (January 01, 1996). Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (methodological), 58, 1, 267-288.

U.S. Census Bureau. (2010). U.S. Census Bureau, Census 2010 Summary File 1. Retrieved 6/26/2012, from http://factfinder2.census.gov

U.S. Census Bureau (2010). U.S. Census Bureau, 2006-2010 American Community Survey. Retrieved 6/26/2012, from http://factfinder2.census.gov

U.S. EPA. (1995). User's Guide to CAL3QHC Version 2.0: A Modeling Methodology for Predicting Pollutant Concentration Near Roadway Intersections . U.S. Environmental Protection Agency, Research Triangle Park, NC.

U.S. EPA. (2012). Report to Congress on Black Carbon (pp. 115–137). U.S. Environmental Protection Agency, Research Triangle Park, NC.

Van den Braken-van Leersum, A. M.; Tintel, C.; van't Zelfde, M.; Cornelisse, J.; Lugtenburg, J. Recl. Trav. Chim. Pays−Bas 1987, 106, 120−128.

Wilton, D., Szpiro, A., Gould, T., & Larson, T. (2010). Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. Science of the Total Environment, 408(5), 1120–1130.

Yang, M., Howell, S. G., Zhuang, J., & Huebert, B. J. (May 05, 2008). Attribution of aerosol light absorption to black carbon, brown carbon, and dust in China - Interpretations of atmospheric measurements during EAST-AIRE. Atmospheric Chemistry and Physics Discussions, 8, 3, 10913-10954.