Spatial analysis and land use regression of VOCs and NO_2 in Dallas, Texas during two seasons

Luther A. Smith[a], Shaibal Mukerjee,[*b] Kuenja C. Chung[c], and Jim Afghani[c]

[a] Alion Science and Technology, Inc., Durham, NC, 27713, USA

[b] National Exposure Research Laboratory, U.S. Environmental Protection Agency (E205-03), Research Triangle Park, NC, 27711, USA

[c] U.S. Environmental Protection Agency, Region 6, 1445 Ross Avenue, Dallas, TX 75202, USA

* Corresponding author. Tel: +1 919 541 1865; fax: +1 919 541 4787

E-mail: mukerjee.shaibal@epa.gov

Passive air sampling for nitrogen dioxide (NO_2) and select volatile organic compounds (VOCs) was conducted at 24 fire stations and a compliance monitoring site in Dallas, Texas, USA during summer 2006 and winter 2008. This ambient air monitoring network was established to assess intra-urban gradients of air pollutants to evaluate impact of traffic and urban emissions on air quality. Ambient air monitoring and GIS data from spatially representative fire station sites were collected to assess spatial variability. Pairwise comparisons were conducted on the ambient data from the selected sites based on city section. These weeklong samples yielded NO_2 and benzene levels that were generally higher during the winter than the summer. With respect to location within the city, the central section of Dallas was generally higher for NO_2 and benzene than north and south. Land use regression (LUR) results revealed spatial gradients in NO_2 and selected VOCs in the central and some northern areas. The process used to select spatially representative sites for air sampling and the results of analyses of coarse- and fine-scale spatial variability of air pollutants on a seasonal basis provide insights to guide future ambient air

exposure studies in assessing intra-urban gradients and traffic impacts.

Introduction

The Dallas, Texas metropolitan area, with a population of over two million, had a growing concern about air quality due to elevated levels of nitrogen oxides and hazardous air pollutants potentially influencing ozone nonattainment. To gain a more complete overview of volatile organic carbon (VOCs) and nitrogen dioxide (NO_2) levels in the City of Dallas, the U.S. Environmental Protection Agency's (EPA) Region 6 and Office of Research and Development conducted monitoring of air toxics. These ambient monitoring data were analyzed to examine differences between sections of the city and combined with variables calculated in a geographic information system (GIS) to develop predicted pollutant levels across the city.

A large number of studies assessing spatial differences of urban air pollutants have employed the exposure prediction technique known as land use regression (LUR) modeling. In these studies, monitoring networks are typically established at a number of sites in an urban area using passive or other field-portable air monitoring devices. Monitored data combined with geographic information system (GIS)-derived variables such as proximity to roadways are used to develop LURs. The LURs can be used to predict ambient levels at residential locations to aid spatially-based epidemiologic health studies[1-5] as well as inform decisions regarding placement of monitoring sites.

Prior to the current study, EPA conducted air exposure monitoring studies at elementary schools in El Paso, Texas[6] and the Detroit, Michigan area[7] and subsequently developed LUR models to assess intra-urban variability of air pollutants for children's asthma studies. Passive air monitors were deployed to measure ambient levels of VOCs and NO_2, and LUR models were developed. Modeled pollutant concentrations were used to assess spatial differences in respiratory health effects among children attending the schools. School sites for monitoring were selected based on sampling convenience in El Paso and statistical analysis of GIS data in Detroit. Traffic-related variables, population density, distance to major point sources, and distance from border crossing, were common explanatory variables in the regression analyses for VOCs and NO_2. Analysis by city section indicated gradients of pollutant levels in El Paso due to elevation and limited NO_2 gradients in Detroit due to industrial/traffic influences. Based on this earlier experience, EPA determined that a similar approach could be applied to examine areas of elevated ambient VOCs and NO_2 in Dallas.

For this study, EPA deployed a passive monitoring network in Dallas during summer 2006 and winter 2008 to explore intra-urban variability and seasonality of hazardous air pollutants. As in the other studies, weeklong sampling periods were used to monitor NO_2 and select VOCs. Monitors were located at City of Dallas fire stations. Overall spatial analyses on a coarse level are presented by comparing city sections. As in El Paso and Detroit, finer scale variability and the influence of different variables on pollutant levels are assessed through the use of LUR models for Dallas. Estimates from the LUR models will be used to assess spatial variation of air quality throughout the city and inform spatial studies being conducted by EPA in other urban areas.

Methods

Selection of ancillary variables and air monitoring sites

The goal of this project was to gain a more complete overview of ambient levels to VOCs and NO_2 levels in Dallas. The study area was defined roughly as the interior of the loop formed by Interstate (I-635) to the north and east, I-20 to the south, State Highway 408 in the southwest, State Highway 12 to the west, and I-35 completing the loop in the northwest. In addition, a buffer of approximately two kilometers was added outside this area. Fig. 1 shows the Dallas fire stations where monitoring was conducted. The fire station numbers are detailed on the City of Dallas Fire Department web site (http://dallasfirerescue.com/sta_list/citymap.html). Use of fire stations offered several advantages. First and foremost, they were well-distributed across the city from a geographic standpoint and representative of ambient exposures in the immediate community. Fire stations typically had enough open accessible physical space to accommodate samplers and the potential for vandalism of the samplers was low since they were continuously staffed.

Spatially representative fire station sites were selected and LUR models developed based on traffic and other urban land-use variables from GIS databases. Based on previous LUR studies[1-3,5,6-9], consideration was given to the following types of ancillary predictor variables: distance to roads carrying certain volumes of vehicles; traffic intensity; population density and distance to point sources. Variables were generated using ArcView 3 and 9 (ESRI, Redlands, CA) with statistical analyses implemented in SAS version 9.1.[10,11] Data sources for variables were: 1) fire station location from City of Dallas Fire Department; 2) modeled traffic count data for Dallas County from the Texas Department of Transportation Travel Demand Forecast Model for 2000; 3) 2000 U.S. Census data; and 4) point source location and emissions data from the EPA 2002 National Emission Inventory database. Ancillary variables generated from these data sources are presented in Tables S1 to S4 of the supplementary data; see also Table1.

From these 51 variables, explanatory variables were selected by performing separate correlation analyses within four types of variable groups: distance to road; traffic intensity; housing unit and population density; distance from point sources. The selected variables had Pearson correlation coefficients > 0.7 with some non-selected variables within the same group and were generally weakly correlated with each other (Table 1). The philosophy behind selecting variables within a group that were weakly correlated was that adding a highly correlated variable to one already selected would not contribute much to the predictive capability. To be useful for predictive purposes, the selected variables also needed to exhibit a reasonable amount of variability across the population of fire stations. Based on these criteria and other considerations such as which data were most reliable and which variables were thought more likely (within their group) to influence the pollutants measured, the following eleven variables were selected as potential explanatory variables: five road distance variables, traffic intensity within one km of the site, population density, distance to two size categories of nitrogen oxide emitters, and distance to one size category each of benzene and ethylbenzene point sources. Table 1 presents these variables and the correlation structure among them for monitored and unmonitored fire stations. The selected variables exhibited a reasonable amount of variability (coefficient of variation, CV > 30%) across the population of fire stations.

The fire stations were ranked on each of the eleven variables and divided into six groups of nine based on these rankings. The groups were designated from 1 (nine lowest ranked sites) to 6 (nine highest ranked sites). These rankings provided the basis for the selection of monitoring sites. Monitoring locations were intentionally spread across Dallas but in such a way that high, medium, and low rankings were present in each part of the city. See Mukerjee *et al.*[7] for more detail on this approach.

This selection process ensured that the spatial analysis results would be representative of the Dallas study area. This was checked in two ways. First, Pearson correlations were calculated between each of the eleven potential predictors; calculations were done separately for selected and nonselected sites. Generally, correlations between variables were weak for both selected and nonselected groups of sites. More importantly, pairs of variables had similar correlations for sites chosen versus remaining sites. Table 1 reports the correlation for both chosen and nonselected sets of sites. A total of 24 sites were chosen from the pool of 55 potential fire stations (see Fig. 1). Finally, results of an eleven dimensional cluster analysis confirmed that the chosen sites were distributed across the various clusters constructed from all 55 fire stations. This site selection process, coupled with actual site visits to confirm feasibility, ensured that the subsequent spatial analysis of ambient data collected would be based on a representative sample of fire stations for Dallas.

Air monitoring

Passive samplers were deployed outdoors at the 24 spatially-representative fire station sites. Passive monitoring was also conducted at a regulatory-based compliance air monitoring station operated by the Texas Commission on Environmental Quality and City of Dallas. All study sites are shown in Fig. 1. The study area shaded in Fig. 1 encompassed almost the entire city limits of Dallas.

Monitoring was done at the compliance station to evaluate LUR model predictions and to compare VOC and NO_2 measurements with corresponding reference method measurements reported in the EPA Air Quality System (AQS) database. The compliance station (referred to here as the Hinton site) was AQS Site 481130069 at 1415 Hinton Street; this site was in a light industrial/office park in northwest Dallas between I-35E and Love Field Airport. Duplicate passive samplers were co-located at Hinton and a fire station in North Dallas to evaluate passive sampler precision.

Ambient monitoring was conducted concurrently at all sites for five weeks from August 1 – September 5, 2006 for summer and January 22 – February 26, 2008 for winter. Weeklong integrated sampling was chosen to represent chronic ambient exposures. Samplers were placed in shelters and suspended at breathing zone height (1.5 to 2 m) in the backyards of the fire stations.

Air sampling was conducted using passive sampler technology. NO_2 was measured using Ogawa Model 3300 passive samplers for NO_2 (Ogawa & Co., Pompano Beach, FL, USA). This two-sided sampler consists of a cylindrical polymeric body (2 cm in diameter and 3 cm long) with a diffusion barrier and two stainless steel screens on each side. The device holds a glass-fiber collection pad coated with triethanolamine (TEA) at each end for sampling. The TEA-coated pads were loaded in the lab just prior to deployment to minimize contamination and degradation. All components, except the collection pad, are re-useable. Analysis of the samplers was conducted using ion chromatography. The Ogawa sampler has been used extensively in other LUR studies.[5] NO_2 is a EPA National Ambient Air Quality Standards (NAAQS) criteria air pollutant and serves as an indicator of mobile and stationary combustion sources.[12] VOCs were measured using PekinElmer (PE) thermal desorption diffusion tubes packed with 40/60 mesh size, unwashed Carbopack X adsorbent for VOC (Supelco, Inc., Bellefonte, PA, USA). After the PE tubes were thermally desorbed, they were ready for re-use and re-deployed in the field. Due to their reusability, the PE tubes used in Detroit[13] were used in this study. Select VOCs such as 1,3-butadiene and BTEX species (benzene, toluene, ethylbenzene, *o*-xylene, and *m,p*-xylene) are reported in this paper. These species are classified as air toxics by EPA and the State of Texas.[14,15] BTEX species and 1,3-butadiene are petroleum-related compounds typically associated with traffic emissions.[16] Evaluation of passive samplers for precision and accuracy was conducted at the Hinton site and a North Dallas fire station (see Passive method evaluation section in Supplementary information). Further details on the air sampling, analyses, and quality assurance methods are discussed elsewhere.[7,13]

Results

Concentrations

Table 2 shows summary statistics of the air pollutants collected at the fire station sites for each season. (Supplemental Table S5 reports mean concentrations for each site.) In general, pollutant levels were higher during winter than summer. In terms of the means, this increase was particularly noticeable for benzene and 1,3-butadiene (67% and 63% increases, respectively) and styrene which in summer was often below its detection limit. This may have been due to colder temperatures affecting atmospheric reaction rates and lower mixing heights resulting in higher concentrations.[17,18]

Monitoring methods, sampling time integrals, and analysis methods in Dallas were the same as those in Detroit[7] and similar to those in El Paso[6], thereby providing an opportunity for comparison. Median pollutant concentrations from Detroit and El Paso were comparable to or higher than Dallas. Complex terrain conditions in El Paso and heavy industrial sources in Detroit may have been factors in higher pollutant concentrations encountered in those cities versus Dallas which was dominated by flat terrain and mobile sources. All data were above method detection limits; summer and winter NO_2 levels were below the annual NAAQS of 53 ppb[12] (Table 2).

Coarse-scale spatial comparisons

Dallas was physically separated by north and south sections and a central, downtown area (see Fig. 1). The city was divided in this manner and median pollutant concentrations from fire station sites in each section were compared. Ten fire stations were located in the north section, nine in the south, and five in the central section.

Table 3 reports median values for each city section and the entire study area for each season and indicates whether the levels in each section were significantly different (at the 5% level) between the two seasons. The Wilcoxon rank sum test[19] was utilized for these comparisons. Wintertime levels were higher in each section for benzene, and for NO_2 in the north and south sections. This also held true when looking at the entire city, and in addition, wintertime *o*-xylene levels were statistically significantly higher when the study area was considered as a whole.

Table 4 reports the results of comparing the city sections to each other within the summer and winter periods. To guard against false positives, these comparisons were done with Dunn's test[19], but modified as suggested by Hochberg and Tamhane.[20] Mukerjee *et al.*[7] provide details of an application of Dunn's test to assess spatial differences. For both summer and winter, the central section had higher NO_2 levels than either the north or south sections. For benzene, the central section was higher than the north in both seasons, but higher than the south section only in summer.

LUR modeling

To determine LUR equations, the observed mean values of the chemicals at each site were plotted against the various potential predictor variables (Table 1). For each chemical, only those predictors for which the chemical appeared to have reasonably consistent behavior were retained for use in developing the LURs. This suggested the use of multiple linear regression to estimate the LURs. In most instances, this was applied with the chemical measurements log-transformed; for the *m,p*-xylene LUR, the predictor variable of distance to a large ethylbenzene source (ETH1) was also log-transformed (see Table 5). For consistency, the summer predictor variables were applied in the winter season in each equation. (Since sampling was for weeklong sampling periods, wind direction was not considered in the LURs.)

When the regressions were first attempted, residual analysis from initial regression attempts indicated large differences between the observed and predicted values at a few sites. These sites varied by chemical. To mitigate this, the regressions were re-run with each site being weighted by the inverse value of Cook's D statistic[21] calculated from the unweighted regression. Thus, the final predictive (LUR) equations downweighted the influence of those sites which departed from the general pattern established by the other locations. See Rawlings[22] for a discussion of multivariate regression including influence diagnostics and weighted regression. In addition to residual analyses, regression diagnostics included cross-validation.

Table 5 presents results of the LUR models for summer and winter data. Predictors which were significant at the 5% level are shown in bold. In each case, the equations show all predictors used, not just those reported as significant. All R^2 values are reported based on the original (not the log-transformed) scale. There were-similarities and differences between summer and winter results in terms of which predictors were found to be significant and performance of the regressions as measured in terms of R^2. Relative to the summer results, benzene and NO_2 "lost" two predictors (in terms of significance at the 5% level) while 1,3-butadiene "gained" two. Similarly, toluene, and *o*-xylene all "added" a significant predictor, while ethylbenzene and *m,p*-xylene both "dropped" one.

In terms of the R^2 values, benzene, ethylbenzene, and *m,p*-xylene, all had noticeably higher R^2's in summer than in winter. On the other hand, NO_2 and 1,3-butadiene all had noticeably lower values in summer. The R^2 values for toluene and *o*-xylene were approximately the same between summer and winter.

Table 6 displays the respective summer and winter results of comparing the measured values observed at the Hinton site to the values predicted there by the regression equations. As an indicator of uncertainty in the predicted values, the table also shows their 95% confidence intervals. (The Hinton site was withheld from the LUR estimation to serve as a validation location.) Relative to summer, the wintertime comparisons at the Hinton site were better for benzene, and *o*-xylene, and a bit worse for toluene. Results for *m,p*-xylene, ethylbenzene, 1,3-butadiene, and NO_2 were similar between seasons.

Figs. 2a-b display the LUR predicted pollutant levels the summer and winter periods for benzene. Similarly, Figs. 3a-b present NO_2 results. (Figs. S2 and S3 in the Supplemental information similarly display the measured concentrations for these pollutants.) Figs. 2 and 3 show generally higher predicted benzene and NO_2 levels in the central section and parts of the north section of Dallas, echoing the results of the statistical comparisons in Table 4. Similar figures were obtained for the other BTEX species. This was expected since the central and north sections of Dallas were more developed than the south.

Discussion and conclusions

Spatially-representative air monitoring sites were established at fire stations in Dallas during two seasons. Week-long sampling using passive air samplers at these sites suggested a temporal difference in concentrations with generally higher levels reported in winter versus summer. City section was also found to have an effect for NO_2 and benzene with the central section exhibiting higher pollutant levels than the north or south areas. Though the concentration differences found here are consistent with the expectations from higher summertime temperatures and lower wintertime mixing heights, these results are not definitive since summer and winter monitoring were conducted more than a year apart. For example, long-term temporal differences may have resulted from such influences as urban growth or increased road construction.

While the formal statistical hypothesis testing detected differences between sections of the city for NO_2 and benzene, differences were not found for toluene, ethylbenzene and the xylenes. However, the figures for these later VOCs suggested differences between parts of Dallas, similar to those seen for benzene. This apparent discrepancy between the hypothesis testing and the figures may be due to a combination of relatively low power (due to the small number of sites within city sections, particularly the central area) and a greater benzene differential than for the other VOCs. Another potential complicating factor might be that benzene is dominated by mobile sources while local sources may play a larger role for the other species.

The LUR results reported in Table 5 suggest which variables are useful for predicting the species examined here. For example, distance to roadways (e.g., DIST75KI and DIST110KI) were useful in predicting certain variables. However, note that there was a seasonal aspect in their utility in that distances to roadways were useful for predicting benzene, ethylbenzene, and NO_2 in summer but not in winter. On the other hand, distance to roadways was useful for predicting 1,3-butadiene and toluene in winter but not in summer. Traffic intensity was important for predicting NO_2 in both seasons and benzene in the summer. This finding for NO_2 is similar to the LUR results reported by Smith *et al.* in El Paso.[6] It is interesting to note that distance to a large benzene source was important for predicting ethylbenzene, both xylenes, 1,3-butadiene, and NO_2 in both seasons. This may be reflective of the fact that the only large benzene source was located in west-central Dallas and this location was relatively near two monitoring sites and the remaining sites were distributed in all distances and directions from it.

Some of the results reported in Table 5 may seem counterintuitive. For example, the results for summertime NO_2 have a positive coefficient for DIST45KI which indicates that NO_2 concentration increases as one moves farther from a roadway that carries between 40,000 and

50,000 vehicles per day. While this may seem puzzling, it may reflect characteristics of the road network in Dallas. Note that the coefficient for DIST75KI is negative indicating that NO_2 concentrations decrease with increasing distance from a roadway carrying between 70,000 and 80,000 vehicles per day. The different signs of these two coefficients may indicate that in Dallas moving away from a moderately traveled roadway may take one nearer to a more heavily traveled roadway. It is also possible that the unexpected positive coefficient for the DIST45KI variable may result from interaction with the larger road network including, for example, local and secondary roads.

Performance of LURs in Dallas were different in comparison to LURs for the same species in El Paso[6] and Detroit[7]. Model R^2 from Dallas Summer versus Detroit (also measured during Summer) were higher for benzene, toluene, and *m,p*-xylene, and lower for *o*-xylene, 1,3-butadiene, and NO_2; ethylbenzene R^2 was the same for both LURs. Winter Dallas LUR R^2s were lower than Detroit LURs for the same species except benzene and toluene. El Paso LURs (measured in Winter) had model $R^2 > 0.9$ and, thus, were higher in comparison to Dallas LURs for both seasons. These differences were puzzling, particularly in comparison to Detroit which used the same sampling and lab analysis methods. Pin-pointing reasons for the different regression performances is difficult but some distinctions exist among the cities. For example, Detroit and El Paso have major border crossings while Dallas does not. Dallas is part of a much larger metropolitan area than the other two. El Paso has complex terrain while Dallas and Detroit do not.

At first glance, the LUR predictions at the Hinton site were disappointing, at least on a percentage basis (Table 6). This is in part due to the low observed levels. For example, the observed value of NO_2 in the summer at Hinton was only 12 ppb and the LUR prediction was 15 ppb leading to a percentage difference of 24%. Another reason for the apparent poor performance was due to the fact that commercial and industrial facilities in the immediate area of the Hinton site were observed to be shuttered or operating at low levels during the monitoring periods. Therefore, local traffic was minimal around Hinton as opposed to the fire stations. Thus, the model overpredicted pollutant levels.

Another perspective on model performance is seen in Table 7 which displays the mean observed levels and average absolute discrepancy between predicted and measured values across the study area. As indicated there, the absolute differences relative to mean measured concentrations are between 10 to 17% for summer and 8 to 25% for winter. Note that the observed pollutant concentrations were low in both winter and summer which partially explains some of the high percentages seen there.

The combination of passive monitoring and GIS and statistical approaches employed here may be useful in identifying local influences on pollutant levels. This in turn could be used to help determine priorities for future monitoring locations. The statistical approach utilized here was multiple linear regression using logarithmic transformations, followed by residual analyses and cross-validation to evaluate adequacy of the models. One might consider alternative approaches such as kriging or neural networks but they were not used here because they are quite data intensive and the limited number of monitoring sites available for this study would likely not adequately support these spatial prediction approaches.

Seasonal differences in the LURs and their predictive power demonstrate the need for caution in developing such models from annual or multi-year averages without considering seasonal or other factors. In fact, in their review Hoek et al.[5] note that seasonal aspects have generally been excluded from LUR modeling efforts either by the nature of the monitoring or averaging out seasonality during the model fitting process. It is worth noting that many LUR models are used as part of a health assessment. If the health issue being studied has a seasonal aspect, then it would be beneficial for the corresponding LUR to account for this. The seasonal consideration discovered here is being further explored in other EPA spatial studies. Thus, the potential should be available in the future to combine these results from Dallas with the other LUR efforts mentioned to obtain a comprehensive analysis of the exposure modeling results across different U.S. cities and seasons.

Acknowledgements

For their roles in various aspects of the study, we thank the following: Casson Stallings, Hunter Daughtrey, Karen Oliver, Dennis Williams, Herb Jacumin, Laura Liao, Mariko Porter, Chris Fortune, and Mike Wheeler of Alion Science and Technology; Carry Croghan, and Mark Sather of EPA; John Koller from the City of Dallas. We also thank the City of Dallas Fire Department for allowing us to use their fire stations. Finally we thank Davyda Hammond and Gary Norris of EPA for reviewing the manuscript. This study was an EPA Region 6 Geographic Initiative (RGI) and Regional Applied Research Effort (RARE) Project. The U.S. Environmental Protection Agency through its Office of Research and Development funded and managed the research described here under contract EP-D-05-065 to Alion. The paper has been subjected to Agency review and approved for publication. Mention of trade names or commercial products does not constitute an endorsement or recommendation for use.

Supplementary information

Tables of ancillary variables considered for use in LUR models, passive method evaluation, mean concentrations at each site, maps of measured NO_2 and benzene concentrations. This material is available on the Electronic Supplementary Information Service at http://www.rsc.org.

References

1 M. Brauer, G. Hoek, P. van Vliet, K. Meliefste, P. Fischer, U. Gehring, J. Heinrich, J. Cyrys, J., T. Bellander, M. Lewne and B. Brunekreef, *Epidemiology*, 2003, **14**, 228-239.

2 M. Jerrett, A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuvaroglu, Morrison J. and C. Giovis, *J. Exp. Anal. Environ. Epidemiol.*, 2005, **15**, 185-204.

3 Z. Ross, P.B. English, R. Scalf, R. Gunier, S. Smorodinsky, S. Wall, M. Jerrett, *J. Exp. Sci. Environ. Epidemiol.*, 2006, **16**, 106-114.

4 D.K. Moore, M. Jerrett, W.J. Mack and N. Künzli, *J. Environ. Monit.*, 2007, **9**, 246-252.

5 G. Hoek, R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer and D. Briggs, *Atmos. Environ.*, 2008, **42**, 7561-7578.

6 L. Smith, S. Mukerjee, M. Gonzales, C. Stallings, L. Neas, G. Norris and H. Özkaynak, *Atmos. Environ.*, 2006, **40**, 3773–3787.

7 S. Mukerjee, L.A. Smith, M.M. Johnson, L.M. Neas and C.A. Stallings, *Sci. Total Environ.*, 2009, **407**, 4642-4651.

8 T. Sahsuvaroglu, A. Arain, P. Kanaroglou, N. Finkelstein, B. Newbold, M. Jerrett, B. Beckerman, J. Brook, M. Finkelstein, M. and N.L. Gilbert, N.L., *J. Air Waste Manage. Assoc.*, 2006, **56**, 1059-1069.

9 S.B. Henderson, B.Beckerman, M. Jerrett, and M. Brauer, *Environ. Sci. Technol.*, 2007, **41**, 2422-2428.

10 SAS Institute, Inc., *Base SAS® 9.1.3 Procedures Guide*, SAS, Cary, NC, 2004.

11 SAS Institute, Inc., *SAS/STAT® 9.1 User's Guide*, SAS, Cary, NC, 2004.

12 U.S. EPA Technology Transfer Network National Ambient Air Quality Standards (NAAQS) Web Site http://www.epa.gov/ttn/naaqs/ (accessed 2010).

13 S. Mukerjee, K.D. Oliver, R.L. Seila, H.H. Jacumin, C. Croghan, E.H. Daughtrey, L.M. Neas and L.A Smith, *J. Environ. Monit.*, 2009, **11**, 220-227.

14 U.S. EPA Technology Transfer Network Air Toxics Web Site http://www.epa.gov/ttn/atw/allabout.html (accessed 2010).

15 State of Texas Commission on Environmental Quality Air Toxics Web Site http://www.tceq.state.tx.us/implementation/tox/AirToxics.html (accessed 2010).

16 E.M. Fujita, *Sci. Total Environ.*, 2001, **276**, 171-184.

17 D.H.F. Atkins and D.S. Lee, *Atmos. Environ.*, 1995, **29**, 223-239.

18 L. Wallace, W. Nelson, R. Ziegenfus, E. Pellizzari, L. Michael, R. Whitmore, H. Zelon, T. Hartwell and R. Perritt, R., *J. Exp. Anal. Environ. Epidemiol.*, 1991, **1**, 157-192.

19 O.J. Dunn, *Technometrics*, 1964, **6**, 241-252.

20 Y. Hochberg and A.C. Tamhane, *Multiple Comparison Procedures*, Wiley, NY, 1987.

21 R.D. Cook, *Technometrics*, 1977, **19**, 15-18.

22 J.O. Rawlings, *Applied Regression Analysis: A Research Tool*, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1988.

23 T.A. Kvålseth, *Am. Statist.,* 1985 **39,** 279-285.

24 A. Scott and C. Wild, *Am. Statist.,* 1991, **45,** 127-129.

TABLE 1 Pearson correlations between explanatory variables considered for site selection and LURs[a]

	POP_DEN[b]	INT_1000[c]	DIST_15KI[d]	DIST_45KI[e]	DIST_75KI[f]	DIST_110KI[g]	DIST_140KP[h]	NOX1[i]	NOX2[j]	BEN1[k]	ETH1[l]
POP_DEN	1	-.20	-.30	.10	.04	-.23	-.09	.02	.15	-.02	.04
INT1000	-.14	1	-.34	-.48	-.41	-.44	-.19	-.19	-.17	-.25	.01
DIST_15KI	-.20	-.48	1	.03	-.08	.47	.08	.26	.03	.50	.27
DIST_45KI	.01	-.57	.34	1	.36	.55	.61	.53	.18	.14	-.34
DIST_75KI	.10	-.35	.36	.34	1	.26	-.08	.09	.08	.51	.32
DIST_110KI	-.29	-.58	.64	.45	.48	1	.73	.64	.31	.41	-.29
DIST_140KP	-.31	-.47	.51	.56	.18	.82	1	.79	.39	-.10	-.73
NOX1	-.31	-.40	.38	.55	.30	.72	.76	1	.38	.37	-.26
NOX2	-.13	-.30	.06	.17	.39	.51	.59	.60	1	.06	-.32
BEN1	.07	-.13	.11	.03	.32	.11	-.27	.26	-.09	1	.68
ETH1	.26	.23	-.29	-.33	-.08	-.59	-.86	-.45	-.60	.65	1

[a] Fire stations with passive sampling data (n=24) appear in the upper triangular portion of the matrix (i.e., above the diagonal of 1s); correlations within the group of unmonitored fire stations (n=21) appear in the lower triangular portion. [b] POP_DEN: population of census tract of location. [c] INT1000: traffic intensity within 1000 m of location. [d] DIST_15KI: Distance to nearest road with 10,000 < traffic volume ≤ 20,000 vehicles/day. [e] DIST_45KI: Distance to nearest road with 40,000 < traffic volume ≤ 50,000 vehicles/day [f] DIST_75KI: Distance to nearest road with 70,000 < traffic volume ≤ 80,000 vehicles/day. [g] DIST_110KI: Distance to nearest road with 100,000 < traffic volume ≤ 120,000 vehicles/day. [h] DIST_140KP: Distance to nearest road with traffic volume ≥ 140,000 vehicles per day. [i] NOX1: Distance to source with NO_x emissions >570,000 lbs/year. [j] NOX2: Distance to source with 21,000 < NOx emissions < 221,000 lbs/year. [k] BEN1: Distance to source with benzene emissions > 270,000 lbs/year. [l] ETH1: Distance to source with ethylbenzene emissions > 4,400 lbs/year.

Table 2 Median pollutant concentrations at Dallas fire stations versus Detroit/Dearborn and El Paso schools[a]

	Dallas summer		Dallas winter		Dallas seasonal difference (%)[c]	Detroit/Dearborn (25 schools; 7/19/2005 – 8/30/2005)[d]	El Paso (22 schools; 11/24/1999 – 12/18/1999)[d]
Pollutant	Concentration (24 fire stations; 8/1/2006 – 9/05/2006)	MDL[b]	Concentration (24 fire stations; 1/22/2008 – 2/26/2008)	MDL[b]			
NO_2	12 (4, 25)	2	14 (2, 22)	2	13	16 (11, 24)	22 (11, 37)
1,3-butadiene	72 (38, 149)	30	117 (48, 314)	33	67	74 (50, 128)	NM[e]
Benzene	232 (83, 388)	27	357 (247, 538)	8	63	466 (338, 698)	777 (489, 1531)
Toluene	539 (162, 1166)	18	617 (232, 1788)	25	20	1401 (980, 1994)	1473 (772, 3306)
Ethylbenzene	86 (31, 190)	11	96 (40, 295)	8	15	186 (126, 360)	250 (152, 558)
o-xylene	86 (30, 218)	13	102 (39, 279)	8	25	200 (120, 338)	298 (177, 672)
m,p-xylene	254 (87, 621)	27	247 (98, 895)	16	10	591 (362, 1228)	838 (474, 1848)

[a] Medians calculated over all sites and weeks. Units for NO_2 in ppbV; VOC in pptV with VOC from El Paso using 3M Organic Vapor Monitors. Minimum and maximum values in parentheses. [b] MDL: Method detection limit. [c] (winter – summer)/summer, based on means. [d] Data summarized from refs. 6 and 7. [e] NM: not measured.

Table 3 Median pollutant concentrations[a] by city section and season based on individual fire stations

	Benzene	Toluene	Ethylbenzene	*m,p*-xylene	*o*-xylene	NO_2
Winter						
North	**344**[b]	552	90	250	99	**14**
Central	**381**	774	136	338	141	16
South	**345**	563	90	238	96	**13**
Entire area	**352**	568	95	267	**103**	**15**
Summer						
North	220	594	92	264	87	11
Central	283	574	111	317	106	15
South	208	411	69	209	71	11
Entire area	225	565	89	260	87	12

[a]VOCs are in pptV and NO_2 is in ppbV. [b] Boldface indicates that the winter values were statistically significantly (5% level) different than the corresponding summer values.

Table 4 Comparison of pollutants at Dallas city sections within seasons

	Benzene	Toluene	Ethylbenzene	*m,p*-xylene	*o*-xylene	NO_2
south vs. central						
summer	C > S[a]	ns[b]	ns	ns	ns	C > S
winter	ns	ns	ns	ns	ns	C > S
south vs. north						
summer	ns	ns	ns	ns	ns	ns
winter	ns	ns	ns	ns	ns	ns
north vs. central						
summer	C > N	ns	ns	ns	ns	C > N
winter	C > N	ns	ns	ns	ns	C > N

[a] C > S means that the central section had statistically significantly (5% level) higher concentrations than the south section. Similarly, C > N means that the central section was higher than the north section. [b] ns: no significant difference at the 5% level.

Table 5 LUR models for Dallas

Summer LURs	R^2 (%)
Benzene = **230.9** – 0.028*DIST15KI – **0.0045*DIST75KI** + 5.9E-4*DIST110KI + **9.1E-5*INT1000**	72
ln Toluene = **6.8** – 3.2E-5*DIST75KI – 8.7E-6*DIST110KI - **1.9E-5*NOX1** + 3.5E-6*NOX2	41
ln Ethylbenzene = **4.9** – 1.1E-5*DIST75KI – **4.0E-5*DIST110KI** – 1.8E-7*INT1000 – **2.6E-5*BEN1** + 7.9E-6*NOX1 – 1.4E-5*NOX2	63
ln *m,p*-xylene = -1.6 + 4.1E-5*DIST15KI – 1.3 E-5*DIST75KI + 3.8 E-6 DIST110KI + 2.5E-7*INT1000 – **5.9E-5*BEN1** + **0.76*ln_ETH1** + **1.4E-5*NOX1**	71
ln *o*-xylene = **4.7** – 2.0E-5*DIST75KI - **3.8E-5*DIST110KI** - **1.5E-5*BEN1** + 5.1E-6*NOX1	46
ln 1,3-butadiene = **4.6** – 4.6E-6*DIST75KI – 1.8E-5*DIST110KI - **1.1E-5*BEN1** – 8.2E-6*NOX2	26
ln NO_2 = **2.6** + **4.2E-5*DIST45KI** – **2.5E-5*DIST75KI** + **1.2E-6*INT1000** – **1.5E-5*BEN1**	34
Winter LURs	
Benzene = **375.5** + 0.001*DIST15KI – 0.004*DIST75KI – 0.002*DIST110KI + 6.1E-7*INT1000	49
ln Toluene = **7.1** – **2.5E-5*DIST75KI** – 8.8E-6*DIST110KI - **2.4E-5*NOX1** - 6.1E-6*NOX2	41
ln Ethylbenzene = **5.0** – 1.2E-5*DIST75KI – 2.3E-5*DIST110KI + 5.0E-8*INT1000 – **2.0E-5*BEN1** + 4.2E-6*NOX1 – 1.3E-5*NOX2	40
ln *m,p*-xylene = 1.1 + 1.4E-5*DIST15KI – 3.3E-5*DIST75KI + 1.3 E-5 DIST110KI + 1.2E-7*INT1000 – **5.7E-5*BEN1** + **0.50*ln_ETH1** + 9.4E-6*NOX1	40
ln *o*-xylene = **5.1** – **9.2E-6*DIST75KI** - **1.7E-5*DIST110KI** - **2.6E-5*BEN1** - 9.8E-7*NOX1	37
ln 1,3-butadiene = **5.3** – **5.0E-6*DIST75KI** – **1.7E-5*DIST110KI** - **2.2E-5*BEN1** – **1.2E-5*NOX2**	40
ln NO_2 = **2.8** - 2.2E-5*DIST45KI – 1.2E-5*DIST75KI + **6.3E-7*INT1000** – **1.4E-5*BEN1**	48

Notes: Bold indicates regression coefficients significant at the 5% level. Log is the natural logarithm. R^2 is reported for the original scale, not the log-transformed scale[23,24].

Table 6 Observed and predicted values at the Hinton site[a]

Pollutant	Measured	Predicted[b]	Difference[c]	Percent difference[d]
Summer				
Benzene	133	243 (236, 249)	110	83
Toluene	494	653 (589, 724)	159	32
Ethylbenzene	61	119 (107, 132)	58	94
m,p-xylene	178	357 (324, 393)	179	101
o-xylene	60	109 (105, 114)	49	83
1,3-butadiene	48	90 (83, 97)	42	89
NO_2	12	15 (14, 17)	3	24
Winter				
Benzene	289	371 (360, 382)	82	28
Toluene	518	792 (749, 839)	275	53
Ethylbenzene	71	133 (112, 157)	61	86
m,p-xylene	194	399 (330, 482)	205	106
o-xylene	67	139 (128, 151)	72	107
1,3-butadiene	84	164 (162, 166)	79	94
NO_2	14	17 (16, 19)	3	25

[a] Units are pptV for all VOCs and ppbV for NO_2. [b] 95% confidence interval in parentheses. [c] Difference = predicted – observed. [d] Percent difference = (difference / observed) x 100

Table 7 Means measured concentrations[a], mean absolute differences[b] and their ratio (%)

Pollutant	Summer			Winter		
	Mean	Absolute difference	Percent absolute difference	Mean	Absolute difference	Percent absolute difference
Benzene	218	34	16	355	27	8
Toluene	536	81	15	642	149	23
Ethylbenzene	88	11	12	101	22	22
m,p-xylene	258	27	10	285	71	25
o-xylene	87	12	14	109	21	19
1,3-butadiene	75	12	16	125	20	16
NO_2	12	2	17	14	2	14

[a] Units are pptV for all VOCs and ppbV for NO_2. [b] Absolute difference = LUR predicted – measured.

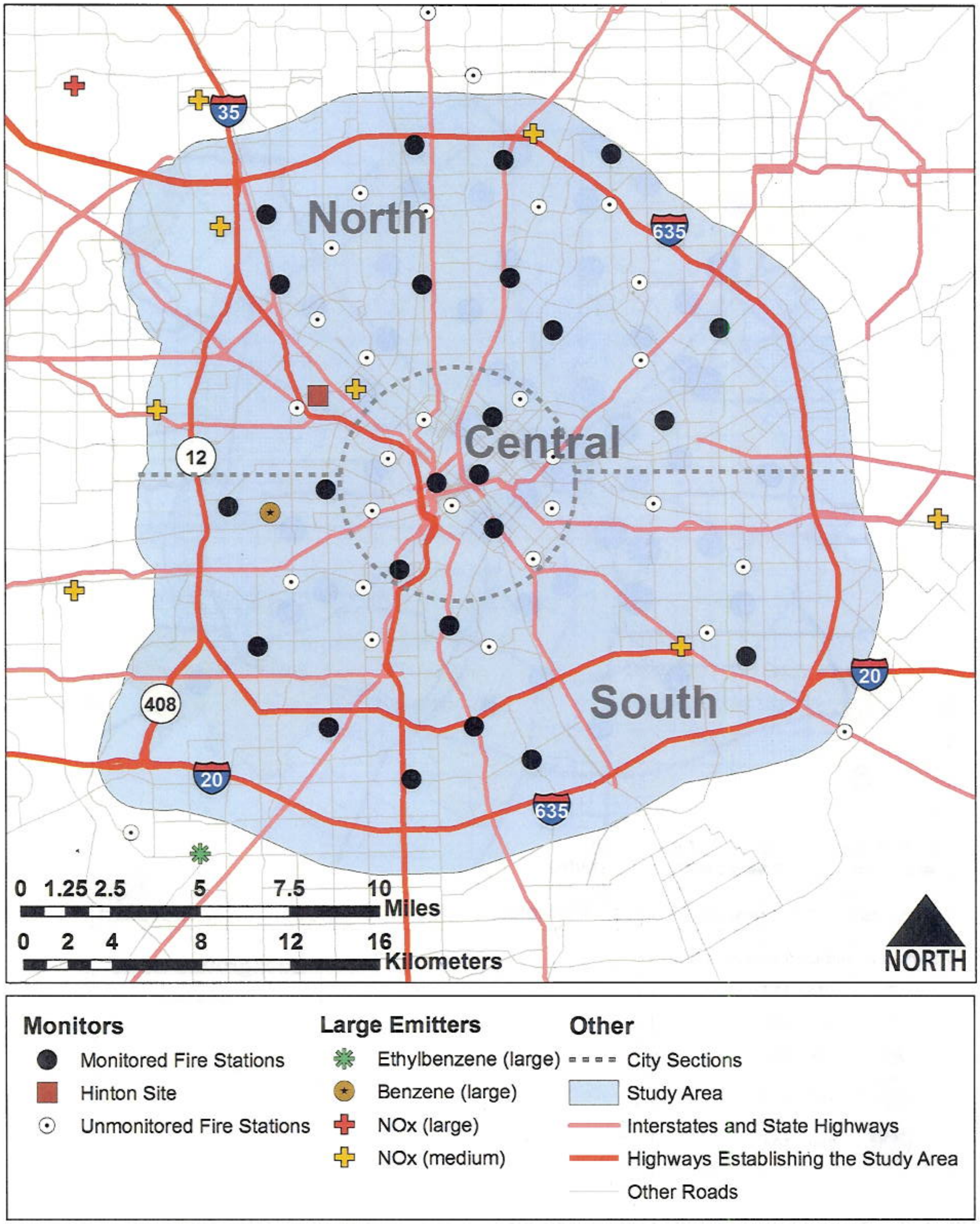


Fig. 1 Locations of fire station and Hinton compliance monitoring sites[a] in Dallas with city sections.

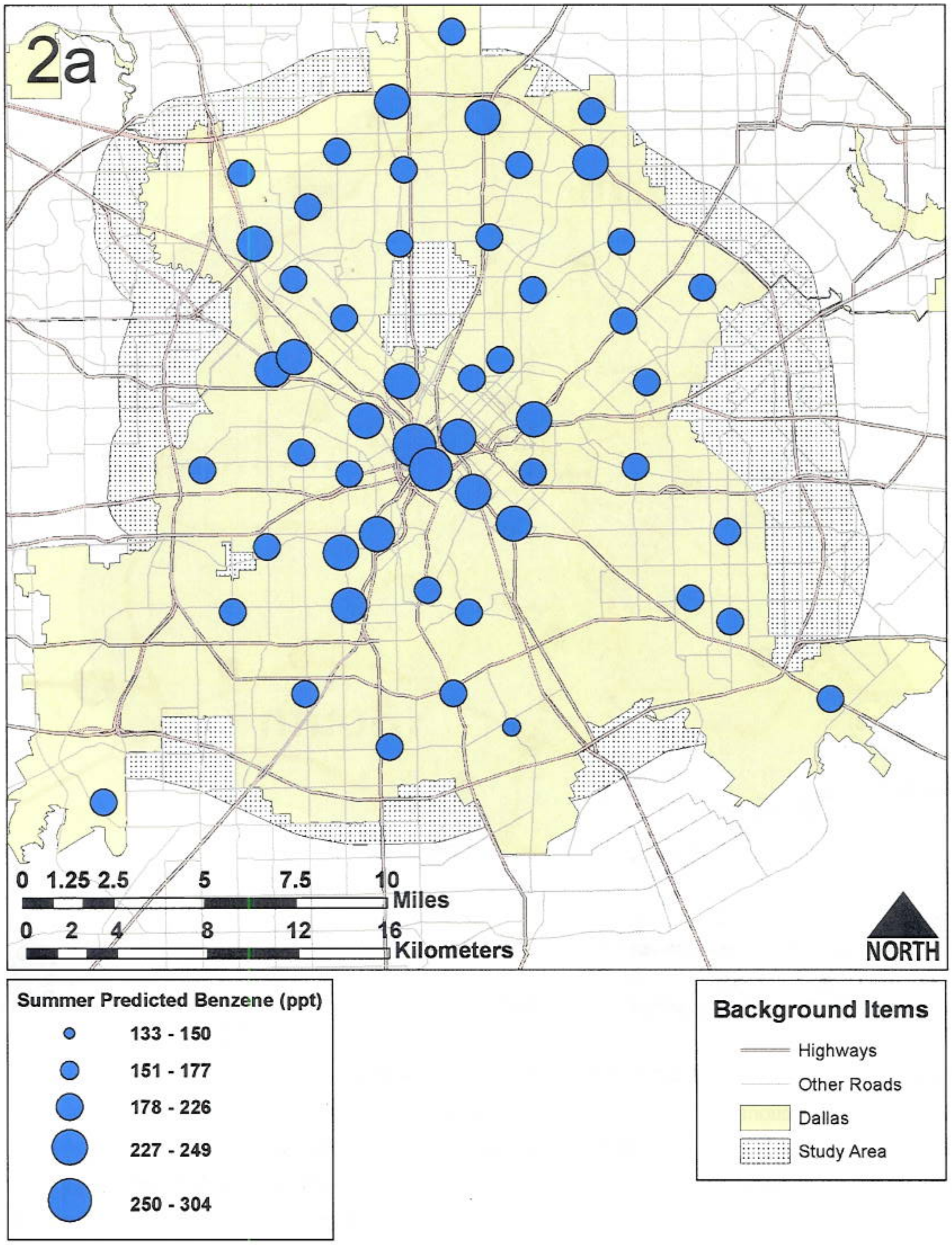
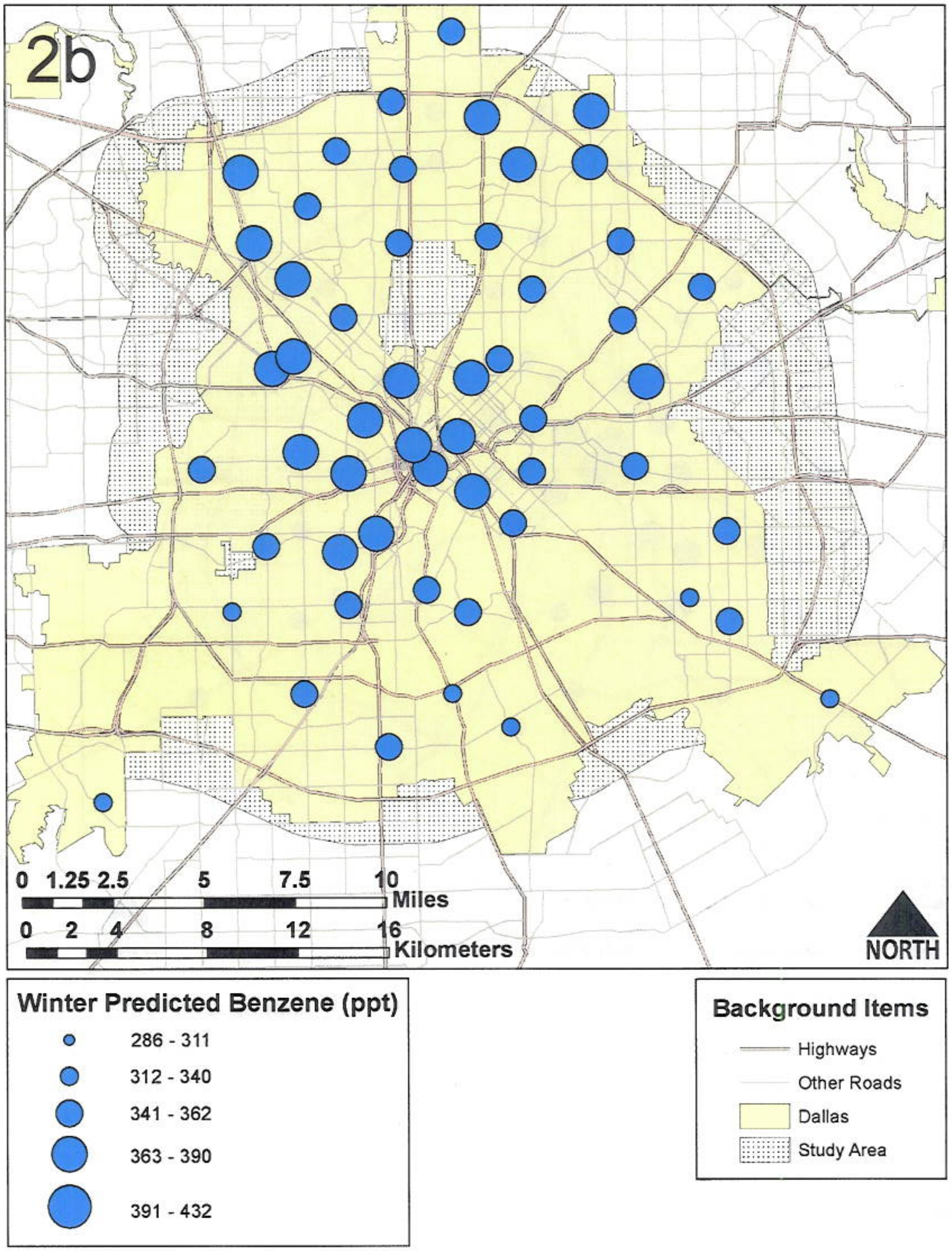


Fig. 2 LUR predicted benzene concentrations in Dallas: (a) summer (b) winter.


2b
0 1.25 2.5 5 7.5 10
Miles
0 2 4 8 12 16
Kilometers
NORTH
Winter Predicted Benzene (ppt)
286 - 311
312 - 340
341 - 362
363 - 390
391 - 432
Background Items
Highways
Other Roads
Dallas
Study Area

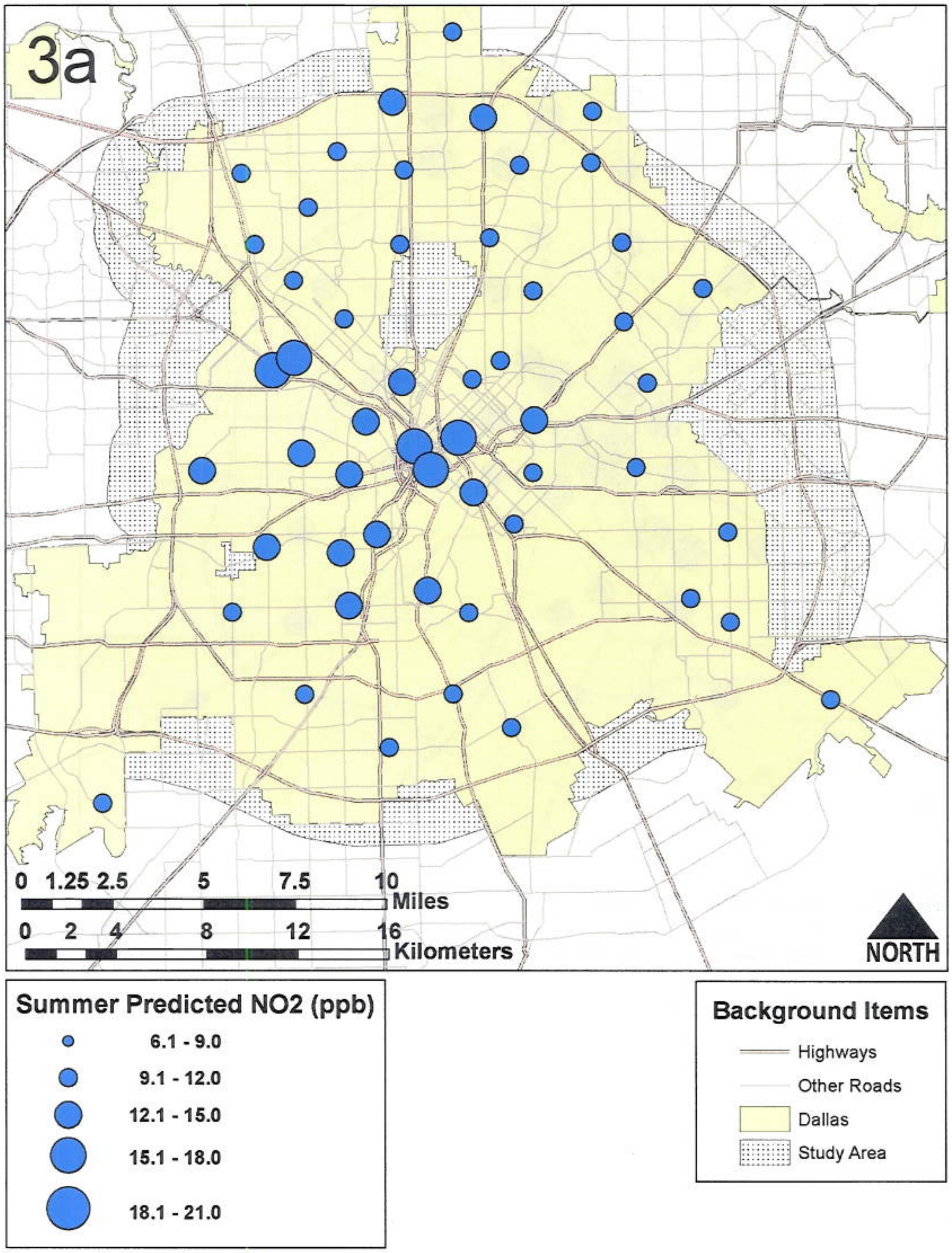
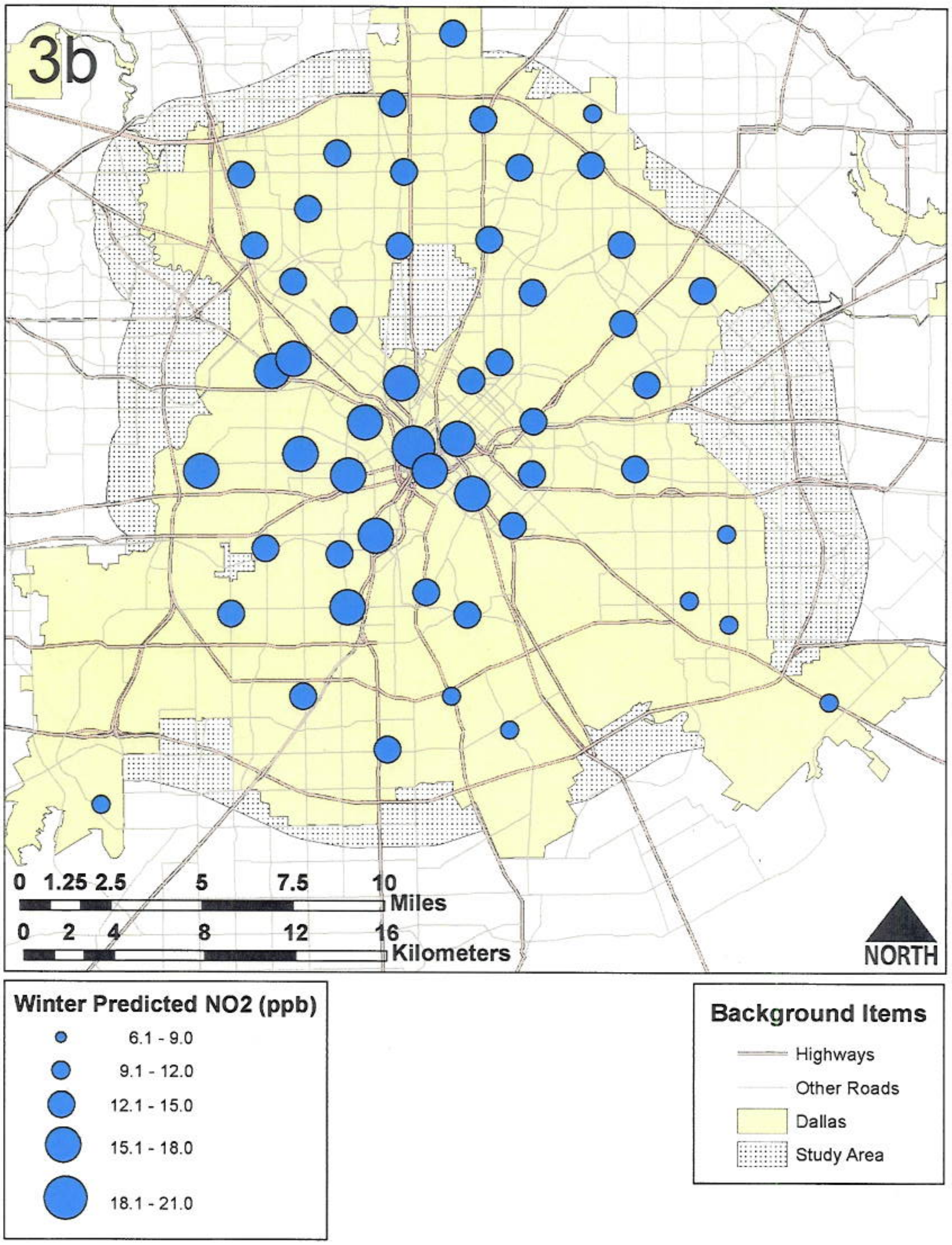


Fig. 3 LUR predicted NO_2 concentrations in Dallas: (a) summer (b) winter.


3b
0 1.25 2.5 5 7.5 10
Miles
0 2 4 8 12 16
Kilometers
NORTH
Winter Predicted NO2 (ppb)
6.1 - 9.0
9.1 - 12.0
12.1 - 15.0
15.1 - 18.0
18.1 - 21.0
Background Items
Highways
Other Roads
Dallas
Study Area