Available online at www.sciencedirect.com
ScienceDirect
Atmospheric Environment 42 (2008) 1359–1369
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Within-urban variability in ambient air pollution: Comparison of estimation methods

Julian D. Marshall[a,*], Elizabeth Nethery[b], Michael Brauer[b]

[a]Department of Civil Engineering, University of Minnesota, Minneapolis, MN 55455, USA
[b]School of Occupational and Environmental Hygiene, The University of British Columbia, Vancouver, BC, Canada V6T 1Z3

Received 15 May 2007; accepted 8 August 2007

Abstract

An important component of air quality management and health risk assessment is improved understanding of spatial and temporal variability in pollutant concentrations. We compare, for Vancouver, Canada, three approaches for estimating within-urban spatiotemporal variability in ambient concentrations: spatial interpolation of monitoring data; an empirical/statistical model based on geographic analyses (“land-use regression”; LUR); and an Eulerian grid model (community multiscale air quality model, CMAQ). Four pollutants are considered—nitrogen oxide (NO), nitrogen dioxide (NO_2), carbon monoxide, and ozone—represent varying levels of spatiotemporal heterogeneity. Among the methods, differences in central tendencies (mean, median) and variability (standard deviation) are modest. LUR and CMAQ perform well in predicting concentrations at monitoring sites (average absolute bias: $<50\%$ for NO; $<20\%$ for NO_2). Monitors (LUR) offer the greatest (least) temporal resolution; LUR (monitors) offers the greatest (least) spatial resolution. Of note, the length scale of spatial variability is shorter for LUR (units: km; 0.3 for NO, 1 for NO_2) than for the other approaches (3–6 for NO, 4–6 for NO_2), indicating that the approaches offer different information about spatial attributes of air pollution. Results presented here suggest that for investigations incorporating spatiotemporal variability in ambient concentrations, the findings may depend on which estimation method is employed.

Keywords: Exposure analysis; Geographic information system (GIS); Traffic emissions; Land-use regression (LUR); Models-3/community multiscale air quality (CMAQ) model

1. Introduction

Improved understanding of spatiotemporal variability in ambient air pollutant concentrations is useful in many contexts, including quantifying air pollution health effects and their distribution among the population; attributing air quality and health impacts to specific emission sources and pollutants; and developing cost-effective impact reduction strategies. Spatial information about ambient concentrations informs discussions in environmental justice (e.g., quantifying and addressing disparities in air pollution exposures); transportation and land-use planning (e.g., whether to allow new childcare facilities to locate near major roadways); and various regulatory needs, including deciding how to meet air quality standards, identifying pollutant hot spots, locating new monitoring

*Corresponding author. Tel.: +1 612 625 2397; fax: +1 612 626 7750.
E-mail address: julian@umn.edu (J.D. Marshall).

1352-2310/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2007.08.012

stations, and providing public information in response to citizen and other stakeholder concern. Temporal information is important in regulatory contexts (e.g., determining compliance with air quality standards), time-series epidemiological studies (e.g., daily mortality and hospital admissions; perinatal and other outcomes where monthly or seasonal variability may be important), and assessment of trends in concentrations. Furthermore, it can shed light on source attribution (e.g., weekday/weekend effects, Marr and Harley, 2002).

In epidemiological studies, improving exposure classification likely reduces bias in exposure-effect analyses. Recent epidemiological research demonstrates the importance of accounting for within-city variability in air pollution concentrations (Jerrett et al., 2005; Miller et al., 2007). Epidemiological studies currently employ various approaches when estimating ambient concentrations. Regulatory monitoring network data offer good temporal resolution, but spatial resolution is limited by the number of monitors and their separation distance (Basu et al., 2004). Thus far, land-use regression (LUR) provides detailed spatial resolution but little information on temporal variability (Brauer et al., 2003; Briggs et al., 1997; Cyrys et al., 2005; Henderson et al., 2007; Moore et al., 2007). Compared to monitoring data and LUR, photochemical dispersion models offer high temporal resolution but intermediate spatial resolution (Knowlton et al., 2004). In general, spatiotemporal heterogeneity in concentrations varies among pollutants and sources (e.g., primary versus secondary pollutants; point source versus distributed emissions).

This paper presents and evaluates three general approaches for estimating spatiotemporal variability in ambient air pollution concentrations: (1) interpolation of monitoring data (four methods are considered below); (2) a relatively new geostatistical/empirical approach called land-use regression (LUR); and (3) an Eulerian grid model. We consider four pollutants: nitrogen oxide (NO), nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3). These pollutants differ in important attributes, including primary (i.e., directly emitted) versus secondary (i.e., formed-in-the-atmosphere) pollutants, and exhibit varying degrees of spatiotemporal heterogeneity.

Despite these three methods being used in health risk assessment and epidemiological analyses (Basu et al., 2004; Bell, 2006; Brauer et al., 2007; Foley et al., 2003; Morgenstern et al., 2007; Ryan et al., 2007; Sanhueza et al., 2003), limited evaluation has been conducted and there have been no formal comparisons among the three methods. We compare these methods for one study area (Vancouver, Canada) and evaluate their characteristics, strengths, and weaknesses in assessing ambient air pollution concentrations, especially for epidemiological research and risk assessment.

2. Methods

Our study location is the Greater Vancouver Regional District (GVRD; land area: 2844 km^2), a coastal area in southwestern British Columbia. The GVRD has a comparatively large population (2.2 million in year-2005, comprising 51% of British Columbia) and average population density (760 km^{-2}—only 20% less than California's South Coast Air Basin, CARB, 2005). Vancouver is a good study location for this investigation because of the costal location, leading to a significant portion of air pollution being emitted within Vancouver, and the comparatively high monitoring station density (see below). In Vancouver, most (~90%) nitrogen oxides (NO_x) and CO emissions are from on-road and non-road transportation sources (GVRD, 2006).

Three general approaches were used to estimate ambient concentrations in the GVRD. The approaches, listed in Table 1 and described in the following subsections, include empirical measurements (spatial interpolation of monitoring data), air dispersion modeling (an Eulerian grid model), and a hybrid empirical/modeling approach (LUR). These approaches were implemented as part of the Border Air Quality Study (BAQS; www.cher.ubc.ca/baqs.htm), a large epidemiological study of air pollution health effects. A specific requirement of BAQS was high spatial resolution of estimated ambient concentrations, in order to reduce exposure misclassification (Ryan et al., 2007).

We estimated concentrations at the centroid location for each of the 56,099 postal codes (PCs) in the GVRD. In Canada, the spatial extent of urban PCs is relatively small, for example, one side of a city block or, in a densely populated area, a single apartment building (GVRD average: 39 people per PC). For LUR, concentration estimates are available for NO and NO_2 only; for the other two methods (interpolation of monitoring data; CMAQ), estimates are available for all four species.

Table 1
Approaches employed in this work

	Abbreviated name	Level of effort[a]	Pollutants[b]				Spatial coverage[c]	Unique values[d]
			CO	NO	NO_2	O_3		
1. Spatial interpolation of ambient monitoring station data								
1a. Nearest monitor within 10 km	Nearest	*	X	X	X	X	49,244–52,121 (88–93%)	57–70 (0.1%)
1b. Average of all monitors within 10 km[e]	–	*	X	X	X	X	49,244–52,121 (88–93%)	69–100 (0.1–0.2%)
1c. Inverse-distance weighted (IDW) average of all GVRD monitors[e]	–	*	X	X	X	X	56,099 (100%)	51,560 (92%)
1d. IDW average of three closest monitors within 50 km	IDW	*	X	X	X	X	56,099 (100%)	51,560 (92%)
2. Land-use regression (empirical/statistical model)	LUR	***		X	X		54,171 (97%)	49,751 (89%)
3. Eulerian grid-cell model (mechanistic mass-balance model)	CMAQ	*****	X	X	X	X	56,099 (100%)	138 (0.2%)

[a]Approximate scale, representing a typical situation: * = ~1–4 person-months; *** = ~1–4 person-years; ***** = 4–10+ person-years, involving significant domain expertise.
[b]CO = carbon monoxide; NO = nitrogen oxide; NO_2 = nitrogen dioxide; O_3 = ozone.
[c]Number (percent) of GVRD postal codes that were assigned an annual average concentration; range among pollutants.
[d]Number (percent) of unique annual average concentrations among GVRD postal codes; range among pollutants. Owing to co-location of PCs, the maximum value for this column is 51,560 (92%).
[e]As described in Section 3.2, approaches 1b and 1c were found to be suboptimal.

In the GVRD, all methods offer 88% or greater coverage of PCs. The nearest-monitor methods (1a and 1b in Table 1) and CMAQ offer comparatively few unique values; inverse-distance weighted (IDW) methods (1c, 1d) and LUR offer many unique values. Level of effort to generate the concentration estimates ranges from a few person-months to several person-years.

2.1. Ambient monitoring station data

GVRD has a large number of monitors (14 stations for NO_2 and NO; 13 for CO; 15 for O_3), and also a high monitor density (roughly one monitor per 160,000 people, or 5 monitors per 1000 km^2). In comparison, for CO (NO_2), the 158 (239) Urbanized Areas (UAs) in the US with monitors have 3 (3) stations per UA on average; only 3% (6%) of these UAs have more than 10 monitors (US EPA, 2006). The South Coast, California, perhaps the most-researched air basin in the world, has less than half the monitoring density (roughly one monitor per 440,000 people; 2 monitors per 1000 km^2, CARB, 2006a, b).

We employed year-2000 daily concentrations for GVRD regulatory ambient monitors, collected by the British Columbia Ministry of Environment (www.env.gov.bc.ca/air). The median number of days per year with missing data is NO: 7, NO_2: 7, CO: 6, O_3: 9. Measurement methods are consistent among monitors: for NO, NO_2, and NO_x, chemiluminescence of gas-phase reaction with ozone; for CO, infrared absorption; and for O_3, ultraviolet absorption (GVRD, 2006). Monitor locations are shown below (Fig. 4).

To estimate concentrations in each PC, we employed four separate spatial interpolation schemes, whereby the PC centroid was assigned the daily average concentration of the (1) nearest monitor within 10 km, (2) average of all monitors within 10 km, (3) IDW average of all GVRD monitors, and (4) IDW average of the three closest monitors within 50 km. These four methods (labeled as approaches 1a–1d in Table 1) cover a range of

spatial interpolation approaches, and each could reasonably be employed in single- or multi-city epidemiological studies. As discussed below, the second and third of these four monitor-based methods (approaches 1b and 1c, Table 1) offered suboptimal results, and thus they were excluded from most analyses.

2.2. Air dispersion model

Three-dimensional Eulerian reactive grid models are common in air quality engineering but are rarely employed in epidemiological studies (Bell, 2006; Jerrett et al., 2005). The US EPA's community multiscale air quality (CMAQ) model incorporates anthropogenic and natural emissions, transport and dilution owing to meteorology, and atmospheric transformations owing to chemical reactions (see www.cmaq-model.org).

CMAQ combines meteorological conditions during 31 March 2004–30 March 2005 with the most recent available emission inventory (year-2000). The simulation domain is 400 km × 480 km and employs 4 km × 4 km grid cells, 1-h time steps, the MC2 meteorological model, and CB4 chemical mechanism (Delle Monache et al., 2006a, b). CMAQ-estimated concentrations are modeled as uniform within each grid cell during each time step. PC centroids were spatially matched to their corresponding CMAQ grid cell, and assigned the annual-average concentration of that cell.

2.3. Land-use regression

LUR uses geographic information systems (GIS), combining measurements and detailed land-use data to predict concentrations throughout an area. LUR is relatively new: the first application to air quality was by Briggs et al. (1997). Jerrett et al. (2005) and Henderson et al. (2007) review the history of LUR. The specific steps are (1) measure long-term concentrations at many locations throughout the study area, (2) obtain land-use measures (e.g., total road length within a 1-km radius) in the vicinity of each monitored location, (3) develop statistical regression equations relating empirical concentrations to nearby land uses, and (4) apply these equations to a dense grid of points throughout an urban area.

To develop the LUR model, NO and NO_2 measurements using passive samplers were conducted at 116 locations throughout the GVRD during 24 February–14 March, 2003, and 8–26 September 2003 (Henderson and Brauer, 2005; Henderson et al., 2007). Locations were selected to capture the full range of concentrations in outdoor air (including high concentrations near roadways) while also placing monitors throughout the study region. (In contrast, regulatory monitoring stations are typically sited to capture broader concentration trends, avoiding roadways and other possible "hot spots".) Sampling dates were selected to estimate annual-average concentrations. Prior analysis of 5 years of monitoring data (1998–2002) revealed that in 70 out of 75 cases, average concentrations during the sampling days-of-year were within 15% of the annual average. In the sampling year (2003), monitoring data indicate that concentrations during the sampling period exhibit a strong 1:1 relationship with annual means; slopes are 1.03 ($R^2 = 0.96$) and 0.89 ($R^2 = 0.98$) for NO and NO_2, respectively (Henderson et al., 2007).

LUR equations employed here use the following geographic variables as concentration predictors; for both NO and NO_2: road density (within 100 and 200 m), elevation, population density (2500 m), and latitude; for NO only: road density (750 m); for NO_2 only: commercial land area (750 m) (Henderson and Brauer, 2005). Model R^2 is 0.62 for NO, 0.53 for NO_2. Because the independent variables are readily available for other urban areas, it would be straightforward to apply the equations to other cities. Thus, LUR offers the possibility for greatly improved efficiency in quantifying within-city concentration variability (Briggs et al., 2000).

LUR output, representing annual average concentrations, was converted to monthly averages using time trends in regulatory ambient monitoring data, as described by Henderson et al. (2007).

3. Results and discussion

Using the approaches in Table 1, we assigned daily-average concentrations at centroid locations for 88–100% of the 56,099 PCs in the GVRD. Results are illustrated in Fig. 1. Below, we compare concentration estimates first at monitoring sites only, then at all PCs. While the first comparison offers a "gold standard" (the monitoring data) against which to compare and validate concentration estimates, monitor locations provide imperfect (incomplete) validation since they incorporate a small number of locations, do not reflect the spatial distribution of the population, and generally avoid

Fig. 1. Annual-average ambient concentration estimates for nitrogen oxide (top panel) and nitrogen dioxide (bottom panel), based on an inverse distance weighting of the three nearest monitoring stations (left), land-use regression (LUR; middle), and the CMAQ air dispersion model (right). The approaches are summarized in Table 1. LUR captures within-neighborhood variability; monitors and CMAQ do not. The concentration color scale employed is identical among the six cases. (Owing to the large number of PCs in the study area, providing visual representation of our results required converting point estimates—the concentrations at each PC centroid—into two-dimensional raster surfaces.)

near-roadway and other high-concentration locations. The second comparison considers all PCs (a roughly population-weighted analysis since areas with higher population density have higher PC density), and therefore offers comparisons relevant to epidemiological investigations. In the second comparison, there is no "gold standard"—concentrations in each PC are unknown—yet comparison among the three approaches yields important insights regarding the utility of the different approaches.

3.1. Comparisons at monitoring station locations

We compared measured and modeled annual-average concentrations (NO, NO_2) at regulatory monitoring station locations. The analyses, summarized in Table 2, suggest reasonable agreement between estimates and data: for CMAQ and LUR, average absolute bias is ~45% for NO, ~17% for NO_2. Estimates are more accurate for NO_2 than for NO, likely because NO_2 is more spatially homogeneous than NO (Spatial coefficients-of-variability are presented in the discussion section). Table 2 suggests better performance for LUR than CMAQ, for NO; for NO_2, the reverse holds. As expected, the monitor-based methods (methods 1a and 1d) do better than LUR and CMAQ, though this comparison, derived at monitor locations, offers little insight regarding other locations.

3.2. Comparisons at postal code centroids

The second comparison employs the following tools: box plots; pair-wise scatter plots; for each scatter plot, the slope of the best-fit line through the origin; Pearson correlation coefficients; weighted and unweighted quartile-based kappa statistics; and, visual displays of results (concentration maps). These results are summarized below.

Two of the four interpolation approaches offered suboptimal results, and therefore were not considered further: (1) Employing the average of all monitors within 10 km (method 1b, Table 1), the location of the highest assigned concentration was not at the highest monitor site. Specifically, this approach is sensitive to monitor placement and in the GVRD inappropriately yields high concentrations southwest of the downtown high-concentration monitor. Analogous problems may arise for other asymmetrical urban areas (e.g., other costal cities). (2) Employing the IDW average of all monitors (method 1c), locations not immediately next to a monitor are assigned a concentration close

Table 2
Corroboration of estimated[a] and measured annual-average concentrations, at regulatory monitoring sites

	Nitrogen oxide				Nitrogen dioxide			
	Nearest	IDW	LUR	CMAQ	Nearest	IDW	LUR	CMAQ
Average difference ($\mu g\,m^{-3}$)	0.25	−1.0	5.4	−11	0.08	−0.88	−4.8	−0.70
Average absolute difference ($\mu g\,m^{-3}$)	0.27	1.1	10.2	14	0.10	0.90	6.0	5.6
Average bias (%)	1	−4	29	−36	0.3	−3	−12	0.4
Average absolute bias (%)	1	5	42	47	0.3	3	17	17
Correlation (%)	99.95	99.7	75	46	99.99	99.6	76	71
t-Test[b]	0.07	0.009	0.31	0.007	0.02	0.002	0.01	0.70

[a]The approaches are summarized in Table 1. Monitor-based methods (nearest; IDW) have non-zero error when compared against monitoring data because (1) they incorporate measurements from neighboring monitors when a site is down, and (2) here, concentration estimates at monitor locations are interpolated from PC centroids, which are near to, but not co-located with, monitors.

[b]Results from Student's *t*-test, indicating whether paired concentrations (method estimates; monitoring data) represent samples likely to have come from underlying populations with the same mean. At $\alpha = 0.1$, differences are statistically significant (i.e., the test accepts the hypothesis that the mean model-measurement difference is non-zero) for two columns: LUR for NO, and CMAQ for NO_2. For the remaining six columns, this test suggests good model–measurement agreement.

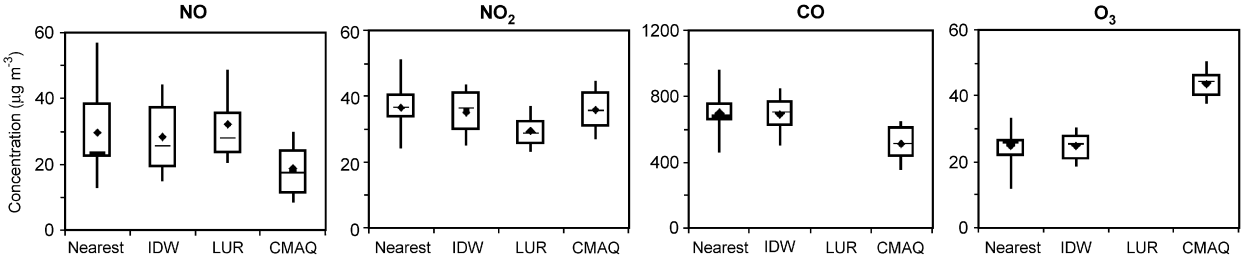

Fig. 2. Box plots of estimated annual-average concentrations. Pollutant and method abbreviations are given in Table 1. LUR estimates are unavailable for CO and O_3. Values shown are the mean (symbol: diamond) and the following percentiles: 10th, 25th, 50th (symbol: bar), 75th, and 90th.

in value to the regional average. The resulting values are likely too spatially homogenous.

Box plots of annual-average (Fig. 2) and monthly-average (supplemental online material (SOM), Fig. A1) concentrations indicate that for each pollutant, the inter-quartile range (IQR; difference between 75th and 25th percentile) is comparable among the methods employed (exception: IQR for monthly-average CO, NO_2, and O_3 concentrations are ~2 times larger for CMAQ as for the other methods). This observation suggests that, for the cases considered here, the range of exposures used in epidemiology studies will not vary significantly depending on which method is chosen (especially for chronic exposures). Reflecting seasonal variations, IQR for a given pollutant and method is ~2–3 times larger for monthly-average than for annual-average concentrations. Comparing the two interpolation methods in Fig. 2, mean and median values are similar, but the 10th-to-90th percentile range is greater for the nearest monitor approach than for the IDW approach. This finding is expected. IDW averages out extreme concentration values but preserves central tendencies.

The extreme upper ends of the concentration distributions (greater than 90th percentile) extend further for LUR than for other methods, because LUR incorporates high concentrations near roadways. For example, the ratio of 99th to 90th percentiles for NO (NO_2) is 1.99 (1.33) for LUR, 1.39 (1.15) for CMAQ, and 1.07 (1.04) averaged for IDW and nearest-monitor methods. Fig. 2 summarizes concentrations at PC centroids only. Because centroids—like residences—are typically not located at road center, many road-center LUR hot-spots in Fig. 1 are not included in Fig. 2.

The three pair-wise comparison metrics considered—annual-average Pearson correlation coefficients (*r*, range: 0.39–0.93), weighted kappas (*K*, range: 0.19–0.69), and slopes of scatter plot

best-fit lines with forced-zero intercept (“slope,” range: 0.57–1.53)—all suggest the following. The methods most similar to each other are the two interpolation methods (r: 0.77–0.88; K: 0.45–0.69; slope: 1.00–1.04). The next most similar methods are CMAQ and the two interpolation methods (r: 0.45–0.66; K: 0.19–0.49; slope: 0.57–1.45). The lowest similarity is between LUR and the three other methods (r: 0.39–0.54; K: 0.19–0.34; slope: 0.81–1.53). For pair-wise comparisons of methods, with NO and NO_2 annual-average concentrations, the proportion of PCs in the same quartile is ~65% for the two interpolation methods, 43–50% for CMAQ versus the two interpolation methods, and 32–42% for LUR versus the three other methods. Scatter plots and correlation coefficients in Fig. 3 illustrate these trends for NO and NO_2; analogous plots for CO and O_3 (SOM, Fig. A2) reveal similar patterns. Correlations and weighted kappas are generally greater for monthly averages (SOM, Tables A1 and A2) than for annual averages.

Each method accounts for different spatial scales of concentration variability. CMAQ and monitor interpolation generally account for urban-scale variations (downtown versus suburbs), whereas LUR also accounts for within-neighborhood variations (e.g., varying distance from a major road). To illustrate these differences, Fig. 4 presents concentrations on a 50-km east–west transect. While all approaches display regional variation in concentrations across the metropolitan area, CMAQ and monitor interpolation (even with GVRD’s high monitor density) do not capture the within-neighborhood concentration variability seen with LUR.

Fig. 5 further illustrates differences in spatial scales among the methods. Here, we randomly selected 100 urban PCs; then, for each selected PC we calculated the (1) distance to all other PCs and (2) concentration difference with all other PCs. For all methods, concentrations are generally less similar (i.e., concentration differences are greater) at increasing distance between two points.

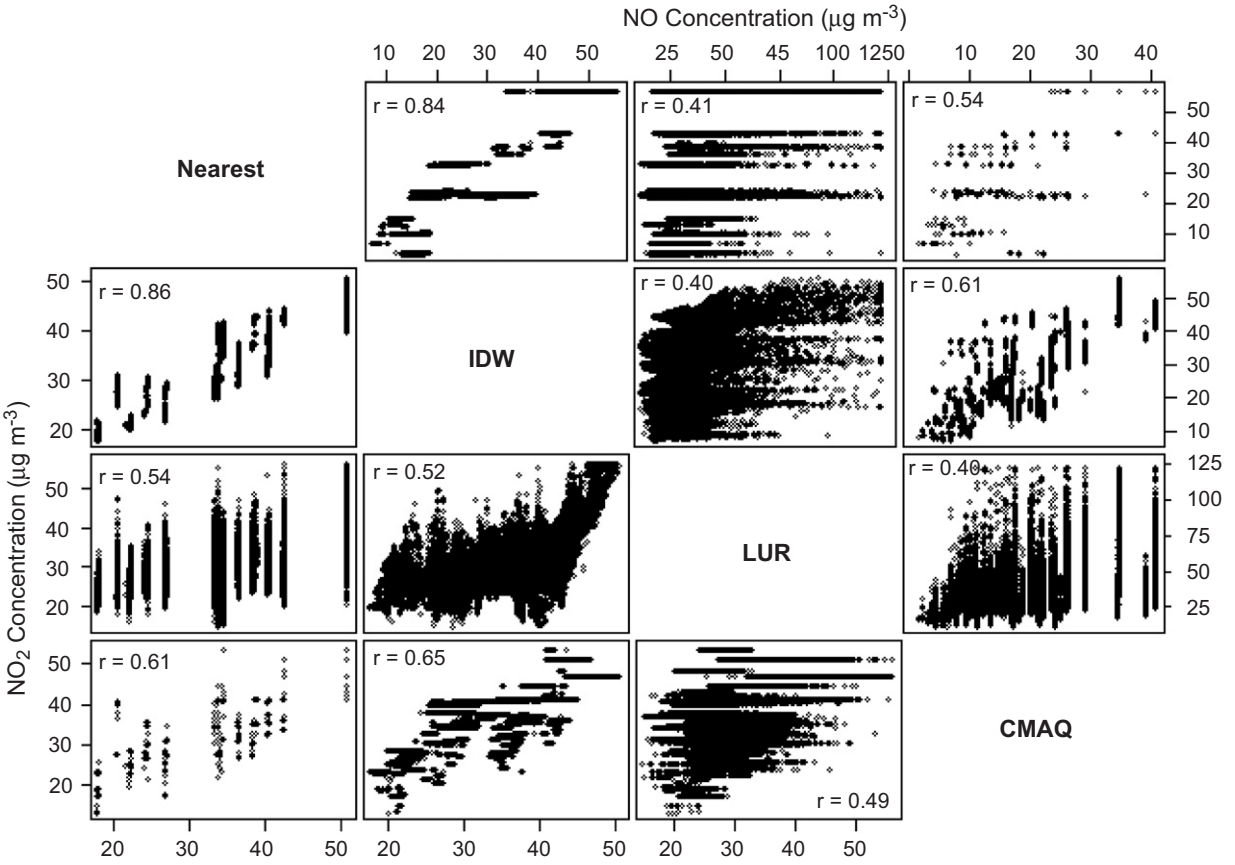

Fig. 3. Pair-wise scatter plots and Pearson correlation coefficients for estimated annual-average NO (upper right) and NO_2 (lower left) concentrations ($\mu g\,m^{-3}$).

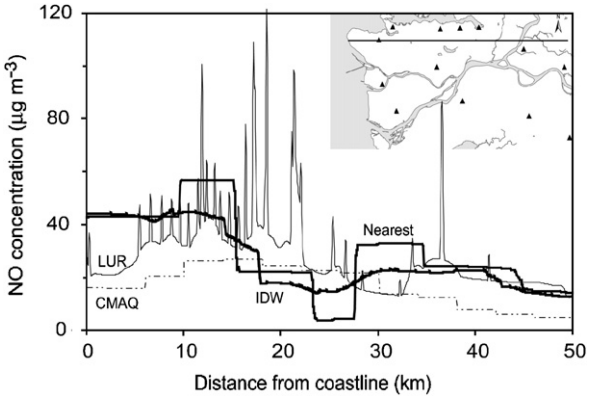

Fig. 4. Estimated nitrogen oxide concentrations along the 50-km transect shown. (For LUR, concentration spikes are generally ~150–300 m wide, i.e., model-estimated "hot spots" extend 75–150 m in each direction from road center.) The map presents NO monitor locations; other locations are mapped in GVRD (2006).

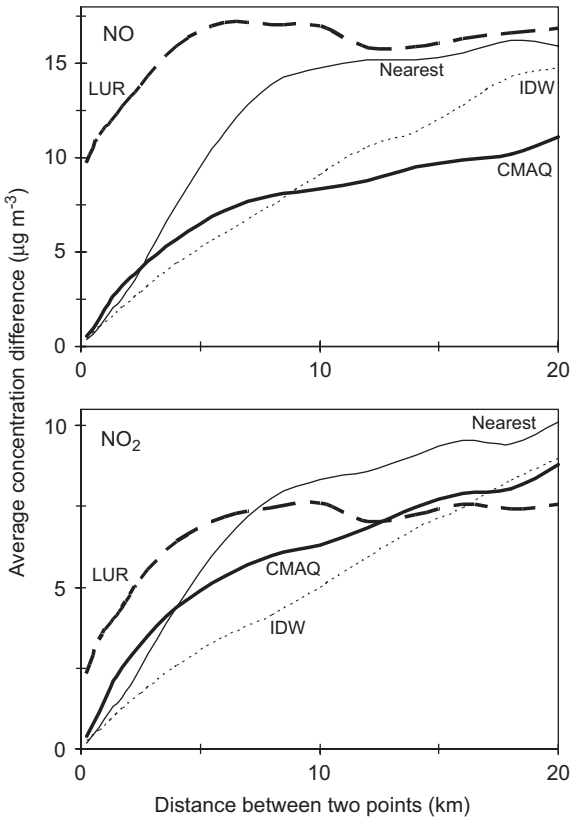

Fig. 5. Average concentration difference between two points, as a function of distance between the two points, for NO (top) and NO_2 (bottom).

At separation distances less than ~12 km for NO (~7 km for NO_2), average concentration difference (Fig. 5 ordinate) is greater for LUR than for the other methods. For example, for locations separated by between 3 and 5 km, the average difference in predicted NO concentration between the two locations is 3 times great for LUR (16 $\mu g\,m^{-3}$) than for the other three methods (average: 5.9 $\mu g\,m^{-3}$). For locations separated by 0.5–1.0 km, the difference in predicted NO concentration is 10 times greater for LUR as for other methods (11 $\mu g\,m^{-3}$, LUR; 1.2 $\mu g\,m^{-3}$, average of other methods).

We define the following characteristic length of spatial variability: the distance where the average concentration difference (Fig. 5 ordinate) equals half of the standard deviation of all concentration estimates for that method. Resulting length scales, in km, for NO (NO_2) are as follows: nearest monitor, 4 (6); IDW, 6 (6); LUR, 0.3 (1); and, CMAQ, 3 (4). These findings are consistent with Fig. 1 and with the inherent differences among methods: LUR offers greater spatial resolution than monitor interpolation or CMAQ. These analyses provide indication of spatial precision, not of accuracy.

Because the methods' concentration estimates account for different phenomena—urban versus neighborhood-scale concentration variations—epidemiological results based on these methods may yield different and potentially *non-overlapping* estimates of the dose–response relationship. Dose–response estimates generated from the different methods might not be directly comparable because they correspond to different aspects of concentration variability. We speculate that the overall dose–response may correspond to a summation of two calculated dose–responses: one computed using within-neighborhood concentration variations; the other using between-neighborhood variations. Further investigation of this hypothesis is needed.

At present, LUR typically includes predictive variables (e.g., traffic density) within a certain radius, typically on the order of a few hundred meters. LUR predictions therefore reflect nearby land uses/emissions, but not necessarily urban-scale trends. Including variables that reflect urban-scale concentration profiles (city center versus suburbs)—for example, distance and direction from the city center, the coast, major industrial areas, or a mountain range—could be considered for including in LUR models. Monitoring station data or CMAQ output could be used as an LUR input. Including a linear relationship with latitude and longitude, as was done here for NO, is likely insufficient for common urban concentration gradients since these

gradients are usually non-monotonic (instead, they first increase, then decrease as one moves across an urban area).

CMAQ grid cells (here, 4 km × 4 km) are too coarse to reveal near-roadway "hot spots" (Fig. 4). Reducing grid cell area by an order of magnitude (~1 km × 1 km—roughly the smallest grid technically possible at present) would increase uncertainty in emissions and meteorology and would dramatically increase computational intensity (Touma et al., 2006), yet model predictions would still lose fidelity in predicting near-roadway concentrations, which vary over spatial scales of ~0.1 km. Thus, we conclude that for the foreseeable future, spatial resolution will be better for LUR than CMAQ. Other dispersion models (e.g., plume models; Lagrangian puff models) can offer excellent spatial resolution; however, applying these models to an urban area (i.e., to all emission sources in the region) has limitations such as (1) high or prohibitive computational intensity (Touma et al., 2006); (2) inability to highlight concentration hotspots (Cook et al., 2007); and (3) reliance on imperfect spatiotemporal emission inventories (Sax and Isakov, 2003; Isakov and Venkatram, 2006). We cannot generalize to all situations based on our investigation of one city and one model, but like Briggs et al. (2000), we expect that estimating long-term spatial variability, including near-roadway hotspots, will typically be more efficient (in terms of level of effort, expertise, necessary funding, and in terms of prediction accuracy) using LUR than using a dispersion model.

Analyses above employ extant estimates. An implicit assumption of our comparisons is that annual-average concentrations and their spatial variations exhibit only minor changes over time scales of a few years. To explore this assumption, we considered spatiotemporal variations among annual-average GVRD monitoring station data during 2000–2002. Between-year variability is small: averages (ranges) for all monitors are 9% (4–13%), NO; 4% (2–9%), NO_2; 6% (1–13%), CO; and, 4% (2–8%), O_3. Key meteorological parameters during those years also exhibit small between-year variability: ~10% or less for mean temperature and annual precipitation (Environment Canada, 2006). For NO, NO_2, and CO, monitors' concentration rankings changed by at most two increments in 82 out of 84 cases (98%). For O_3, ranking changed by at most two increments in 25 out of 31 cases (81%). Spatial coefficients-of-variability (i.e., standard deviation among monitors divided by GVRD mean concentration) are relatively consistent among years (NO: 55%; NO_2: 28%; CO: 21%; O_3: 24%; as expected, NO is more spatially heterogeneous than the other pollutants). Thus, available evidence suggests that urban-scale spatial heterogeneities do not vary significantly year-to-year, though further investigation—especially employing LUR or otherwise considering locations other than monitoring sites—is warranted.

Another important next step is comparing against exposure measurements. One such investigation, employing 1–3 repeated 48-h personal exposure measurements of NO, NO_2, and $PM_{2.5}$ for 62 pregnant women in Vancouver, found only moderate correlation with LUR, IDW, and nearest-monitor methods (Nethery, 2007). As expected, correlations were higher for subjects with low mobility. Personal exposure measurements typically cover only a small fraction of a year (here, at most 6 days) and include time in microenvironments; methods evaluated above describe long-term concentrations and represent ambient concentrations only.

4. Conclusions

We implemented and compared three general approaches for estimating concentrations at each PC centroid for use in an epidemiological study: spatial interpolation of ambient monitoring data; LUR; and, the CMAQ dispersion model. Model–measurement comparisons at monitoring station locations indicate reasonable model performance. Two of four spatial interpolation methods employed (methods 1b and 1c, Table 1) were found to be suboptimal in this setting.

In general, the three approaches reflect different spatial scales: urban-scale variations for interpolated ambient monitoring data and CMAQ; neighborhood-scale variations for LUR. Differences in means and standard deviations among the methods are modest. LUR exhibits higher spatial resolution (precision) than the other methods, a finding we quantified using a characteristic length of spatial variability.

We suggest that LUR can be modified to better account for urban-scale concentration variations. Comprehensively characterizing within-neighborhood variations using monitoring data or CMAQ output would require a substantial effort level and is therefore unlikely to occur in the near future.

However, nesting models could offer an efficient approach, e.g., using CMAQ output or monitoring data as an independent (input) variable in a LUR model.

There is currently increasing interest in incorporating within-urban concentration variability in epidemiological, environmental justice, and other investigations. As shown here, available methods for estimating ambient concentrations differ from each other, for example, reflecting different spatial scales of concentration variability. As a result, findings from those investigations may depend on which method is employed.

Acknowledgments

We gratefully acknowledge the following funding sources: the Bridge Program (supported by the Michael Smith Foundation for Health Research (MSFHR) and the Canadian Institutes for Health Research) at The University of British Columbia (UBC); the Centre for Health and Environmental Research (funded by MSFHR) at UBC; the Border Air Quality Study (funded by Health Canada and the British Columbia Center for Disease Control); and, faculty start-up funds from University of Minnesota. The following UBC researchers generated, or provided important modifications to, datasets employed here: Sarah Henderson (land-use regression); Thomas Nipen and Roland Stull (CMAQ); and, Lillian Tamburic and Cornel Lencar (ambient monitoring data). Monitoring data were obtained from the British Columbia Ministry of Environment.

Appendix A. Supporting information

Supplementary on-line materials (SOM) associated with this article can be found in the online version at doi:10.1016/j.atmosenv.2007.08.012.

References

Basu, R., Woodruff, T.J., Parker, J.D., Saulnier, L., Schoendorf, K.C., 2004. Comparing exposure metrics in the relationship between PM2.5 and birth weight in California. Journal of Exposure Analysis and Environmental Epidemiology 14, 391–396.

Bell, M.L., 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: a case study of ozone in the Northern Georgia region of the United States. Environment International 32, 586–593.

Brauer, M., Hoek, G., Van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., Brunekreef, B., 2003. Prediction of long term average particulate air pollution concentrations by traffic indicators for epidemiological studies. Epidemiology 14, 228–239.

Brauer, M., Hoek, G., Smit, J., de Jongste, J.C., Postma, D.S., Gerritsen, J., Kerkhof, M., Brunekreef, B., 2007. Air Pollution and the development of asthma, allergy and respiratory infections in a birth cohort. European Respiratory Journal 29, 879–888.

Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., van Reeuwijk, H., Smallbone, K., Van der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. International Journal of Geographical Information Science 11, 699–718.

Briggs, D.J., de Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., Smallbone, K., 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. Science of the Total Environment 253, 151–167.

CARB, 2005. Surface area, population, and average daily vehicle miles traveled. 1 February 2005. In: The California Almanac of Emissions and Air Quality—2006 Edition (Appendix D). Available: 〈http://www.arb.ca.gov/aqd/almanac/almanac05/pdf/appd05.pdf〉 (accessed 9 May 2007).

CARB, 2006a. The California Almanac of Emissions and Air Quality—2006 Edition. California Air Resources Board. November 3, 2006. Available: 〈http://www.arb.ca.gov/aqd/almanac/almanac.htm〉 (accessed 3 April 2007).

CARB, 2006b. Quality Assurance Air Monitoring Site Information Interactive Map. California Air Resources Board. 12 April 2006. Available: 〈http://www.arb.ca.gov/qaweb/mapdemo/map_module.php〉 (accessed 3 April 2007).

Cook, R., Strum, M., Touma, J.S., Palma, T., Thurman, J., Ensley, D., Smith, R., 2007. Inhalation exposure and risk from mobile source air toxics in future years. Journal of Exposure Science and Environmental Epidemiology 17, 95–105.

Cyrys, J., Hochadel, M., Gehring, U., Hoek, G., Diegmann, V., Brunekreef, B., Heinrich, J., 2005. GIS-based estimation of exposure to particulate matter and NO_2 in an urban area: stochastic versus dispersion modeling. Environmental Health Perspectives 113, 987–992.

Delle Monache, L., Deng, X., Zhou, Y., Stull, R., 2006. Ozone ensemble forecasts: 1. A new ensemble design. Journal of Geophysical Research 111, D05307.

Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., Stull, R., 2006. Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. Journal of Geophysical Research 111, D05308.

Environment Canada, 2006. National Climate and Data Archive. 6 April 2006. Available: 〈http://www.climate.weatheroffice.ec.gc.ca〉 (Accessed 18 July 2007).

Foley, G.J., Georgopoulos, P.G., Lioy, P.J., 2003. Accountability within new ozone standards. Environmental Science and Technology 37, 392A–399A.

GVRD, 2006. Lower Fraser Valley Ambient Air Quality Report 2005. Greater Vancouver Regional District and Fraser Valley Rural District. October 2006. Available: 〈http://www.gvrd.bc.ca/air/pdfs/AmbientAirQualityReport2005.pdf〉 (accessed 8 March 2007).

Henderson, S., Brauer, M., 2005. Measurement and Modeling of Traffic-related Air Pollution in the British Columbia Lower

Mainland for Use in Health Risk Assessment and Epidemiological Analysis. 29 June 2005. Available: 〈www.cher.ubc.ca/PDFs/traffic_report_full.pdf〉 (accessed 28 June 2007).

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

Isakov, V., Venkatram, A., 2006. Resolving neighborhood scale in air toxics modeling: a case study in Wilmington, CA. Journal of the Air and Waste Management Association 56, 559–568.

Jerrett, M., Burnett, R.T., Ma, R., Pope, C.A., Krewski, D., Newbold, K.B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E.E., Thun, M.J., 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16, 727–736.

Knowlton, K., Rosenthal, J.E., Hogrefe, C., Lynn, B., Gaffin, S., Goldberg, R., Rosenzweig, C., Civerolo, K., Ku, J.Y., Kinney, P.L., 2004. Assessing ozone-related health impacts under a changing climate. Environmental Health Perspectives 112, 1557–1563.

Marr, L.C., Harley, R.A., 2002. Spectral analysis of weekday–weekend differences in ambient ozone, nitrogen oxide, and non-methane hydrocarbon time series in California. Atmospheric Environment 36, 2327–2335.

Miller, K.A., Siscovick, D.S., Sheppard, L., Shepherd, K., Sullivan, J.H., Anderson, G.L., Kaufman, J.D., 2007. Long-term exposure to air pollution and incidence of cardiovascular events in women. New England Journal of Medicine 356, 447–458.

Moore, D.K., Jerrett, M., Mack, W.J., Kunzli, N., 2007. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. Journal of Environmental Monitoring 9, 246–252.

Morgenstern, V., Zutavern, A., Cyrys, J., Brockow, I., Gehring, U., Koletzko, S., Bauer, C.-P., Reinhardt, D., Wichmann, H.-E., Heinrich, J., 2007. Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. Occupational and Environmental Medicine 64, 8–16.

Nethery, E., 2007. From measures to models: predicting exposure to air pollution among pregnant women. M.Sc. Thesis, University of British Columbia, Vancouver, British Columbia, Canada.

Ryan, P.H., LeMasters, G.K., Biswas, P., Levin, L., Hu, S., Lindsey, M., Bernstein, D.I., Lockey, J., Villareal, M., Hershey, G.K.K., Grinshpun, S.A., 2007. A comparison of proximity and land use regression traffic exposure models and wheezing in infants. Environmental Health Perspectives 115, 278–284.

Sanhueza, P.A., Reed, G.D., Davis, W.T., Miller, T.L., 2003. An environmental decision-making tool for evaluating ground-level ozone-related health effects. Journal of the Air and Waste Management Association 53, 1448–1459.

Sax, T., Isakov, V., 2003. A case study for assessing uncertainty in local-scale regulatory air quality modeling applications. Atmospheric Environment 37, 3481–3489.

Touma, J.S., Isakov, V., Ching, J., Seigneur, C., 2006. Air quality modeling of hazardous pollutants: current status and future directions. Journal of the Air and Waste Management Association 56, 547–558.

US EPA, 2006. AirData: Access to Air Pollution Data. United States Environmental Protection Agency. 〈www.epa.gov/air/data/index.html〉 (accessed 3 March 2007).