Reviewer's report

Title: Land Use Regression Modeling of Intra-Urban Residential Variability in Multiple Traffic-Related Air Pollutants

Version: 1 **Date:** 26 October 2007

Reviewer: Jeff Brook

Reviewer's report:

General

This paper describes the analysis of a large number of nitrogen dioxide (NO2), black carbon (BC) and fine particulate matter (PM2.5) measurements in Boston. These are each hypothesized to have some relationship with vehicular emissions, but potentially a different relationship, thereby requiring closer examination. The results presented are of interest to researchers concerned with how best to assign exposures to cohort members residing in the same city with traffic representing one of the main sources of the exposure. Health and exposure researchers (i.e., those with closely related research interests) will find this paper useful as it examines some details of the factors that influence exposure in more detail than past work. In undertaking this analysis the method of land-use regression (LUR) modeling was utilized. Separate models were developed and compared for NO2, BC and PM2.5.

One of the underlying questions the authors indicated they were attempting to address through the LUR approach was which traffic indicators best capture variability in different pollutants or within different settings. Comparisons of the different LUR models provided some insight regarding this issue and the effect that season and modifiers of the traffic effects (e.g., wind speed, obstructions) had on the LUR models. This paper offers some new insights on these issues and its relatively systematic sensitivity analysis and the parameters tested in that analysis make this paper a useful contribution. However, given inherent uncertainties in the predictors and the relatively small differences in the R2 between some of the different models that were compared it does not seem likely that all the findings here can be generalized. For example, although the authors show that road length within the 50 m buffer was best for NO2 and it was 200 m for BC with adjustment for obstructions or still winds, respectively, I would not recommend that future studies simply use these distances for their buffers (i.e., without their own sensitivity analysis). Similarly, road length within the buffer was found to be as good, if not better, than other traffic indicators (e.g., average daily traffic), but given limitations in the data available, I would not expect that a different location with more detailed traffic density data would necessarily also come to the same conclusion. Overall, the authors should be commended for beginning to look at such issues, but ultimately more attention would need to be paid to the reasons why different predictors were identified and what the physical meanings could be behind these differences before a truly general set of results could be put forward. For example, why were BC results interchangeable when using still winds or mean daytime wind speed or percent of day downwind from road? More attention would also need to be paid to the true underlying reliability or completeness of the different variables that were obtained or derived before they could be discounted. Hopefully, more research on these topics, addressing these limitations will be done in the future.

One of the main aspects of exposure that their LUR models desire to characterize in this paper is spatial differences among neighborhoods. However, the dataset available for this purpose was limited by the fact that in collecting the measurements the authors could only sample at 3 homes simultaneously and the homes changed weekly. Near the end of the discussion the authors indicate that this is desired, "… temporal covariates are needed for long-term exposure estimation models". However, this needs to be explained more clearly. What does it mean that "A temporally-staggered sampling design is necessary for long-term residential exposure estimation of multiple pollutants;"? Certainly, one would not want to rely solely on one short multi-site campaign to characterize the long-term exposure pattern. Sampling in different seasons is highly desired, but moving locations weekly and only 3 locations at a time seems to bring in much more temporal variability than is needed. Consequently, there was a significant temporal signal (week to week and season to season) in their data. The authors accounted for this using an existing, continuously operating central site and "temporal corrections", but these 'signals' likely did not capture all the temporal variability. The data from this site may also have different abilities to provide a correction based upon pollutant.

Temporal corrections have been necessary in past European work although more locations were sampled simultaneously and several of the Canadian models have been based upon more measurements all collected simultaneously. In the paper, the authors carry out some sensitivity analysis of the selection of central site, but not much detail is provided regarding which sites and the number of central sites available for each pollutant (e.g., there was probably only one site for BC and potentially several for PM2.5 and NO2, although this information is not forthcoming in the paper). This component of sensitivity analysis is important and, if possible, should be expanded. The authors should include more discussion in the paper on weaknesses in only being able to obtain data at 3 sites at a time. Furthermore, when they isolate and discuss the temporal and spatial signals and the amount of the variability explained by the traffic indicators it would be instructive for them to contrast their resulting models with those developed elsewhere where more sites were sampled simultaneously. This could provide useful insight for future efforts to obtain data for LUR model development.

A key conclusion of the paper is that to disentangle the differential effects on health endpoints separate exposure models (i.e., for NO2, PM2.5 and BC) are needed. This brings up a couple of critical issues related to research to identify "causal agents". Firstly, if an epidemiological study has separate models for different pollutants (i.e., PM2.5, BC and NO2) and uses these estimated exposures there are clearly significant errors associated with each exposure estimate and these vary by pollutant (differential exposure error). This is evidenced by the different model R2 values. At the next level of exposure error is the fact that the exposure estimates from these models are just for the subject's home outdoors. Differential indoor penetration and each individual's mobility add more error. Thus, given these exposure errors, which are significant, to what extent can health studies using these separate models truly disentangle the different pollutant's effects on respiratory or cardiovascular health? How should such results, if they point towards one of those pollutants over the others, be interpreted given these exposure errors? New insights may be obtained, but definitive results will probably continue to be illusive and ideally the authors should raise these challenging issues in the paper.

In addition, even if differential risks among NO2, BC and PM2.5 were identified with sufficient confidence, there still remains the question as to what each of them truly represents. For example, BC is associated with combustion and in cities motor vehicles are clearly an important source, but just because BC is more strongly related with a health endpoint does not mean BC is the causal agent because BC is almost always coated with a range of non-volatile and semi-volatile organics and can also have metals bound to it. So if the ultimate question is, what pollutants are causal, then it must be made clear that BC is also complex and thus, only an indicator, albeit somewhat more specific than PM2.5. Much research is now being motivated by trying to learn more about "causal agents". In motivating this study, it would be useful for the authors to define what they mean by "causal agents" and how separate NO2, BC and PM2.5 models will help resolve the issue.

I recommend that the authors pay some attention to the points I have raised above, and also address the minor essential revisions below. Then with the paper will warrant publication.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Background

The authors should provide a little more extensive literature review/background of what has been done in LUR, its strengths and weaknesses, including more citations. This recent paper (Arain M.A. et al., 2006) provides some insight on modification of the traffic-pollutant concentration relationship by meteorology.

There also appears to be considerable similarities in the objectives of the exposure model for BC utilized by Maynard et al. (2007). This was for the same city (Boston) and was also interested in both spatial and temporal variations. Thus, an overview of this work and how this new paper differs and improves upon that work should be included.

Arain M.A., Blair R., Finkelstein N., Brook J.R., Sahsuvaroglu T., Beckerman B., Zhang L., and Jerrett M., 2007. The use of wind fields to improve empirical chronic air pollution exposure models. Atmospheric Environment, 41:3453–3464.

Maynard D., Coul B.A., Gryparis A., Schwartz J., 2007. Mortality risk associated with short-term exposure to traffic particles and sulfates. Environmental Health Perspectives 115, 751-755.

Methods

Describe the quality of the reflectance data.

"Integrated measurements for each pollutant were collected for one week per season per home whenever feasible, in two sessions of 3 to 4 days duration."

- This is not clear. Were two samples in a week obtained at each location? If yes, were they then averaged to one value?

Give more details of the central site data and measurement methods and any collocated data comparing the methods at the home to the central site monitor method.

Show the degree of temporal variability vs. the spatial variability

"Traffic counts were collected using the Trax I Plus (JAMAR Technologies, Horsham, PA), on the highest-density road within 100 m of the home."

- were counts collected over the same time period as the PM and NO2 sampling, ie 3-4 days?

Table 3 – please express BC in the same units

Define the months in the two seasonal periods and be consistent using either summer/winter or warmer/cooler.

Pg8 "mean concentrations reported at a central site monitor (Massachusetts Department of Environmental Protection (DEP) monitor in the central Roxbury neighborhood)"

- What were the measurement methods at this site and how do they compare, directly, to the methods used at the homes and how were any systematic differences adjusted for?

"Finally, we anticipated that residential EC may display a different relationship with central site EC by season, and allowed for season-specific slopes in the model. Key sources (i.e., traffic, wood smoke, home heating fuel) may increase spatial variability during winter, when lower atmospheric mixing height may increase their influence."

- given the speculative nature of this additional adjustment it seems logical to test whether or not these factors could also influence NO2 and PM2.5 spatial variability. I don't follow the rationale for only assuming that EC could be affected by this.

“EC from different sources may also be detected differently by the reflectance-based method we used and the optical method (aethelometer) at the central site.”

- Give references to support this statement.

What evidence do you have that it is really season that drives this in Boston?

Sensitivity analyses: state reasons for the selected evaluations. IE, what scientific issues are you trying to examine? Bring this issues forward more clearly at the start of the paper as well.

Results

First paragraph of the results:

“PM2.5 and EC were significantly correlated during winter and summer ($p < 0.05$), while EC and NO2 were marginally correlated in both seasons, and PM2.5 and NO2 were not. ”

- how were these correlations done to separate time and space effects?

Figure 2a shows that EC tends to be similar in magnitude among the two seasons based upon the central site monitor with more peaks in the cooling season while at the homes there were clearly higher concentrations in the non-cooling (winter?) season.

- This needs to be investigated more closely. Either the authors picked, systematically dirtier locations in the winter (a spatial effect) or the measurement method drifted between the two periods or there is truly a seasonal effect in local exposure (at homes) that is not picked up at the selected central site.

Fig. 2b

Central site range was 0.25-1.5 µg/m3 while home range was 0.2-0.9 1/m – put these all in the same units and comment on if the data from the aethelometer can really be used to adjust for the temporal component in the reflectance data without first making sure they are compatible. This comment could go for PM2.5 and NO2 as well since the central site used different measurement technologies. Intercomparison data are needed.

Fig. 2c

NO2 – what was the measurement technique at the central site? It was more-variable across the homes compared to the central site (unlike EC). Is the greater range at the homes compared to the central site logical to expect? Why different than EC and how does this vary by season?

Sensitivity Analysis: Use of the central site monitor for temporal variability

- Be more specific regarding how many different sites were available for EACH pollutant. Were the results the same for each pollutant or was there only one site with EC so in effect a sensitivity analysis was not done for EC.

Sensitivity Analysis: “Sensitivity analyses indicated that other wind variables (mean daytime wind speed, percent of day downwind from road) were significant and may be substituted for percent of low wind speed hours, losing only marginal explanatory power"

- how can a directional variable replace a speed variable? They should be telling you something different so if they are both offering the same information in the model and can't both be included then there are some problems with what they are actually measuring.

Define what an obstruction would be and physically how it would affect NO_2 more than EC.

"Finally, we tested the addition of a season term to the $PM_{2.5}$ model, found no effect on the central site term or overall fit, and opted to leave it out. The effect of other combustion sources (i.e. smoking or grilling near outdoor monitor) would be increased by approximately 35%, however, potentially indicating seasonal differences in these sources."

- This is unclear. Does including the season term then allow you to include the other combustion sources in the mode as well? If so, maybe this is more physically-meaningful and so should be discussed.

Pg 15

"Thus, although NO_2 is higher during summer months"

- conventional knowledge is that NO_2 is higher in winter months and the central site data in Table 3 also shows this. Why are the home and central site data showing a different seasonal trend and how might this be affecting the changes in the results between including and not including season impacts the NO_2 model?

"Finally, a one-at-a-time exclusion cross-validation was performed to assess the internal consistency of model results. The Pearson correlation between estimated and measured log $PM_{2.5}$ was 0.84, 0.63 for log EC, and 0.66 for NO_2 ($p < .0001$ in all cases). Each correlation is in keeping with model R^2's, and indicates acceptable internal validity."

- this paper jumps from R^2 to Spearman to Pearson. Be more consistent, if possible. What does it mean "in keeping with model R^2's"? Find a way to compare these cross validation and actual model results directly and discuss differences and how much skill the model terms, beyond central site temporal signal, are offering.

Discussion:

"(4) modification of traffic-concentration relationships by site characteristics and meteorology."

- Couldn't this have also been due modification of home – central site relationships by site characteristics and meteorology – especially given that removing the seasonal indicator, which was impacting upon the central site to home relationship, altered the effect that meteorology had for one of the pollutants?

“potentially because of our focus within dense urban neighborhoods, while most prior intra-urban studies have actually examined metropolitan regions.”

- I disagree with the characterization that this study’s focus was variability within dense urban neighborhoods. Sites were approaching 20 km apart from SW to NE, which is the scale of many metropolitan areas and which means you were more-likely looking at variability between dense urban neighborhoods.

- Part of the reason there was greater correlation in the European studies could be because of the greater use of diesel cars so there is a strong connection between EC and PM2.5 due to more mass coming from these vehicles and there is also more NO2 from diesel vs. gasoline engines.

“A slightly larger buffer length of 200 m was effective for EC, as expected, as EC generally lies in the smaller range of PM2.5 and thus may deposit further from the source under similar conditions.”

- this is speculative and should be removed or the authors should present the available literature on deposition velocity vs. particle size for particles in the ranges know for EC and PM2.5 mass.

Discretionary Revisions (which the author can choose to ignore)

What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.