Approaches for Assessing Exposure to Traffic-Related Pollution for Tracking

Presented at the Second Annual Environmental Public Health Tracking Conference, Atlanta, GA, 2005

Paul English*
Rusty Scalf*
Svetlana Smorodinsky*
Robert Gunier*
Eric Roberts*
Zev Ross †
Michael Jerrett ‡

* CA Dept of Health Services, Oakland, CA
† Zev Ross Spatial Associates, Ithaca, NY
‡ University of Southern CA, Dept. of Biostatistics
Acknowledgements

• Funding provided by:

• National Cancer Institute Grant ##R21CA094723-03

• Centers for Disease Control and Prevention Cooperative Agreement #U50/CCU922449-01
The significance of traffic-related pollutants

– Traffic contributes to high proportion of overall air pollution burden
  • e.g.: 99% of CO burden and 76% of NO₂ burden in London attributable to traffic

– Major source of greenhouse gas emissions (CO₂)

– Increasing recognition of association with wide spectrum of health effects
Health Effects Associated with Motor-Vehicle Emissions

Motor-Vehicle Emissions → CO, PM → Childhood Leukemia

Benzene → Fetal Hypoxia, Growth, Birth weight, Prematurity, Birth defects

NO₂, Ozone, PM, DEPs → Respiratory Illness, Lung Function

PM, DEPs → Lung Cancer

Cardiovascular Disease, Premature death
Issues for routine surveillance of traffic exposures

- Routinely collected data
- Standardized collection of data
- Ease of obtaining accurate data inputs
- Indicator or marker is accurate proxy for exposure – specificity for pollutants?
Exposure Approaches

- Traffic Indicators
  - Self-reported traffic density
  - Distance from residence to roadway
  - Traffic density by census block group
  - Buffering

- GIS w/dispersion or regression modeling

- Geostatistical kriging/other interpolation methods

- Integrated meteorological-emission models

- Personal/Environmental monitoring

- Model Evaluation/Comparison
  - Which approach performs best with least cost of resources?
## Exposure Assessment of Traffic-Related Pollutants

<table>
<thead>
<tr>
<th>Complexity</th>
<th>Methods</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex</td>
<td>More complex modeling (e.g. ADMS-Urban)</td>
<td>Computationally complex, cost</td>
</tr>
<tr>
<td></td>
<td>GIS w/ dispersion modeling</td>
<td>More data inputs (e.g. meteorology, building configurations, emissions, etc.)</td>
</tr>
<tr>
<td></td>
<td>GIS w/ regression-based models</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residence near fixed air monitors</td>
<td>Assumes homogenous exposure</td>
</tr>
<tr>
<td></td>
<td>Distance-weighted traffic volume</td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>Census block-group traffic density</td>
<td>Misclassification (e.g. wind)</td>
</tr>
</tbody>
</table>
Traffic Density computed as:
(Kilometers of Road) X AADT
Area of Blockgroup in Hectares
## Traffic density – Evaluation results

<table>
<thead>
<tr>
<th>Air Pollutant</th>
<th>CO</th>
<th>1,3 – Butadiene</th>
<th>Benzene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Density (log) r</td>
<td>0.65</td>
<td>0.56</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Source: Hertz et al, 2000
Legend

- Residence of Case
- Residence of Control
- ADT Road Segment and associated volume (cars/day)
- Distance of Residence from ADT Segment
- 550 ft Buffer Boundary
Risk of 2 or more medical care visits vs. 1 visit for asthma by quintile of traffic volume on nearest street (girls)

Odds Ratio

Traffic volume (cars/day)

(adjusted for race and type of visit)

Source: English et al. 1999
Passive Monitoring Methods

- **Gradko tubes** – Palmes style diffusion tubes with triethanolamine coated metal screens.
- 37 locations
- Two-week sampling period (10 – 20 days).
- Located at libraries, churches and police stations.
- Weather covers used to protect from rain and limit sampling rate.
### Precision of Diffusion Tubes by Monitoring Period

<table>
<thead>
<tr>
<th></th>
<th>San Diego 2003</th>
<th>Alameda 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Duplicates</td>
<td>37</td>
<td>50</td>
</tr>
<tr>
<td>Intraclass Correlation Coefficient</td>
<td>0.97</td>
<td>0.86</td>
</tr>
<tr>
<td>Mean Coefficient of Variation</td>
<td>3.3%</td>
<td>6.4%</td>
</tr>
<tr>
<td>Precision (Sc)</td>
<td>0.5 ppb</td>
<td>1.0 ppb</td>
</tr>
<tr>
<td>Mean Relative Deviation</td>
<td>4.7%</td>
<td>9.1%</td>
</tr>
</tbody>
</table>
Define Buffers

Field sampling location
Traffic measurements

Within varying buffer sizes:

• Maximum traffic counts
• Distance-weighted traffic w/Gaussian dispersion
• Traffic volume (cars-km/hr)
Correlation of indicator ranks with ranks of NO$_2$ values

<table>
<thead>
<tr>
<th></th>
<th>$r$ (spearman)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. traffic in 300 m buffer</td>
<td>0.57</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gaussian adjusted max traffic in 1000 m buffer</td>
<td>0.48</td>
<td>0.002</td>
</tr>
<tr>
<td>Sum of all volumes in 300 m buffer</td>
<td>0.69</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Sum of all volumes between 40 and 300 m</td>
<td>0.68</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
Correlation of ranks of NO2 values in center of buffer and traffic volume within 300 m

Rank of NO2 Values

Rank of sum of traffic volume (cars-km/hr) in 300 m buffer
Modeled total NOx for 2000, San Diego County
Land Use Regression: Sampling Locations
Twelve Removed for Validation
Clip to Buffers
### Final Model

*R-Squared 79%*

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.3051</td>
<td>1.1039</td>
<td>4.81</td>
<td>0.0000</td>
<td>-</td>
</tr>
<tr>
<td>Road Length (40m)</td>
<td>29.4083</td>
<td>7.0382</td>
<td>4.18</td>
<td>0.0002</td>
<td>1.05</td>
</tr>
<tr>
<td>Traffic Volume (40-300m)</td>
<td>0.0017</td>
<td>0.0004</td>
<td>4.23</td>
<td>0.0002</td>
<td>1.29</td>
</tr>
<tr>
<td>Traffic Volume (300-1000m)</td>
<td>0.0002</td>
<td>0.0001</td>
<td>3.72</td>
<td>0.0007</td>
<td>1.08</td>
</tr>
<tr>
<td>Distance to Coast</td>
<td>0.0003</td>
<td>0.0001</td>
<td>4.62</td>
<td>0.0001</td>
<td>1.25</td>
</tr>
</tbody>
</table>
Predictions

All Validation Samples

*On average predicts to within 2.1 ppb*
Evaluation/Comparison of Models

• Factor of 2: (percent of predicted values between 0.5 and 2 of observed values)
• Fractional Bias (between -2 and 2; 0 is complete agreement between obs and predicted values)
### Comparison of approaches

<table>
<thead>
<tr>
<th></th>
<th>ADMS-Urban (n=38) (all sources plus background)</th>
<th>Land Use Regression (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.60</td>
<td>0.79</td>
</tr>
<tr>
<td>Fraction of 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Fractional Bias</td>
<td>17.8%</td>
<td>11.9%</td>
</tr>
<tr>
<td>% within 5 ppb</td>
<td>68.4%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Comparison of approaches

• Collins, 1998 - UK
  – Compared kriging, regression, and hybrid models to NO2 measurements (Palmes tubes)
  Adjusted R2

• Kriging: 43.9%
• GIS w/dispersion (CALINE3): 62%
• Land Use Regression (traffic volume, land cover, altitude): 81.7%
Maximum/Minimum values for residential areas (Collins, 1998)

<table>
<thead>
<tr>
<th>Method</th>
<th>Max</th>
<th>Min</th>
<th>% &gt;monitored average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kriging</td>
<td>21.64</td>
<td>41.67</td>
<td>35%</td>
</tr>
<tr>
<td>GIS w/dispersion</td>
<td>18.15</td>
<td>82.45</td>
<td>9%</td>
</tr>
<tr>
<td>GIS w/Regression</td>
<td>23.06</td>
<td>58.16</td>
<td>17%</td>
</tr>
</tbody>
</table>
Factors influencing exposure estimation:
South Coast Air Basin, CA: Based on time-activity data and CAMx model (Marshall, et al. unpublished data)
Discussion

• Each approach has limitations/strengths
  – High-end modeling is costly, requires highly trained staff, requires many data inputs
    • Potential for error if model inputs are not specified correctly or are not current
    • Models need to be evaluated, updated
    • Cost can be lower once model is running and staff is in place
    • Models that measure intra-urban variation at small scale (not regional models) necessary to capture spatial variation from traffic
Discussion, cont.

• Need for evaluation at
  – different measurement heights and wind speeds
  – Wind direction parallel to roads
  – Range of pollution concentrations

• Pollution specific models
  – Is NO\textsubscript{2} good proxy for other pollutants?
    • cost
Discussion, cont.

• Acute vs. chronic disease:
  – Hourly estimated data for acute conditions (modeling)
  – Average annual for chronic conditions (average annual traffic, monthly/seasonal pollutant levels)

• Cross-sectional vs. cohort study design
• Primary data collection – representation of sampling period
Sherlock Holmes and the case of D.C. area asthma.

Eliminate all the other factors, and the one that remains, however improbable, must be air.
Approaches

• Distance of residence from road
  – Easy to calculate (errors inherent in geocoding dependent on accuracy of street network)
  – Associated with …
  – Does not take into account varying levels of traffic density

• Traffic Density on nearest road to residence
  – Can be coupled with distance to compute distance-weighted traffic density
  – Rjinders et al. finds personal and env. measurements of NO2 related to distance and traffic volume at nearest road
  – Associated with repeated medical visits for asthma (English, et al)
  – Automated distance-weighted traffic density service for CA developed by CEHTP
  – Does not take into account prevailing upwind/downwind (misclassification)
  – Can partition out car and truck traffic
  – Could model home and workplace address
  – No specificity of pollutant