Land Use Regression of Particulate Matter in Calgary, Canada

S. Bertazzon[1], F. Underwood[1], M. Johnson[2], J. Zhang[2]

[1]University of Calgary, Department of Geography, 2500 University Dr. NW, Calgary, AB, Canada, T2N 1N4
Email: {bertazzs, feunderw}@ucalgary.ca

[2]Health Canada, Air Health Science Division, 269 Laurier Ave West, Ottawa, ON, Canada, K1A 0K9
Email: markey.johhnson@hc-sc-gc.ca; joyce.zhang@alumni.utoronto.ca

Abstract

Two-week integrated samples of particulate matter ($PM_{1.0}$, $PM_{2.5}$, PM_{10}) were collected in summer and winter in Calgary, Canada. PM concentrations were higher in summer for all size fractions. In both seasons, spatial variation and clustering were moderate. Land use regression (LUR) models were estimated for each PM size fraction and season, yielding $R^2 > 0.75$ for $PM_{2.5}$ and PM_{10} in summer, and $R^2 > 0.45$ for $PM_{1.0}$ in summer and for all winter models. Summer models yielded consistent predictors across size fractions, representing industrial emissions, local traffic, and major arterial traffic. Winter predictors included industrial emissions, major arterial traffic, and distance from open, snow-covered parks. The models suggest industrial pollution covered large areas in both seasons, and was affected by prevailing winds in summer, whereas traffic-related pollution decayed rapidly as distance from roads increased.

1. Introduction

Particulate matter (PM) is a mixture of small particles: acids, organic chemicals, metals, and dust particles (EPA 2016). Coarse particles (PM_{10}) are 2.5−10 micrometers in diameter; fine particles ($PM_{2.5}$) are less than 2.5 micrometers. Particulate pollution is associated with reduced visibility, environmental degradation, and adverse health effects, e.g., respiratory and cardiovascular morbidity and mortality (Rückerl *et al.* 2011), with evidence that health impacts and chemical composition vary by size fraction (Kelly and Fussell 2012). Land use regression (LUR) yields air pollution estimates at fine spatial resolution based on the relationship between air pollution values and land use variables observed at sampled points (Henderson *et al*., 2007). Most LUR literature focuses on NO_2, with a few studies modelling $PM_{2.5}$, ultrafine particles, and PM components (e.g., Henderson *et al*., 2007, Zhang *et al*., 2015). This paper is the first study comparing models for three PM size fractions. Further novel elements in the well-established LUR literature are the inclusion of prevailing winds and the use of GIScience to advance spatial understanding of air pollution: an example of best practice for a spatial turn in health and environmental research (Richardson *et al*., 2013).

2. Methods

Air monitoring campaigns were conducted in Calgary in August 2010 and January-February 2011. A network of 50 monitors was deployed in each campaign (Bertazzon *et al*. 2015). Due to power outages and equipment failures, the campaigns yielded 27 valid summer PM samples and 29 winter samples. Predictor variables were defined on circular buffers from each sampling point. In addition, windrose variables were defined on buffers modified according to the prevailing winds in each season (Zhang *et al*. 2015).

Getis G and Moran's I spatial statistical tests were conducted to assess spatial clustering and autocorrelation in the variables, based on a row-standardized 3-nearest-neighbours spatial weights matrix. Model selection was conducted on each PM size fraction: cross-correlation analysis selected one predictor from each category in Table 1, followed by backward variable selection (Bertazzon *et al*. 2015).

Table 1: Model Variables

Response variables	*Unit*
PM1.0, PM2.5, PM10	ug/m3

Land use variables	*Name*	*Unit or description*	*Circular buffers (meters)*	*Windrose buffer*	*dstnc*
Local roads	LRD	Total length of road segments within buffer, in meters	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	√
Major (arterial) roads	MRD		100, 200, ..., 500, 750, 1000	1500, 3000, 5000	√
Primary highways	PHW		100, 200, ..., 500, 750, 1000	1500, 3000, 5000	√
Expressways	EXPW		100, 200, ..., 500, 750, 1000	1500, 3000, 5000	√
Sum:MRD+PHW+EXPW	SMRD	Sum of segments	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Sum: PHW + EXPW	EXPHW	Sum of segments	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Traffic volume	TV	Year avg weekday T	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Population density	POP_den	Pop.in DB×DB buff. prt/ inters. area	100, 200, ..., 500, 750, 1000, 1500, 2000, 2500	1500, 3000, 5000	
Dwelling density	DWL_den	Dwl.in DB×DB buff. prt/ inters. area	100, 200, ..., 500, 750, 1000, 1500, 2000, 2500	1500, 3000, 5000	
Land use: residential	LU_res	Zoning category	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Land use: parks	LU_park	Zoning category	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Land use: institutional	LU_inst	Zoning category	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Land use: commercial	LU_com	Zoning category	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Land use: industrial	LU_ind	Zoning category	100, 200, ..., 500, 750, 1000	1500, 3000, 5000	
Industr. PM emissions	PM_EM	Report emitting pts	1000, 2000, 3000, ..., 6000	1500, 3000, 5000	√

Environmental var.s	*Name*	*Unit*
Elevation	Elev	meters
Wind speed & direction	WS_N, WS_E, WS_S, WS_W	km/hr at 10 m heigth

3. Results

Descriptive statistics for the three sets of independent variables are summarized in Table 2.

Table 2: Standard and Spatial Descriptive Statistics

		Sample	**Min.**	**Max.**	**Range**	**Mean**	**S. D.**	**S-W**	**p (SW)**	**Moran**	**p (I)**	**Getis G**	**p (G)**
PM1.0	summer	27	4.86	7.35	2.49	6.37	0.53	0.96	0.41	-0.05	0.84	0.12	0.30
	winter	29	1.50	7.36	5.86	4.18	1.23	0.98	0.84	0.06	0.09	0.11	0.12
PM2.5	summer	27	7.03	10.74	3.71	8.41	0.85	0.94	0.11	-0.01	0.81	0.04	0.26
	winter	29	2.32	9.75	7.43	5.48	1.65	0.98	0.85	0.07	0.42	0.04	0.06
PM10	summer	27	11.30	23.76	12.46	15.16	2.69	0.90	0.01	0.11	0.25	0.04	0.16
	winter	29	4.17	16.33	12.16	9.13	3.09	0.97	0.60	0.15	0.25	0.04	0.03

Particulate matter levels exhibited higher mean values in the summer. Spatial autocorrelation and clustering were never significant according to Moran's I and Getis G tests. The negative sign of Moran's I for $PM_{1.0}$ and $PM_{2.5}$ in summer suggests a dispersed, rather than clustered, spatial pattern. The Shapiro-Wilks test indicated normality for most distributions, except for summer PM_{10}. Histograms and q-q plots did not indicate large anomalies; therefore, after analyzing the log-transformed variables, all models were run on the raw variables. Seasonal LUR models for each pollutant are summarized in Table 3.

Table 3: Summer and Winter LUR Models for $PM_{1.0}$, $PM_{2.5}$, and PM_{10}

Summer PM1.0	std. β	t value	partial R^2
Intercept	6.60	40.35	
LU_indwr_3000	0.48	3.13	0.21
LRD_dist	-0.47	-3.08	0.19
EXPW_dist	-0.05	-1.56	0.02
R^2	0.49	Adj. R^2	0.42
AIC	33.20	Res. SE	0.40
Res Moran I	-0.15	p (RI)	0.78
BP test	0.44	p (BP)	0.93

Summer PM2.5	std. β	t value	partial R^2
Intercept	7.87	39.49	
LU_indwr_3000	0.78	7.35	0.56
LRD_dist	-0.34	-3.13	0.10
SMRD750	0.28	2.60	0.09
R^2	0.75	Adj. R^2	0.72
AIC	39.71	Res. SE	0.45
Res Moran I	0.02	p (RI)	0.62
BP test	2.81	p (BP)	0.42

Summer PM10	std. β	t value	partial R^2
Intercept	14.19	33.47	
LU_indwr_5000	0.64	5.53	0.48
LRD_dist	-0.24	-2.28	0.06
EXPW400	0.34	2.90	0.22
R^2	0.75	Adj. R^2	0.72
AIC	101.25	Res. SE	1.42
Res Moran I	-0.12	p (RI)	0.57
BP test	2.15	p (BP)	0.54

Winter PM1.0	std. β	t value	partial R^2
Intercept	3.38	8.62	
PM_EM6000	0.38	2.45	0.21
MRDwr_3000	0.29	1.87	0.13
LU_park200	-0.32	-2.13	0.13
R^2	0.47	Adj. R^2	0.41
AIC	84.87	Res. SE	0.95
Res Moran I	-0.11	p (RI)	0.62
BP test	1.52	p (BP)	0.68

Winter PM2.5	std. β	t value	partial R^2
Intercept	4.25	11.34	
PM_EM6000	0.31	1.84	0.17
LU_ind300	0.46	2.76	0.28
MRD200	0.27	1.95	0.06
R^2	0.51	Adj. R^2	0.45
AIC	99.68	Res. SE	1.22
Res Moran I	-0.04	p (RI)	0.46
BP test	2.59	p (BP)	0.46

Winter PM10	std. β	t value	partial R^2
Intercept	8.31	10.90	
PM_EM6000	0.30	1.82	0.16
LU_ind300	0.32	1.91	0.20
LU_park200	-0.35	-2.40	0.18
R^2	0.54	Adj. R^2	0.48
AIC	134.3	Res. SE	2.22
Res Moran I	0.00	p (RI)	0.33
BP test	3.19	p (BP)	0.36

Summer models yielded better results for coarser particulate, with $R^2 > 0.75$ for $PM_{2.5}$ and PM_{10}, and $R^2 = 0.49$ for $PM_{1.0}$. These models contained very similar sets of predictors. *Industrial-land-use* was the largest contributor to all models, on very large buffers, ranging from 3,000- to 5,000-meter radii, their shape affected by the prevailing wind (i.e., windrose). The second contributor, local traffic, was represented by the same variable in all models: *Distance-from-local-roads*. The third contributor was *Expressways* for $PM_{1.0}$ and PM_{10}, and *Sum-of-major-roads* for $PM_{2.5}$, on circular buffers ranging from 400- to 750-meter radii. The rank-order of local vs. arterial traffic was reversed in the PM_{10} model.

Winter models yielded R^2 values between 0.47 and 0.54, with slightly higher values for coarser particulate. The R^2 value was consistent with the summer value of $PM_{1.0}$, and substantially lower for $PM_{2.5}$ and PM_{10}. Industrial emissions were the main contributor to all three models, represented by *Particulate-matter-emissions*, constantly on very large, 6,000-meter radius buffers. Two of the three models featured a second prominent predictor representing industrial activities: *Industrial-land-use*, on a much smaller 300-meter radius buffer. Major arterial traffic was significant for $PM_{2.5}$ and marginally significant for $PM_{1.0}$. As in the summer models, its buffer was small for $PM_{2.5}$, but large and affected by prevailing winds for $PM_{1.0}$. *Park-land-use-within-200-meter-buffer* was significant for $PM_{1.0}$ and PM_{10}.

Standard regression diagnostics and residual tests for all models provided no evidence that any model assumptions were violated. Spatial clustering or autocorrelation in all model residuals were not significant according to the Lagrange multipliers and Breusch-Pagan tests.

4. Discussion

Spatial analyses confirmed $PM_{1.0}$, $PM_{2.5}$, and PM_{10} as regional pollutants, characterized by moderate spatial variation, with non-significant spatial clustering and autocorrelation in both seasons. Recorded particulate concentrations were lower in the winter. Summer models yielded higher goodness of fit, but winter models were more consistent across size fractions. Analytical results were consistent and interpretable, despite the low sample size.

Although model selection was conducted independently for each pollutant, it led to a remarkably consistent set of predictors, particularly for the summer models. Predictors of the summer models indicate significant association of particulate matter with industrial activities and with traffic, at the local and arterial levels. The correlation of PM with large windrose industrial buffers suggested that particulate matter of industrial origin was found at large distances from the source, with movement affected by prevailing summer winds. Conversely, correlation with relatively small circular traffic buffers suggested traffic-related PM, on local and major roads, decays rapidly as distance from roads increases.

Winter models suggested the association with industrial emissions was even stronger, particularly for coarser sizes, as $PM_{2.5}$ and PM_{10} models contained two predictors representing industrial activities. Like in summer, industrial predictors were selected on very large buffers. By contrast, winter buffers were circular, suggesting a lesser role of the wind on the widespread pattern of PM pollution of industrial origin. Association of PM with traffic was somewhat weaker in the winter, as local traffic was never significant, whereas arterial traffic was only significant for $PM_{2.5}$ and marginally significant for $PM_{1.0}$. Nonetheless, the spatial pattern of traffic pollution was consistent with the summer, with small buffers indicating rapid pollution decay as distance from roads increased. Distance from parks and open spaces, on very small buffers, was significant in the winter for $PM_{1.0}$ and PM_{10}. With most areas of the city typically covered by snow, this may indicate that particulate levels were lower over snow-covered open spaces in the winter.

5. Conclusion

Recorded PM concentrations were higher in the summer. LUR models yielded $R^2 > 0.75$ for $PM_{2.5}$ and PM_{10} in the summer, and $R^2 > 0.45$ for summer $PM_{1.0}$ and for all PM size fractions in the winter. Summer predictors were industrial emissions, local traffic, and major arterial traffic. Winter predictors included industrial emissions, industrial land use, major arterial traffic, and distance from open, snow-covered spaces. For all size fractions, the models suggested that industrial pollution extended over large areas in both winter and summer, and was affected by prevailing winds in summer; whereas traffic-related pollution, both on local roads and on major roads, decayed rapidly as distance from roads increased, in both seasons. These results are being shared with clinicians and used to inform in the creation of more environmentally-advanced models in a second study currently underway.

References

Bertazzon, S., Johnson, M., Eccles, K., & Kaplan, G. (2015). Accounting for spatial effects in land use regression for urban air pollution modelling. *Spatial and Spatio-temporal Epidemiology,* **14–15**: 9–21.

EPA, United States Environmental Protection Agency (2016). Criteria Air Pollutants. https://www.epa.gov/criteria-air-pollutants. Retrieved May 2, 2016.

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

Kelly, F. and Fussell, J. (2012). Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment*, **60**: 504–526.

Richardson, D., Volkow, N., Kwan, M.-P., Kaplan, R., Goodchild, M., & Croyle, R. (2013). Spatial turn in health research. *Science*, **339**(6126), 1390-1392.

Rückerl, R., Schneider, A., Breitner, S., Cyrys, J., & Peters, A. (2011). Health effects of particulate air pollution: A review of epidemiological evidence. *Inhalation Toxicology*, **23**(10): 555–592,

Zhang, J., Sun, L., Barrett, O., Bertazzon, S., Underwood, F., & Johnson, M. (2015). Development of land-use regression models for metals associated with airborne particulate matter in a North American city. *Atmospheric Environment,* **106**: 165–177.