



Spatial regression modelling of particulate pollution in Calgary, Canada

Stefania Bertazzon · Isabelle Couloigner · Mojgan Mirzaei

Accepted: 25 November 2020 / Published online: 5 January 2021
© Springer Nature B.V. 2021

Abstract The study presents a spatial analysis of particulate pollution, which includes not only particulate matter, but also black carbon, a pollutant of growing concern for human health. We developed land use regression (LUR) models for two particulate matter size fractions, $PM_{2.5}$ and PM_{10} , and for δC , an index calculated from black carbon (BC)—a component of $PM_{2.5}$ —which indicates the portion of organic versus elemental BC. LUR models were estimated over Calgary (Canada) for summer 2015 and winter 2016. As all samples exhibited significant spatial autocorrelation, spatial autoregressive lag (SARlag) and error (SARerr) models were computed. SARlag models were preferred for all pollutants in both seasons, and yielded goodness of fit aligned with or higher than values reported in the literature. LUR models yielded consistent sets of predictors, representing industrial activities, traffic, and elevation. The obtained model coefficients were then combined with local land use variables to compute fine-scale concentration predictions over the entire city. The predicted concentrations were slightly lower and less

dispersed than the observed ones. Consistent with observed pollution records, prediction maps exhibited higher concentration over the road network, industrial areas, and the eastern quadrants of the city. Lastly, results of a corresponding study of PM in summer 2010 and winter 2011 were considered. While the small size of the 2010–2011 sample hampered a multi-temporal analysis, we cautiously note comparable seasonal patterns and consistent association with land use variables for both PM fine fractions over the 5-year interval.

Keywords Spatial land use regression (LUR) · Spatially autoregressive lag and error models · Air pollution fine-scale prediction maps · Air pollution and human health · Particulate matter ($PM_{2.5}$ and PM_{10}) · Black carbon (BC, delta-C, organic vs. elemental)

Introduction and rationale

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 in diameter. Black carbon (BC) is a component of $PM_{2.5}$, formed by the incomplete combustion of biomass and fossil fuels,

S. Bertazzon (✉) · I. Couloigner · M. Mirzaei
Department of Geography, University of Calgary, Canada,
2500 University Dr. NW, Calgary T2N 1N4, Canada
e-mail: bertazzs@ucalgary.ca

I. Couloigner
e-mail: icouloig@ucalgary.ca

M. Mirzaei
e-mail: mojgan.mirzaei@ucalgary.ca

and part of the complex mixture often referred to as soot (Bond et al. 2013). BC is an indicator of a variable mixture of particulate material from a large variety of combustion sources, and can be separated into: elemental carbon (EC), mainly an indicator of fossil fuel combustion; and organic carbon (OC), mainly an indicator of biomass burning. Particulate pollution is associated not only with reduced visibility, environmental degradation, and climate change (Ramanathan and Carmichael 2008), but also with adverse health effects, not limited to respiratory and cardiovascular morbidity and mortality (Ruckerl et al. 2011; Janssen et al. 2012). Particle mixture and chemical composition—as well as health impacts—vary by size fraction, and health impacts vary across individuals, e.g., by age and susceptibility (Kelly and Fussell 2012; Tian et al. 2018). Scientific evidence of the environmental and health effects of BC pollution has grown in recent years, along with increasing attention to its spatial variability. Its sources range from diesel engines and coal-fired power plants to residential wood burning, agricultural waste burning, and forest fires (Janssen et al. 2012). Health concerns have sparked debate on the role of wood-burning fires and stoves as BC emitters in developed countries, both in the scientific community (Rokoff et al. 2017) and among the general public (e.g., The Guardian¹), and Canada is no exception (The Globe and Mail²).

Despite growing awareness of the spatial heterogeneity of air pollution, air quality monitoring tends to remain sparse over space, due to monitoring cost. For example, the regulatory network of Calgary, a city of over 800 km², consists of only three continuous monitoring stations, which collect PM records (CRAZ 2016). With respect to BC, Canada (ECCC 2018) has released yearly Black Carbon Inventories since 2015 (with data from 2013), yet data are aggregated at the Country level. Despite promising advances, satellite and image analysis technologies allow only for particle pollution estimates at relatively coarse spatial resolution (Hu et al. 2014; Zhang et al. 2018); hence, fine scale air quality measurements continue to be hampered by high monitoring costs. To date, land use

regression (LUR) models remain a valuable method to estimate air pollution at high spatial resolution, even at the intra-urban level, based on the observed relationship between pollution records and land use variables at sampled sites (Henderson et al. 2007).

This study presents a spatial analysis of particulate pollution in the city of Calgary, Alberta, Canada. It focusses on the analysis of two particulate matter size fractions, PM_{2.5} and PM₁₀, as well as δC (i.e., OC-EC), a black carbon index that is used as indicator of the proportion of pollution associated with fossil fuel versus biomass combustion. The analysis uses data drawn from a monitoring campaign conducted with Health Canada in summer 2015 and winter 2016 (Couloigner et al. 2017). Both PM size fractions as well as δC exhibit significant spatial variation over the area; therefore, a spatial version of standard linear LUR models, i.e., spatial autoregressive (SAR) models are estimated for each pollutant in each season. LUR model coefficients are then used to compute air pollution estimates at fine spatial resolution over the whole city. Finally, we consider a similar study of PM conducted over the same area in 2010–2011 (Bertazzon et al. 2016): based on analyses conducted over the 5-year interval, we offer some preliminary thoughts on the temporal spatial pollution pattern of PM over the area.

The use of spatial methods in LUR models, in comparison with standard LUR, i.e., linear models, reduces the spatial error associated with parameter estimates; therefore, this study improves our understanding of particulate pollution over space, thus increasing the reliability of fine-scale pollution estimates. These fine-scale pollution estimates can feed into models of health risk and environmental exposure, to aid the definition of spatially-aware health policies and healthy behavior guidelines, potentially reducing the health risks associated with particulate pollution. Furthermore, the LUR analysis of δC can contribute to our understanding of the role of domestic wood burning in black carbon pollution.

The LUR literature presents relatively few analyses of particulate matter, the majority of which, despite some exceptions (e.g., Bertazzon et al. 2016; Henderson et al. 2007; Xu et al. 2018; Zhang et al. 2015), are concerned with PM_{2.5}. Studies focusing on black carbon have only begun to emerge in recent years (Clougherty et al. 2013; Dons et al. 2013; Saraswat et al. 2013; Hankey and Marshall 2015; van Nunen

¹ <https://www.theguardian.com/money/2018/may/26/wood-burner-open-fire-pollution-cleaning-up-air-quality>.

² <https://www.theglobeandmail.com/life/home-and-garden/design/the-quest-for-cleaner-fire-why-its-time-to-rethink-our-favourite-way-to-get-warm/article28780027/>.

et al. 2017; Weichenthal et al. 2016; Lee et al. 2017). The main contribution of this paper to the LUR fine particle literature is the development of explicitly spatial LUR models, that is, spatial lag autoregressive models. Furthermore, the paper analyzes an array of particles, and models the difference between organic and elemental carbon (δC), providing insights on the main contributor of BC pollution: residential biomass burning in winter versus forest fires in summer, or fossil fuel combustion. Finally, the study addresses the seasonal and temporal dimension of fine particle pollution over a large urban area: it presents models estimated for two consecutive diametric seasons (summer 2015 and winter 2016), as well as their corresponding analyses for summer 2010 and winter 2011 over the same area for $PM_{2.5}$ and PM_{10} . Due to sample limitations of the 2010–2011 study, this paper cannot provide a multi-temporal spatial analysis.

Methods

Study site and monitoring data

Calgary is located on the eastern foothills of the Rocky Mountains, with an elevation range of 300 m around the 1050 m of the downtown core. Its climate is cold and dry, with prevailing cold and dry strong Arctic (northern) winds, with periods of warmer, moister western currents from the Pacific Ocean (see Fig. 1). The metropolitan area extends over more than 825 square kilometers; with a population of almost 1.4 million, it ranks as the fourth largest metropolitan area in Canada (Statistics Canada 2017). Industrial land use (light manufacturing) lies mostly in the eastern quadrants of the city. Commercial and urban transportation occurs on a network of roads and two railroads across the city.³ Lying on the transition between the foothills and the prairie, the urban area is surrounded by livestock operations, hay growing and agriculture, satellite communities, and First Nations territories (Tesar 2018).

The sampling campaign was conducted in summer 2015 (August 5–19) and winter 2016 (January 20–February 3), with the deployment of 125 monitors over an extended regional study area; of these sites, 84 were

located within the city of Calgary. In addition to the 84 urban sites, this study used sites located within a 10-km buffer around the city, for a total of 108 monitoring sites. Using a modified version of the location–allocation (L–A) method suggested by Kanaroglou et al. (2005), sampling sites were identified to optimize spatial coverage and site representativeness, with a minimum 1-km distance between any two sites, as detailed by Bertazzon et al. (2015, 2019). The campaign made extensive use of volunteered geographic information (VGI) (Goodchild 2007). Volunteer hosts contributed to the campaign success in many ways, e.g., by allowing for optimal monitor location, by avoiding disturbances to the recordings, such as mowing the lawn or smoking near the monitors, and by promptly alerting the research team of power outages and any other malfunction or potential interference with the recording.

For each pollutant ($PM_{2.5}$, PM_{10} , and BC), the analyses yielded 85–86 valid samples in summer 2015 and 93–100 valid samples in winter 2016 (Couloigner et al. 2017). All these sample sites are shown in Fig. 1. Health Canada analyzed the recordings for each pollutant and provided validated 2-week integrated pollutant concentrations for each season (Couloigner et al. 2017). Gravimetric $PM_{2.5}$ and PM_{10} measurements were collected using Harvard Cascade Impactors developed by Lee et al. (2006) with 37 mm Teflon filters. BC was measured via optical scanning of gravimetric $PM_{2.5}$ samples using a SootScan Model OT21 transmissometer, for which the analysis was conducted at two wavelengths: 880 nm (measuring EC) and 370 nm (measuring OC), thus generating 2-week integrated measures of EC and OC. $\delta C (= OC - EC)$ was calculated to provide an indication of fossil fuel combustion (when $\delta C < 0$) versus biomass burning (when $\delta C > 0$) (Wang et al. 2012).

Predictors

Predictors were computed from information acquired from several sources, detailed by Bertazzon et al. (2019), which included land use (Calgary Region 2016), topography (AltaLIS lidar), industrial emissions (Environment and Climate Change Canada 2014; Environment and Natural Resources Canada 2015), and transportation network (Natural Resources Canada 2016) data. Predictors were calculated on

³ Light rail transit lines run alongside some of the main roads.

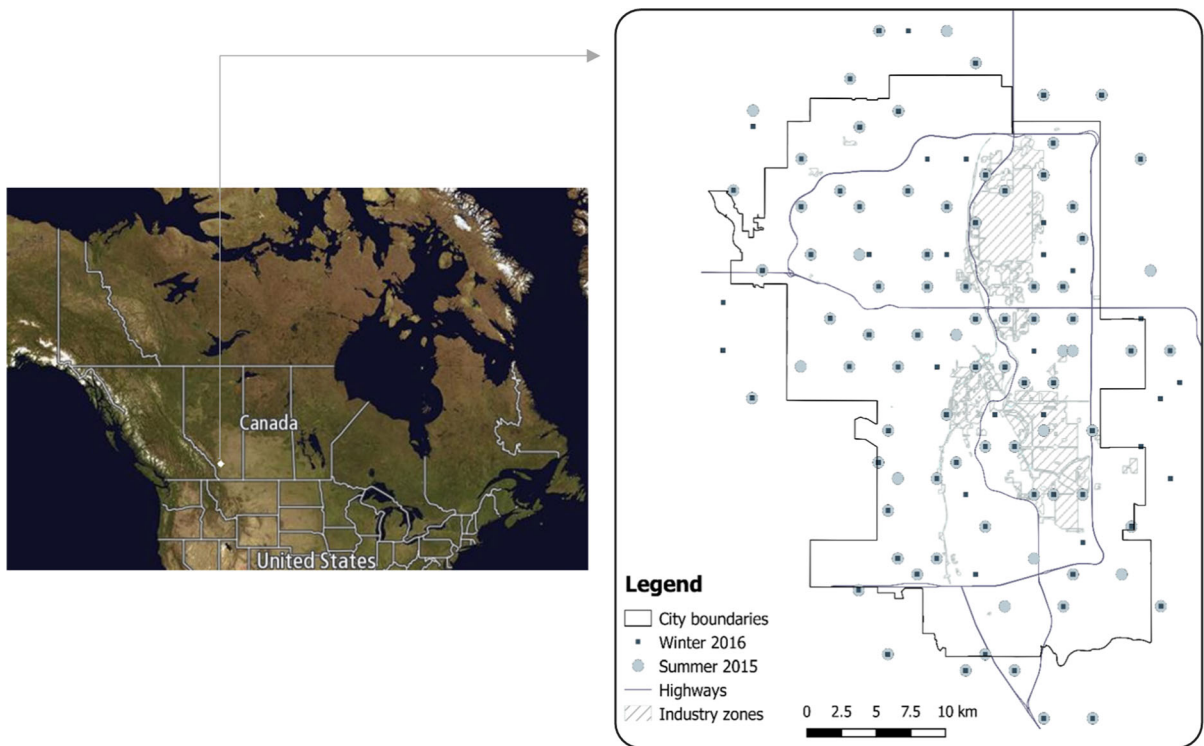


Fig. 1 Monitoring sites for summer 2015 and winter 2016 for the city of Calgary, Alberta, Canada

circular buffers defined on each monitoring site, listed in Table 1.

LUR and spatial methods

Following the methodology described by Bertazzon et al. (2015, 2019), standard descriptive statistics and exploratory spatial data analyses were conducted on the response and predictor variables. Moran's I spatial statistical tests (Getis 2008) were conducted to assess spatial autocorrelation and clustering in the response variables. Predictors were identified using cross-correlation analysis as well as a combination of expert knowledge, stepwise selection, and subsets regression methods, as detailed by Bertazzon et al. (2019). Model selection was conducted independently for each model. Models were assessed by conventional regression diagnostics for spatial autoregressive modelling. The goodness of fit of spatially autoregressive models was assessed by a pseudo R^2 , calculated as the squared correlation between observed and predicted values, following Anselin (1988). The adjusted pseudo R^2 was computed with the standard adjusted R^2 formula (Burt et al. 2009).

Standard linear LUR models (Hoek et al. 2008) are described by Eq. 1 and are generally estimated by ordinary least squares (OLS). Preliminary LUR models were estimated for each pollutant and season using this method. The residuals of each model were tested via Lagrange Multiplier and residual Moran's I tests (Getis 2008).

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \epsilon_i \quad (1)$$

In light of the spatial pattern exhibited by all the response variables and the OLS model residuals, lag and error spatially autoregressive (SAR) models (Anselin 1988; Bivand et al. 2013) were estimated. Lag SAR models (referred to as SARlag throughout the paper) are described by Eq. 2, and error SAR models (referred to as SARerr throughout the paper) are described by Eq. 3.

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \rho W y_i + \epsilon_i \quad (2)$$

$$y_i = \sum_k \beta_k x_{ik} + u_i, \bar{u} = \lambda W \bar{u} + \bar{\epsilon} \quad (3)$$

Table 1 Land use regression predictors. Adapted from Bertazzon et al. (2016)

Variables	Unit or description	Circular buffers (m)
Elevation	Elevation at the sampling site in meters	
Local roads	Total length of road segments within buffer, in meters	100, 200, ..., 500, 750, 1000
Collector roads		
Arterial roads		
Expressways and highways		
Major roads	Sum of Collector and Arterial road segments within buffer, in meters	
Land use: residential	Percentage of area covered by zoning category within buffer	100, 200, ..., 500, 750, 1000
Land use: industrial		
Land use: commercial		
Land use: institutional		
Land use: parks		
Industrial PM2.5 emissions	Reported emitting points	1000–6000 every 1000
Industrial PM10 emissions		

where y_i is the response variable at location i ; β_k are the regression coefficients; x_{i1} through x_{ik} , are the predictor variables; ρ and λ are the spatial autoregressive coefficients (for lag and error model respectively); W is a spatial weights matrix (which was created using 3 nearest neighbors); and ε is the error.

Spatial autocorrelation in OLS model residuals is conceptualized and addressed differently in the SARlag model (Eq. 2) versus the SARerr model (Eq. 3): the SARlag specification models a spatially autocorrelated dependent variable, whereas the SARerr specification models a spatially autocorrelated error, which may be associated with model misspecification, e.g., a missing independent variable.

Land use and environmental variables were created using ESRI ArcGIS. LUR models and fine level predictions were calculated in R (R Core Team 2018) using mainly the ‘spdep’ (Bivand et al. 2013; Bivand and Piras 2015), ‘car’ (Fox and Weisberg 2011) and ‘lmtest’ packages (Zeileis and Hothorn 2002). All maps were created using ESRI ArcGIS.

Results

Descriptive statistics

Descriptive statistics for all the pollutants are summarized in Table 2: the measures of central tendency are median and mean; the measures of dispersion are

standard deviation (SD) and interquartile range (IQR); the measure of normality is the Shapiro–Wilk (S–W) test; and the measure of spatial autocorrelation and clustering is global Moran’s I.

The two PM size fractions exhibited similar values in both seasons (mean and median), with greater variation in the winter (SD and IQR), whereas δC exhibited larger values and greater variance in the winter. δC exhibited negative values, indicating that elemental carbon was always greater than organic carbon. The distributions could be considered statistically normal ($S-W \geq 0.95$), and both PM fine fractions, as well as δC , exhibited significant spatial autocorrelation (Moran’s I) and clustering (Florax et al. 2003) in both seasons. Spatial autocorrelation was lower for both PM size fractions in the summer.

Summer LUR models

Lag and error SAR models for summer 2015 are summarized in Table 3.

All three summer models yielded identical sets of significant predictors. These sets comprised: two industrial indicators, i.e., industrial land use (within a 1000-m buffer) and the respective PM emitters (in all cases within a 3000-m buffer); one traffic indicator, i.e., collector roads within a 500-m buffer; and elevation. In terms of goodness of fit, the adjusted pseudo R^2 lay around 0.60 for both $PM_{2.5}$ models, around 0.70 for the PM_{10} models, and around 0.80 for the δC models. Consistent with AIC, these results

Table 2 Descriptive statistics of valid PM samples and delta C index for the 2015–2016 campaign

	N	Min	Median	IQR	Max	Mean	SD	S–W test	p(S–W)	Moran's I	p(z(I))
PM _{2.5}											
Summer	86	4.50	6.60	1.28	8.90	6.61	0.86	0.98	0.17	0.45	7.0E+09
Winter	100	1.60	6.75	4.30	14.30	6.66	2.66	0.98	0.09	0.79	<2.2e–16
PM ₁₀											
Summer	85	8.10	12.60	4.40	19.90	13.31	2.88	0.96	5.3E–03	0.43	1.3E–09
Winter	93	2.70	11.80	7.30	23.40	11.59	5.00	0.98	0.14	0.71	<2.2e–16
δC											
Summer	86	– 0.86	– 0.12	0.27	0.15	– 0.17	0.20	0.95	1.5E–03	0.68	<2.2e–16
Winter	96	– 1.34	– 0.58	0.47	– 0.12	– 0.61	0.33	0.95	1.8E–03	0.79	<2.2e–16

indicate a slightly better performance of the lag model for PM_{2.5} and δC, whereas the error model performed slightly better for PM₁₀. RSS exhibited consistent values for PM_{2.5} and δC in contrast with higher values for PM₁₀. The intercept for PM_{2.5} was high in SARerr (+ 29.7% compared to the maximum observed PM_{2.5} concentration) while it was similar to the median observed value (Table 2) for SARlag. Likewise, the intercept for PM₁₀ in SARerr was higher than SARlag (+ 56.4% vs + 9% compared to the maximum observed PM₁₀ concentration). For δC, the intercept of SARerr was as well high compared to the maximum observed δC. For these reasons, SARlag was the preferred model in all cases, and will be used to estimate fine-scale concentration levels.

Winter LUR models

Lag and error SAR models for winter 2016 are summarized in Table 4.

The winter models for all pollutants yielded similar sets of predictors, which were also consistent with the summer models. Significant predictors were: industrial indicators, i.e., land use within 1000 m; the respective PM emitters within 6000 m; traffic indicators, i.e., local roads within 100-m or major roads within a 750-m buffer; and elevation; in addition, commercial land use within 500 m was marginally significant only for PM_{2.5}. The rank order of significance of the predictors varied across models.

Overall, the winter models achieved greater goodness of fit than the summer ones, with adjusted pseudo R² around 0.86 for PM_{2.5} and δC, and in the high 0.70 s for PM₁₀. The two PM_{2.5} models were similar,

with slightly lower AIC and RSS for SARlag. The intercept of SARerr was again high (+ 83% compared to the maximum observed PM_{2.5} concentration) while it was + 7% for SARlag. For PM₁₀, the SARlag model achieved lower AIC and RSS, as well as a lower intercept. For δC, the two models were very similar, with lower RSS, AIC, and intercept for SARlag. Hence, SARlag was, again, the preferred model in all cases, and will be used to estimate fine-scale concentration levels for the winter.

Estimated concentration surfaces

Using the coefficients yielded by the SARlag models summarized in Tables 3 and 4, PM_{2.5}, PM₁₀, and δC concentrations were estimated, for summer 2015 and winter 2016, for the whole city. The estimation scale was the Dissemination Block (DB) level (Statistics Canada 2011), i.e., more than 7100 points within the urban area.

Descriptive statistics were calculated on the estimated DB level concentrations (Table 5) in order to compare them with the observed ones (Table 2) as a map accuracy assessment.

With respect to central tendency, the tables show that the medians of the estimated values were similar to those of the observed ones, with the estimated values slightly lower (< 5%), with the exception of winter PM₁₀ (+ 8%) and summer δC (– 42%). Estimated mean values were similar to the observed ones, again with the single exception of summer δC (– 29%). The interquartile range was generally lower in the estimated values (by 19% to 38%), with the exception, again, of winter PM₁₀, where it was 8%

Table 3 Summer 2015 PM_{2.5}, PM₁₀, and δC spatially autoregressive models

	SARlag		SARerr		PM ₁₀		SARlag		SARerr	
	Beta	z value	Beta	z value	Beta	z value	Beta	z value	Beta	z value
Intercept	6.39	4.03	11.54	7.10	Intercept	13.66	3.18	31.13	5.73	
Industrial 1000 m	0.01	3.69	0.02	3.92	Industrial 1000 m	0.04	3.73	0.05	4.02	
PM2.5 emitters 3000 m	0.06	2.36	0.07	2.12	PM10 emitters 3000 m	0.13	2.19	0.23	2.72	
Collector 500 m	1.8E-04	3.15	1.7E-04	3.00	Collector 500 m	4.8E-04	2.73	3.8E-04	2.30	
Elevation	- 2.7E-03	- 2.48	- 4.9E-03	- 3.30	Elevation	- 0.01	- 2.41	- 0.02	- 3.55	
(p)-R ² /adjusted	0.61	0.57	0.61	0.58	(p)-R ² /adjusted	0.69	0.66	0.71	0.68	
AIC	147.9		148.8		AIC	320.9		318.75		
RSS	22.8		22.8		RSS	183.7		172.28		
ρ or λ/p(lrt)	0.43	5.8E-05	0.46	9.2E-05	ρ or λ/p(lrt)	0.55	2.2E-08	0.64	7.4E-09	
Moran I/p(z(I))	0.02	0.36	- 0.01	0.50	Moran I/p(z(I))	0.03	0.3	- 0.01	0.50	
LMerr Res	0.04	0.84	0.02	0.88	LMerr Res	0.15	0.7	0.02	0.88	
δC	SARlag		SARerr		SARlag		SARerr			
	Beta	z value	Beta	z value	Beta	z value	Beta	z value		
Intercept	- 0.31	- 1.52	- 1.02	- 2.90						
Industrial 1000 m	- 3.1E-03	- 5.50	- 3.1E-03	- 4.65						
PM2.5 Emitters 3000 m	- 8.6E-03	- 2.08	- 0.01	- 1.34						
Collector 500 m	- 2.8E-05	- 2.90	- 3.4E-05	- 3.40						
Elevation	3.0E-04	1.63	8.7E-04	2.72						
(p)-R ² /adjusted	0.81	0.80	0.80	0.78						
AIC	- 152.61		- 138.05							
RSS	0.63		0.71							
ρ or λ/p(lrt)	0.65	1.0E-12	0.71	1.8E-09						
Moran I/p(z(I))	0.02	0.34	- 0.04	0.62						
LMerr Res	0.07	0.79	0.19	0.66						

Table 4 Winter 2016 PM2.5, PM10 and δC spatially autoregressive models

PM _{2.5}	SARlag		SARerr		PM ₁₀		SARlag		SARerr	
	Beta	z value	Beta	z value	Beta	z value	Beta	z value	Beta	z value
Intercept	15.11	5.32	26.17	7.84	Intercept	32.38	5.42	32.38	50.41	6.57
Industrial 1000 m	0.03	6.05	0.03	5.35	Industrial 1000 m	4.30E-02	2.93	4.30E-02	0.04	2.21
PM2.5 emitters 6000 m	0.05	2.91	0.12	4.46	PM10 emitters 3000 m	0.59	3.74	0.59	0.63	3.23
Local road 100 m	7.7E-04	1.48	5.5E-04	0.97	Local road 100 m	3.50E-03	2.92	3.50E-03	2.4E-03	1.7
Elevation	- 0.01	- 5.03	- 0.02	- 6.33	Elevation	- 0.03	- 5.15	- 0.03	- 0.04	- 5.38
Commercial 500 m	0.02	1.65	0.02	1.25						
(p)-R ² /adjusted	0.87	0.86	0.87	0.86	(p)-R ² /adjusted	0.8	0.78	0.8	0.76	0.75
AIC	301.31		308.42		AIC	435.37		435.37	455.81	
RSS	93.24		94.82		RSS	469.2		469.2	566.36	
ρ or $\lambda/p(\text{Irt})$	0.53	8.2E-12	0.65	3.1E-10	ρ or $\lambda/p(\text{Irt})$	0.5	1.90E-08	0.5	0.58	8.2E-04
Moran I/p(z(I))	- 0.01	0.49	- 0.03	0.62	Moran I/p(z(I))	- 0.14	0.95	- 0.14	- 0.04	0.64
LMerr Res	0.01	0.91	0.19	0.67	LMerr Res	2.97	0.08	2.97	0.24	0.62
δC	SARlag		SARerr		SARlag		SARerr			
	Beta	z value	Beta	z value	Beta	z value	Beta	z value		
Intercept	- 1.27	- 4.1	- 2.77	- 6.61						
Industrial 750 m	- 4.0E-03	- 6.01	- 3.2E-03	- 4.73						
PM2.5 emitters 6000 m	- 8.6E-03	- 4.41	- 1.5E-02	- 4.32						
Major road 750 m	- 2.4E-05	- 3.18	- 1.8E-03	- 2.61						
Elevation	1.1E-03	4.05	2.2E-03	5.7						
(p)-R ² /adjusted	0.87	0.86	0.87	0.86						
AIC	- 117.07		- 106.76							
RSS	1.29		1.33							
ρ or $\lambda/p(\text{Irt})$	0.58	5.1E-13	0.72	9.8E-11						
Moran I/p(z(I))	0.03	0.32	- 0.01	0.52						
LMerr Res	0.1	0.75	0.03	0.85						

Table 5 Descriptive statistics of fine-scale estimated PM and δC from the 2015–2016 models

	N	Min	Median	IQR	Max	Mean	SD
PM_{2.5}							
Summer 15	7137	5.10	6.42	0.80	10.03	6.57	0.74
Winter 16	7137	0.31	6.69	3.06	15.86	6.90	2.49
PM₁₀							
Summer 15	7137	7.79	12.58	2.84	25.58	13.08	2.69
Winter 16	7137	− 2.64	12.76	7.89	39.24	14.04	6.50
δC							
Summer 15	7137	− 1.19	− 0.07	0.18	0.17	− 0.12	0.21
Winter 16	7137	− 1.90	− 0.55	0.38	0.09	− 0.60	0.33

higher. As well, the standard deviation of the estimated values was similar to that of the observed values. Overall, the pollutants’ estimations exhibited slightly lower values than the observed concentrations and with lower dispersion. Notably, the sample size of the estimated values was over 70 times larger than the

observed ones, which may affect the comparability of the statistics.

DB-level PM_{2.5} and PM₁₀ estimated concentration maps obtained from SARlag model coefficients are shown in Fig. 2 for summer 2015 and in Fig. 3 for winter 2016. DB-level δC estimated concentration

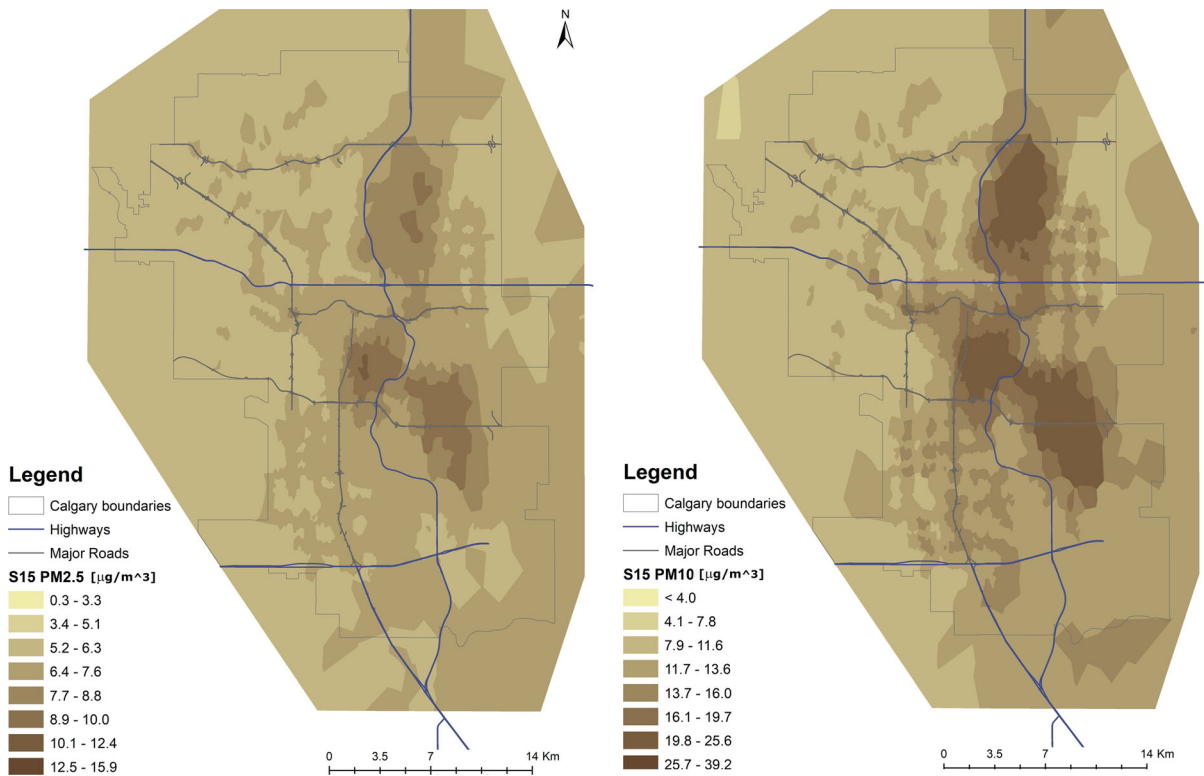


Fig. 2 Estimated concentration maps of summer PM_{2.5} and PM₁₀ at fine scale using the SARlag models

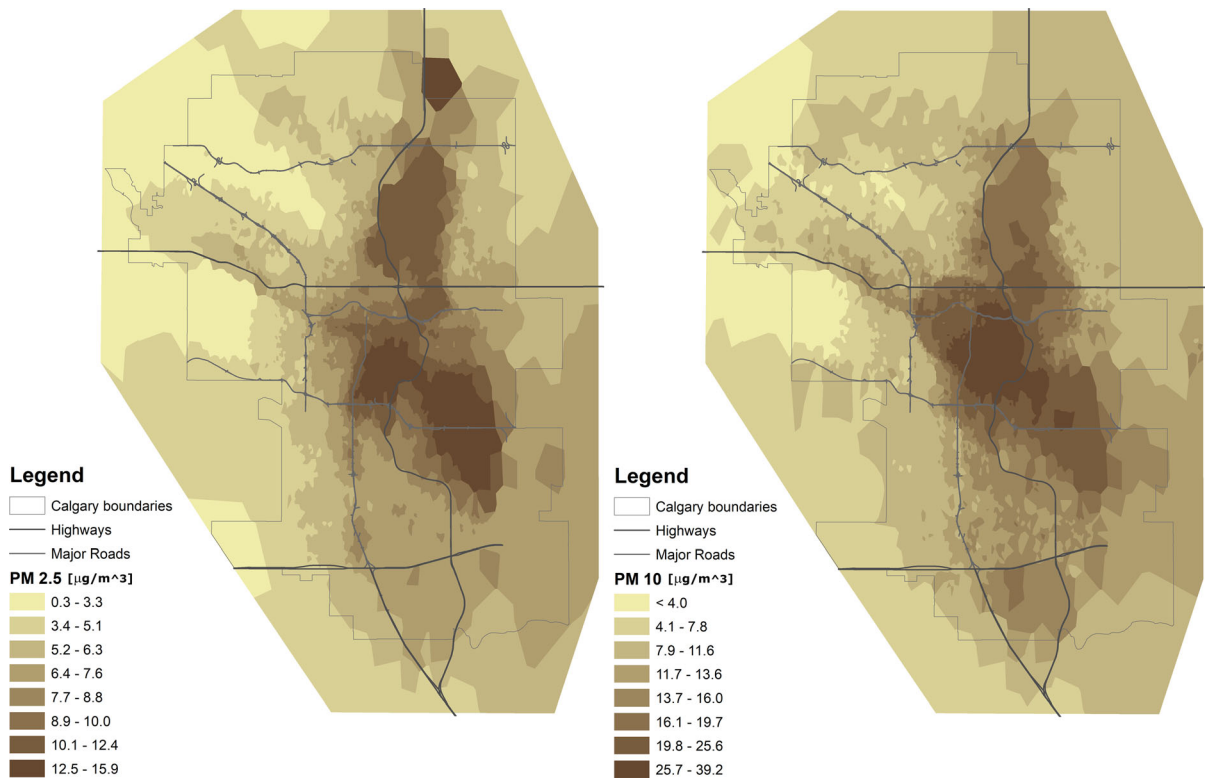


Fig. 3 Estimated concentration maps of winter $PM_{2.5}$ and PM_{10} at fine scale using SARlag model coefficients

maps, obtained with the same method, are presented in Fig. 4.

The estimated concentration surfaces for the two PM size fractions exhibited consistent patterns, with elevated values over the industrial areas and the road network. $PM_{2.5}$ exhibited a higher background level outside these areas, with a more diffused pattern. For both fine fractions, maps show slightly higher concentrations in the eastern part of the city, according to the prevailing winds, captured by the spatial autoregressive term, which are westerly in the summer and northwesterly in the winter.

Winter and summer surfaces are mapped using a consistent classification for each pollutant. For both fine fractions, the estimated winter spatial patterns exhibited association with industrial zones and with the road network. The association with the local road network was more pronounced in winter than summer. They both exhibited higher concentrations over the east.

The summer estimated concentration surface of δC presented slightly positive values and exhibited a sharp contrast between low concentrations in the west

quadrants versus high concentrations in the east quadrants: the industrial areas emerged clearly even from the eastern polluted background. The winter map exhibited a more consistent pattern of pollution over all quadrants. In the winter, pollution also radiated more gradually from industrial areas.

PM models for summer 2010 and winter 2011

A corresponding campaign was conducted for $PM_{2.5}$ and PM_{10} in summer 2010 (August 4–18) and winter 2011 (January 29–February 11), deploying 50 monitors within the city limits with the allocation strategy described above. Due to power outages, equipment failures, and other interferences, the campaign yielded only 27 valid samples in the summer and 29 in the winter. Due to the unpredictable nature of the malfunctions, the spatial sample was more random than planned (Zhang et al. 2015; Bertazzon et al. 2016). Descriptive statistics are summarized in Table 6.

In summer 2010, both PM fractions exhibited greater (mean and median) values than in winter 2011,

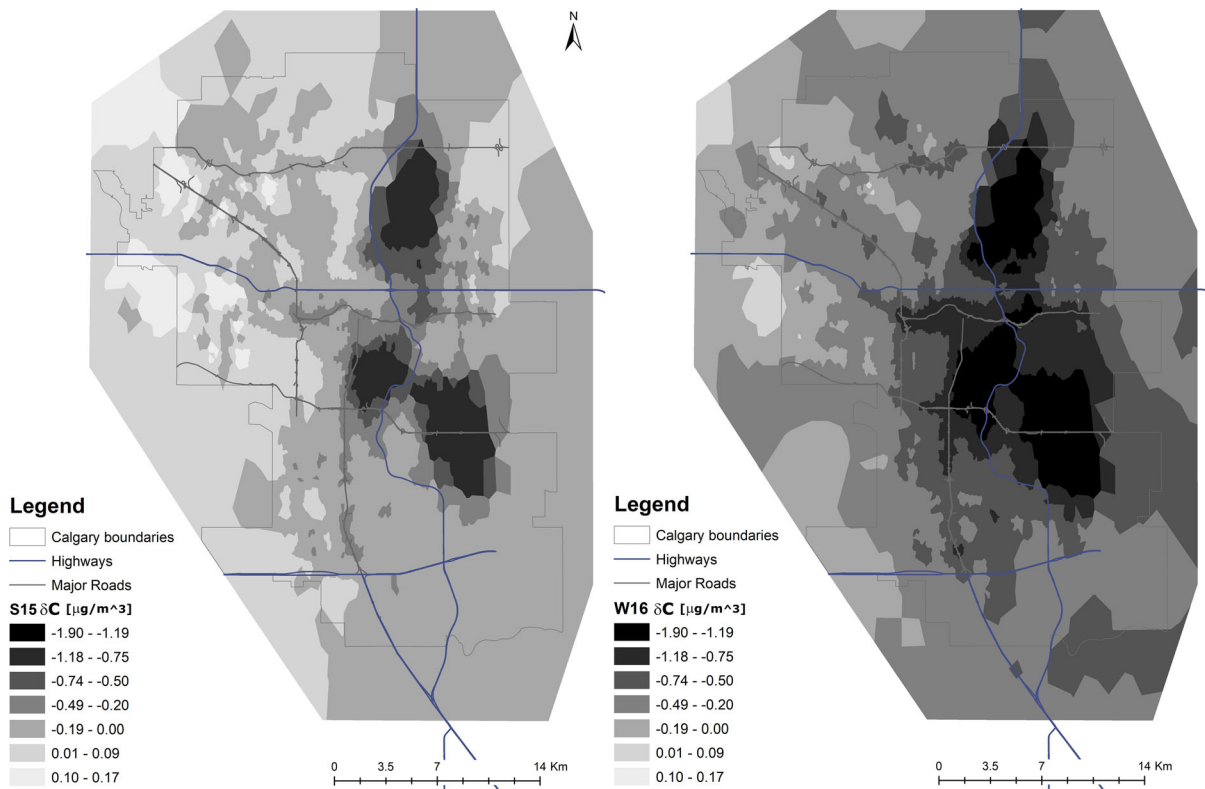


Fig. 4 Estimated concentration map of summer and winter δC at fine scale using SARlag model coefficients

Table 6 Descriptive statistics of valid PM samples for the 2010–2011 campaigns. Adapted from Bertazzon et al. (2016)

	N	Min	Median	IQR	Max	Mean	SD	S–W test	p(S–W)	Moran’s I	p(z(I))
PM_{2.5}											
Summer	27	7.03	8.29	1.03	10.74	8.41	0.85	0.94	0.11	0.00	7.8E–01
Winter	29	2.32	5.29	1.61	9.75	5.48	1.65	0.98	0.85	0.07	0.42
PM₁₀											
Summer	27	11.30	14.26	2.97	23.76	15.16	2.69	0.90	0.01	0.11	2.3E–01
Winter	29	4.17	8.66	4.36	16.33	9.13	3.09	0.97	0.60	0.13	0.21

yet both fine fractions exhibited greater variability (IQR, standard deviation) in the winter. Both fine fractions could be considered normal in the winter ($S-W \geq 0.95$) but not in the summer ($S-W \leq 0.95$). Neither fine fraction exhibited significant spatial autocorrelation in either season.

In light of these results, standard regression methods, i.e., OLS, were employed to estimate the LUR models, summarized in Table 7.

Summer 2010 LUR models yielded one industrial indicator and two traffic indicators. For both PM size fractions, adjusted R^2 lied in the low 0.70 s and residual spatial autocorrelation was not significant. RSS for PM_{10} was approximately twice as large as for $PM_{2.5}$. Similarly, the intercept of PM_{10} was approximately twice as large as that of $PM_{2.5}$.

Winter 2011 LUR models were less consistent. For both fine fractions they featured two industrial indicators, i.e., industrial land use (over buffers ranging

Table 7 Standard LUR models for PM_{2.5} and PM₁₀ in summer 2010 and winter 2011. Adapted from Bertazzon et al. (2016)

PM _{2.5}	Beta	t value	PM ₁₀	Beta	t value
Summer 2010					
Intercept	7.87E+00	39.49	Intercept	1.42E+01	33.47
Industrial WR ¹ 3000 m	5.86E−07	7.35	Industrial WR ^a 5000 m	6.82E−07	5.53
Local roads distance	− 5.11E−03	− 3.13	Local roads distance	− 1.16E−02	− 2.28
Sum major roads 750 m	7.25E−05	2.60	Expressways 400 m	2.33E−03	2.90
(p)-R ² /adjusted	0.75	0.72	(p)-R ² /adjusted	0.75	0.72
AIC	39.71		AIC	101.25	
RSS	4.75		RSS	46.41	
Moran I/p(z(I))	0.01	0.64	Moran I/p(z(I))	− 0.12	0.57
LMerr Res	0.25	0.62	LMerr Res	0.74	0.39
Winter 2011					
Intercept	4.25E+00	11.34	Intercept	8.31E+00	10.90
Industrial WR ¹ 3000 m	1.00E−05	2.76	Industrial 300 m	1.31E−05	1.91
PM emitters 6000 m	2.43E−01	1.84	PM emitters 6000 m	0.44	1.82
Major roads 200 m	1.55E−03	1.95	Parks 200 m	− 5.05E−05	− 2.40
(p)-R ² /adjusted	0.51	0.45	(p)-R ² /adjusted	0.54	0.48
AIC	99.68		AIC	134.33	
RSS	37.41		RSS	123.57	
Moran I/p(z(I))	− 0.04	0.46	Moran I/p(z(I))	− 1.43E−03	0.33
LMerr Res	0.06	0.80	LMerr Res	1.00E−04	0.99

^aWR: wind rose buffers, i.e., shapes created from circular buffers to represent prevailing winds

greatly in size) and PM emitters over identical buffers; the third predictor was major roads for PM_{2.5}, and park land use, with negative coefficient, for PM₁₀. For both fine fractions, adjusted R² lied around 0.50 and residual spatial autocorrelation was not significant. RSS of PM₁₀ was approximately four times larger than that of PM_{2.5}, and the intercept of PM₁₀ was approximately twice that of PM_{2.5}.

Discussion

Seasonal and spatial distribution of PM and δC

The two PM size fractions exhibited consistent seasonal and spatial patterns. Seasonally, concentrations were similar in summer and winter, with greater variability in the winter. As well, spatial autocorrelation was significant in both seasons, yet greater in the winter. δC exhibited similar seasonal and spatial

patterns, yet with greater values in winter than summer.

Winter PM pollution is generally associated with heating and winter driving; summer PM pollution in a prairie city with dry climate and strong winds may be associated with dust, dirt, and soil particles, largely originating naturally in open spaces in and around the urban area. In the winter, the snow cover might keep particles on the ground, reducing the suspended amount. Summer particulate pollution may also be associated with smoke from forest fires, which occur in the late summer in the forested mountain areas west and southwest of Calgary and have become increasingly frequent and severe in recent years. Smoke and particles are transported over the city by westerly winds, at times posing visibility and health hazards. However, no major forest fire events were recorded in the nearby mountain areas during the summer campaign (Mirzaei et al. 2018).

δC exhibited consistently negative values, indicating that the elemental carbon fraction was always

larger than the organic carbon fraction. Specifically, the negative sign of δC in the winter suggests that domestic biomass burning, through wood-fire burning stoves and fireplaces, provides no more than a modest contribution to BC pollution. These results thus appear to challenge the widespread narrative that ascribes much of the BC pollution to Calgarians' wood burning in their fireplaces during the winter. Further, δC is lower in winter than in summer, indicating that the portion of elemental carbon pollution is larger in the winter, which is consistent with higher PM pollution in the winter, associated with gas heating and winter traffic, with vehicle engines running harder and idling more, therefore polluting more during the cold months (i.e., when daily temperatures average -7.5 °C, versus 15.2 °C summer average daily temperatures).

Spatial LUR models and predictors

All models achieved a goodness of fit aligned with—or greater than—results reported in the literature (e.g., Hankey and Marshall 2015; Lee et al. 2017; van Nunen et al. 2017; Weichenthal et al. 2016). Spatially autoregressive models often achieve goodness of fit greater than the corresponding standard models, owing to the additional autoregressive term (refer to Eqs. 1, 2, and 3). While varying in value, the autoregressive ρ and λ coefficients consistently exhibited high and significant values. These results, along with residual spatial autocorrelation tests, suggest that the spatial autoregressive models effectively address the spatial autocorrelation observed in the response variables. The better performance of the SARlag models, compared with the SARerror models, further suggests that the observed spatial autocorrelation is indeed associated with the spatial clustering in the response variables, rather than an indication of misspecification, e.g., missing variable. Overall, SARlag models appear to be adequate tools for all three response variables in both seasons. This said, both SARlag and SARerr PM_{10} models exhibit high RSS in both seasons, higher than for the corresponding $PM_{2.5}$ models, despite R^2 and AIC values that are substantially aligned across the two size fractions. This residual variability does not appear to be associated with spatial clustering, as residual Moran's I and LM tests are never significant. It may be, therefore, simply associated with greater variability in the response

variable, as PM_{10} exhibits higher IQR and standard deviation than $PM_{2.5}$ in both seasons.

Summer models feature an identical set of four predictors across the three pollutant species, almost in the same rank order of significance: two industrial pollution indicators, one transportation indicator, and elevation. Winter models feature more diverse sets of predictors. Elevation ranks among the most significant predictors in all three winter models. Notably, the winter buffer of industrial land use for $PM_{2.5}$ is twice as large as the corresponding summer buffer. The winter traffic pollution indicator is local roads within a 100-m buffer, as opposed to collector roads within 500 m in the summer, which suggests a more local dimension of traffic PM pollution in this season, as noted above. As discussed elsewhere (Bertazzon et al. 2015), this may also be a spurious indicator, representing residential pollution, i.e., heating. In addition, the $PM_{2.5}$ model contains commercial land use within a 500-m buffer, possibly indicating an association with parking, starting, and idling vehicles, which may occur both with delivery trucks and consumer vehicles on shopping trips. For δC , industrial buffers are smaller, elevation is relatively less significant, and the traffic indicator is major roads, over a relatively large 750-m buffer.

Larger buffers and more significant elevation in the winter suggest the effect of stronger and more variable winds, in conjunction with lower temperatures. Indeed, these meteorological variables may be associated with the greater variability and spatial variability observed for all particles in the winter.

Estimated concentration surfaces

Estimated surfaces were calculated from model coefficients and local land use variables. These variables remain constant across seasons. With consistent summer models, differences across surfaces result only from differences in coefficients, whereas in the winter they emerge also from a more diverse set of predictors and buffer sizes.

Descriptive statistics of the estimated concentrations (map analysis) indicate that the estimated concentrations are slightly lower than the observed ones, with slightly lower dispersion. The only notable exception is summer δC , which exhibits mean and median estimated values approximately 30–40% lower than observed values. This difference may be

related to natural particles, as well as smoke from forest fires, as discussed above. The prediction map suggests that the western part of the city, closer to meadows and open land, is slightly more affected by dust than by transportation and industrial pollution. Dirt particles and forest fire smoke would enter the study area from the less-polluted west and southwest, forming a local anomaly, possibly associated with elevation, but not with traffic or industrial activities. Given the prevalence of other associations across the sampled points, this anomaly may have escaped the LUR models, resulting in under-prediction in the city's western quadrants.

Summer estimated concentration surfaces portray a more intense pattern of PM_{10} compared with $PM_{2.5}$, with higher concentrations over the industrial areas and the road network. Pollution appears to decline gradually from industrial areas and medium-size roads. Higher concentrations over the eastern quadrants are consistent with industrial land use and prevailing winds. The winter surfaces present sharper contrast between high- and low-pollution areas. Visually, industrial areas and road network merge forming highly polluted blotches that also coincide with valley bottoms where roads run, and pollution may stagnate in the winter. Larger buffers of model predictors form a more continuous winter pollution pattern over the road network. Conversely, δC exhibits more contrast in the summer and a more diffused pattern in the winter. Higher pollution over the east quadrants is more a feature of summer than of winter. The road network exhibits association with pollution only locally in the winter, e.g., NW, NE, and SW of the industrial areas. Pollution is not discernibly higher over major roads, which is consistent with model predictors, and suggests that heavy traffic travels on the partially completed ring road (Bertazzon et al. 2019), as recommended by the city of Calgary truck route bylaw (City of Calgary 2020).

Preliminary considerations: 2015–2016 versus 2010–2011 models

Despite the smaller sample size of the 2010–2011 study (approximately 1/3 of the 2015–2016 study), PM seasonal patterns were consistent across the two campaigns, although concentrations were higher in summer 2010 than winter 2011. In 2010–2011, spatial autocorrelation was not significant, suggesting that the

spatial properties detected in a large and regular sample could not be detected in a smaller and more random spatial sample.

The spatial properties of the sample led to the estimation of standard, rather than spatial, LUR models for 2010–2011, which exhibited no significant residual spatial autocorrelation. Some predictors used in that study were not the same as in 2015–2016, yet they were consistent indicators of the same phenomena (Bertazzon et al. 2016, 2019). This said, the sets of predictors were consistent, especially for the summer models. Winter 2011 models featured industrial predictors, along with major roads for $PM_{2.5}$ and park land use, with negative sign, for PM_{10} . The latter predictor may be associated with the snow cover, as discussed, or may be a spurious predictor, possibly related to the high variability of larger particles in the winter. Despite the use of standard models in 2010–2011 and spatially autoregressive models in 2015–2016, goodness of fit was comparable, with the exception of winter 2011, which was lower than winter 2016, as well as lower than summer 2010. Interestingly, despite the noted differences in sample size, model specification, and predictors, the 2010–2011 PM_{10} models exhibited higher intercept and RSS than the corresponding $PM_{2.5}$ models, as did the 2015–2016 models.

The limitations of the 2010–2011 sample prevented us from conducting a multi-temporal analysis across the 5-year interval; nonetheless, we note, with due caution, consistency between seasonal patterns and association with land use variables both for $PM_{2.5}$ and PM_{10} . Building on these preliminary, yet promising results, the research team was planning a new monitoring campaign for summer 2020 and winter 2021, which was halted by the COVID-19 pandemic. New campaigns will be conducted as soon as the conditions permit it.

Conclusion

The study analyzed two particulate matter size fractions, $PM_{2.5}$ and PM_{10} , along with black carbon (BC), a component of $PM_{2.5}$. Specifically, the BC index δC was calculated, as an indicator of the portion of organic versus elemental BC. Land use regression (LUR) models were estimated for particulate samples collected over the city of Calgary (Canada) in summer

2015 and winter 2016. All samples, $PM_{2.5}$, PM_{10} , and δC , exhibited significant spatial autocorrelation in both seasons. Therefore, spatial LUR models were computed, to reduce the spatial error associated with standard LUR models. Spatial autoregressive lag (SARlag) and error (SARerr) models were compared, and SARlag models were preferred for all pollutants in both seasons. The SARlag LUR models yielded goodness of fit aligned with or higher than values reported in the literature. Both seasonal PM_{10} models exhibited relatively high intercept and residual sum of squares, compared with the corresponding $PM_{2.5}$ models. All three summer models yielded identical sets of predictors, representing industrial activities, traffic, and elevation. The three winter models exhibited more diversity in predictors and related buffer sizes; in all models, elevation was more significant in winter than in summer. The model coefficients were used in combination with local land use variables to compute fine-scale concentration predictions over the city, i.e., at the dissemination block level, or for over 7100 points. The predicted surfaces exhibited seasonal variation, as well as differences across size fraction and pollutants. Descriptive statistics of estimated values indicated that predicted concentrations were slightly lower (mean and median) and less dispersed than the observed concentrations. Summer δC exhibited the largest difference between observed and estimated concentrations. The visual pattern discernible on the maps shows higher concentration over the road network and industrial areas, as well as the down-wind eastern quadrants. Lastly, the results of a corresponding study of the two PM size fractions in summer 2010 and winter 2011 were considered. Due to fewer monitors and random recording malfunctions, the 2010–2011 sample size was approximately 1/3 of the 2015–2016 sample size, which hampered a multi-temporal LUR analysis of PM. Preliminary thoughts drawn from the two studies suggest consistency of PM seasonal patterns and their association with land use variables over the 5-year interval.

Acknowledgements We wish to thank the Canadian Institutes for Health Research, CIHR Institute for Population and Public Health, and the O'Brien Institute for Public Health for funding the project “Walk smart, breathe smart”. We acknowledge our colleagues of Health Canada for their contribution to the monitoring campaigns, along with student volunteers, field crews, members of the OIPH Geography of Health study group, and all the individuals and organizations that hosted air

monitors. We also acknowledge the University of Calgary's Spatial and Numeric Data Service, the City of Calgary, Rockyview County, Statistics Canada, Environment Canada, and the Calgary Region Airshed Zone for data used in this study. Mojgan Mirzaei wishes to thank “Eyes High Doctoral Recruitment Scholarship” for supporting her doctoral work.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human animal rights The research does not involve human participants or animals.

References

- Anselin, L. (1988). *Spatial econometrics: Methods and models*. New York: Kluwer.
- Bertazzon, S., et al. (2015). Accounting for spatial effects in land use regression for urban air pollution modelling. *Spatial and Spatio-temporal Epidemiology*, 14–15, 9–21.
- Bertazzon, S., Underwood, F., Johnson, M., & Zhang, J. (2016). Land use regression of particulate matter in Calgary, Canada. In *International conference on GIScience short paper proceedings* (vol. 1).
- Bertazzon, S., Couloigner, I., & Underwood, F. E. (2019). Spatial land use regression of nitrogen dioxide over a 5-year interval in Calgary, Canada. *International Journal of Geographical Information Science*, 33(7), 1335–1354.
- Bivand, R., & Piras, G. (2015). Comparing implementations of estimation methods for spatial econometrics. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v063.i18>
- Bivand, R., et al. (2013). Computing the Jacobian in Gaussian spatial autoregressive models: An illustrated comparison of available methods. *Geographical Analysis*, 45(2), 150–179.
- Bond, T. C., et al. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11), 5380–5552.
- Burt, J., Barber, G., & Rigby, D. L. (2009). *Elementary statistics for geographers* (3rd ed.). New York: Guilford Press.
- Calgary Region. (2016). ‘*Calgary Region Open Data*’. Retrieved December 9, 2020, from <https://data.calgary.ca/widgets/uumr-spb3>.
- City of Calgary. (2020). Retrieved October 1, 2020 from <https://www.calgary.ca/transportation/roads/truck-and-dangerous-goods/truck-route-map.html>.
- Clougherty, J. E., et al. (2013). Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: The New York City Community Air Survey (NYCCAS). *Journal of Exposure Science & Environmental Epidemiology*, 23(3), 232–240.
- Couloigner, I., et al. (2017). *Spatial modelling of air pollutants in the city of Calgary and surrounding areas, spatial knowledge and information Canada 2017*. Alberta: Banff.
- CRAZ. (2016). *Calgary region Airshed Zone Society*. Retrieved December 9, 2020, from <http://craz.ca/monitoring/what-we-monitor/>.

- Dons, E., et al. (2013). Modeling temporal and spatial variability of traffic-related air pollution: Hourly land use regression models for black carbon. *Atmospheric Environment*, 74, 237–246.
- ECCC. (2014). *Environment and climate change Canada 'National Pollutant Release Inventory'*. Retrieved December 9, 2020, from <http://www.ec.gc.ca/inrp-npri/default.asp?lang=En&n=4A577BB9-1>.
- ECCC. (2015). *Environment and natural resources Canada 'Environmental indicators'*. Retrieved December 9, 2020, from <https://www.canada.ca/en/environment-climate-change/services/environmental-indicators.html>.
- ECCC. (2018). *Canada's black carbon inventory 2018. Report for environment and climate change Canada, June 2018*. Retrieved December 9, 2020, from <https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/air-pollutant-emissions.html>.
- EPA. (2016). *Particulate matter (PM) pollution*. Retrieved December 9, 2020, from <https://www.epa.gov/pm-pollution>.
- Florax, R. J. G. M., et al. (2003). Specification searches in spatial econometrics: The relevance of Hendry's methodology. *Regional Science and Urban Economics*, 33(5), 557–579.
- Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*. Thousand Oaks, CA: SAGE.
- Getis, A. (2008). A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical Analysis*, 40(3), 297–309.
- Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221.
- Hankey, S., & Marshall, J. D. (2015). Land use regression models of on-road particulate air pollution (Particle number, black carbon, PM_{2.5}, particle size) using mobile monitoring. *Environmental Science and Technology*, 49(15), 9194–9202.
- Henderson, S. B., et al. (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.
- Hoek, G., et al. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42(33), 7561–7578.
- Hu, X., et al. (2014). 10-year spatial and temporal trends of PM_{2.5} concentrations in the southeastern US estimated using high-resolution satellite data. *Atmospheric Chemistry and Physics*, 14, 6301–6314.
- Janssen, N. A. H., et al. (2012). *Joint World Health Organization (WHO)/Convention Task Force on Health Aspects of Air Pollution, 'Health Effects of Black Carbon' World Health Organization*. Copenhagen: Regional Office for Europe.
- Kanaroglou, P. S., et al. (2005). Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. *Atmospheric Environment*, 39(13), 2399–2409.
- Kelly, F. J., & Fussell, J. C. (2012). Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment*, 60, 504–526.
- Lee, M., et al. (2017). Land use regression modelling of air pollution in high density high rise cities: A case study in Hong Kong. *Science of the Total Environment*, 592, 306–315.
- Lee, S. J., Demokritou, P., Koutrakis, P., & Delgado-Saborit, J. M. (2006). Development and evaluation of personal respirable particulate sampler (PRPS). *Atmospheric Environment*, 40(2), 212–224.
- Mirzaei, M. et al. (2018). Modeling wildfire smoke pollution by integrating land use regression and remote sensing data: Regional multi-temporal estimates for public health and exposure models. *Atmosphere*, 9, 335. Special issue: "Impacts of Air Pollution on Human Health". <https://doi.org/10.3390/atmos9090335>.
- NRC. (2016). 'Free Data—GeoGratis' from *Natural Resources Canada*. Retrieved December 9, 2020, from <https://www.nrcan.gc.ca/topographic-information/10785>.
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ramanathan, V., & Carmichael, G. (2008). Global and regional climate changes due to black carbon. *Nature Geoscience*, 1(4), 221–227.
- Rokoff, L. B., et al. (2017). Wood stove pollution in the developed world: A case to raise awareness among pediatricians. *Current Problems in Pediatric and Adolescent Health Care*, 47(6), 123–141.
- Ruckerl, R., et al. (2011). Health effects of particulate air pollution: A review of epidemiological evidence. *Inhalation Toxicology*, 23(10), 555–592.
- Saraswat, A., et al. (2013). Spatiotemporal land use regression models of fine, ultrafine, and black carbon particulate matter in New Delhi, India. *Environmental Science and Technology*, 47(22), 12903–12911.
- Shi, Y., et al. (2016). Developing street-level PM_{2.5} and PM₁₀ land use regression models in high-density Hong Kong with urban morphological factors. *Environmental Science and Technology*, 50(15), 8178–8187.
- Statistics Canada. (2011). *Population and dwelling count highlight tables, 2016 census*. Retrieved December 9, 2020, from <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hltfst/pd-pl/index-eng.cfm>.
- Statistics Canada. (2017). *Focus on geography series, 2016 census*. Statistics Canada Catalogue no. 98-404-X2016001. Ottawa, Ontario: Data products, 2016 Census.
- Tesar, A., Treaty 7. (2018). *The Canadian encyclopedia*. Retrieved December 9, 2020, from <https://www.thecanadianencyclopedia.ca/en/article/treaty-7>.
- Tian, L., et al. (2018). Spatiotemporal changes in PM_{2.5} and their relationships with land-use and people in Hangzhou. *International Journal of Environmental Research and Public Health*, 15(10), 2192.
- van Nunen, E., Vermeulen, R., Tsai, M.-Y., Probst-Hensch, N., Ineichen, A., Davey, M., et al. (2017). Land use regression models for ultrafine particles in six European areas. *Environmental Science & Technology*, 51(6), 3336–3345.
- Wang, Y., et al. (2012). Multiple-year black carbon measurements and source apportionment using Delta-C in Rochester, New York. *Journal of the Air and Waste Management Association*, 62(8), 880–887. <https://doi.org/10.1080/10962247.2012.671792>.
- Weichenthal, S., et al. (2016). Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: A

- land use regression model. *Environmental Pollution*, 208, 241–248.
- Xu, S., et al. (2018). A hybrid Grey-Markov/LUR model for PM₁₀ concentration prediction under future urban scenarios. *Atmospheric Environment*, 187(8), 401–409.
- Zeileis, A., & Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News*, 2(3), 3.
- Zhang, G., et al. (2018). Critical review of methods to estimate PM_{2.5} concentrations within specified research region. *ISPRS International Journal of Geographic Information*, 7(9), 368.
- Zhang, J. Y., et al. (2015). Development of land-use regression models for metals associated with airborne particulate matter in a North American city. *Atmospheric Environment*, 106(7), 165–177.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.