



High-resolution spatial and spatiotemporal modelling of air pollution using fixed site and mobile monitoring in a Canadian city[☆]

Sierra Nicole Clark^{a,b}, Ryan Kulka^b, Stephane Buteau^{c,d,e}, Eric Lavigne^f, Joyce J.Y. Zhang^b, Christian Riel-Roberge^g, Audrey Smargiassi^{c,d,e}, Scott Weichenthal^{b,h}, Keith Van Ryswyk^{b,*}

^a Environmental and Social Epidemiology Section, Population Health Research Institute, St. George's, University of London, London, UK

^b Air Pollution Exposure Science Section, Water and Air Quality Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Government of Canada, Ottawa, Ontario, Canada

^c Institut National de sante publique du Quebec (INSPQ), Quebec, Canada

^d École de santé publique, Département de santé environnementale et santé au travail, Université de Montréal, Québec, Canada

^e Centre of Public Health Research, University of Montreal and CIUSSS du Centre-Sud-de-l'île-de-Montréal, Montreal, Canada

^f Populations Studies Division, Environmental Health Science and Research Bureau, Health Canada, Ottawa, Canada

^g Direction de santé publique, Centre intégré universitaire de santé et de services sociaux (CIUSSS) de la Capitale-Nationale, Quebec City, Quebec, Canada

^h Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada

ARTICLE INFO

Keywords:

Land use regression
UFP
Mobile monitoring
Spatiotemporal
Air pollution
Canada

ABSTRACT

The development of high-resolution spatial and spatiotemporal models of air pollutants is essential for exposure science and epidemiological applications. While fixed-site sampling has conventionally provided input data for statistical predictive models, the evolving mobile monitoring method offers improved spatial resolution, ideal for measuring pollutants with high spatial variability such as ultrafine particles (UFP).

The Quebec Air Pollution Exposure and Epidemiology (QAPEE) study measured and modelled the spatial and spatiotemporal distributions of understudied pollutants, such as UFPs, black carbon (BC), and brown carbon (BrC), along with fine particulate matter (PM_{2.5}), nitrogen dioxide (NO₂), and ozone (O₃) in Quebec City, Canada. We conducted a combined fixed-site (NO₂ and O₃) and mobile monitoring (PM_{2.5}, BC, BrC, and UFPs) campaign over 10-months. Mobile monitoring routes were monitored on a weekly basis between 8am–10am and designed using location/allocation modelling. Seasonal fixed-site sampling campaigns captured continuous 24-h measurements over two-week periods. Generalized Additive Models (GAMs), which combined data on pollution concentrations with spatial, temporal, and spatiotemporal predictor variables were used to model and predict concentration surfaces.

Annual models for PM_{2.5}, NO₂, O₃ as well as seven of the smallest size fractions in the UFP range, had high out of sample predictive accuracy (range r^2 : 0.54–0.86). Varying spatial patterns were observed across UFP size ranges measured as Particle Number Counts (PNC). The monthly spatiotemporal models for PM_{2.5} ($r^2 = 0.49$), BC ($r^2 = 0.27$), BrC ($r^2 = 0.29$), and PNC ($r^2 = 0.49$) had moderate or moderate-low out of sample predictive accuracy. We conducted a sensitivity analysis and found that the minimum number of 'n visits' (mobile monitoring sessions) required to model annually representative air pollution concentrations was between 24 and 32 visits dependent on the pollutant.

This study provides a single source of exposure models for a comprehensive set of air pollutants in Quebec City, Canada. These exposure models will feed into epidemiological research on the health impacts of ambient UFPs and other pollutants.

1. Introduction

Exposure to ambient air pollution is associated with adverse health

impacts throughout the life course, including respiratory conditions, cardiometabolic diseases, cancers, neurodegenerative diseases, birth outcomes, as well as premature mortality (Boogaard et al., 2022; Buteau

[☆] This paper has been recommended for acceptance by Prof. Pavlos Kassomenos.

* Corresponding author.

E-mail address: Keith.VanRyswyk@hc-sc.gc.ca (K. Van Ryswyk).

et al., 2018, 2020, COMEAP, 2022, GBD 2019 Diabetes and Air Pollution Collaborators, 2022; Crouse et al., 2010; Goldberg et al., 2001, 2006, 2013; Parent et al., 2013; Smargiassi et al., 2020; Tétreault et al., 2016; Turner et al., 2020; World Health Organization, 2021). These impacts, whether acute or chronic (World Health Organization, 2021), can vary in magnitude across cities and neighbourhoods (Bennett et al., 2019; Kheirbek et al., 2016; Liu et al., 2017; Pascal et al., 2013; Stieb et al., 2021). This is a result of spatial and temporal variations in air pollutant levels, which is influenced by locations of emissions (e.g., road-traffic networks, shipping ports, airports, industrial sites), atmospheric transport of emitted pollutants (e.g., prevailing wind direction), factors which can attenuate or reduce pollutant levels (e.g., urban parks/trees/green infrastructure) (Gourdji, 2018), as well as population characteristics and behaviours which can influence exposure and susceptibility to health effects. As such, air pollution epidemiological studies, particularly those conducted within cities, require highly spatially resolved pollution data to capture local scale variations.

The majority of the evidence of the health impacts of ambient air pollution are from studies of PM_{2.5}, NO₂, and O₃ (Boogaard et al., 2022; COMEAP, 2022; GBD 2019 Diabetes and Air Pollution Collaborators, 2022; Turner et al., 2020; World Health Organization, 2021) which are routinely monitored in many cities around the world. However, there is ongoing concern regarding ultra-fine particles (UFPs, typically with aerodynamic diameter <100 nm) as these are not typically regulated by governments and we don't know whether they might induce more toxic effects than PM_{2.5} mass concentrations. It is of public health importance to study the health effects of UFPs, especially by investigating the different size ranges that might induce health effects. Therefore, there is a need for dense urban monitoring campaigns using refined methods measuring several pollutants, including UFPs, to develop high resolution exposure models that can improve risk estimates in epidemiological studies.

Air pollution monitoring campaigns for epidemiological applications have traditionally used fixed-site approaches, where monitors are mounted at discrete locations for defined periods of time (Xie et al., 2017). More recently, mobile monitoring (MM) campaigns have been undertaken to increase the spatial resolution of urban air pollution measurement, in a cost-effective and efficient way. This is done by attaching air pollution monitoring equipment to mobile apparatus (e.g., vehicles, bicycles, people) travelling along defined routes (Anand and Phuleria, 2021; Chambliss et al., 2020; Deshmukh et al., 2020; Dionisio et al., 2010; Hankey et al., 2019; Kerckhoffs et al., 2021; Liu et al., 2019; Messier et al., 2018; Xie et al., 2017). These routes are often travelled on multiple days or seasons of the year (called 'visits' or 'drive days') (Kerckhoffs et al., 2024) to obtain temporally representative estimates (Kerckhoffs et al., 2021; Messier et al., 2018). An important but still understudied question, which has implications for improving the cost-effectiveness of MM campaign design, is how many visits are needed to accurately characterize the spatial pattern of air pollution in different urban environments (Blanco et al., 2023; Kerckhoffs et al., 2024). The answer to this question can vary by the pollutant, the temporal context of the study's research question (seasonal, annual, pre/post intervention), and the spatial context of a study (entire city, particular urban area of interest).

Statistical-based modelling approaches have frequently been used to predict the spatial and/or temporal variability in ambient air pollution concentrations based on measurements from fixed or MM campaigns (Hoek et al., 2008) within cities in high resolution. The land use regression (LUR) model is based on regressing spatially, and sometime temporally, structured geospatial and temporal predictors that represent a range of sources and the propagation and attenuation of their emissions in the urban environment (Hoek et al., 2008; Xie et al., 2017). LUR-based models have been successfully implemented with a variety of modelling approaches, including linear regression, generalized additive models (GAM), machine learning approaches, and Bayesian as well (e.g. (Hankey et al., 2019; Kerckhoffs et al., 2021; Lim et al., 2020; Liu et al.,

2019; Lloyd et al., 2023; Messier et al., 2018; Minet et al., 2018; Simon et al., 2018; Van den Bossche et al., 2020; Van den Hove et al., 2020).).

The Quebec City Air Pollution Exposure and Epidemiology (QAPEE) study sought to characterize a diverse set of understudied air pollutants in Quebec City for applications in epidemiological studies. We leveraged exposure measurements of PM_{2.5}, black and brown carbon, and UFPs collected in a 10-month MM campaign (collected between 8am–10am) as well as NO₂ and O₃ collected continuously throughout the 24-h day over two-week periods with fixed-site seasonal sampling. We then took a spatial and spatiotemporal land use regression (LUR) approach to model and predict annual and monthly ambient air pollutant concentrations in high-spatial resolution in Quebec City in a consistent way. Lastly, we conducted a sensitivity analysis to estimate the MM data collection sessions ('n visits') required to achieve robust spatial annual models for several pollutants.

2. Materials and methods

2.1. Study location

Quebec City is the capital of the province of Quebec, Canada with an estimated population of ~530,000 people (Statistics Canada, 2017). Situated on the north bank of the Saint Lawrence River, the city experiences cold winters and warm summers, with average temperatures dropping down between -10 and -25 °C in January, and up to around 20 °C in July (Government of Canada, 2022).

Residents of Quebec City may be exposed to a mixture of air pollutants varying in magnitude across the city. There is a busy city-centre, a historic district atop the Cap-Diamant hill, an international airport (Jean Lesage Airport), and a seaport in close proximity to several residential areas (Limoilou, Vanier and Québec – Basse-Ville). Port activities include shipping, cruise line traffic and major industries including an incinerator and paper mill (Institut national de santé publique du Québec, 2023). Furthermore, some of the population still burns solid fuels in winter for heating, which contributes to wintertime PM_{2.5} concentrations.

2.2. Data collection

2.2.1. Air pollution data

Our monitoring campaign consisted of 10-months of MM and two seasonal (two-week long) fixed-site measurements across Quebec City. The study area was established as a 372 km² zone surrounding the urban area of Quebec City. Mobile and fixed-site sampling locations were selected using a location allocation algorithm that took into account spatial variability in land use parameters and the distribution of study population (Kanakoglou et al., 2005). A pre-estimated pollutant surface was created using suitability analysis which incorporates common influential factors identified from the literature (Weichenthal et al., 2016a, 2016b), including highway, major road, local road, railway, parks, commercial zoning, and industrial zoning. Next, a surface of spatial variation of these factors was created using a semivariogram function. Also, as we were interested in the areas where the density of the population was high, a weighting scheme (Kanakoglou et al., 2005) was applied to the variability surface. Lastly, monitoring locations and routes were selected using a maximum attendance location allocation algorithm based on the population-weighted semi-variance (Kanakoglou et al., 2005). The location allocation algorithm was used to specify 60 sampling locations for the fixed-site sampling (FSS) campaign and 100 locations to inform route design for the MM campaign.

The FSS campaign was carried out during 14-day continuous periods in the summer (September 2019) and winter (March 2020, prior to the closure of businesses as a result of the COVID-19 pandemic) to capture concentrations of NO₂ (ppb) and O₃ (ppb). Our approach was similar to previous work in another Canadian city employing passive badges in a multi-season FSS study to create annual NO₂ models (Crouse et al.,

2009). The later summer and winter periods were chosen over mid-season periods to avoid several complications such as fireworks, holidays, high snowbanks, and adequate sunlight for solar powered equipment. Having gained permission from the municipal authorities, sampling setups were mounted to lamp poles at a height of ~3 m. Setups consisted of a rain shelter, passive Ogawa samplers (Ogawa, Florida, USA) for NO₂ and O₃. For quality assurance and control (QA/QC) purposes, we collected six duplicate and blank NO₂ and O₃ samples each season for blank corrections and to derive estimates of precision. Measured seasonal concentrations were averaged together to develop an estimate of an annual average concentration.

Three routes, 179 km in total length, were designed to cover the 100 locations specified by the location allocation algorithm, each requiring 2 h to complete (Appendix 1). Residential areas were prioritized in the design of the MM routes. The use of local roads was prioritized, using arterial roads to transit from neighbourhood to neighbourhood. Highways represented ~5 km of the MM network. With the conservative assumption of a mean 65 km/h driving speed based on a previous MM campaign (Weichenthal et al., 2016a) and a minimum road segment size of 100 m, we set the objective of conducting ~54 MM sessions to yield 5 min of data for each road segment thus providing a rich, seasonally balanced dataset to develop spatial and spatiotemporal air pollution models. Thus, the design of weekly monitoring sessions for the 1-year study period from June 2019–June 2020 was determined. Data collected after March 26, 2020 were removed from the analysis due to the COVID-19 pandemic and provincial lockdowns. A hydrogen fuel cell vehicle was equipped with monitoring equipment and an elutriator mounted on the back window (Fig. A1 in Appendix). Each monitoring week consisted of randomly assigning each route to either a Tuesday, Wednesday or Thursday and conducting each session from 8:00–10:00. For each route, six starting points were established and used in an alternating fashion to ensure the timing of each route was even over the standard 2-h time period.

Pollutants measured in the MM campaign included PM_{2.5} (µg/m³), black carbon (ng/m³), brown carbon (ng/m³), and UFPs (measured as Particle Number Count (PNC; pts/cm³)) and particle size mode (nm). These measures were collected at 1-s recording intervals. PNC and particle size mode were measured with the DiscMini from Testo (Lenzkirch, Germany) and covered a size range between 10 nm–700 nm, however the median annual particle size mode captured across MM locations was 38 nm with a range of 24.5 nm–50.2 nm which is within the typical UFP size range (<100 nm). One-minute PNCs were also estimated separately for the following 13 particle size modes: 365.2, 273.8, 205.4, 154.0, 115.5, 86.6, 64.9, 48.7, 36.5, 27.4, 20.5, 15.4, 11.5 nm by the Nanoscan from TSI, Inc. We measured PM_{2.5} concentrations with the DustTrak 8530 from TSI (Minnesota, USA), and black carbon (BC) and brown carbon (BrC) concentrations using the AE-200 Aethalometer from Magee Scientific (California, USA). Coordinate locations were measured with a DG-100 GlobalSat global positioning system (GPS) (WorldCom Group). For information on QA/QC methods please refer to Appendix 1.

Following MM data collection, the routes were divided into 500 m road segments. Each second of MM data was assigned to the 500 m road segment to which it was closest based on the road segment centreline. The median value for each combination of monitoring week and road segment was then calculated, giving the median value of one pass. For the annual models, we further aggregated the data by taking the median value across the 40 passes of each road-segment from June 12th 2019 to March 26th 2020 and for the monthly models, we aggregated the data by taking the median value across the 4 passes of each road-segment each month. The ‘median of medians’ approach was chosen as the metric of central tendency to reduce the impact of influential observations (Hankey and Marshall, 2015).

2.2.2. Predictor variables

A list of the spatial GIS (Geographic Information Systems) and temporal meteorological predictor variables which were candidates for the

models, their transformations, and sources of information, are included in Appendix 2. To process predictor variable data, we created circular buffers with radius of 50, 100, 200, 300, 400, 500, 750 and 1000 m around the latitude and longitude coordinate location of the fixed measurement site (FSS campaign), or the centroid of the 500 m long road-segment (MM campaign). We then mapped the spatial predictor variables to each buffer, centred by the coordinate location of the measurement site, through spatial overlay. Zonal statistics (e.g., average, sum, area, depending on spatial predictor type – see Appendix 2) of each spatial predictor overlapping within each buffer were calculated. We additionally calculated the Euclidean distances of the measurement sites to predictor variable locations. Meteorological variables did not have any spatial features (i.e., one value applies to the whole area) and thus did not undergo the aforementioned processing steps.

2.3. Modelling

Prior to model building, we assessed the univariate distributions of our pollution and predictor data with histograms. Due to the right skew of our pollution data, we applied log transformations to achieve normal distributions for modelling. We then assessed the linearity/non-linearity of the bivariate distributions between all combinations of pollution and predictor data with scatter plots (two continuous variables) and box and whisker plots (one continuous and one categorical variable).

For our predictive modelling framework, we used Generalized Additive Models (GAM) as they allow for modelling linear and non-linear associations and can also flexibly account for spatial, temporal, and hierarchical structures within the data (Ross, n.d.). GAMs have also been used previously for similar applications achieving high predictive accuracy (Harper et al., 2021; Lloyd et al., 2023). Using the FSS and MM data, we built two types of LUR models:

- i) models of **annual** median ambient concentrations of PM_{2.5}, BC, BrC, PNC (total and for 13 size fractions), particle size mode and **annual** average ambient concentrations of NO₂, and O₃ (i.e., spatial models);
- ii) models of **monthly** median ambient concentrations of PM_{2.5}, BC, BrC, and PNC (i.e., spatiotemporal models).

2.3.1. Spatial annual models

We used FSS data to model ambient average concentrations of NO₂ (ppb), and O₃ (ppb) and MM data to model ambient median concentrations of PM_{2.5} (µg/m³), BC (ng/m³), BrC (ng/m³), PNC (pts/cm³), particle size mode (nm), and PNCs specific to 13 size fractions (pts/cm³).

We initially identified the ‘best’ combination of predictor variables for each pollutant model by running up to 4017 models each, representing all combinations of predictors, and selecting combinations which maximized model predictive accuracy and minimized error (more details are in Appendix 2). After the ‘best’ combination of predictor variables were identified for each pollutant LUR model (regardless of the associated p-value of the coefficient), we continued to build the GAM models by testing and applying non-linear smooth functions to the predictor variables. Specifically, we first assessed the linearity between each continuous predictor variable and each log-transformed pollutant with bivariate scatterplots, without adjusting for the other variables. If the bivariate relationships appeared non-linear, we then applied non-linear spline functions to the corresponding predictor variables in the multivariable models. The optimal number of knots was determined through graphical analysis of GAMs to ensure that modelled relationships did not overfit the data and key aspects of relationships were not missed.

We evaluated whether model assumptions (e.g., residual random and constant variance) were upheld using diagnostic plots of residuals. Furthermore, given the nature of our air pollution data and the data collection campaign, it was highly likely that the pollution data were correlated across space, which could lead to residual spatial

autocorrelation in the models. We thus checked whether the models had residual spatial autocorrelation using the Global Moran's I statistic of spatial randomness (Dormann et al., 2007). We also visually assessed the potential spatial patterning of residuals by mapping them across the study area. If spatial autocorrelation appeared present, similar to the approach taken by Harper et al. (2021), we then explicitly modelled this spatial dependence by including the coordinate locations of road-segment centroids (MM) or sampling sites (FSS) into the model as interaction terms (longitude x latitude) with Gaussian Process (GP) smooth functions with 10–100 knots, opting for fewer numbers of knots where possible to do so. We call these the *smooth spatial terms* in the model.

We evaluated and compared the predictive accuracy of the final selected spatial annual GAM models with 10-fold cross validation with 10% random sites held out (10%_{SiteCV}). Metrics of predictive accuracy on the held-out data included the Root Mean Squared Error (RMSE) (measure of dispersion) and Mean Error (ME) (measure of bias), and square of the correlation coefficient between the observed and model predicted values (r^2) (approximates R^2) as well as the slope of the line between the observed and model predicted values. As a stricter cross-validation test which accounted for more of the spatial correlation of the dependent (pollution) data, we additionally conducted cross-validation of spatially independent folds (sometimes called strategic block cross-validation) (Roberts et al., 2017). We created a grid over the study area with 10 spatially independent grid cells. We then assigned each of the MM routes/fixed-monitoring site coordinate locations to their corresponding spatially overlapping grid cell (1–10) and additionally ran the cross-validation process with 10% of spatially independent sites held out ten times (10%_{SpatialCV}). These model evaluation procedures were also applied to the spatiotemporal models described below.

2.3.2. Spatiotemporal models (monthly models)

Lastly, we built spatiotemporal GAM LUR models for the prediction of monthly ambient concentrations of PM_{2.5}, BC, BrC, and PNC. The inputs for these models included monthly median concentrations from MM data collected over a 10-month period and spatial, temporal, and spatiotemporal predictor variables.

2.3.2.1. Spatial component. As a first step in the spatiotemporal model building process, we built up the spatial component of each model by first identifying the spatial GIS predictors used in the annual median models. If a spatial GIS predictor could also vary over time and we had data for that (e.g., monthly NDVI levels), we used the temporally varying version. We then fit models of monthly aggregated pollution concentrations and spatial GIS predictors and extracted the residuals for each site and month of the year. Furthermore, if the model residuals were spatially autocorrelated, we added in a smoothed spatial term.

2.3.2.2. Temporal component. As a second step, we identified the essential temporal components of the models by fitting models with the residuals from the spatial models developed in step one with monthly averaged meteorological data as predictors. This was done to identify which meteorological variables could account for variation in pollution concentrations, that was not accounted for by the spatial variables (Harper et al., 2021). Meteorological variables which were considered were temperature, relative humidity, wind speed, atmospheric pressure, and wind direction. The meteorological variables with significant p-values (<0.05) in the models were brought forward to the next step in the model building process, as was done in (Harper et al., 2021).

2.3.2.3. Spatiotemporal components. As a third and final step, we built up the full spatiotemporal models of monthly median pollution concentrations by adding in the selected spatial GIS and temporal meteorological variables as predictors. Additionally, to capture overall

monthly trends in pollution concentrations, we fitted a spline function to a continuous variable representing 'month of the year' in the models. Lastly, we hypothesized that the underlying spatial relationship between pollution concentrations was non-stationary and could vary over time, and so we conditioned the smoothed spatial term (interaction between longitude and latitude) on season ($n = 4$ seasons).

Model equations are included in Appendix 2.

2.3.3. Sensitivity analysis

A sensitivity analysis was conducted to determine the minimum number of 'n visits' required to attain stable models for spatial pattern of PM_{2.5}, PNC, particle size mode, BC and BrC. The annual models for these pollutants were rerun reflecting MM designs consisting of 8, 12, 16, 20, 24, 28, 32, and 36 visits (weekly MM sessions). For each of these 'n visit' categories, 15 random samples from the total 40 visits were taken. Boxplots presenting the adjusted r^2 variability of the 15 models for each 'n visits' group were plotted to investigate the relationship between the number of visits on model prediction. Further, model beta coefficients and spatial patterns were compared across 'n visit' categories to explore the number of visits required to attain stable models.

3. Results

3.1. Measured ambient concentrations

From the FSS campaign, annual NO₂ concentrations were measured across sites with a median value of 3.1 ppb (interquartile range (IQR): 1.9–4.9) and ozone of 54.1 ppb (IQR: 50.4–58.8). From the three MM routes, we measured annual median concentrations of corrected PM_{2.5} at 6.8 $\mu\text{g}/\text{m}^3$ (IQR: 6.1–7.7) with levels varying in magnitude across sites (Fig. 1) and between seasons (Figure A3.1). Of the 10 months captured, February had the highest median monthly concentration of PM_{2.5} (9.1 $\mu\text{g}/\text{m}^3$) and September the lowest (4.8 $\mu\text{g}/\text{m}^3$). BC and BrC had median measured annual concentrations at 1164 ng/m^3 (IQR: 1047–1349) and 881 ng/m^3 (IQR: 719–1064), which correspond to ~1.16 (BC) and 0.88 (BrC) in units of $\mu\text{g}/\text{m}^3$. Monthly median BC (1476 ng/m^3) and BrC (1148 ng/m^3) concentrations were similarly elevated in February (Figure A3.1).

Median total PNC were 8376 pt/cm^3 (IQR: 5790–10,999), while the specific size fractions with the highest median concentrations were within the range of 27.4–115.5 nm and the lowest concentrations were found for larger size fractions of 205.4–273.8 nm (Fig. 2). PNC varied seasonally, with the highest median monthly concentrations in January and February and lowest in March 2020 (Figure A3.1).

3.2. Model performance

The majority of models had relatively high or moderate-high predictive accuracy, even when predicting to new sites held out of the model training process. The predictive accuracy of the models, expressed as the r^2 , slope, RMSE, and the ME from 10-fold cross-validation of 10% random sites (10-fold 10%_{Site CV}), is shown for all models in Table 1. Models of annual PM_{2.5}, NO₂, PNC, and for six of the smallest PNC size fractions within the UFP range (11.5 nm–48.7 nm), had 10-fold 10%_{Site CV} r^2 which ranged from 0.70 to 0.86, while models of O₃, BC, and BrC had more moderate r^2 ranging from 0.50 to 0.54. Models of the larger PNC size fractions had worse 10-fold 10%_{Site CV} predictive accuracy compared with the smaller size fractions, with r^2 ranging from 0.29 to 0.56 for 64.9–154.0 nm size fractions. We were unable to successful model PNC size fractions larger than 205.4 nm. Almost all models tended to slightly over-estimate the mean predicted values compared with the measured, based on the ME statistics (Table 1). The spatiotemporal models of monthly median concentrations had moderate-high 10-fold 10%_{Site CV} predictive accuracy for PM_{2.5} (r^2 0.49; RMSE: 5.6 $\mu\text{g}/\text{m}^3$) and PNC (r^2 0.49; RMSE: 5245.8 pt/cm^3), while more moderate-low for BC and BrC (r^2 : 0.27–0.29). When undertaking

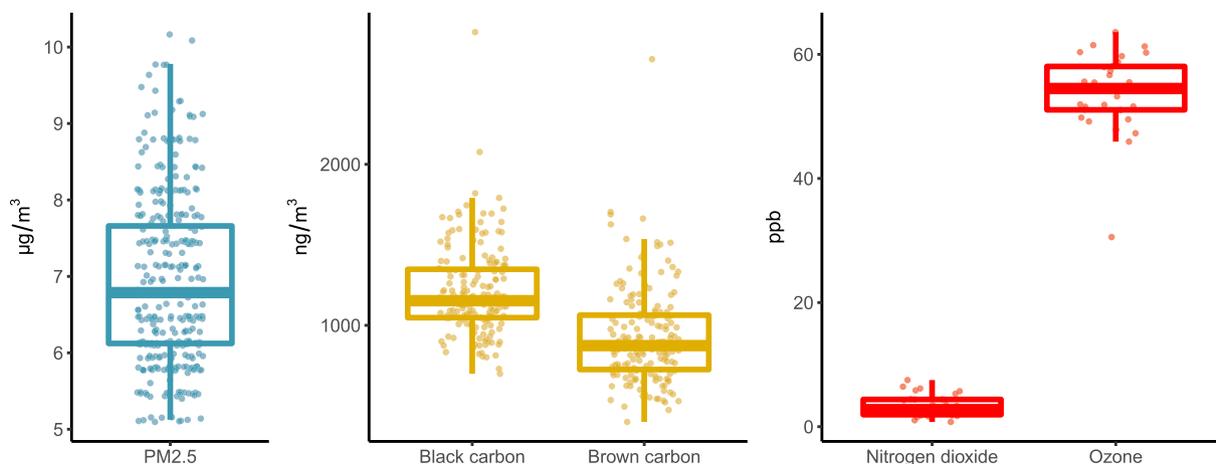


Fig. 1. Distributions of measured ambient annual concentrations of $PM_{2.5}$ ($\mu\text{g}/\text{m}^3$) (corrected), BC (ng/m^3), BrC (ng/m^3), nitrogen dioxide (NO_2 , ppb), and ozone (O_3 , ppb). Box plots show the median (centre line through box) and interquartile (IQR) ranges of the distribution (upper and lower bounds of the box ('hinges')) between measurement locations. Data beyond the end of the whiskers (i.e., "outlying" points) are not shown. Note the different units and minimum and maximum values between plots.

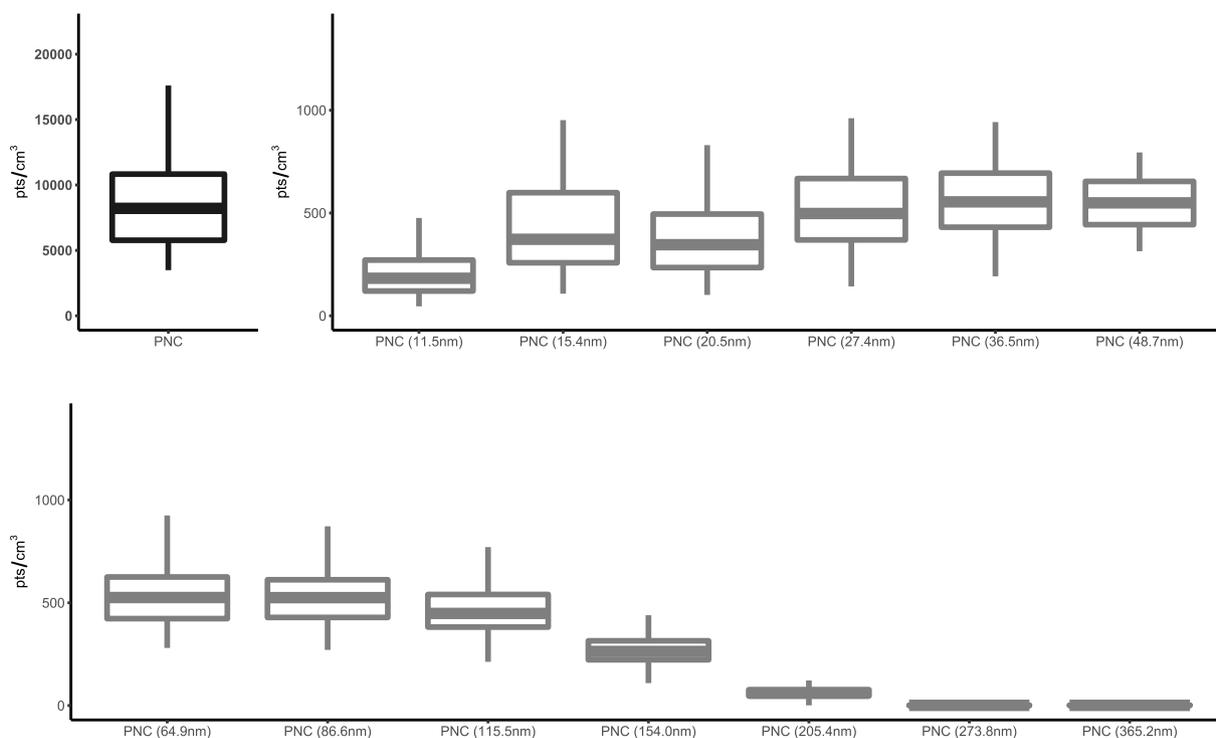


Fig. 2. Distributions of measured ambient annual concentrations of PNC for all particles and 13 particle size ranges from the mobile monitoring campaign. Box plots show the median (centre line through box) and interquartile (IQR) ranges of the distribution (upper and lower bounds of the box ('hinges')) between measurement locations. Data beyond the end of the whiskers (i.e., "outlying" points) are not shown. Note the different units and minimum and maximum values between plots and the total PNC was measured with the DiscMini and 13 particle size ranges with the Nanoscan.

the model building process, in almost all cases, adding splines with 3–5 knots to one or more continuous predictor variables to allow for non-linear associations, increased model predictive accuracy. Model tables are provided in the appendices and indicate which predictor variable associations were modelled as splines (Spatial models: Appendix 4; Spatiotemporal models: Appendix 5).

Considering that some degree of spatial autocorrelation was present in almost all of the air pollution data collected from the MM campaign (as would be expected based on the MM method), and in the subsequent models developed, we added in spatial terms based on coordinate location into those models to account for this effect. In addition to

increasing model predictive accuracy, this also removed most or all of the residual spatial autocorrelation present in the models (evaluated with Moran's I statistic). However, given the spatial nature of our input data, we also conducted an additional assessment of model predictive accuracy with 10-fold CV of 10% spatially independent sites (10-fold 10%_{Spatial} CV). The models with the highest r^2 from 10-fold 10%_{Spatial} CV were of annual median concentrations of $PM_{2.5}$ (0.77), PNC (0.75), particle size mode (0.63), and six of the smallest measured PNC size fractions 11.5 nm–36.5 nm (0.48–0.64), while r^2 for BC and BrC was fairly moderate at 0.34 (BC) – 0.40 (BrC). For the spatiotemporal models, r^2 from 10-fold 10%_{Spatial} CV was moderate for $PM_{2.5}$ (0.39),

Table 1

Evaluation of annual and monthly air pollution concentration model predictive accuracy with 10-fold cross validation holding out 10% of random sites.

Air pollutant (units)	Obs. (n)	N preds ^b	Measured concentration across sites (median)	r ²	10-fold 10% _{site} CV r ²	10-fold 10% _{site} CV slope of the line	10-fold 10% _{site} CV RMSE	10-fold 10% _{site} CV ME
Spatial (annual) models								
NO ₂ (ppb) ^a	60	8	3.1	0.84	0.76	1.02	1.06	0.13
O ₃ (ppb) ^a	59	6	54	0.65	0.54	0.77	6.65	0.09
BC (ng/m ³)	355	9	1164	0.57	0.50	0.93	186.52	11.73
BrC (ng/m ³)	355	10	881	0.61	0.54	0.76	192.92	14.69
PM _{2.5} (µg/m ³)	355	6	6.8	0.87	0.86	1.00	1.25	0.05
PNC (pts/cm ³) ^c	355	8	8376	0.88	0.81	1.00	1656.71	107.22
Particle size mode (nm) ^c	355	9	38	0.81	0.78	0.78	2.01	0.06
PNC size ranges^d								
PNC _{11.5nm} (pts/cm ³)	353	9	201	0.78	0.67	1.03	78.24	9.12
PNC _{15.4 nm} (pts/cm ³)	353	9	407	0.82	0.74	1.03	130.98	14.46
PNC _{20.5 nm} (pts/cm ³)	355	9	360	0.81	0.71	1.00	115.98	12.71
PNC _{27.4 nm} (pts/cm ³)	355	10	509	0.81	0.75	1.03	151.64	15.06
PNC _{36.5 nm} (pts/cm ³)	355	8	555	0.82	0.73	1.01	151.99	12.35
PNC _{48.7 nm} (pts/cm ³)	354	7	552	0.75	0.70	0.99	107.81	8.05
PNC _{64.9 nm} (pts/cm ³)	355	9	525	0.63	0.56	0.97	105.23	7.99
PNC _{86.6 nm} (pts/cm ³)	355	11	524	0.45	0.38	0.97	111.20	9.35
PNC _{115.5 nm} (pts/cm ³)	355	7	448	0.46	0.44	1.01	90.89	8.50
PNC _{154.0 nm} (pts/cm ³)	355	7	263	0.32	0.29	0.96	62.43	6.20
PNC _{205.4 nm} (pts/cm ³)	351	5	61	0.09	0.05	0.77	23.10	3.93
PNC _{273.8 nm} (pts/cm ³)	355	5	1	0.07	0.00	0.32	0.68	0.04
PNC _{365.2 nm} (pts/cm ³)	355	6	1	0.29	0.00	–	–	–
Spatiotemporal (monthly) models								
BC (ng/m ³)	3549	14	Range across months: 922–1476	0.31	0.27	0.97	450.00	67.61
BrC (ng/m ³)	3550	14	Range across months: 722–1148	0.39	0.29	0.92	462.77	78.77
PM _{2.5} (ug/m ³)	3538	11	Range across months: 4.8–9.1	0.63	0.49	0.86	5.61	0.71
PNC ^c (pts/cm ³)	3542	13	Range across months: 6096–13,336	0.62	0.49	0.97	5245.79	767.23

RMSE: Root mean squared error; ME: Mean Error; N preds: Number of predictors.

^a FSS campaign.^b Number of predictors include spatial (latitude and longitude) and temporal terms (e.g., month, season).^c Instrumentation (DiscMini from Testo (Lenzkirch, Germany), range 10 nm–700 nm).^d Instrumentation (Nanoscan from TSI Inc. (Minnesota, USA), range 10 nm–450 nm).

and PNC (0.40), but lower for BC (0.18). For BrC, the r^2 from 10%_{Spatial} CV was actually higher (0.32) than from 10-fold 10%_{Site} CV (0.29).

3.3. Spatial variations and predictors of ambient pollutant concentrations

Predicted concentrations of pollutants had significant spatial variability across the city (Figs. 3 and 4). Each prediction surface shown in Figs. 3 and 4 also present some of their key predictors such as highways, bus routes, rail lines and ports. The spatial variation in many of the pollutants' ambient concentrations was heavily influenced by the road-network in the city. NO₂ in particular had concentrations which were notably elevated along the city's bus routes. PM_{2.5}, BC, BrC, PNC, and PNC size fractions between 11.5 and 15.4 nm and 115.5 nm had elevated concentrations along the highway network. Ambient concentrations of BC and BrC were additionally significantly associated with the location of the bus network, and additionally proximity to gas stations for the smallest PNC size fractions within the UFP range. Proximity to the seaports, namely Quai 106 (south of the city) was a significant predictor of higher concentrations of annual BC and PM_{2.5}. Furthermore, aviation related variables had an influence in some models. Annual median concentrations of O₃ were predicted by altitude of avgas fuel, with lower altitudes corresponding to higher O₃.

The presence of locations of interest dotted around the city sometimes predicted ambient pollutant concentrations in local areas. In particular, the location of restaurants was a consistent predictor in the BC, BrC, PNC, and PNC 20.5–86.6 nm models. Furthermore, the spatial patterns of predicted mid-range UFP size fractions 20.5–86.6 nm appeared to be highly influenced by the spatial distribution of this variable, showing elevated concentrations in the city centre and commercial areas. The degree of urbanicity, greenness, and elevation had an influence in some of the models. Higher NDVI levels in the NO₂, PNC, and PNC 11.5–15.4 nm models for example was a strong predictor of lower ambient concentrations. Annual average concentrations of O₃ were predicted by elevation, with the highest concentrations appearing towards the outlying western and northern parts of the city.

The spatial patterning of the monthly predicted concentrations from the spatiotemporal models of PM_{2.5}, BC, BrC, and PNC was influenced by a combination of the GIS spatial predictor variables (which were shared with the annual median models) as well as the spatial terms (latitude x longitude) which were allowed to vary by season. As well, the temporally varying nature of some of the GIS variables, such as NDVI, could have an additional influence on changes to the spatial patterning of predicted concentrations throughout the year. The magnitude of the predicted concentrations in each month were further influenced by temporally varying meteorological variables as well as the month of the year. Examples of the monthly predicted maps can be seen in Figure A5.1.

3.4. 'n visits' sensitivity analysis

In the sensitivity analysis, for all pollutants (PM_{2.5}, PNC, particle size mode, BC, and BrC) the interquartile ranges of the adjusted r^2 values overlap with that of the 36-visit category at 16 visits or less, suggesting minimal improvement in model prediction with additional visits beyond 16. However, the subsequent examination of the impact of increasing visits on the stability of beta coefficients (Figures A6.2–6) and prediction surfaces (Figures A6.7–11) suggest more visits are required to achieve robust models. For PM_{2.5}, 32 visits were required for all predictors to achieve beta coefficients that were consistent across the 15 models and similar to those of the main 40-visit model (Figure A6.1). This is reflected in the examination of the changes in the spatial pattern of prediction surfaces with increasing 'n visits' (laps). Overall, to achieve stable models 32 visits were required for PM_{2.5} and BC, 28 for PNC and BrC, and 24 for particle size mode. Further detail on the results of the 'n visits' sensitivity analysis are presented in Appendix 6.

4. Discussion

In this study, we conducted a unique 10-month monitoring campaign designed to capture both temporal and spatial variations in particles within the ultrafine size range, BC, BrC, and other more commonly measured air pollutants (PM_{2.5}, NO₂, O₃). Leveraging these data, we developed 24 different land use regression models to predict high-resolution annual concentrations (spatial models), and for a subset of pollutants, monthly concentrations (spatiotemporal models) in Quebec City, Canada. Annual models of PM_{2.5}, PNC, PNC_{11.5}, PNC_{15.4}, PNC_{20.5}, PNC_{27.4}, PNC_{36.5}, PNC_{48.7}, particle size mode, and NO₂, had the highest out of sample predictive accuracy, with r^2 ranging between 0.67 and 0.86. The predicted pollutant concentrations had significant spatial variability across the city, and these spatial patterns sometimes differed between pollutants and pollutant size fractions. In the context of this city, the sensitivity analysis on the number of visits required to produce robust spatial annual models indicated that 32 visits were required for PM_{2.5} and BC, 28 for PNC and BrC, and 24 for the size mode.

There are several previous studies conducted mostly in Europe, North America, and Asia which have used MM data, or in combination with fixed-site sampling data, to build LUR predictive models of ambient pollutant concentrations. These models represent annual averages, seasonal averages, and for a few studies, average concentrations which represent different periods of a day, for PM_{2.5}, PM₁, UFP (and UFP size), inorganic species (SO₄, NO₃, NH₄), and organic aerosols, PNC, BC, NO, and NO₂ (Blanco et al., 2023; Hankey et al., 2019; Kerckhoffs et al., 2021, 2017; Lim et al., 2020; Liu et al., 2019; Messier et al., 2018; Minet et al., 2018; Robinson et al., 2019; Shairsingh et al., 2021; Simon et al., 2018; Van den Bossche et al., 2020; Van den Hove et al., 2020; Xu et al., 2021; Ye et al., 2020). Across these studies, LUR model R² (sometimes reported from cross-validation and other times not) ranged from 0.08 to 0.80. Comparison of model performance between studies can be challenging, due to studies being conducted in a variety of settings (rural-/urban/national), data collected from a range of number of locations and passes at each location, and studies which used a variety of predictor variables and different types of statistical models such as linear regression (Hankey et al., 2019; Kerckhoffs et al., 2021, 2017; Lim et al., 2020; Liu et al., 2019; Robinson et al., 2019; Shairsingh et al., 2021; Simon et al., 2018; Van den Hove et al., 2020; Xu et al., 2021; Ye et al., 2020), LASSO (Kerckhoffs et al., 2021), Generalized Additive Models (GAMs) (Lloyd et al., 2023), Random Forest (Kerckhoffs et al., 2021; Lim et al., 2020; Van den Hove et al., 2020), Kriging covariate models (Blanco et al., 2023; Messier et al., 2018), Support Vector Regression (Van den Bossche et al., 2020), deep Convolution Neural Network (CNN) (Lloyd et al., 2023), stacked ensemble combining multiple machine learning algorithms together (Lim et al., 2020), as well as combined models of GAMs and deep CNNs (Lloyd et al., 2023). Across this sample of MM LUR studies, we could not identify any clear indication of a certain type of statistical model outperforming the others, though worth mentioning that the stacked ensemble model in Lim et al. (2020) achieved a cross-validation R² of 0.80. For our study, we chose to use GAM models because they allow for flexible modelling of linear and/or non-linear associations between predictor and dependent variables and can flexibly account for spatial, temporal, and/or hierarchical structures within the data. A recent study by Lloyd et al. (2023) showed that combining GAM and deep CNN models trained on satellite-view images did not significantly improve over the GAM LUR models on their own.

Unlike most previous MM LUR studies, we accounted for spatial residual autocorrelation within our LUR models built with MM data by including smoothed spatial terms in our models, capturing the underlying influence of space on pollutant concentrations, following a similar approach as (Harper et al., 2021) for models developed for Chongqing, China. We additionally evaluated these models by randomly holding out 10% of sites in a spatially independent block pattern. This is because model accuracy evaluated from cross-validation where sites are randomly distributed across the study area may be over-inflated if data

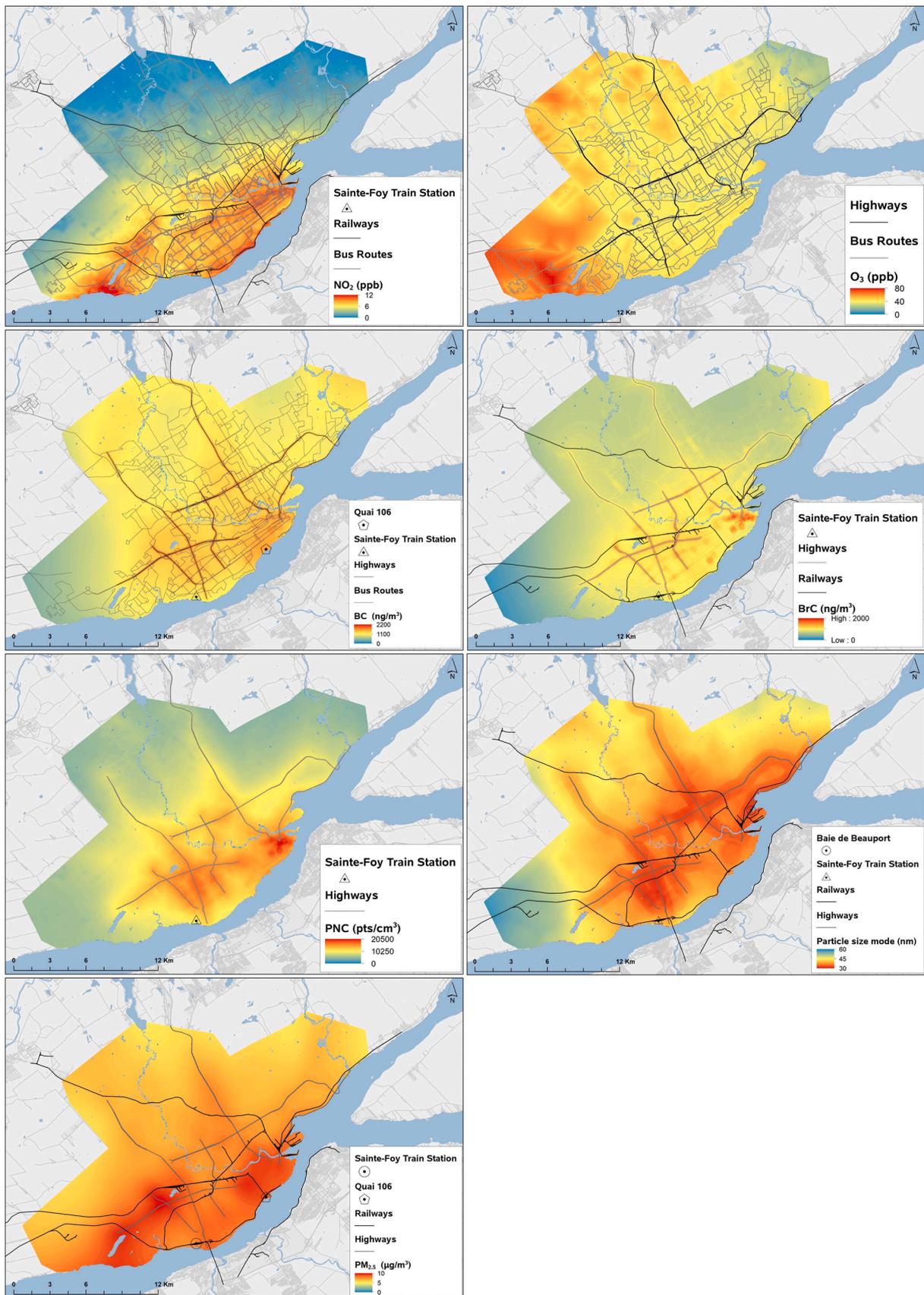


Fig. 3. Predictions from annual spatial models of NO₂, O₃, BC, BrC, total PNCs, particle size mode, and PM_{2.5}.

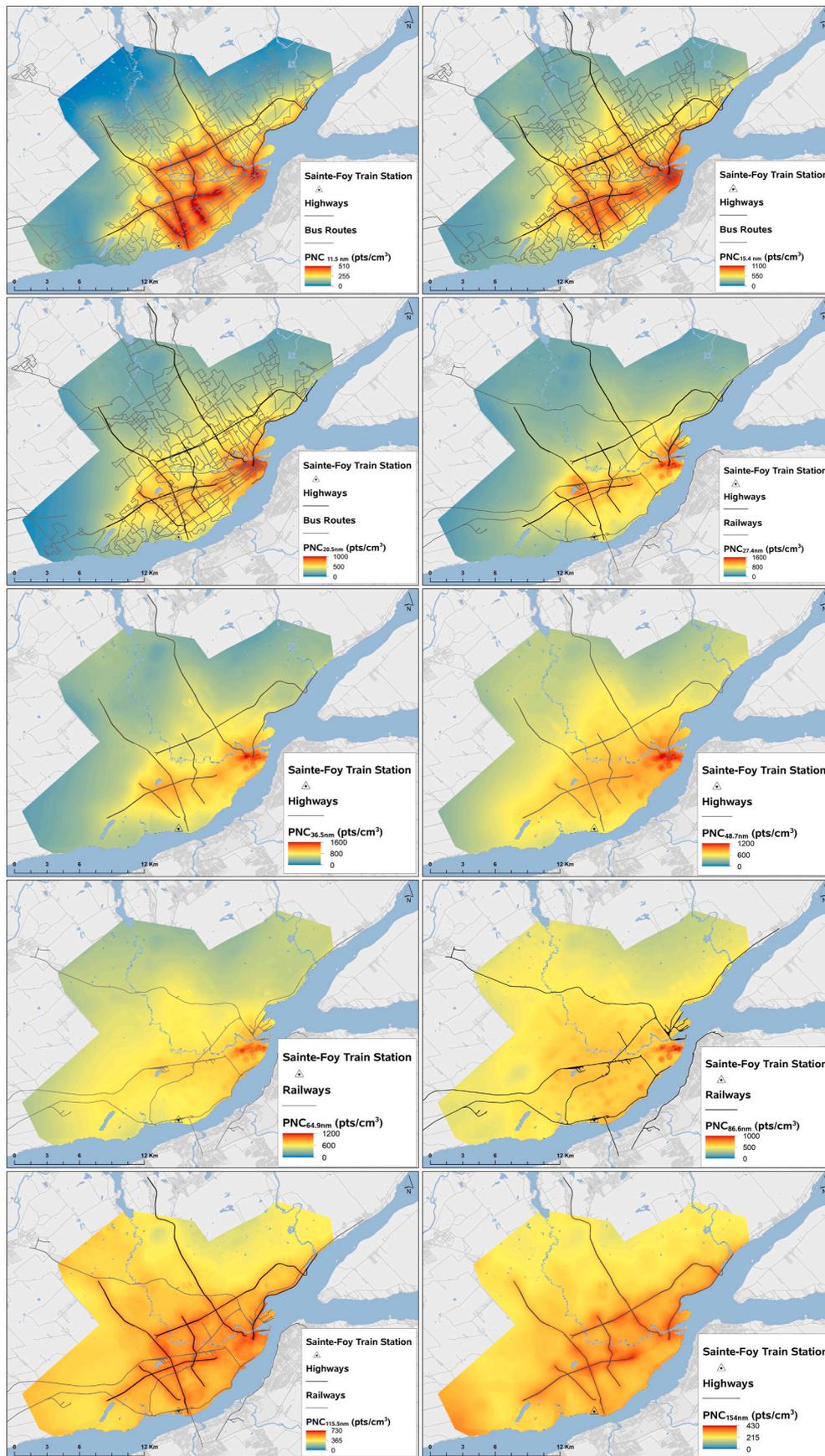


Fig. 4. Predictions from annual models of PNC size ranges.

are spatially autocorrelated (Harper et al., 2021). We found that the model accuracy statistics were not substantially different when evaluating with 10-fold spatially independent sites and 10-fold 10% random sites. Annual models of PM_{2.5}, PNC, particle size mode, and NO₂ (10-fold 10%site CV r^2 range: 0.76–0.86), as well as models of smaller UFP size fractions between PNC_{11.5} – PNC_{48.7} (10-fold 10%site CV r^2 range: 0.67–0.75), were all high performing, particularly when compared to previous LUR MM studies (see paragraph above for previous studies).

Our annual spatial model performance outperformed the predictive accuracy of the monthly spatiotemporal models. While the LUR spatiotemporal models included temporal (meteorological) predictor variables, temporal terms (month of the year), and allowed for the smoothed spatial terms to vary by season, the majority of predictors identified for the best set for the models were static in time (except for NDVI), and therefore unable to explain temporal variations in pollutant concentrations occurring across space. Furthermore, taking the median pollutant concentration across 10-months of the year (equal to 40 passes) provides for greater stability in the estimate that is more easily explained by the spatial covariates in the model. Previous MM LUR studies have been extended to develop spatiotemporal models to predict pollution concentrations for different days (Van den Bossche et al., 2020) or periods of the day (Hankey et al., 2019), as well as seasons of the year (Xu et al., 2021). While our spatiotemporal models had comparable performance to these previous studies, our out-of-sample predictive accuracy was slightly lower than what was found in the study by Xu et al., (2021) (PM_{2.5}, BC), indicating that further work is needed to improve the accuracy of these spatiotemporal models, such as including additional time-resolved spatial predictors (e.g., hourly traffic count surfaces).

The MM approach can increase the spatial resolution of data collection from the traditional FSS method. However, this comes at a cost to the temporal coverage for each data point. The data from FSS campaigns represent concentrations integrated over continuous two week-long periods covering 24-h of each day. MM data are far more limited in time, with each MM session providing only a ‘snapshot’ of the temporal variability of air pollution. This limitation is compensated for by conducting repeated MM sessions (or visits). The assumption inherent in each MM study is that the spatial pattern of the pollutant does not vary temporally beyond the chosen number of monitoring sessions (or visits to each location). Thus, the chosen number is sufficient to give a robust characterization of the spatial variability of air pollution for the study’s area of interest (city, country, etc.) and time frame (annual, seasonal, etc.). Our ‘n visits’ sensitivity analysis attempted to address this assumption for an annually representative spatial pattern of five pollutants in Quebec City, Canada representing the rush hour and post-rush hour period of the morning (8am–10am). The reliability of the assumption of a temporally invariant spatial pattern may vary between pollutant due to the seasonal and spatial nature of their sources. Our sensitivity analysis suggested PM_{2.5} and BC to have the most variable spatial pattern, therefore needing more visits than PNC, BrC, and particle size mode to obtain a robust model. These results were comparable to similar work conducted in Seattle, USA; Blanco et al. conducted a study in a 1200 km² area where the point of diminishing return and stable model results for PNC and BC was attained at 100–278 sites with 8–26 visits each. For PM_{2.5}, this point was reached at 150–278 sites with 16–24 visits each (Blanco et al., 2023). Hatzopoulou et al. (2017) conducted a sensitivity analysis investigating the effect of number of visits on the stability of beta coefficients based on a MM study covering three seasons in Montreal, Canada. The study concluded that between 10 and 12 visits for 150–200 sites were required per site for UFP models. Further, they demonstrated that widely different model surfaces can result from subsets of their main dataset both in number of sites and visits per site. In our sensitivity analyses, we demonstrated this as well. Most importantly, we observed that while similar adjusted r^2 values can be obtained with fewer visits, the spatial patterns of the prediction surfaces between these models can still differ.

Therefore, it is wise to examine beta coefficient stability and spatial pattern consistency when comparing models reflecting difference choices in MM methodology. Further, this growing field of work emphasizes the value added in including sensitivity analyses in MM-based air pollution modelling studies.

Strengths: A key strength of our study was our unique 10-month MM campaign designed to capture both temporal and spatial variation in UFPs and several other pollutants measured simultaneously. Our MM was conducted using a hydrogen fuel cell vehicle to avoid the monitoring vehicle emissions impacting our measurements. Other strengths include FSS and MM data collection design informed by location allocation modelling, our rigorous model building and evaluation approach, which produced city-wide estimates of air pollution concentrations covering both space and time, with fairly high predictive accuracy. Our modelling approach also explicitly dealt with potential and real spatial autocorrelation in the MM air pollution data, which is often present but rarely accounted for in air pollution LUR modelling studies. Our study also features the creation of spatial models for various particle sizes within the UFP size range, allowing us to investigate the spatially heterogeneous health effects of exposure to different UFP size fractions across the city in novel epidemiological studies.

Limitations: The MM component of our study involved routine collocation at a regulatory monitoring station to validate our monitoring methods for PM_{2.5}, PNC, BC, and BrC. We observed good agreement for PNC (slope = 0.77, R^2 = 0.92). For PM_{2.5}, our measurements over predicted (slope = 0.35), however, good linear agreement (R^2 = 0.88) allowed us to correct for the bias in our data. We had poor comparison for BC (R^2 = 0.28), which may result in an overall under or over-estimation for the city, though spatial variations (e.g., differences between locations) are not expected to be impacted, or only at random. While the MM campaign was seasonally balanced, it was not diurnally balanced and therefore our resulting annual surfaces based on MM data are only representative of exposures during the morning rush hour, and post-rush hour period (8–10 AM) when road-traffic flows are generally higher than other times of the day (Blanco et al., 2023). Our MM-based models also are subject to a weekday bias. As weekends can be distinctly lower in PM_{2.5} relative to weekday concentrations (Elansky et al., 2020; Tavella et al., 2023) our monitoring exclusive to weekdays may have resulted in a positive bias. Although measurements may not represent annual concentrations, documenting spatial contrasts when sources of pollutants peak remain pertinent for health. Also, in general, previous work has been shown that MM-based models can result in measured concentrations which are systematically biased high relative to short-term stationary monitoring design (Kerckhoffs et al., 2016). We also had limited FSS monitoring sites on the outer edges of our study area and in the more rural areas, particularly near the airport. The sparse FSS monitoring network in these areas may have led to a greater inaccuracy in the predictions in these areas as the sites may not capture a fully representative dataset of measurements, such as near the airport where both NO₂ and ozone were elevated. However, the model residuals in these areas were no higher or lower than in other, more urbanised and densely monitored, areas of the city, suggesting a more complex chemical relationship that annual LUR models are unable to capture.

5. Conclusions

In this study, we conducted a 10-month long MM and fixed-site campaign to measure particulate-based air pollutants (PM_{2.5}, BC, BrC, UFPs) and two gaseous pollutants (O₃ and NO₂) across Quebec City, Canada. Leveraging these data, we produced predictive statistical models of the spatial and spatiotemporal pattern of these air pollutants. Our annual models for NO₂, PM_{2.5}, and particles within the UFP size fraction range achieved cross-validated r^2 of 0.76–0.86. Models for O₃, BC and BrC had lower r^2 between 0.65 and 0.57. These exposure models will enable air pollution epidemiological research which will benefit from a common source of modelling and a range of UFP metrics to

advance the understanding of the health impacts of UFPs and other urban air pollutants.

Funding sources

This study was funded by the Addressing Air Pollution Horizontal Initiative of the Government of Canada and the City of Quebec *Mon Environment Ma Sante* program of the City of Quebec's Capitale-Nationale Public Health department.

CRedit authorship contribution statement

Sierra Nicole Clark: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ryan Kulka:** Writing – review & editing, Validation, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Stephane Buteau:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Eric Lavigne:** Writing – review & editing, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Joyce J.Y. Zhang:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Christian Riel-Roberge:** Writing – review & editing, Validation, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation. **Audrey Smargiassi:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Scott Weichenthal:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Keith Van Ryswyk:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This study was funded by the Addressing Air Pollution Horizontal Initiative of the Government of Canada and the *Mon Environment Ma Sante* program of the City of Quebec's Capitale-Nationale Public Health department. The authors would like to acknowledge the contribution of Liu Sun for location allocation modelling for site selection, Katarina Kunarac for GIS support, and Phillip Pham for analytical support. S.N.C. conducted the modelling while working for Health Canada but co-drafted the manuscript while at St George's, University of London.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envpol.2024.124353>.

References

Anand, A., Phuleria, H.C., 2021. Spatial and seasonal variation of outdoor BC and PM_{2.5} in densely populated urban slums. *Environ. Sci. Pollut. Res.* 28, 1397–1408. <https://doi.org/10.1007/s11356-020-10564-y>.

- Bennett, J.E., Tamura-Wicks, H., Parks, R.M., Burnett, R.T., Pope III, C.A., Bechle, M.J., Marshall, J.D., Danaei, G., Ezzati, M., 2019. National and county life expectancy loss from particulate matter pollution in the USA. *PLoS Med.* 1–18.
- Blanco, M.N., Bi, J., Austin, E., Larson, T.V., Marshall, J.D., Sheppard, L., 2023. Impact of mobile monitoring network design on air pollution exposure assessment models. *Environ. Sci. Technol.* 57, 440–450. <https://doi.org/10.1021/acs.est.2c05338>.
- Boogaard, H., Patton, A.P., Atkinson, R.W., Brook, J.R., Chang, H.H., Crouse, D.L., Fussell, J.C., Hoek, G., Hoffmann, B., Kappeler, R., Kutlar Joss, M., Ondras, M., Sagiv, S.K., Samoli, E., Shaikh, R., Smargiassi, A., Szpiro, A.A., Van Vliet, E.D.S., Vienneau, D., Weuve, J., Lurmann, F.W., Forastiere, F., 2022. Long-term exposure to traffic-related air pollution and selected health outcomes: a systematic review and meta-analysis. *Environ. Int.* 164, 107262 <https://doi.org/10.1016/j.envint.2022.107262>.
- Buteau, S., Goldberg, M.S., Burnett, R.T., Gasparrini, A., Valois, M.F., Brophy, J.M., Crouse, D.L., Hatzopoulou, M., 2018. Associations between ambient air pollution and daily mortality in a cohort of congestive heart failure: case-crossover and nested case-control analyses using a distributed lag nonlinear model. *Environ. Int.* 113, 313–324. <https://doi.org/10.1016/j.envint.2018.01.003>.
- Buteau, S., Shekarrizfard, M., Hatzopoulou, M., Gamache, P., Liu, L., Smargiassi, A., 2020. Air pollution from industries and asthma onset in childhood: a population-based birth cohort study using dispersion modeling. *Environ. Res.* 185, 109180 <https://doi.org/10.1016/j.envres.2020.109180>.
- Chambliss, S.E., Preble, C.V., Caubel, J.J., Cados, T., Messier, K.P., Alvarez, R.A., Lafranchi, B., Lunden, M., Marshall, J.D., Szpiro, A.A., Kirchstetter, T.W., Apte, J.S., 2020. Comparison of mobile and fixed-site black carbon measurements for high-resolution urban pollution mapping. *Environ. Sci. Technol.* 54, 7848–7857. <https://doi.org/10.1021/acs.est.0c01409>.
- COMEAP, 2022. *Cognitive Decline, Dementia and Air Pollution: a Report by the Committee on the Medical Effects of Air Pollutants*. London, UK.
- Crouse, D.L., Goldberg, M.S., Ross, N.A., 2009. A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal, Canada. *Atmos. Environ.* 43 (32), 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>.
- Crouse, D.L., Goldberg, M.S., Ross, N.A., Chen, H., Labrèche, F., 2010. Postmenopausal breast cancer is associated with exposure to traffic-related air pollution in Montreal, Canada: a case-control study. *Environ. Health Perspect.* 118, 1578–1583. <https://doi.org/10.1289/ehp.1002221>.
- Deshmukh, P., Kimbrough, S., Krabbe, S., Logan, R., Isakov, V., Baldauf, R., 2020. Identifying air pollution source impacts in urban communities using mobile monitoring. *Sci. Total Environ.* 715, 136979 <https://doi.org/10.1016/j.scitotenv.2020.136979>.
- Dionisio, K., Rooney, M.S., Arku, R.E., Friedman, A.B., Hughes, A.F., Vallarino, J., Agyei-mensah, S., Spengler, J.D., Ezzati, M., 2010. Within-neighborhood patterns and sources of particle pollution: mobile monitoring and geographic information system analysis in four communities in Accra, Ghana. *Environ. Health Perspect.* 607–613. <https://doi.org/10.1289/ehp.0901365>.
- Dormann, C.F., McPherson J B, M., Araujo, M., Bivand, R., Bolliger, J., Carl, G., DaviesR, G., Hirtzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, R., Reineking, B., Schröder, B., Schurr, M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30, 609–628. <https://doi.org/10.1111/j.2007.0906-7590.05171.x>.
- Elansky, N.F., Shilkin, A.V., Ponomarev, N.A., Semutnikova, E.G., Zakharova, P.V., 2020. Weekly patterns and weekend effects of air pollution in the Moscow megacity. *Atmos. Environ.* 224 (January), 117303 <https://doi.org/10.1016/j.atmosenv.2020.117303>.
- GBD 2019 Diabetes and Air Pollution Collaborators, 2022. Estimates, trends, and drivers of the global burden of type 2 diabetes attributable to PM_{2.5} air pollution, 1990–2019: an analysis of data from the Global Burden of Disease Study 2019. *Lancet Planet. Health* 6, e586–e600. [https://doi.org/10.1016/S2542-5196\(22\)00122-X](https://doi.org/10.1016/S2542-5196(22)00122-X).
- Goldberg, M.S., Burnett, R.T., Bailar, J.C., Brook, J., Bonvalot, Y., Tamblyn, R., Singh, R., Valois, M.F., 2001. The association between daily mortality and ambient air particle pollution in Montreal, Quebec. 1. Nonaccidental mortality. *Environ. Res.* 86, 12–25. <https://doi.org/10.1006/enrs.2001.4242>.
- Goldberg, M.S., Burnett, R.T., Yale, J.F., Valois, M.F., Brook, J.R., 2006. Associations between ambient air pollution and daily mortality among persons with diabetes and cardiovascular disease. *Environ. Res.* 100, 255–267. <https://doi.org/10.1016/j.envres.2005.04.007>.
- Goldberg, M.S., Burnett, R.T., Stieb, D.M., Brophy, J.M., Daskalopoulou, S.S., Valois, M.F., Brook, J.R., 2013. Associations between ambient air pollution and daily mortality among elderly persons in Montreal, Quebec. *Sci. Total Environ.* 463–464, 931–942. <https://doi.org/10.1016/j.scitotenv.2013.06.095>.
- Gourdji, S., 2018. Review of plants to mitigate particulate matter, ozone as well as nitrogen dioxide air pollutants and applicable recommendations for green roofs in Montreal, Quebec. *Environ. Pollut.* 241, 378–387. <https://doi.org/10.1016/j.envpol.2018.05.053>.
- Government of Canada, 2022. *Canadian Climate Normals 1981–2010 Station Data [WWW Document]*.
- 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. *Environ. Sci. Technol.* 49, 9194–9202. <https://doi.org/10.1021/acs.est.5b01209>.
- Hankey, S., Sforza, P., Pierson, M., 2019. Using mobile monitoring to develop hourly empirical models of particulate air pollution in a rural Appalachian community. *Environ. Sci. Technol.* 53, 4305–4315. <https://doi.org/10.1021/acs.est.8b05249>.

- Harper, A., Baker, P.N., Xia, Y., Kuang, T., Zhang, H., Chen, Y., Han, T.L., Gulliver, J., 2021. Development of spatiotemporal land use regression models for PM_{2.5} and NO₂ in Chongqing, China, and exposure assessment for the CLIMB study. *Atmos. Pollut. Res.* 12, 101096 <https://doi.org/10.1016/j.apr.2021.101096>.
- Hatzopoulou, M., Valois, M.F., Levy, I., Mihele, C., Lu, G., Bagg, S., Minet, L., Brook, J., 2017. Robustness of land-use regression models developed from mobile air pollutant measurements. *Environ. Sci. Technol.* 51 (7), 3938–3947.
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* 42, 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>.
- Institut national de santé publique du Québec, 2023. Portrait de la pollution de l'air à Québec et de certains de ses impacts sur la santé des résidents des territoires des clsc de Limoilou-Vanier et de Québec-basse-ville: INSPQ.
- Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., Brook, J.R., 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. *Atmos. Environ.* 39, 2399–2409. <https://doi.org/10.1016/j.atmosenv.2004.06.049>.
- Kerckhoffs, J., Hoek, G., Messier, K.P., Brunekreef, B., Meliefste, K., Klompmaker, J.O., Vermeulen, R., 2016. Comparison of ultrafine particle and black carbon concentration predictions from a mobile and short-term stationary land-use regression model. *Environ. Sci. Technol.* 50, 12894–12902. <https://doi.org/10.1021/acs.est.6b03476>.
- Kerckhoffs, J., Hoek, G., Vlaanderen, J., van Nunen, E., Messier, K., Brunekreef, B., Gulliver, J., Vermeulen, R., 2017. Robustness of intra urban land-use regression models for ultrafine particles and black carbon based on mobile monitoring. *Environ. Res.* 159, 500–508. <https://doi.org/10.1016/j.envres.2017.08.040>.
- Kerckhoffs, J., Hoek, G., Gehring, U., Vermeulen, R., 2021. Modelling nationwide spatial variation of ultrafine particles based on mobile monitoring. *Environ. Int.* 154, 106569 <https://doi.org/10.1016/j.envint.2021.106569>.
- Kerckhoffs, J., Hoek, G., Vermeulen, R., 2024. Mobile monitoring of air pollutants; performance evaluation of a mixed-model land use regression framework in relation to the number of drive days. *Environ. Res.* 240, 117457 <https://doi.org/10.1016/j.envres.2023.117457>.
- Khairbek, I., Haney, J., Douglas, S., Ito, K., Matte, T., 2016. The contribution of motor vehicle emissions to ambient fine particulate matter public health impacts in New York City: a health burden assessment. *Environ. Heal. A Glob. Access Sci. Source* 15. <https://doi.org/10.1186/s12940-016-0172-6>.
- Lim, C.C., Kim, H., Vilcassim, M.J.R., Thurston, G.D., Gordon, T., Chen, L., Lee, K., Heim binder, M., Kim, S., 2020. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea. *Environ. Int.* 131, 1–26. <https://doi.org/10.1016/j.envint.2019.105022>.
- Liu, M., Huang, Y., Ma, Z., Jin, Z., Liu, X., Wang, H., Liu, Y., Wang, J., Jantunen, M., Bi, J., Kinney, P.L., 2017. Spatial and temporal trends in the mortality burden of air pollution in China: 2004–2012. *Environ. Int.* 98, 75–81. <https://doi.org/10.1016/j.envint.2016.10.003>.
- Liu, M., Peng, X., Meng, Z., Zhou, T., Long, L., She, Q., 2019. Spatial characteristics and determinants of in-traffic black carbon in Shanghai, China: combination of mobile monitoring and land use regression model. *Sci. Total Environ.* 658, 51–61. <https://doi.org/10.1016/j.scitotenv.2018.12.135>.
- Lloyd, M., Ganji, A., Xu, J., Venuta, A., Simon, L., Zhang, M., Saeedi, M., Yamanouchi, S., Apte, J., Hatzopoulou, M., Weichenthal, S., 2023. Predicting spatial variations in annual average outdoor ultrafine particle concentrations in Montreal and Toronto, Canada: integrating land use regression and deep learning models. *Environ. Int.* <https://doi.org/10.1016/j.envint.2023.108106>. The Author(s).
- Messier, K.P., Chambliss, S.E., Gani, S., Alvarez, R., Brauer, M., Choi, J.J., Hamburg, S.P., Kerckhoffs, J., Lafranchi, B., Lunden, M.M., Marshall, J.D., Portier, C.J., Roy, A., Szpiro, A.A., Vermeulen, R.C.H., Apte, J.S., 2018. Mapping air pollution with google street view cars: efficient approaches with mobile monitoring and land use regression. *Environ. Sci. Technol.* 52, 12563–12572. <https://doi.org/10.1021/acs.est.8b03395>.
- Minet, L., Liu, R., Valois, M.F., Xu, J., Weichenthal, S., Hatzopoulou, M., 2018. Development and comparison of air pollution exposure surfaces derived from on-road mobile monitoring and short-term stationary sidewalk measurements. *Environ. Sci. Technol.* 52, 3512–3519. <https://doi.org/10.1021/acs.est.7b05059>.
- Parent, M.É., Goldberg, M.S., Crouse, D.L., Ross, N.A., Chen, H., Valois, M.F., Liataud, A., 2013. Traffic-related air pollution and prostate cancer risk: a case-control study in Montreal Canada. *Occup. Environ. Med.* 70, 511–518. <https://doi.org/10.1136/oemed-2012-101211>.
- Pascal, M., Corso, M., Chanel, O., Declercq, C., Badaloni, C., Cesaroni, G., Henschel, S., Meister, K., Haluza, D., Martin-Olmedo, P., Medina, S., 2013. Assessing the public health impacts of urban air pollution in 25 European cities: results of the Apekom project. *Sci. Total Environ.* 449, 390–400. <https://doi.org/10.1016/j.scitotenv.2013.01.077>.
- Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guiller-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40, 913–929. <https://doi.org/10.1111/ecog.02881>.
- Robinson, E.S., Shah, R.U., Messier, K., Gu, P., Li, H.Z., Apte, J.S., Robinson, A.L., Presto, A.A., 2019. Land-use regression modeling of source-resolved fine particulate matter components from mobile sampling. *Environ. Sci. Technol.* 53, 8925–8937. <https://doi.org/10.1021/acs.est.9b01897>.
- Ross, N., n.d. GAMS N R: 3 - Spatial GAMS and Interactions [WWW Document].
- Shairsingh, K.K., Brook, J.R., Mihele, C.M., Evans, G.J., 2021. Characterizing long-term NO₂ concentration surfaces across a large metropolitan area through spatiotemporal land use regression modelling of mobile measurements. *Environ. Res.* 196, 111010 <https://doi.org/10.1016/j.envres.2021.111010>.
- Simon, M.C., Patton, A.P., Naumova, E.N., Levy, J.I., Kumar, P., Brugge, D., Durant, J.L., 2018. Combining measurements from mobile monitoring and a reference site to develop models of ambient ultrafine particle number concentration at residences. *Environ. Sci. Technol.* 52, 6985–6995. <https://doi.org/10.1021/acs.est.8b00292>.
- Smargiassi, A., Sidi, E.A.L., Robert, L.E., Plante, C., Haddad, M., Gamache, P., Burnett, R., Goudreau, S., Liu, L., Fournier, M., Pelletier, E., Yankoty, I., 2020. Exposure to ambient air pollutants and the onset of dementia in Québec, Canada. *Environ. Res.* 190 <https://doi.org/10.1016/j.envres.2020.109870>.
- Statistics Canada, 2017. Québec, V [Census subdivision], Quebec and Canada [Country] (table). *Census Profile. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. Released November 29, 2017.* <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E> (accessed June 14, 2023).
- Stieb, D.M., Evans, G.J., To, T.M., Lakey, P.S.J., Shiraiwa, M., Hatzopoulou, M., Minet, L., Brook, J.R., Burnett, R.T., Weichenthal, S.A., 2021. Within-City variation in reactive oxygen species from fine particle air pollution and COVID-19. *Am. J. Respir. Crit. Care Med.* 204, 168–177. <https://doi.org/10.1164/rccm.202011-41420C>.
- Tavella, R.A., Galeao da Rosa Moraes, N., Maciel Aick, C.D., Ramires, P.F., Pereira, N., Soares, A.G., da Silva Júnior, F.M.R., 2023. Weekend effect of air pollutants in small and medium-sized cities: the role of policies stringency to COVID-19 containment. *Atmos. Pollut. Res.* 14 (2) <https://doi.org/10.1016/j.apr.2023.101662>.
- Tétreault, L.F., Doucet, M., Gamache, P., Fournier, M., Brand, A., Kosatsky, T., Smargiassi, A., 2016. Childhood exposure to ambient air pollutants and the onset of asthma: an administrative cohort study in Québec. *Environ. Health Perspect.* 124, 1276–1282. <https://doi.org/10.1289/ehp.1509838>.
- Turner, M.C., Andersen, Z.J., Baccarelli, A., Diver, W.R., Gapstur, S.M., Pope, C.A., Prada, D., Samet, J., Thurston, G., Cohen, A., 2020. Outdoor air pollution and cancer: an overview of the current evidence and public health recommendations. *CA A Cancer J. Clin.* 70, 460–479. <https://doi.org/10.3322/caac.21632>.
- Van den Bossche, J., De Baets, B., Botteldooren, D., Theunis, J., 2020. A spatio-temporal land use regression model to assess street-level exposure to black carbon. *Environ. Model. Software* 133, 104837. <https://doi.org/10.1016/j.envsoft.2020.104837>.
- Van den Hove, A., Verwaeren, J., Van den Bossche, J., Theunis, J., De Baets, B., 2020. Development of a land use regression model for black carbon using mobile monitoring data and its application to pollution-avoiding routing. *Environ. Res.* 183, 108619 <https://doi.org/10.1016/j.envres.2019.108619>.
- Weichenthal, S., Van Ryswyk, K., Goldstein, A., Bagg, S., Shekarrizfard, M., Hatzopoulou, M., 2016a. A land use regression model for ambient ultrafine particles in Montreal, Canada: a comparison of linear regression and a machine learning approach. *Environ. Res.* 146, 65–72. <https://doi.org/10.1016/j.envres.2015.12.016>.
- Weichenthal, S., Van Ryswyk, K., Goldstein, A., Shekarrizfard, M., Hatzopoulou, M., 2016b. Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: a land use regression model. *Environ. Pollut.* 208, 241–248. <https://doi.org/10.1016/j.envpol.2015.04.011>.
- World Health Organization, 2021. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide: annex 3. Summaries of systematic reviews of evidence informing the air quality guideline levels [WWW Document]. URL. <https://www.ncbi.nlm.nih.gov/books/NBK574588/>.
- Xie, X., Semanski, I., Gautama, S., Tsiligianni, E., Deligiannis, N., Rajan, R.T., Pasveer, F., Phillips, W., 2017. A review of urban air pollution monitoring and exposure assessment methods. *ISPRS Int. J. Geo-Inf.* 6, 1–21. <https://doi.org/10.3390/ijgi6120389>.
- Xu, X., Qin, N., Qi, L., Zou, B., Cao, S., Zhang, K., Yang, Z., Liu, Y., Zhang, Y., Duan, X., 2021. Development of season-dependent land use regression models to estimate BC and PM₁₀ exposure. *Sci. Total Environ.* 793, 148540 <https://doi.org/10.1016/j.scitotenv.2021.148540>.
- Ye, Q., Li, H.Z., Gu, P., Robinson, E.S., Apte, J.S., Sullivan, R.C., Robinson, A.L., Donahue, N.M., Presto, A.A., 2020. Moving beyond fine particle mass: high-spatial resolution exposure to source-resolved atmospheric particle number and chemical mixing state. *Environ. Health Perspect.* 128, 1–10. <https://doi.org/10.1289/EHP5311>.