Original Research Article

ISEE ENVIRONMENTAL EPIDEMIOLOGY

OPEN

Predicting within-city spatiotemporal variations in daily median outdoor ultrafine particle number concentrations and size in Montreal and Toronto, Canada

Alessya Venuta[a], Marshall Lloyd[a], Arman Ganji[b], Junshi Xu[b], Leora Simon[a], Mingqian Zhang[b], Milad Saeedi[b], Shoma Yamanouchi[b], Eric Lavigne[c], Marianne Hatzopoulou[b], Scott Weichenthal[a,d,*]

Background: Epidemiological evidence suggests that long-term exposure to outdoor ultrafine particles (UFPs, <0.1 μm) may have important human health impacts. However, less is known about the acute health impacts of these pollutants as few models are available to estimate daily within-city spatiotemporal variations in outdoor UFPs.

Methods: Several machine learning approaches (i.e., generalized additive models, random forest models, and extreme gradient boosting) were used to predict daily spatiotemporal variations in outdoor UFPs (number concentration and size) across Montreal and Toronto, Canada using a large database of mobile monitoring measurements. Separate models were developed for each city and all models were evaluated using a 10-fold cross-validation procedure.

Results: In total, our models were based on measurements from 12,705 road segments in Montreal and 10,929 road segments in Toronto. Daily median outdoor UFP number concentrations varied substantially across both cities with 1st–99th percentiles ranging from 1389 to 181,672 in Montreal and 2472 to 118,544 in Toronto. Outdoor UFP size tended to be smaller in Montreal (mean [SD]: 34 nm [15]) than in Toronto (mean [SD]: 44 nm [25]). Extreme gradient boosting models performed best and explained the majority of spatiotemporal variations in outdoor UFP number concentrations (Montreal, R^2: 0.727; Toronto, R^2: 0.723) and UFP size (Montreal, R^2: 0.823; Toronto, R^2: 0.898) with slopes close to one and intercepts close to zero for relationships between measured and predicted values.

Conclusion: These new models will be applied in future epidemiological studies examining the acute health impacts of outdoor UFPs in Canada's two largest cities.

Keywords: ultrafine particles, exposure model, machine learning

[a]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada; [b]Department of Civil and Mineral Engineering, University of Toronto, Toronto, Canada; [c]Environmental Health Science Research Bureau, Health Canada, Ottawa, Canada; and [d]Air Health Science Division, Health Canada, Ottawa, Canada

The research described in this article was conducted in part under a contract to the Health Effects Institute (HEI), an organization jointly funded by the United States Environmental Protection Agency (EPA) and certain motor vehicle and engine manufacturers. The contents of this article do not necessarily reflect the views of HEI or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers. This study also received funding from Health Canada.

The data used to develop our models is available upon reasonable request.

SDC *Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.environepidem.com).*

**Corresponding Author. Address: Department of Epidemiology, Biostatistics, and Occupational Health, School of Population and Global Health, McGill University, 2586 St. Zotique East, Montreal, QC H1Y 1C7, Canada. E-mail: scott.weichenthal@mcgill.ca (S. Weichenthal).*

Environmental Epidemiology (2024) 8:e323

Received 15 March, 2024; Accepted 17 June, 2024

Published online 22 July 2024

DOI: 10.1097/EE9.0000000000000323

Introduction

Outdoor air pollution is a global public health concern and contributes to millions of deaths around the world each year.[1] While fine particulate matter ($PM_{2.5}$, <2.5 μm) is known to contribute to cardiovascular and respiratory morbidity and mortality, less is known about other outdoor air pollutants such as ultrafine particles (UFP, <0.1 μm) that can be present in high number concentrations in urban environments and vary greatly across both space and time.[2–4] While existing evidence suggests potentially important health impacts of long-term exposures to outdoor UFPs,[5–8] less is known about the acute health impacts of these pollutants given the relative absence of high-resolution exposure models capable of estimating daily spatiotemporal variations across major urban centers. Indeed, most existing studies of short-term exposures to outdoor UFPs have relied on a small number of fixed-site monitors

What this study adds:
Epidemiological evidence to date suggests that long-term exposures to outdoor UFPs may contribute to increased risks of adverse cardiovascular and respiratory outcomes. However, less is known about the acute health impacts of outdoor UFPs as few models have been developed to estimate high-resolution within-city spatiotemporal variations in these pollutants. In this study, we developed and evaluated new models to predict daily outdoor UFP number concentrations and UFP size in Canada's two largest cities using a unique database of mobile measurements. These models will be applied in future epidemiological analyses to examine the acute health impacts of these pollutants.

that cannot capture high-resolution spatiotemporal variations and have produced inconsistent results.[9–12]

Over the past decade, mobile monitoring has been used to collect data for outdoor UFP number concentrations for the purpose of developing exposure models for use in epidemiological studies.[4,13,14] However, many previous exposure modeling studies have important limitations with regard to the spatial domain covered and/or the duration and time periods captured by mobile monitoring (e.g., focusing only on rush hour periods).[13,14] Moreover, existing studies primarily focus on modeling spatial variations in annual average outdoor UFP concentrations with little attention paid to modeling daily spatiotemporal variations. In this study, we developed new high-resolution exposure models for daily spatiotemporal variations in outdoor UFP number concentrations and UFP size in Canada's two largest cities (Montreal and Toronto) using a unique mobile monitoring database covering all days of the week, months of the year, and most times of the day.[2] We examined several different machine learning models previously used to estimate spatiotemporal variations in environmental pollutants[15] and our final models are available to support future epidemiological analyses examining the acute health impacts of outdoor UFPs.

Methods

Mobile monitoring data

This study was conducted in Montreal and Toronto, Canada, which have populations of 1.9 million and 2.9 million, respectively. Data collection took place on the island of Montreal (431.5 km^2) and within the formal city boundaries of Toronto (630.2 km^2). The mobile monitoring campaign designed to collect outdoor UFP data is described in detail elsewhere.[2] Briefly, mobile monitoring was conducted on predefined routes in both cities using gasoline vehicles equipped with either a naneos Partector 2 nanoparticle dosimeter (Montreal) or a Testo DiSCmini handheld nanoparticle counter (Toronto). These devices each sampled UFP number concentrations (particles/cm^3) and mean UFP size (nm) at 1-second intervals. Each instrument has an internal pump and the sampling tube for each instrument was positioned out the rear passenger window of the vehicle used for mobile monitoring. Both instruments operate using the same underlying measurement principle and all instruments were factory calibrated prior to field data collection and were zero checked weekly to verify proper instrument function. All UFP measurements were time-synchronized to a Geographic Positioning System monitor that recorded the latitude and longitude of the vehicles during mobile monitoring. Predefined monitoring routes were selected using a clustering algorithm designed to capture the range of land use and traffic characteristics present across each city.[2] These routes were monitored over a 14-month period between June 2020 and August 2021. Additional mobile monitoring data were also collected in Toronto during summer 2023 (May–August) using the same routes and study design. Monitoring was carried out from Monday through Sunday (i.e., all days of the week were sampled over the study period), between 7 am and 11 pm. All monitoring routes were randomized by route location, day of the week, and time of day to balance weather conditions across routes.

Land use and meteorological variables

Information on land use and traffic data (including distance to railways, length of major roads/area of land use types within a buffer [100 m, 200 m, or 300 m], number of bus routes or bus stops within a buffer, etc.) was extracted for each road segment along monitoring routes using ArcMap 10.8.1 (ESRI, Redlands, CA). These data were obtained from DMTI Spatial (Richmond Hill, CA), EMee (INRO, Montreal, Canada), City of Montreal, City of Toronto, Canadian National Pollution Release Inventory, Statistics Canada, Toronto Transit Commission, and Société de Transport de Montreal. Night light levels[16] (higher traffic areas tend to have more light at night) and land coverage[17] (urban and trees) data were obtained from satellite data using Google Earth Engine. Using an approach from a previous air pollution modeling study,[3] the number of bird species spotted in a given area was obtained from a Cornell Lab of Ornithology database (areas with greater human activity typically have greater air pollution and may have fewer distinct species of birds).[18–20] Additionally, meteorological variables (hourly averages) during mobile monitoring were obtained from central monitoring stations located in each city including temperature (°C), wind speed (m/s), relative humidity (%), atmospheric pressure (kPa), and precipitation (mm). Outdoor $PM_{2.5}$ concentrations (μg/m^3) were also compiled from fixed-site monitors in each city during mobile monitoring (using the closest hourly average to the time of mobile monitoring) to capture potential temporal patterns in regional air pollution episodes that could influence outdoor UFPs on a broad spatial scale.

Data processing

All monitoring routes were divided into 100 m road segments and median UFP measurements (i.e., number concentration and size) were aggregated to the centroids of each 100 m road segment for each monitoring day. UFP number concentrations were log-transformed to achieve a normal distribution. To eliminate extreme daily values, all UFP number concentration data were trimmed to exclude data outside the 1st and 99th percentiles. In addition, road segments with fewer than 5 samples on a given day were also excluded (<5 seconds of monitoring time). This was a pragmatic decision meant to balance spatial coverage across each city and the number of samples available for any given road segment. The mean number of daily samples per road segment was 32 in Montreal and 17 in Toronto. Geographic information system variables were linked to road segment centroids (based on latitude and longitude coordinates) and daily meteorological variables were joined by matching on date and time rounded to the nearest hour. All predictor variables were standardized for model development.

Statistical analysis

In total, 78 land use variables were extracted for potential inclusion in the final models. We first examined single-variable linear models and excluded variables that were not associated with UFPs (i.e., 95% confidence interval [CI] included the null). Next, highly correlated predictors ($r > 0.7$) were identified, and we retained the predictor that was most strongly correlated with UFP measurements from each correlated pair. A list of the variables retained for the final models is provided in Table S1; http://links.lww.com/EE/A292. All models were developed separately for each city.

Machine learning models

Three machine learning algorithms were examined in predicting daily spatiotemporal variations in outdoor UFP number concentrations (particles/cm^3) and UFP size (nm) including generalized additive models (GAMs), random forest models, and extreme gradient boosting models (XGBoost). All machine learning models were developed using the mlr (v2.19.0) package in R (version 4.3.1). The GAM model was developed using a gamboost algorithm, which learns an ensemble of GAMs to make final predictions and is designed to exclude variables that are not informative as part of the modeling process.[21] The selection of model hyperparameters for random forest and XGBoost models was based on 1000 random iterations over a predefined hyperparameter search space (Tables S2 and S3; http://links.lww.com/EE/A292) with optimal hyperparameters and model predictors selected based on a 10-fold cross-validation procedure. Feature importance plots were examined for random forest and XGboost models and plots of model errors versus number of trees were examined to verify that a sufficient number of trees were employed during training. Feature importance plots provide a relative ranking of variable importance in terms of how useful each variable was in constructing decision trees used in the model. The fact that a given variable was important in constructing the decision tree model does not imply a strong linear correlation between the two variables. All final models were evaluated using a 10-fold cross-validation procedure across the entire database within each city (i.e., not stratified by specific times/locations within each city) and R^2 and root mean square error values from this procedure are reported. Within each city, the same dataset was used for training all models to facilitate model comparisons. Model residuals were mapped for all final models to examine the potential spatial clustering of model errors.

Results

In total, our analytical database included approximately 12,705 road segments in Montreal (51,174 samples) and 10,929 road segments in Toronto (70,402 samples). Descriptive statistics for daily median outdoor UFP number concentrations and UFP size are shown in Table 1. Overall, daily median outdoor UFP number concentrations and UFP size varied substantially across Montreal and Toronto with a slightly larger range of outdoor UFP number concentrations observed across Montreal and a larger range of outdoor UFP sizes observed in Toronto. UFP size was inversely correlated with UFP number concentrations in both cities (Montreal: $r = -0.69$; Toronto: $r = -0.53$). Outdoor $PM_{2.5}$ mass concentrations were weakly correlated with UFP number concentrations in Montreal ($r = 0.086$) and Toronto ($r = -0.0532$) whereas stronger correlations were observed between outdoor $PM_{2.5}$ and UFP size (Montreal: $r = 0.30$; Toronto: $r = 0.43$). Ambient weather conditions captured during mobile monitoring reflected seasonal variations typical of these cities with daily outdoor temperatures ranging from −24 to 33 °C in Montreal and −1 to 31 °C in Toronto. Daily mean outdoor $PM_{2.5}$ concentrations were typically low in both cities but tended to be higher in Toronto (2020–2021: mean = 7.3 μg/m^3; SD = 3.2 μg/m^3; summer 2023: mean = 23.3 μg/m^3; SD = 14.1 μg/m^3) than in Montreal (mean: 6.4 μg/m^3; SD: 4.6 μg/m^3). Outdoor $PM_{2.5}$ concentrations in Toronto during summer 2023 were higher than normal because of wildfire smoke occasionally impacting the region.

Machine learning model performance is summarized in Table 2. The XGBoost models performed best for both UFP number concentrations and UFP size with similar model performance observed in both cities. Scatter plots of measured versus predicted values from the XGBoost models are shown for Montreal and Toronto in Figure 1 and illustrate the strong relationships between measured and predicted values in both cities. Feature importance plots for the XGBoost models for UFP number concentrations and UFP size are shown in Figures S1–S4; http://links.lww.com/EE/A292. For both UFP number concentrations and size, the most important predictors included a combination of spatial (e.g., highways within 300 m, NOx emissions within 100 m) and temporal (weather/$PM_{2.5}$) variables. Maps of XGBoost model residuals are shown in Figures S5–S8; http://links.lww.com/EE/A292 and did not suggest any obvious spatial clustering of model errors. Additional scatter plots of measured and predicted values (Figures S9–S16; http://links.lww.com/EE/A292) and maps of model residuals (Figures S17–S24; http://links.lww.com/EE/A292) are available in the Supplemental Material for the GAM and random forest models that did not perform as well in predicting outdoor UFP number concentrations or UFP size. The shapes of relationships between standardized predictor variables and outdoor UFP number concentrations and UFP size in the GAM models are also shown in Figures S25–S28; http://links.lww.com/EE/A292. While similar predictors were generally selected by the gamboost algorithms in each city, the shapes of associations were sometimes different (although the magnitudes of associations for any individual predictor were generally small). In general, lower wind speeds and higher NOx emissions within 100 m were associated with higher UFP number concentrations (although a "U" shaped relationship was observed for temperature in Toronto) whereas higher outdoor $PM_{2.5}$ concentrations, higher relative humidity, higher temperature, and fewer highways within 300 m were associated with larger UFP size.

Maps of predicted daily median outdoor UFP number concentrations from the final XGBoost models are shown in Figure 2 for days across all four seasons. Figure 2 highlights substantial spatial variations in outdoor UFP number concentrations across each city and how these variations are expected to change across a range of seasonal weather conditions. In particular, outdoor UFP concentrations in Montreal tended to be greatest in winter whereas Toronto had higher concentrations in summer. This pattern is consistent with relationships observed between temperature and UFP number concentrations in each city in the GAM models (Figures S25 and S26; http://links.lww.com/EE/A292) whereby an inverse linear relationship was apparent between temperature and outdoor UFP number concentrations in Montreal and a U-shaped relationship was observed in Toronto. Maps of predicted spatial variations in daily median outdoor UFP size from final XGBoost models are shown in Figure S29; http://links.lww.com/EE/A292 and highlight substantial spatial variations across a range of seasonal weather conditions.

Discussion

Increasing evidence suggests that outdoor UFPs may have important long-term health impacts,[5–8] but less is known about the short-term consequences of population exposures to these pollutants. In this study, we developed new machine learning models to predict high-resolution spatiotemporal variations

Table 1.

Descriptive data for daily median outdoor UFP number concentrations (particles/cm^3) and UFP size (nm) in Montreal (2020–2021) and Toronto, Canada (2020–2021, summer 2023)

Pollutant	No. road segments	No. samples	Mean (SD)	Median	1st–99th percentile
Montreal					
UFPs	12,705	51,174	22,562 (36,177)	12,743	1389–181,672
UFP size			34 (15)	31	10–77
Toronto					
UFPs	10,929	70,402	18,352 (24,162)	11,878	2472–118,544
UFP size			44 (25)	38	14–154

Table 2.

Machine learning model performance in predicting daily median outdoor UFP number concentrations and UFP size in 10-fold cross-validation procedures

			Results of 10-fold cross-validation			
City	Pollutant	Model	Slope (95% CI)	Intercept (95% CI)	R^2	Root mean square error
Montreal	Log(UFPs)	GAM	1.15 (1.14, 1.15)	−1.42 (−1.55, −1.28)	0.312	0.841
		Random forest	1.24 (1.23, 1.25)	−2.26 (−2.39, −2.13)	0.377	0.801
		XGBoost	1.06 (1.06, 1.07)	−0.572 (−0.626, −0.519)	0.727	0.530
	UFP size (nm)	GAM	1.21 (1.20, 1.23)	−7.17 (−7.69, −6.65)	0.327	12.4
		Random forest	1.40 (1.39, 1.42)	−13.7 (−14.2, −13.2)	0.436	11.3
		XGBoost	1.04 (1.03, 1.04)	−1.25 (−1.40, −1.10)	0.823	6.33
Toronto	Log(UFPs)	GAM	1.42 (1.40, 1.43)	−3.91 (−4.09, −3.73)	0.229	0.739
		Random forest	1.41 (1.40, 1.43)	−3.88 (−4.02, −3.73)	0.322	0.693
		XGBoost	1.07 (1.07, 1.08)	−0.684 (−0.731, −0.638)	0.723	0.443
	UFP size (nm)	GAM	1.08 (1.07, 1.08)	−3.33 (−3.69, −2.97)	0.518	17.4
		Random forest	1.05 (1.05, 1.06)	−2.38 (−2.62, −2.15)	0.719	13.3
		XGBoost	1.02 (1.02, 1.03)	−0.899 (−1.03, −0.773)	0.898	8.01

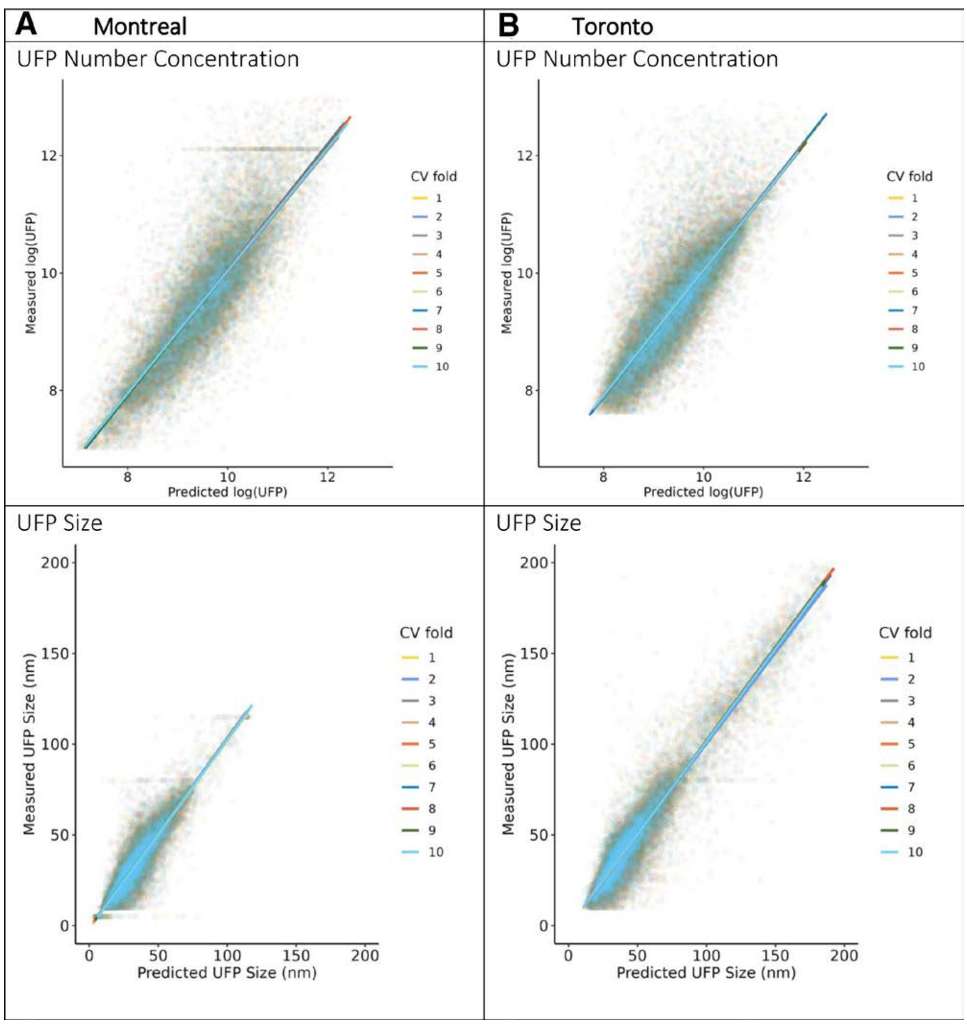

Figure 1. Measured versus predicted values in 10-fold cross-validation procedures for outdoor UFP number concentrations (log [particles/cm^3]) and UFP size (nm) in Montreal (A) and Toronto (B), Canada using the XGBoost models.

in outdoor UFP number concentrations and UFP size across Canada's two largest cities. These models explained the majority of spatiotemporal variations in outdoor UFP number concentrations and UFP size without obvious spatial clustering of model errors in either city. In particular, the XGBoost models performed best likely owing to the use of a series of decision trees with each subsequent tree in the series trying to "correct the errors" of the previous tree. Spatial patterns in daily median outdoor UFP number concentrations were apparent in both cities with higher concentrations in the eastern portion of Montreal and the western portion of Toronto, which is consistent with the spatial distribution of major combustion sources in each city (i.e., industry, traffic). Seasonal differences in outdoor UFPs varied between cities, with Montreal having higher

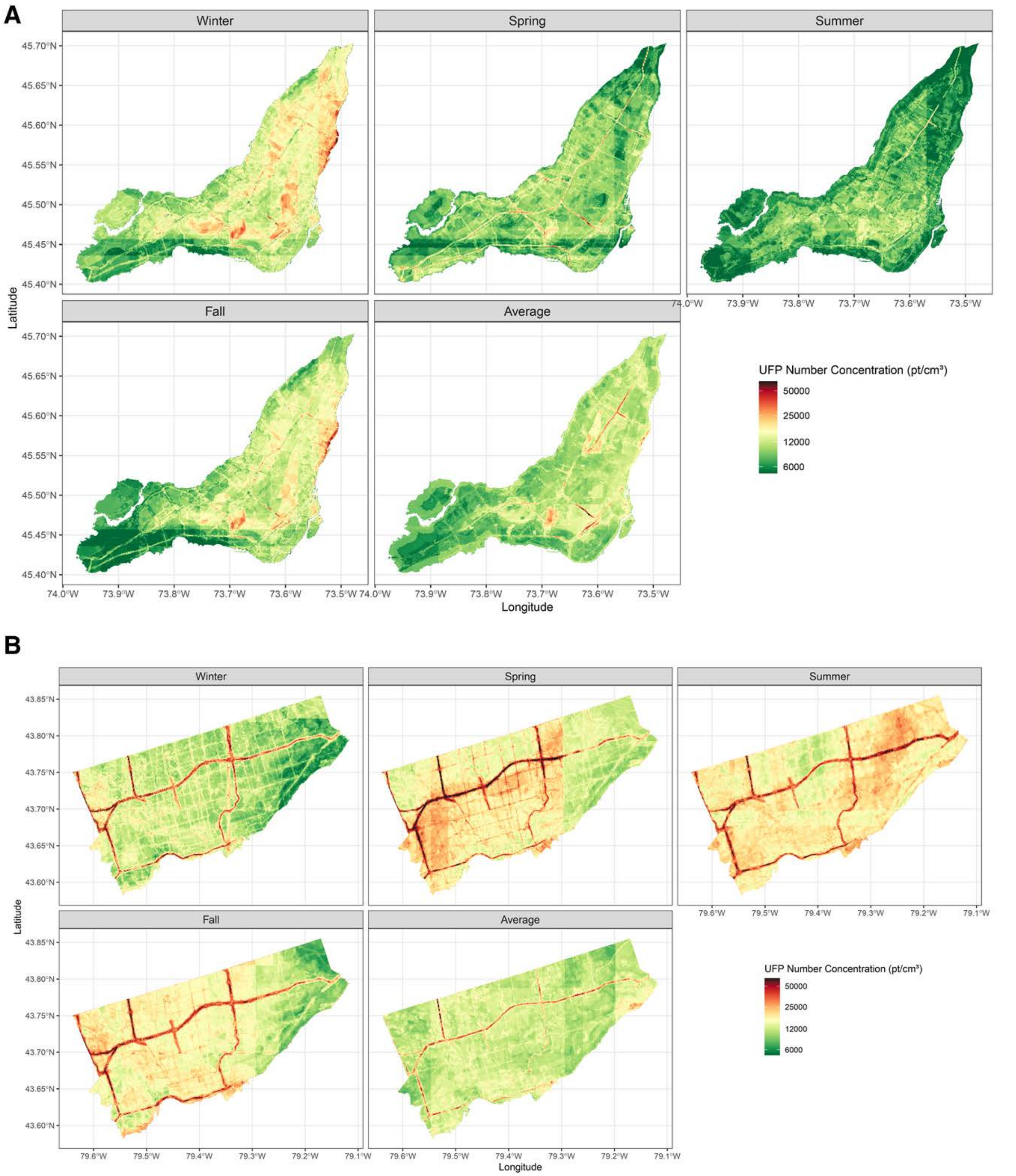

Figure 2. Predicted daily median outdoor UFP number concentrations across Montreal (A) and Toronto (B) for weather conditions observed in different seasons (winter: temperature = 0 °C; relative humidity = 30%; wind speed = 0 m/s; pressure = 101 kPa; precipitation = 5 mm; $PM_{2.5}$ = 7 μg/m^3; spring: temperature = 10 °C; relative humidity = 80%; wind speed = 0 m/s; pressure = 101 kPa; precipitation = 5 mm; $PM_{2.5}$ = 7 μg/m^3; summer: temperature = 30 °C; relative humidity = 55%; wind speed = 3 m/s; pressure = 101 kPa; precipitation = 5 mm; $PM_{2.5}$ = 7 μg/m^3; fall: temperature = 30 °C; relative humidity = 30%; wind speed = 0 m/s; pressure = 101 kPa; precipitation = 0 mm; $PM_{2.5}$ = 7 μg/m^3; average: temperature = 15 °C; relative humidity = 65%; wind speed = 3 m/s; pressure = 101 kPa; precipitation = 0 mm; $PM_{2.5}$ = 7 μg/m^3).

concentrations in winter and Toronto having higher concentrations in summer. These patterns were consistent with the shapes of concentration–response relationships observed between UFPs and weather variables in GAM models and could be partially explained by more severe winters in Montreal, which promote higher UFP concentrations.[22] It is also possible that differences exist between cities/seasons in the secondary formation of outdoor UFPs, but we did not specifically examine this question and this should be the focus of future work. Nevertheless, our new models address an important knowledge gap in predicting short-term population exposures to UFPs and will serve as a resource for future epidemiological studies examining the acute cardiorespiratory health impacts of outdoor UFPs in Canada.

To date, several studies have evaluated the acute cardiorespiratory health impacts of outdoor UFPs and have reported inconsistent results. For example, Stafoggia et al[9] examined the association between short-term variations in outdoor UFP concentrations and mortality in eight European cities using data from fixed-site monitors and generally reported null results once other air pollutants were included in the models. Similar results were also reported for five cities in Germany, Czech Republic, Slovenia, and Ukraine based on fixed-site monitors in each location.[12] Conversely, Bergmann et al[10] examined associations between short-term variations in outdoor UFP concentrations and cardiorespiratory morbidity/mortality using data from a single background monitor in Copenhagen and observed positive associations, particularly for chronic obstructive pulmonary disease mortality and hospital admissions for asthma. Similarly, a time-series study of daily outdoor UFP concentrations and mortality in three German cities (including data from six fixed-site monitors) reported positive associations between outdoor UFPs and respiratory mortality but not for cardiovascular mortality.[11]

Finally, a meta-analysis of studies of short-term variations in outdoor UFP concentrations and respiratory morbidity noted substantial heterogeneity in existing studies and recommended expanded research and harmonized exposure assessment procedures in future investigations.[16] Studies of the acute health impacts of outdoor UFPs utilizing high-resolution exposure models to capture within-city spatiotemporal variations were not identified, but our new models address this important knowledge gap and can be applied in future studies to improve our understanding of the acute-term health impacts of these pollutants. Importantly, we also developed models to predict daily spatial variations in outdoor UFP size, which has not traditionally been considered in epidemiological studies of outdoor UFPs. Indeed, it is possible that outdoor UFP size is independently associated with morbidity/mortality (i.e., different UFP sizes could have different health impacts for a given UFP number concentration) and our new models can be used to explore this question in future studies.

Our study had several important strengths including use of a large database of mobile UFP measurements collected across two major Canadian cities reflecting all seasons of the year, all days of the week, and most times of the day. However, it is important to recognize several limitations. First, while our best-performing models slightly overestimated observed UFP concentrations, clear spatial patterns in model errors were not identified suggesting that spatial contrasts in exposures used in future epidemiological analyses should not be impacted by systematic differences in model performance across the spatial domain. In addition, our monitoring data did not include overnight periods (11 pm–7 am) and omission of this data could contribute to bias in our exposure estimates. In particular, if outdoor UFP number concentrations tend to be lower during overnight periods (and mean UFP sizes tend to be larger), our predictions may tend to overestimate daily median outdoor UFP number concentrations (and underestimate UFP size). Moreover, the number of daily measurements per road segment was small, and thus our estimates of daily median values on each road segment are likely imprecise. However, this tradeoff is characteristic of mobile data whereby increased spatial coverage comes at the cost of fewer measurements in any individual location. While this limitation certainly contributes to errors in our predictions, there was no clear spatial clustering of model errors across each city. This is an important consideration as it suggests that the magnitude of exposure measurement error will be similar for populations across each city in future applications of our models in epidemiological analyses. Finally, as noted above, our models did not include predictors to explicitly capture secondary UFP formation or terms for total daily traffic counts or heavy-duty vehicles. Instead, we used road-segment level NO_x emissions to estimate spatial variations in vehicle emissions, but future studies should aim to include a wider range of traffic characteristics as well as potential predictors of secondary UFP formation when possible. In summary, we developed new models to predict daily spatiotemporal variations in outdoor UFP number concentrations and UFP size across Canada's two largest cities. These models are now available to support epidemiological analyses in these locations to improve our understanding of the acute health impacts of outdoor UFPs.

Conflicts of interest statement

The authors declare that they have no conflicts of interest with regard to the content of this report.

References

1. Weichenthal S, Pinault L, Christidis T, et al. How low can you go? Air pollution affects mortality at very low levels. *Sci Adv*. 2022;8:eabo3381.
2. Lloyd M, Ganji A, Xu J, et al. 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*. 2023;178:108106.
3. Lloyd M, Carter E, Guzman F, et al. Predicting within-city spatial variations in outdoor ultrafine particle and black carbon concentrations in Bucaramanga, Colombia: a hybrid approach using open source geographic data and digital images. *Environ Sci Technol*. 2021;55:12483–12492.
4. Kerckhoffs J, Hoek G, Gehring U, Vermeulen R. Modelling nationwide spatial variation of ultrafine particles based on mobile monitoring. *Environ Int*. 2021;154:106569.
5. Bouma F, Janssen NA, Wesseling J, et al. Long-term exposure to ultrafine particles and natural and cause-specific mortality. *Environ Int*. 2023;175:107960.
6. Bai L, Chen H, Hatzopoulou M, et al. Exposure to ambient ultrafine particles and nitrogen dioxide and incident hypertension and diabetes. *Epidemiology*. 2018;29:323–332.
7. Bai L, Weichenthal S, Kwong JC, et al. Associations of long-term exposure to ultrafine particles and nitrogen dioxide with increased incidence of congestive heart failure and acute myocardial infarction. *Am J Epidemiol*. 2019;188:151–159.
8. Weichenthal S, Olaniyan T, Christidis T, et al. With-city spatial variations in ambient ultrafine particle concentrations and incident brain tumours in adults. *Epidemiology*. 2020;31:177–183.
9. Stafoggia M, Schneider A, Cyrys J, et al; UF&HEALTH Study Group. Association between short-term exposure to ultrafine particles and mortality in eight European urban areas. *Epidemiology*. 2017;28:172–180.
10. Bergmann M, Andersen ZJ, Massling A, et al. Short-term exposure to ultrafine particles and mortality and hospital admissions due to respiratory and cardiovascular disease in Copenhagen, Denmark. *Environ Pollut*. 2023;336:122396.
11. Schwarz M, Schneider A, Cyrys J, Bastian S, Breitner S, Peters A. Impact of ambient ultrafine particles on cause-specific mortality in three German cities. *Am J Respir Crit Care Med*. 2023;207:1334–1344.
12. Lanzinger S, Schneider A, Breitner S, et al; UFIREG study group. Associations between ultrafine and fine particles and mortality in five central European cities – results from the UFIREG study. *Environ Int*. 2016;88:44–52.
13. Weichenthal S, Van Ryswyk K, Goldstein A, Shekarrizfard M, Hatzopoulou M. Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: a land use regression model. *Environ Pollut*. 2016;208:241–248.
14. Weichenthal S, Van Ryswyk K, Goldstein A, et al. A land use regression model for ambient ultrafine particles in Montreal, Canada: a comparison of linear regression and a machine learning approach. *Environ Res*. 2016;146:65–72.
15. Ren X, Mi Z, Georgopoulos PG. Comparison of machine learning and land use regression for fine scale spatiotemporal estimation of ambient air pollution: modelling ozone concentrations across the contiguous United States. *Environ Int*. 2020;142:105827.
16. Samoli E, Rodopoulou S, Schneider A, et al. Meta-analysis on short-term exposure to ambient ultrafine particles and respiratory morbidity. *Eur Respir Rev*. 2020;29:200116.

17. Buchhorn M, Lesiv M, Tsendbazar NE, Herold M, Bertels L, Smets B. Copernicus global land cover layers—collection 2. *Remote Sens*. 2020;12:1044.
18. Chace JF, Walsh JJ. Urban effects on native avifauna: a review. *Landsc Urban Plann*. 2006;74:46–69.
19. Clergeau P, Croci S, Jokimäki J, Kaisanlahti-Jokimäki ML, Dinetti M. Avifauna homogenisation by urbanisation: analysis at different European latitudes. *Biol Conserv*. 2006;127:336.
20. Aronson MFJ, La Sorte FA, Nilon CH, et al. A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proc Biol Sci*. 2014;281:20133330.
21. Hofner B, Mayr A, Robinzonov N, Schmid M. Model-based boosting in R. A hands-on tutorial using the R package mboost. *Comp Stat*. 2014;29:3–35.
22. Sioutas C, Delfino RJ, Singh M. Exposure assessment for atmospheric ultrafine particles (UFPs) and implications in epidemiologic research. *Environ Health Perspect*. 2005;113:947–955.