A Mobile Air Pollution Monitoring Data Set

Matthew D. Adams ^{1,*} and Denis Corr ²¹ Department of Geography, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada² Corr Research, Hamilton, ON L8P 1X7, Canada; denis.corr@gmail.com

* Correspondence: md.adams@utoronto.ca; Tel.: +1-905-569-4761

Received: 8 November 2018; Accepted: 17 December 2018; Published: 22 December 2018

Abstract: Air pollution was observed in Hamilton, Ontario, Canada using monitors installed in a mobile platform from November 2005 up to November 2016. The dataset is an aggregation of several project specific monitoring days, which attempted to quantify air pollution spatial variation under varying conditions or in specific regions. Pollutants observed included carbon monoxide, nitric oxide, nitrogen dioxide, total nitrogen oxides, ground-level ozone, particulate matter concentrations for size cuts of 10 μm , 2.5 μm and 1 μm , and sulfur dioxide. Observations were collected over 114 days, which occurred in varying seasons and months. During sampling, the mobile platform travelled at an average speed of 27 km/h. The samples were collected as one-minute integrated samples and are prepared as line-segments, which include an offset for instrument response time. Sampling occurred on major freeways, highways, arterial and residential roads. This dataset is shared in hopes of supporting research on how to best utilize air pollution observations obtained with mobile air pollution platforms, which is a growing technique in the field of urban air pollution monitoring. We conclude with limitations in the data capture technique and recommendations for future mobile monitoring studies.

Dataset: Data are provided as a supplemental document.

Dataset License: CC-BY-SA

Keywords: air pollution; mobile monitoring; particulate matter; nitrogen oxides; ozone; sulfur dioxide; carbon monoxide; Hamilton; Ontario; Canada

1. Summary

Air pollution is well established to have negative human health effects [1–3]. By the 1970s it was apparent that measuring pollution in an urban setting required more than one monitoring unit to capture the spatial variation [4,5]. Monitoring programs began to use multiple monitors within a city; however, this can quickly become expensive as the number of pollutants and locations observed increases. Another approach that researchers have applied to assess the spatial variability of air pollutants is mobile air pollution monitoring, where air pollution monitoring equipment is mounted on a mobile platform, such as a motor-vehicle. Two general approaches have emerged, which include the use of lightweight lower cost sensors that are small and portable, which may be attached to an individual or on a bike [6]. The second approach is to install reference grade air monitoring equipment in a mobile platform (e.g., commercial van) [7]. By reference grade, we are referring to the equipment that a national or regional environmental agency has approved for use in their monitoring programs (e.g., EPA Reference Equivalent Methods). Mounting monitors on a mobile platform allows for observations in various locations.

Mobile data collection results in a unique dataset, which is spatially and temporally discontinuous. Unlike a fixed location monitor that measures change across time; mobile data may not repeat sampling

at a location with regularity. Repetition may be staggered, which introduces periods of unobserved data. Researchers have begun to address the challenge of determining a long-term mean from incomplete samples, which has involved adjusting the mobile data based on the value at a fixed location monitor [8], or incorporating both mobile and fixed location data into the air pollution model [9]. Attempts have been made to understand the sampling saturation necessary in mobile campaigns [10–12]. We believe that investigating challenges with mobile collection is timely in air pollution research.

This dataset contains some of the earliest examples of mobile air pollution monitoring [7]. The data have supported many research outputs that have included the identification of elevated air pollution concentrations during localized temperature inversion episodes in the City of Hamilton [13], land use regression modelling of SO₂ [14], comparison of air quality change between two time-periods [15], examining the effectiveness of road dust mitigation approaches [16], real-time air pollution mapping of multiple pollutants [9], and estimating air pollution exposure by mode choice of students [17].

Our motivation to provide this data to the research community, is that it is not possible for us to address all the challenges that surround the use of mobile air pollution data; we hope by providing this rich dataset to the research community that others can use this data for simulation studies in mobile sampling campaigns, to compare their results to another geographic region, or supplement their studies. This data set consists of eleven years of mobile air pollution monitoring conducted in Hamilton, Ontario, Canada between November 2005 and up to November 2016. The observations were collected during 114 days of air pollution monitoring.

2. Data Description

The data are provided as a single plain text comma separated values file. The file contains 46,080 rows of observations (air pollution observations) and 16 columns (attributes) for each observation. The data were collected as one-minute integrated samples and are prepared as line segments, i.e., the start to end points of the one-minute integrated sample. The column headers and description of each attribute are included in Table 1. NA data are indicated by a value of NA.

Table 1. Air pollution file attributes.

Column Name	Description
ID	Unique ID for each observation.
CO	Carbon monoxide concentrations (ppm).
NO	Nitric oxide concentrations (ppb).
NO ₂	Nitrogen dioxide concentrations (ppb).
NO _x	Total nitrogen oxides concentration (ppb).
O ₃	Ground level ozone concentration (ppb).
PM1	Particulate matter concentrations for PM < 1 µm. Units µg/m ³ .
PM2	Particulate matter concentrations for PM < 2.5 µm. Units µg/m ³ .
PM10	Particulate matter concentrations for PM < 10 µm. Units µg/m ³ .
SO ₂	Sulfur dioxide concentrations (ppb).
POINT_X_START	Start Node of Line - X-coordinate, UTM 17N NAD83 projection.
POINT_Y_START	Start Node of Line - Y-coordinate, UTM 17N NAD83 projection.
POINT_X_END	End Node of Line - X-coordinate, UTM 17N NAD83 projection.
POINT_Y_END	End Node of Line - Y-coordinate, UTM 17N NAD83 projection.
START_TIME	Start of collection for line segment. Format: YYYY-MM-DD HH:MM:SS
END_TIME	End of collection for line segment. Format: YYYY-MM-DD HH:MM:SS

3. Methods

The study area is Hamilton, Ontario, Canada, which has received many studies on its air quality [7,13,18–21]. Hamilton is developed around an industrial core, which has traditionally focused on steel production. In recent years, these operations have slowed. The 2006 population for Hamilton was 504,560, increasing to 519,955 in 2011 and to 536,930 in 2016 [22]. During the period of study, the population increased by 6.4%. The Niagara Escarpment transects Hamilton, separating the region

into an upper and a lower city. Temperature inversions have been demonstrated to elevate air pollution concentrations in the lower city [13]. The city is intersected with two freeways and two other major highways, which are sources of urban pollution. The city is located on the western tip of Lake Ontario. In Figure 1, we present the land use, transportation network, the escarpment that separates the upper and lower city, and the boundary for the City of Hamilton.

Figure 1. Panel A (Top), presents the extent of the City of Hamilton with the land use, transportation network, and the escarpment. Panel B (Bottom), overlays the air pollution line segment start points on the land use layer.

The mobile pollution data were collected between November 2005 and November 2016, covering a span of eleven years. In Figure 2, we present the periods when observations were conducted, and which pollutants were measured. During the data collection there was no standard location

revisit approach (i.e., a location may only have been sampled once). The data were collected for specific studies, but they have all been combined into this larger dataset. During collection a GPS unit obtained locational information that was paired with the air pollution observations. The average speed of the mobile platform was 27 km per hour; the intent was to drive as slowly as possible, but during highway data collection this was limited to speeds that would not impede traffic flow. The operation of the mobile platform including instrument calibration and operation was overseen by technicians from the Ontario Ministry of the Environment, Conservation and Parks. The monitoring instrumentation is presented in Table 2. Following October 31, 2011, a few of the instruments were replaced. Monitoring instruments were installed in a commercial van with the air intake centered on the vehicle roof. In Figure 3, we present images of the mobile platform including the installation of monitors, the air intake, and the sampling manifold.

Figure 2. Air Pollution Observations by Pollutant and Time. Black lines indicate monitoring periods. The red line indicates the change in instrumentation.

Figure 3. Mobile Monitoring Unit. Panel A: Rack mounted monitors. Panel B: Air intake highlighted by red rectangle. Panel C: Sampling Manifold connected to air intake.

Table 2. Air pollution Monitor Attributes.

Pollutant	Instrument	Usage Dates	Response Time
Carbon Monoxide	Thermo 48C	Nov 2005–Oct 2011	60 s
Carbon Monoxide	Thermo 48i	Nov 2011–End	60 s
Nitrogen Oxides (NO, NO ₂ , NO _x)	Thermo 42C	Nov 2005–Oct 2011	80 s
Nitrogen Oxides (NO, NO ₂ , NO _x)	Teledyne 200 EU	Nov 2011–End	20 s
Ozone	Thermo 49C	Nov 2005–End	20 s
Particulate Matter (1, 2.5, 10)	Grimm 1.107	Nov 2005–End	6 s
Sulfur Dioxide	Monitor Labs 8850	Nov 2005–Oct 2011	240 s
Sulfur Dioxide	Thermo 43C	Nov 2011–End	110 s

We report one-minute integrated samples, which are the average concentration obtained over one-minute. Because the vehicle is usually moving during the one-minute observation the samples are not point samples but are better represented as line segments. The line segment length varies based on the speed of the vehicle over the minute. The air pollution instruments report in near-real time, but there is a lag-time from when the air sample enters the instrument until the measurement is completed. This lag period is presented in Table 2 for each instrument and ranges from 6 s up to 240 s. The lag period was corrected by offsetting the GPS coordinates based on the length of each instruments' response time. An additional lag of 20 s was added to account for the air intake time, which was estimated from giving a puff test at the air intake. However, we recognize that this value may be over or under estimated as the sampling system received modifications during the period of study. The potential locational error is greatest for the longest segments, which correspond to faster speeds. For example, the CO instruments have a lag of 60 s (+20 system lag), for an air pollution measurement completed at 10:00:00, we would have selected the line segment end coordinates (POINT_X_END & POINT_Y_END) obtained for the GPS at 9:58:40 (80 s earlier). The line segment start location (POINT_X_START & POINT_Y_START) would have been the GPS coordinates from 9:57:40, which is when the observation began. In Table 3, we present summary statistics for the observation data.

Table 3. Air pollution Observation Summary Statistics.

Pollutant	Min	Mean	Median	Max	S.D.
CO	0	0.74	0.5	64.13	1.68
NO	0	28.43	13	916.2	44.12
NO ₂	0	15.88	12	474.5	16.04
NO _x	0	43.33	27.44	978.75	51.87
O ₃	0	24.19	24.5	237.5	11.63
PM ₁₀	0	12.15	10	291	9.02
PM _{2.5}	0	15.08	12.73	731	12.87
PM ₁	0	46.45	25.5	2640	97.36
SO ₂	0	8.53	4.1	283	10.92

4. User Notes

Other open data we have found useful include the Government of Canada's Historical Climate Data (URL: <http://climate.weather.gc.ca/>), which includes hourly meteorological observations within and adjacent to the study area. Depending on the station the data may include temperature (°C), dew point temperature (°C), relative humidity (%), wind direction, wind speed (km/h), visibility (km), station pressure (kPa), humidex, and wind chill (°C) and a weather description. Fixed location air pollution data can be obtained from the Ontario Ministry of the Environment Conservation and Parks' Air Quality Ontario Webpage (URL: <http://www.airqualityontario.com/>), which includes three monitors in Hamilton and one adjacent in Burlington. The pollutants observed at each monitor may include O₃ (ppb), PM_{2.5} (µg/m³), NO (ppb), NO₂ (ppb), NO_x (ppb), SO₂ (ppb) and CO (ppm), which are reported at hourly intervals. Additional pollutants beyond criteria air pollutants can be obtained from Environment and Climate Change Canada's National Air Pollution Surveillance

Program (NAPS) (URL: <http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx>), such as VOCs and trace elements at varying intervals from hourly, daily, monthly, or irregular intervals. Industrial point source emission data are available from the Government of Canada's National Pollutant Release Inventory (URL: <https://www.canada.ca/en/services/environment/pollution-waste-management/national-pollutant-release-inventory.html>), which tracks over 320 pollutants at 7,000 facilities in Canada.

5. Limitations and Recommendations

Reflecting on this data collection, we would be remiss if we did not address the many challenges we encounter in the use of these data. First, the challenge of capturing a spatiotemporal phenomenon with a spatiotemporal platform. Because air pollution releases will move throughout the region, capturing short spatiotemporal events is problematic. For example, if elevated concentrations are obtained, does this represent an infrequent release passing through the airshed, or is it representative of the general conditions? This may be improved with a regular monitoring schedule, such as that employed by Apte et al. [12]. Second, understanding the limitations of the equipment. In this approach we have assumed that instrument response time is equal to a lag period, which is not completely the case, but is our best approximation in rapidly changing conditions. The lag period is affected by the flow in the air intake, the length of the sample tubing, and the operation mode of the instrument among other factors.

Based on our experience in working with these data, we have two recommended data collection approaches that include: (1) If the objective is to develop an air pollution surface of ambient air pollution concentrations, our recommendation is a drive and park approach. With this approach, the researcher will select their sample locations based on their sampling criteria, then drive to each site and sample for a period while operating from battery power, for example 15 min. The length of time will be determined by the research objective and local temporal variability, e.g., understanding a day's spatial variation will require shorter parking periods compared to understanding long-term conditions. Using an approach where the vehicle is parked, provides an opportunity for collecting additional parameters, such as wind-speed and wind-direction. (2) If the goal is to obtain in-traffic concentrations, then we recommend that researchers drive the road segment multiple times over their period of interest. The multiple segments can be averaged to provide a more representative estimate of local in-traffic pollution. For example, if we were interested in measuring freeway concentrations, we would drive a loop of the freeway segments between two interchanges. It may be helpful to capture video while driving, which could be processed using machine learning techniques to identify potential high emitting vehicles [23].

Prior to data collection, we recommend conducting a test for mobile vehicle self-pollution, which is important for the drive and park approach. The self-pollution may occur from driving to the location, or if the vehicle must be running to supply power for the monitoring units. We suggest setting up the monitors in a "clean" environment with the same air intake height, and then driving the mobile vehicle beside the unit and recording the change to background conditions. If the vehicle is turned off, the time it takes to return to background conditions should be calculated, which can be used to remove observations from the time-series. If the vehicle is running during collection, the change observed to background conditions could be subtracted from the data.

Supplementary Materials: The air pollution dataset is available online at <http://www.mdpi.com/2306-5729/4/1/2/s1>.

Author Contributions: Conceptualization, D.C.; Methodology, M.D.A. & D.C.; Formal Analysis, M.D.A.; Writing-Original Draft Preparation, M.D.A. & D.C.; Writing-Review & Editing, M.D.A. & D.C.; Visualization, M.D.A.; Project Administration, D.C. & M.D.A.; Funding Acquisition, D.C. & M.D.A.

Funding: We would like to acknowledge the many funders and supporters for this research that include Hamilton Public Health, Clean Air Hamilton, Ontario Ministry of the Environment, Conservation and Parks.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rckerl, R.; Schneider, A.; Breitner, S.; Cyrys, J.; Peters, A. Health effects of particulate air pollution: A review of epidemiological evidence. *Inhal. Toxicol.* **2011**, *23*, 555–592. [[CrossRef](#)] [[PubMed](#)]
2. Hoek, G.; Krishnan, R.M.; Beelen, R.; Peters, A.; Ostro, B.; Brunekreef, B.; Kaufman, J.D. Long-term air pollution exposure and cardio- respiratory mortality: A review. *Environ. Health* **2013**, *12*, 43. [[CrossRef](#)] [[PubMed](#)]
3. Brunekreef, B.; Holgate, S.T. Air pollution and health. *Lancet* **2002**, *360*, 1233–1242. [[CrossRef](#)]
4. Goldstein, I.; Landovitz, L. Analysis of air pollution patterns in New York City—I. Can one station represent the large metropolitan area? *Atmos. Environ.* **1977**, *11*, 47–52. [[CrossRef](#)]
5. Goldstein, I.; Landovitz, L. Analysis of air pollution patterns in New York City—II. Can one aerometric station represent the area surrounding it? *Atmos. Environ.* **1977**, *11*, 53–57. [[CrossRef](#)]
6. Elen, B.; Peters, J.; van Poppel, M.; Bleux, N.; Theunis, J.; Reggente, M.; Standaert, A. The Aeroflex: A bicycle for mobile air quality measurements. *Sensors* **2013**, *13*, 221–240. [[CrossRef](#)] [[PubMed](#)]
7. Wallace, J.; Corr, D.; Deluca, P.; Kanaroglou, P.; McCarry, B. Mobile monitoring of air pollution in cities: The case of Hamilton, Ontario, Canada. *J. Environ. Monit.* **2009**, *11*, 998–1003. [[CrossRef](#)] [[PubMed](#)]
8. Larson, T.; Henderson, S.B.; Brauer, M. Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environ. Sci. Technol.* **2009**, *43*, 4672–4678. [[CrossRef](#)]
9. Adams, M.D.; Kanaroglou, P.S. Mapping real-time air pollution health risk for environmental management: Combining mobile and stationary air pollution monitoring with neural network models. *J. Environ. Manag.* **2016**, *168*, 133–141. [[CrossRef](#)]
10. Xu, X.; Brook, J.R.; Guo, Y. A statistical assessment of saturation and mobile sampling strategies to estimate long-term average concentrations across urban areas. *J. Air Waste Manag. Assoc.* **2007**, *57*, 1396–1406. [[CrossRef](#)]
11. Adams, M.D.; Kanaroglou, P.S. A Method for Reducing Classical Error in Long-Term Average Air Pollution Concentrations from incomplete Time-Series Data. In *Edited Volume: Spatial Analysis in Health Geography*; Kanaroglou, P.S., Delmelle, E.M., Paez, A., Eds.; Ashgate Publishing: Farnham, UK, 2015; pp. 183–196.
12. Apte, J.S.; Messier, K.P.; Gani, S.; Brauer, M.; Kirchstetter, T.W.; Lunden, M.M.; Marshall, J.D.; Portier, C.J.; Vermeulen, R.C.H.; Hamburg, S.P. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environ. Sci. Technol.* **2017**, *51*, 6999–7008. [[CrossRef](#)] [[PubMed](#)]
13. Wallace, J.; Corr, D.; Kanaroglou, P. Topographic and spatial impacts of temperature inversions on air quality using mobile air pollution surveys. *Sci. Total Environ.* **2010**, *408*, 5086–5096. [[CrossRef](#)] [[PubMed](#)]
14. Kanaroglou, P.S.; Adams, M.D.; De Luca, P.F.; Corr, D.; Sohel, N. Estimation of sulfur dioxide air pollution concentrations with a spatial autoregressive model. *Atmos. Environ.* **2013**, *79*, 421–427. [[CrossRef](#)]
15. Adams, M.D.; De Luca, P.F.; Corr, D.; Kanaroglou, P.S. Mobile Air Monitoring: Measuring Change in Air Quality in the City of Hamilton, 2005–2010. *Soc. Indic. Res.* **2012**, *108*, 351–364. [[CrossRef](#)]
16. DeLuca, P.F.; Corr, D.; Wallace, J.; Kanaroglou, P. Effective mitigation efforts to reduce road dust near industrial sites: Assessment by mobile pollution surveys. *J. Environ. Manag.* **2012**, *98*, 112–118. [[CrossRef](#)] [[PubMed](#)]
17. Adams, M.D.; Yiannakoulis, N.; Kanaroglou, P.S. Air pollution exposure: An activity pattern approach for active transportation. *Atmos. Environ.* **2016**, *140*, 52–59. [[CrossRef](#)]
18. Arain, M.A.; Blair, R.; Finkelstein, N.; Brook, J.R.; Sahsuaroglu, T.; Beckerman, B.; Zhang, L.; Jerrett, M. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. *Atmos. Environ.* **2007**, *41*, 3453–3464. [[CrossRef](#)]
19. Radisic, S.; Newbold, K.B.; Eyles, J.; Williams, A. Factors influencing health behaviours in response to the air quality health index: A cross-sectional study in Hamilton, Canada. *Environ. Health Rev.* **2016**, *59*, 17–29. [[CrossRef](#)]
20. Simone, D.; Eyles, J.; Newbold, K.B.; Kitchen, P.; Williams, A. Air Quality in Hamilton: Who is Concerned? Perceptions from Three Neighbourhoods. *Soc. Indic. Res.* **2012**, *108*, 239–255. [[CrossRef](#)]
21. Jerrett, M.; Burnett, R.T.; Kanaroglou, P.; Eyles, J.; Finkelstein, N.; Giovis, C.; Brook, J.R. A GIS–environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environ. Plan. A* **2001**, *33*, 955–973. [[CrossRef](#)]

22. Hamilton. *2001 to 2016 Census Population and Dwelling Counts by Ward for Hamilton*; Hamilton, ON, Canada, 2017. Available online: <https://www.hamilton.ca/moving-hamilton/community-profile/census-data-hamilton> (accessed on 22 December 2018).
23. Weichenthal, S.; Hatzopoulou, M.; Brauer, M. A picture tells a thousand ... exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology. *Environ. Int.* **2018**. [[CrossRef](#)] [[PubMed](#)]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).