Supplementary Information (SI) for:

Air Quality Monitoring Case Study Using Mobile Low-cost Sensors mounted on Trash-Trucks: Methods Development and Lessons Learned

Authors: Priyanka deSouza[1†], Amin Anjomshoaa[1], Fabio Duarte[1,2], Ralph Kahn[3] , Prashant Kumar[3], Carlo Ratti[1]

[1]Senseable City Lab, Massachusetts Institute of Technology, Cambridge MA, United States
[2] Pontifícia Universidade Católica do Paraná, Brazil
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, United States
[4]Global Centre for Clean Air Research (GCARE), Department of Civil and Environmental Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, United Kingdom
[†] Corresponding author

Section S1: Experiment design in Cambridge, MA

S1.1: OPC-N2

In this experiment, the sampling frequency used is 60 Hz, the response time of the Alphasense OPC-N2 is negligible, and the lag of pumping air through the OPC is less than one second. We chose this frequency as it is the highest that the OPC is capable of measuring at, to maximize the number of measurements.

Other studies have used different sampling frequencies (Elen et al., 2012). Van de Bossche et al., (2015), in their validation of mobile air quality measurements of black carbon using stationary sensors found that the longer the measurements were averaged over, the closer the mobile and stationary measurements agreed. The sensitivity of the methods used to different temporal resolutions needs to be tested in future experiments.

S1.2 GPS sensor

The GPS sensor used is an Adafruit Ultimate GPS module, which has a location accuracy of better than 3 meters. This uncertainty varies depending on signals reflected by buildings and other obstructions in the urban environment. The OPC-N2 air quality monitor and the collocated GPS sensor were synchronized using the data-collection procedure outlined in Anjomshoaa et al. (2018). No GPS data post-processing is required because all the GPS points fall on road segments, and small variations in the data can be neglected[1].

S1.3 Measurement Platform

The air quality monitors were deployed along with a set of non-intrusive sensors, including thermal cameras, temperature, humidity, accelerometer, and GPS sensors, on the tops of City of Cambridge trash-trucks, just over the cab, to minimize truck-related contamination (Figure S1). Anjomshoaa et al. (2018) provide additional information about the sensor package used.

Although trash-trucks provide complete spatial coverage of a city, we sacrifice temporal coverage. A limitation of using these vehicles is that in Cambridge, they operate only between 0700 and 1400 h (local time) on weekdays, and thus the results from our experiment are not generalizable beyond these time periods. Another disadvantage is that the trash-trucks generate emissions themselves, making it important to separate these from the pollution produced from other sources at a given location.

We were careful to position our monitors facing away from the truck exhaust outlet and as distant from it as possible, to minimize contamination by emissions from the truck itself. However, depending on wind speed and wind direction, truck emissions, or those from immediately adjacent vehicles, could impinge on our monitors and affect the results.

In our experiment, 64.4% of the measurements made by the trash-trucks occurred when the vehicles were stationary. The mean speed is 9 km/h (excluding the measurements when the trash-trucks are stationary) due to frequent trash-collection stops. The correlation of $PM_{2.5}$ with speed is negligible (-0.02), indicating that the two are likely unrelated or weakly related, and any relationship is overwhelmed by other factors. This is important, as it indicates that emissions from a trash-truck when it is moving do not have a major impact on the measured $PM_{2.5}$. When the trash-trucks are stationary, i.e., when the majority of measurements occur, we detect large variations in pollution values.

However, we cannot distinguish times when the trash-trucks were idling and when they were at a halt with the engine off. We therefore cannot exclude the possibility that that high pollution values measured when the trucks were stationary are associated with idling truck engines. From observations of the trucks, we note that when drivers stop on streets to collect garbage, they typically turn their engines off. Therefore, we believe that the former hypothesis is reasonable.

[1] http://senseable.mit.edu/cityscanner/

Aerosol emissions from vehicles usually have diameters in the range of 0.02 - 0.13 μm for diesel engines and 0.02 - 0.06 μm for gasoline engines (Zhu et al., 2002). As our low-cost monitors are incapable of detecting particles < 0.38 μm, these emissions likely do not affect our measurements significantly, which favors this assumption. In future deployments, it will be important to test this assumption explicitly, and to distinguish between times when the trucks are idling with their engines on or off, in order to make definitive claims.

Different experiments in the past have used different techniques to account for self-sampling. Some studies have used electric vehicles to avoid self-sampling. As an important aspect of the current initiative is to use existing urban fleets, to remain applicable to similar configurations that might be deployed in other cities, and therefore, we did not consider using dedicated fleets as an option. Other studies, such as Apte et al., (2018) also sampled using gasoline vehicles. Like in our study, they positioned their instrument inlets away from the exhaust pipe to avoid self-sampling. They found that self-emissions mattered only in rare circumstances when the car was reversed into its own exhaust plume after periods of idling in low wind conditions. Tunno et al. (2012) only took samples when the car engine was turned off during their mobile air quality monitoring experiment in Pittsburgh to avoid self-sampling. Padró-Martínez (2012) and Patton et al., (2014) removed data at low wind-speeds, when the wind was behind the car, and at high wind speeds, to avoid self-sampling in Somerville, Massachusetts (~ 14% of the data collected). Jiang et al., (2005) used methanol (emitted by the mobile laboratory but not found in Mexico City's fleet) as a tracer to remove data contaminated by self-sampling. In summary, there is no consensus in the literature about how to remove self-sampling and how important it is. More work needs to be done in the future to propose a protocol to deal with this issue. To mitigate against this issue, in the current study we focus on robust hotspots, for which there are multiple measurements on multiple days that are substantially above average values.

S1.4 Study area and sampling protocol

Our sampling area consisted of all street segments in Cambridge, MA, and some streets in Somerville and Boston, an area of approximately 6 km x 8 km in size (Figure S2, Supplementary Information). The sampling area is bounded by the Charles river to the south, and Cambridge Street, Beacon Street and Massachusetts Avenue to the north.

Our air quality monitors were deployed on the trash-trucks between April 21, 2017 and August 14, 2017 on 27 separate days /sampling runs. (In this study, a "run" is the aggregate of data collected on one day from one truck.) OPC measurements were made every second, totaling 575,800 data points during the experiment. Each trash-truck followed a different, fixed route on each day of the week. As we did not sample different routes on the same weekdays, we did not obtain daily measurements for fixed routes, but for a set of different routes that spanned the entire city over the five-day work-week. As such, we obtained uneven temporal sampling of the trash-truck routes. Figure S2 shows the number of unique days when each road was sampled. Each road was sampled on an average of 4 days. 47 roads were sampled more than 15 times (one road was sampled 25 times), 21 roads were sampled between 10 and 15 times, and the remaining 616 roads were sampled fewer than 10 times.

For a future pilot, we aim to combine the measurements made by deploying OPCs on trash-trucks with other vehicle fleets, such as buses, offering routine sampling, to get multiple measurements for each road on our sampling routes.

Tables

Table S1: *Minimum and maximum diameters for the 16 bins that the Alphasense OPC-N2 measures particle counts. The OPC-N2 measurement sensitivity range is 0.1-150,000 μg/m^3.*[2]

Bin number	**Minimum Diameter (μm)**	**Maximum Diameter (μm)**
1	0.38	0.54
2	0.54	0.78
3	0.78	1
4	1	1.3
5	1.3	1.6
6	1.6	2.1
7	2.1	3
8	3	4
9	4	5
10	5	6.5
11	6.5	8

[2] https://www.iscapeproject.eu/wp-content/uploads/2017/09/iSCAPE_D1.5_Summary-of-air-quality-sensors-and-recommendations-for-application.pdf

12	8	10
13	10	12
14	12	14
15	14	16
16	16	17.5

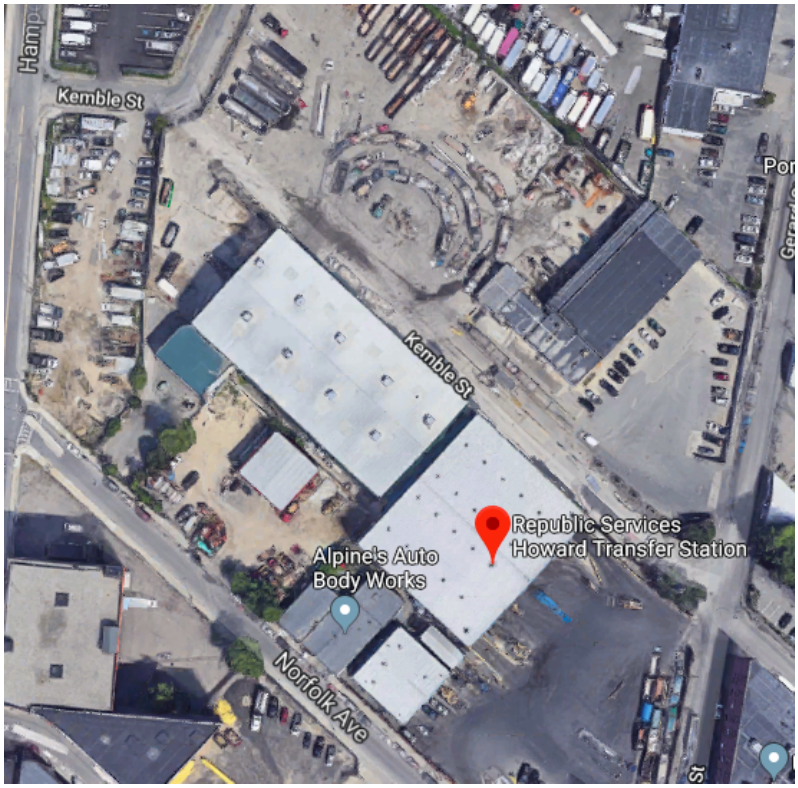
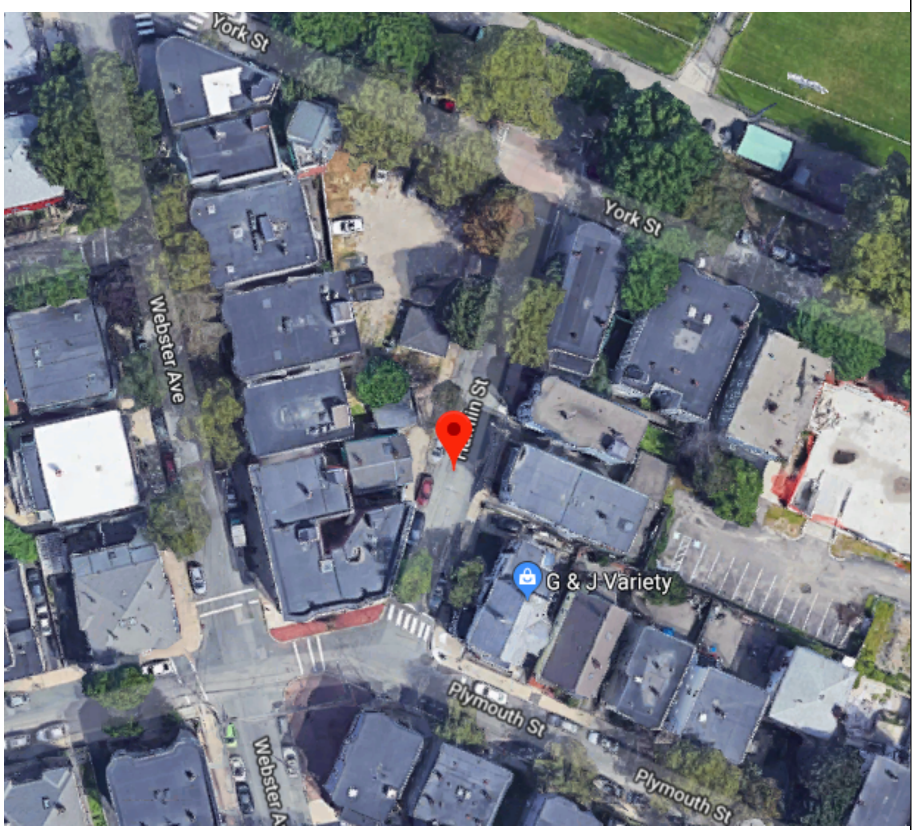
***Table S2**: Comparison between the different background correction techniques*

Background correction methods	**Mean Difference between corrected and uncorrected $PM_{2.5}$ measurements**	**Mean Difference between corrected and uncorrected N1 measurements**	**Mean Difference between corrected and uncorrected N12 measurements**
Spline of Minimums	-4%	-3%	4.1%
Reference monitor	0.6% (note there are some NA values where no data was available from the reference monitor)	NA	NA
Percentile	2.4%	1.5%	2%

***Table S3**: The locations of the hotspots identified in Figure 2 (number of points in the cluster are > 10 and number of unique days of measurement > 1) are described, and enlarged images of these sites are provided.*

Summary	**Name**	**Google Image**

Two clusters (more than 100 meters apart) were found here. The total measurements for which $PM_{2.5}$ > 100 µg/m^3 are 44. The trashtruck only went here on one day, during the period of measurement	Rocky Hill Farm and Rock Hill Transport and Service Corporation, Saugus, MA	
1875 measurements where $PM_{2.5}$ > 100 µg/m^3 make up the cluster here. The number of unique days that comprise this cluster are 3.	Cambridge Public Works Depot	

218 measurements where $PM_{2.5}$ > 100 μg/m^3 make up the cluster here. The number of unique days that comprise this cluster are 10.	Roxbury Garbage Collection Station	
34 measurements where $PM_{2.5}$ > 100 μg/m^3 make up the cluster here. The number of unique days that comprise this cluster are 2.	Hamlin Street	

Figures

***Figure S1**: Sensor package mounted on the top of a trash truck.*

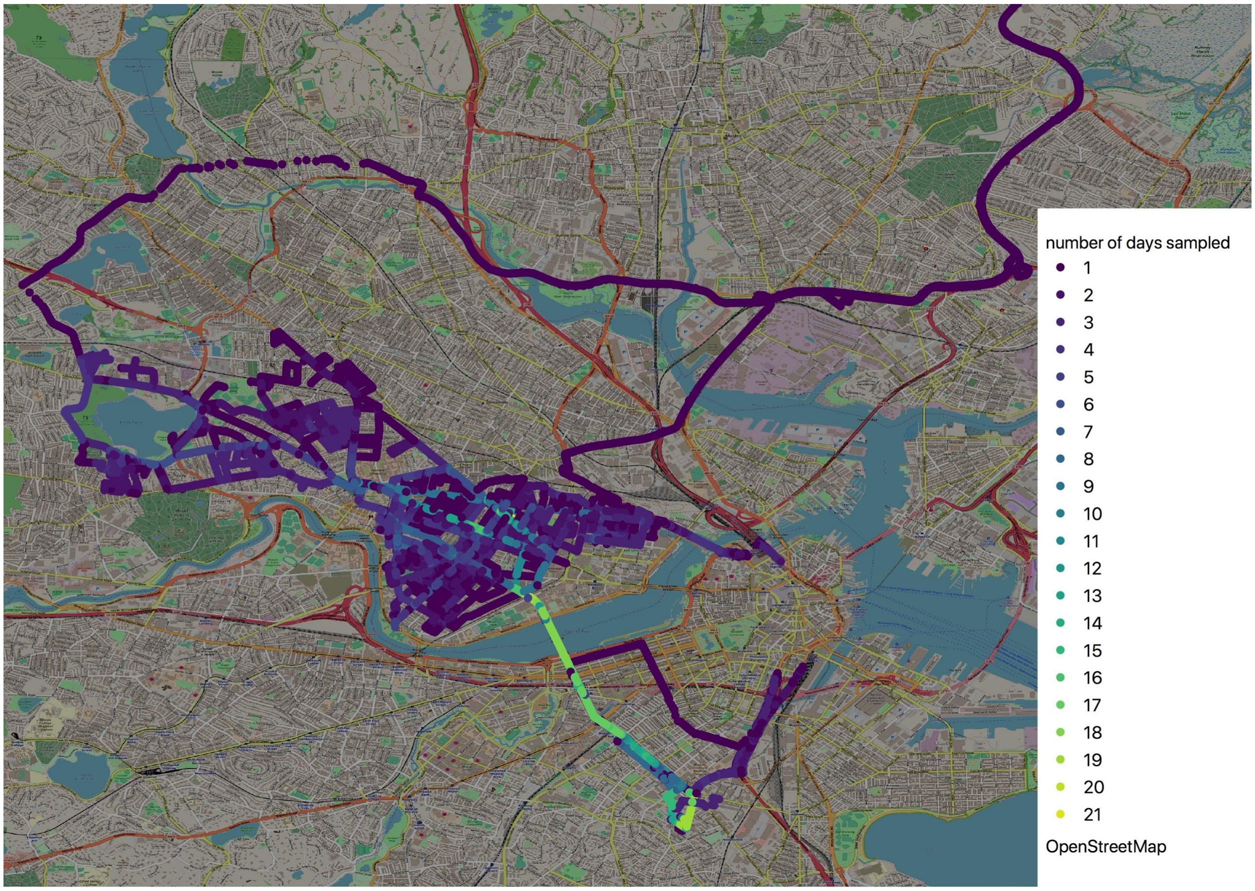


Figure S2*: The number of unique days over which each road on the sampling route was sampled.*

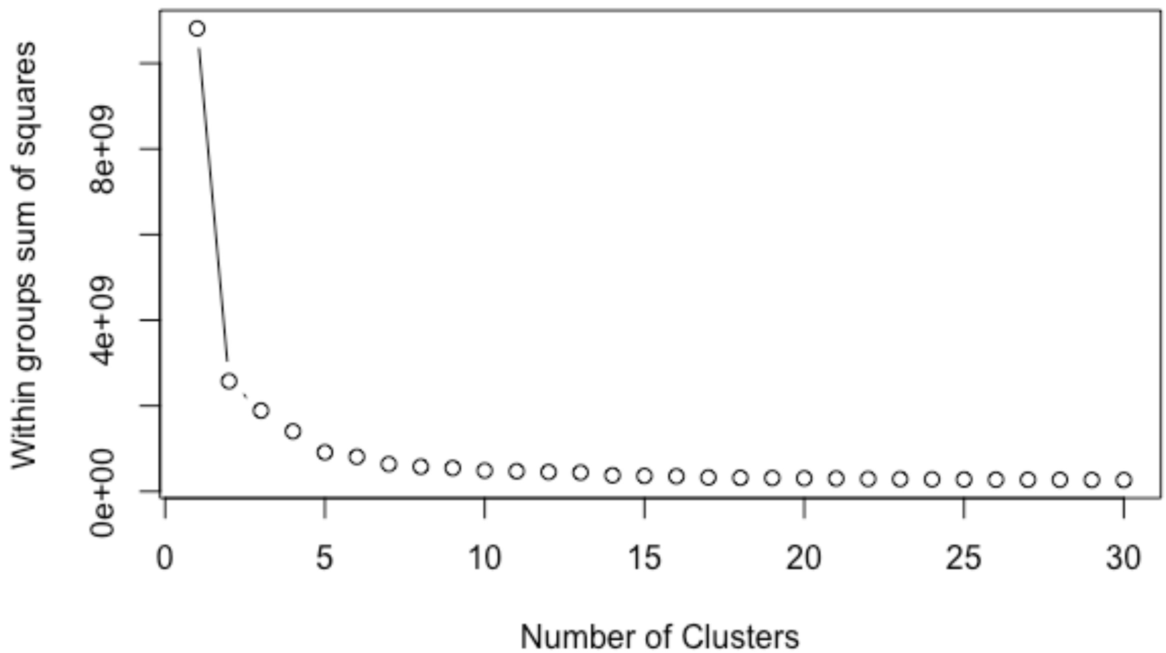
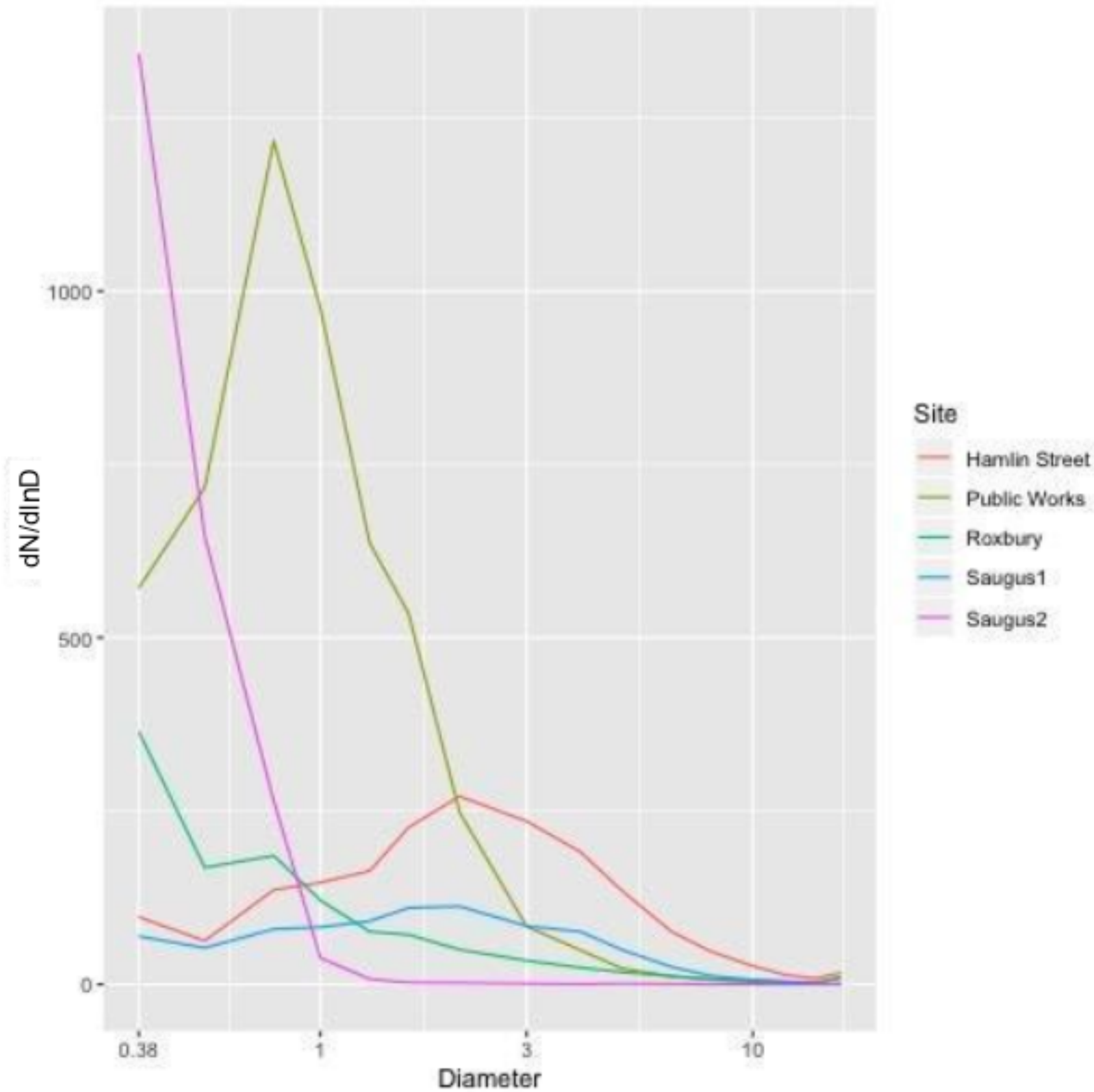


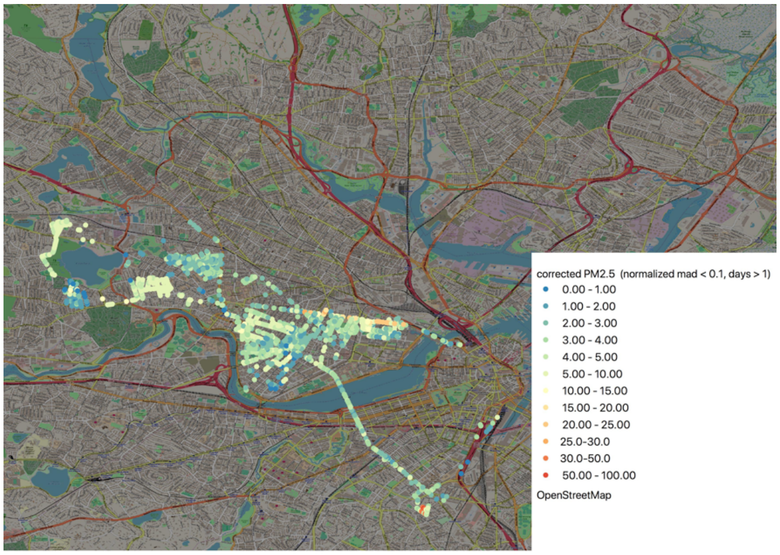
Figure S3: Within group error as a function of cluster size.



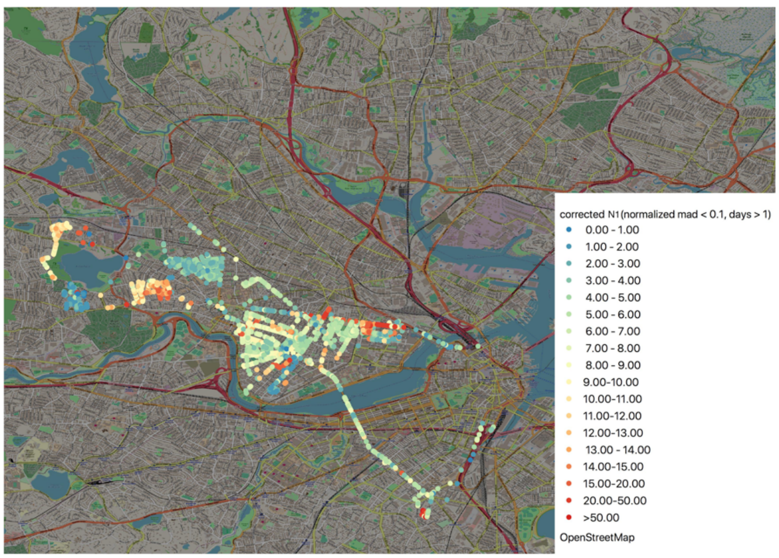
Site	PM_1 (µg/m^3)	$PM_{2.5}$ (µg/m^3)	PM_{10} (µg/m^3)	N1(#/ml)	N12(#/ml)	Number of points in cluster	Number of unique days
Saugus 1	21	114	1522	28	76	11	1 (Hours: 9:00 am, Month: May)

Saugus 2	161	194	213	337	6	30	1 (Hours: 9:00 am, Month: May)
Cambridge Public Works	242	554	1555	333	290	1875	3 (Hours: 8:00 am, 9:00 am, 12:00, Months: May, July, August)
Roxbury	58	993	993	103	48	218	10 (Hours: 8:00 am, 10:00 am, 11:00 am; Months: May, July, August)
Hamlin Street	33	255	5025	39	177	34	2 (Hours: 9:00am, Month: July)

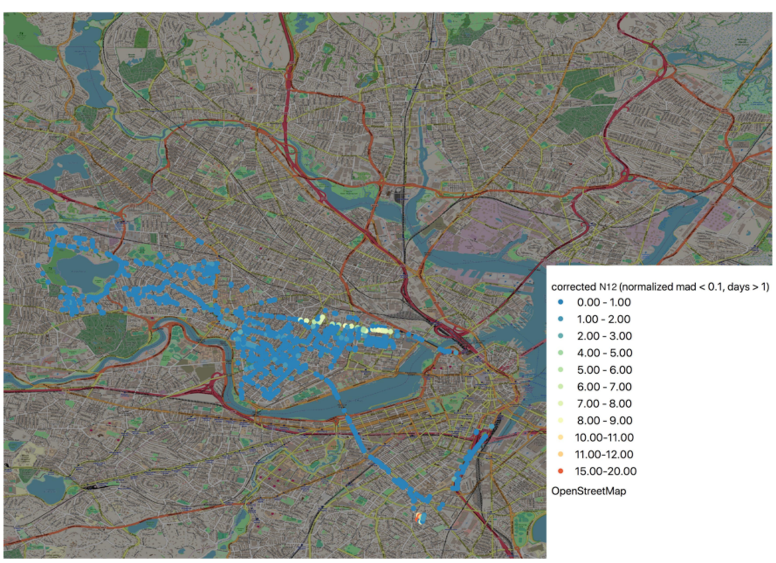
***Figure S4**: Average size distributions of each cluster of hotspots described in Figure S5. dN/dlnDp on the y-axis is in units of #number of particles/ ml for each ln(Dp) and d(lnDp). The table below the figure corresponds to the average PM1, PM2.5, PM10, N1 and N12 corresponding to each cluster*


corrected PM2.5 (normalized mad < 0.1, days > 1)
0.00 - 1.00
1.00 - 2.00
2.00 - 3.00
3.00 - 4.00
4.00 - 5.00
5.00 - 10.00
10.00 - 15.00
15.00 - 20.00
20.00 - 25.00
25.0-30.0
30.0-50.0
50.00 - 100.00
OpenStreetMap

a)


corrected N1 (normalized mad < 0.1, days > 1)
0.00 - 1.00
1.00 - 2.00
2.00 - 3.00
3.00 - 4.00
4.00 - 5.00
5.00 - 6.00
6.00 - 7.00
7.00 - 8.00
8.00 - 9.00
9.00-10.00
10.00-11.00
11.00-12.00
12.00-13.00
13.00 - 14.00
14.00-15.00
15.00-20.00
20.00-50.00
>50.00
OpenStreetMap

b)


corrected N12 (normalized mad < 0.1, days > 1)
0.00 - 1.00
1.00 - 2.00
2.00 - 3.00
4.00 - 5.00
5.00 - 6.00
6.00 - 7.00
7.00 - 8.00
8.00 - 9.00
10.00-11.00
11.00-12.00
15.00-20.00
OpenStreetMap

c)

Figure S5: *a) Map of median* $PM_{2.5}$ *for each 30 m segment, b) Map of the median background-corrected number concentration of particles having diameters between 0.38 µm and < 1 µm (N1 or* $N_{0.38\text{-}1}$*) c) Map of the median background-corrected number concentration of particles having diameters between 1 µm and < 11 µm (N12 or* $N_{1\text{-}12}$*), all of which had been sampled on more than one day, and where the normalised error in the median* $PM_{2.5}$ *derived from the bootstrapping is ≤ 20%*

Figure S6: *Map of most frequent cluster for each 30-meter road segment in Cambridge.*

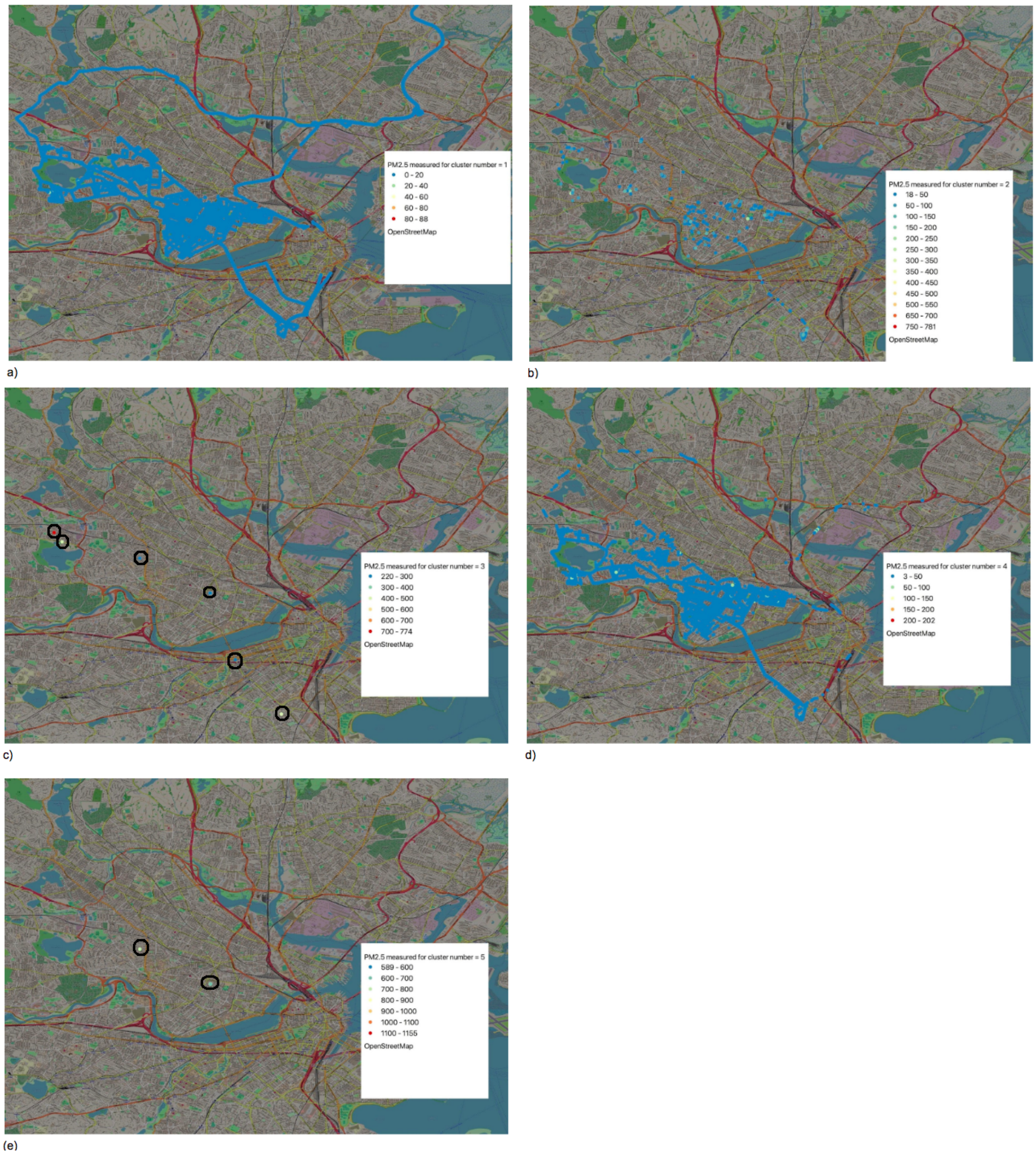


Figure S7*: Top: Map of $PM_{2.5}$ corresponding to a) cluster number 1, b) cluster number 2, c) cluster number 3, d) cluster number 4, e) cluster number 5. The clusters corresponding to cluster 3 and 5 are highlighted with circles, as the lone points are hard to see*