Opportunistic mobile air pollution monitoring: A case study with city wardens in Antwerp

Joris Van den Bossche [a, b, *], Jan Theunis [a], Bart Elen [a], Jan Peters [a], Dick Botteldooren [c], Bernard De Baets [b]

[a] VITO — Flemish Institute for Technological Research, 2400 Mol, Belgium
[b] KERMIT, Dept. of Mathematical Modelling, Statistics and Bioinformatics, Faculty of Bioscience Engineering, Ghent University, 9000 Ghent, Belgium
[c] INTEC, Dept. of Information Technology, Faculty of Engineering and Architecture, Ghent University, 9000 Ghent, Belgium

HIGHLIGHTS

- An opportunistic mobile monitoring campaign is conducted with the collaboration of city employees.
- The opportunistic approach can identify broad spatial trends of the urban air quality.
- It is a challenge to collect sufficient data to cover both spatial and temporal variability.
- The evaluation of this set-up is relevant for participatory and crowd-sourcing campaigns.

GRAPHICAL ABSTRACT
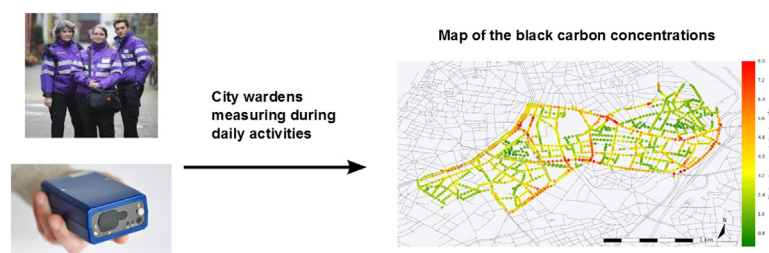
ABSTRACT

The goal of this paper is to explore the potential of opportunistic mobile monitoring to map the exposure to air pollution in the urban environment at a high spatial resolution. Opportunistic mobile monitoring makes use of existing mobile infrastructure or people's common daily routines to move measurement devices around. Opportunistic mobile monitoring can also play a crucial role in participatory monitoring campaigns as a typical way to gather data.

A case study to measure black carbon was set up in Antwerp, Belgium, with the collaboration of city employees (city wardens). The Antwerp city wardens are outdoors for a large part of the day on surveillance tours by bicycle or on foot, and gathered a total of 393 h of measurements. The data collection is unstructured both in space and time, leading to sampling bias. A temporal adjustment can only partly counteract this bias. Although a high spatial coverage was obtained, there is still a rather large uncertainty on the average concentration levels at a spatial resolution of 50 m due to a limited number of measurements and sampling bias. Despite of this uncertainty, large spatial patterns within the city are clearly captured.

This study illustrates the potential of campaigns with unstructured opportunistic mobile monitoring, including participatory monitoring campaigns. The results demonstrate that such an approach can indeed be used to identify broad spatial trends over a wider area, enabling applications including hotspot

---

* Corresponding author. KERMIT, Dept. of Mathematical Modelling, Statistics and Bioinformatics, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.
E-mail address: jorisvandenbossche@gmail.com (J. Van den Bossche).

identification, personal exposure studies, regression mapping, etc. But, they also emphasize the need for repeated measurements and careful processing and interpretation of the data.

1. Introduction

Air quality is of increasing concern, given its impact on human health (Pope et al., 2009; Lim et al., 2012). In the urban environment, the exposure to air pollution is largely influenced by the activity patterns of the population and the high spatial and temporal variability in pollutant concentrations (Setton et al., 2011; Dons et al., 2012). Especially traffic-related pollutants such as NOx, ultrafine particles (UFP) and black carbon (BC) show large differences in concentration levels in space and time on a small scale (Vardoulakis et al., 2011; Peters et al., 2014). It is important to take this within-city variability into account for accurate exposure estimation (Fruin et al., 2014). Traditional central monitoring stations may not always accurately characterize the spatial variability in the surrounding area and may thus not be representative for the whole city (Wilson et al., 2005). In contrast, mobile platforms are able to acquire air quality data at a high spatial resolution. As such, mobile data are increasingly used to assess the variability within the urban environment and to map the concentration levels people are effectively exposed to (Peters et al., 2014). Given the high temporal variability of urban air quality, a limited number of mobile measurements may only represent a snapshot and may thus not be representative. To map the urban air quality in a reproducible way and at a high spatial resolution, a sufficient number of repeated measurements is required (Van den Bossche et al., 2015). However, this is generally much more labour intensive compared to stationary measurements. One way to collect the large amounts of data that are needed, is to take advantage of existing mobile infrastructure or people's common daily routines to move measurement devices around through the city, without specifically designing the travelled route of the mobile carrier for the measurement campaign. We will further call this *opportunistic mobile monitoring*. Different forms of opportunistic data collection can be distinguished, and they will be described in more detail in the next section. Opportunistic mobile monitoring can also play a crucial role in participatory monitoring campaigns as a typical way to gather data.

In this paper, we describe a case study using such an opportunistic data collection scheme, consisting of mobile measurements of black carbon (BC) and carried out with the collaboration of city wardens of the city of Antwerp. As a consequence of their working routines, there is no control over the followed route and the campaign has an unstructured set-up. The data collection is unevenly spread in space and time, leading to sampling bias as different locations will possibly be measured at different days and/or hours of the day and complicating the data interpretation. This is a major problem which has to be addressed to be able to compare the results between these locations and to use the results for air quality mapping. The issues discussed in Van den Bossche et al. (2015) concerning the representativeness of mobile measurements, the sensitivity to peaks and varying background concentrations, will now be even more important, and will have to be re-evaluated.

The goal of this paper is to explore the potential of an opportunistic mobile monitoring approach with an unstructured set-up, as could for instance be the case in an unstructured participatory monitoring campaign, to obtain a reliable, high-resolution map of the urban air quality (not a city-wide map but restricted to the measured locations). It will be evaluated whether or not the recommendations for mapping the urban air quality with a structured mobile monitoring campaign as outlined in Van den Bossche et al. (2015) can be applied here. The focus of this paper is on the characterization of the sampling bias, its impact on the result and ways to counteract the sampling bias through the application of temporal adjustment. The resulting air pollution concentrations will be compared to concentrations from a targeted campaign.

2. Opportunistic mobile monitoring

We define *opportunistic mobile monitoring* as data collection making use of existing carriers to move measurement devices around. The movement of the carriers (the travelled route) is uncontrollable from the point of view of the researcher, as it is not designed and performed with the data collection in mind as primary goal. The data collection takes advantage of existing mobile infrastructure or people's common daily routines. This contrasts with *targeted mobile monitoring*, which is a coordinated, goal-driven approach in which the mobile measurements are deliberately planned and carried out with a specific purpose in mind (see e.g. case studies described in Peters et al. (2013, 2014) and Van den Bossche et al. (2015)). Opportunistic mobile monitoring is a promising approach to collect large data sets that give useful additional information at a reasonable cost compared to classical data collection methods. But, depending on the set-up of the data collection, such new data can lead to new challenges in data processing and interpretation.

This is closely related to opportunistic people-centric sensing as presented by Campbell et al. (2008). They describe a new sensing paradigm leveraging humans as part of the sensing infrastructure. By using small computational devices carried by individuals in their daily activities, information related to human activity and to the environment around them can be sensed opportunistically (Campbell et al., 2008; Kapadia et al., 2009; Kumar et al., 2015). As the data originates from sensors carried by people, new challenges for information security and privacy have to be addressed (Kapadia et al., 2009). However, in this paper, we do not restrict it to people and focus more on the uncontrolled aspect of the data collection and its implications for the processing and interpretation of the results.

Opportunistic data collection can take different forms. We can consider two axes along which opportunistic data collection campaigns can vary. Firstly, they can vary according to the degree of human interaction they need. Possible human interactions are related to carrying the measurement system, the operation and maintenance of the measurement system and to the data collection and handling. Examples of campaigns that can run independently for long periods without human interaction after initial set-up are those based on sensors mounted on vehicles such as cars, buses or trams. The more human interaction the data collection needs, the more the user-friendliness of the instrument and the motivation of the people involved become important issues. Secondly, the data collection can follow a repeated structure along the same routes and/or within the same time frame or can be rather unstructured. The studies of Hasenfratz et al. (2015) and Hagemann et al.

(2014) are examples of opportunistic data collection with a fixed structure, as they performed mobile measurements with sensors installed on the roof of public transport vehicles (trams). In these cases, the measurements are restricted to the route of the bus or the tram tracks. Another example of fixed routes are commuters performing mobile measurements (e.g. Weichenthal et al., 2008), although in this case the route taken is more flexible. Aoki et al. (2009), who built an environmental air quality sensing system and deployed it on street sweeping vehicles, can be situated somewhat in the middle between fixed or unstructured set-up. The street sweepers most likely still follow a rather fixed route, but will already cover a larger number of streets. The case study described in this paper with city wardens carrying measurement devices during their daily surveillance tours results in unstructured measurements without distinct patterns in space or time. The studies of Hasenfratz et al. (2015) and Hagemann et al. (2014) require limited human interaction as the sensor nodes are supplied with power from the vehicles and data are transmitted automatically (Hasenfratz et al., 2015). When using humans as the mobile carrier (Weichenthal et al., 2008; this paper), more human interaction is needed. Other possibilities to gather data opportunistically would be to work with postmen, taxi drivers, parking wardens, and so on.

Opportunistic data collection is not restricted to the air quality examples above. Noise mapping using smartphone apps (e.g. NoiseTube, D'Hondt et al., 2013) can also be done in an opportunistic way. Floating car data coming from so-called probe vehicles (i.e. vehicles that are equipped with the necessary devices to transmit data to a data centre at regular time intervals) are used to map the congestion and speed. These probe vehicles can be deployed specifically (targeted or irregular), but this information can also be retrieved from existing uncontrolled cars equipped with a GPS.

The characteristics of opportunistic monitoring as outlined above hold some consequences for the processing and interpretation of the data. From the perspective of the researcher, there is no control over the specific location and time of the measurements. This could result in sampling bias where certain urban microenvironments are underrepresented or absent in the data. For example, a postman or a parking warden will not frequently enter green zones such as parks. Similarly, there could also be a bias due to the time of the measurements. In the case of data collection by commuters, the measurements are mainly limited to the rush hours. The sampling can also be biased by the weather conditions, e.g. when the data collection stops when it rains, or when the commuter takes the car instead of the bicycle on rainy days. As a consequence, there will be different measurement conditions for different locations, hindering the comparability of the results between these locations. This is a major problem, as it complicates the data interpretation (the comparison of the measured concentrations at the different locations), making it less evident to use the results for air quality mapping.

2.1. Participatory air quality monitoring

Opportunistic mobile monitoring is also important in the context of the growing interest in participatory air quality monitoring. Already in 2006, Burke et al. mentioned the theoretical potential of participatory sensing to investigate the relationship between air quality, urban traffic and public health. In Snyder et al. (2013), US-EPA scientists give an overview of possible changes in air quality monitoring due to the materialization of lower-cost, easy-to-use, portable air pollution monitors (sensors) that provide high-time resolution data in near real-time. Following successful examples of large-scale collaborative efforts such as OpenStreetMap or Wikipedia, bottom-up Do-It-Yourself (DIY) monitoring approaches based on low-cost sensors and smart phone apps are appearing in domains such as noise, air quality or radiation monitoring. Often the idea of pervasive or ubiquitous sensing is put forward, relying on a multitude of sensors with which data are collected in an un-coordinated almost effortless way, and on intelligent data post processing and mining.

Several research groups have developed portable devices, integrating commercially available low-cost gas sensors, GPS and mobile phones (Dutta et al., 2009; Zappi et al., 2012). However, none of these sensors can be used as such to measure outdoor air quality. Snyder et al. (2013) indicate that many commercially available sensors have not been challenged rigorously under ambient conditions, including both typical concentrations and environmental factors. Accurate portable instruments are available for components such as ultrafine particles and black carbon. They can be used by non-specialist users (Buonocore et al., 2009; Dons et al., 2011), but they are still too expensive to use at a large scale. Hence, the hypothesis that the aggregation of a multitude of data that are collected in an un-coordinated way will lead to a useful result cannot be tested at this point in time. The case we present in this paper somehow mimics an unstructured participatory monitoring campaign using actually available sensors. In the future, such campaigns could potentially manifest much of the same characteristics.

3. Materials and methods

3.1. Opportunistic measurement campaign with city wardens

A case study was set up to map the average exposure to black carbon (BC) in streets and public spaces in Antwerp (51°12′ N, 4°26′ E, medium-sized city of 480,000 inhabitants, 985 inhabitants km^{-2}), Belgium, during a one-year period (from July 2012 until June 2013). The case study was carried out with the collaboration of city wardens and the Environmental Services of the city of Antwerp. The Antwerp city wardens are city employees with a surveillance task. They are outdoors for a large part of the day carrying out surveillance tours by bicycle or on foot. These surveillance tours do not follow fixed routes or times. Three teams of 2 city wardens each were equipped with a measurement unit, using the VITO airQmap platform[1] (see Fig. 1a). The measurement unit consists of a micro-aethalometer (MicroAeth Model AE51, AethLabs), a lightweight sensor that allows to measure BC at a high (1 s) frequency, and a GPS (Locosys Genie GT-31 GPS). The airQmap platform further consists of a home station, a netbook with custom-made, easy to use software to read out the measurement devices, keep their clocks synchronized and to transmit the measurements over the internet. The city wardens were asked to read out and charge the measurement devices on a daily basis, and to replace the filters of the micro-aethalometers every second day. There was limited follow-up apart from technical support to the instrumentation. Because of the risk of damage to the micro-aethalometer, the city wardens were asked not to take measurements when it was raining. Each measurement is automatically linked to its geographic location by time-synchronizing the BC and GPS data. Further processing steps including the filtering for unreliable GPS data, BC noise reduction and calibration and spatial aggregation are described in more detail below. The study area (i.e. the surveillance area of the three teams together) covers approximately 3.6 km² (a quarter of the inner city), and is pictured in Fig. 2.

During the measurement campaign, 393 h of raw 1 s measurements were recorded for the three teams combined (459 h of measurements before filtering for GPS quality), spread over 110

---

[1] http://www.airqmap.com.

Fig. 1. (a) Micro-aethalometer and a GPS together in a belt bag as used in airQmap and (b) members of the 3 teams of city wardens measuring air quality in Antwerp during 1 year (© De Nieuwe Antwerpenaar).
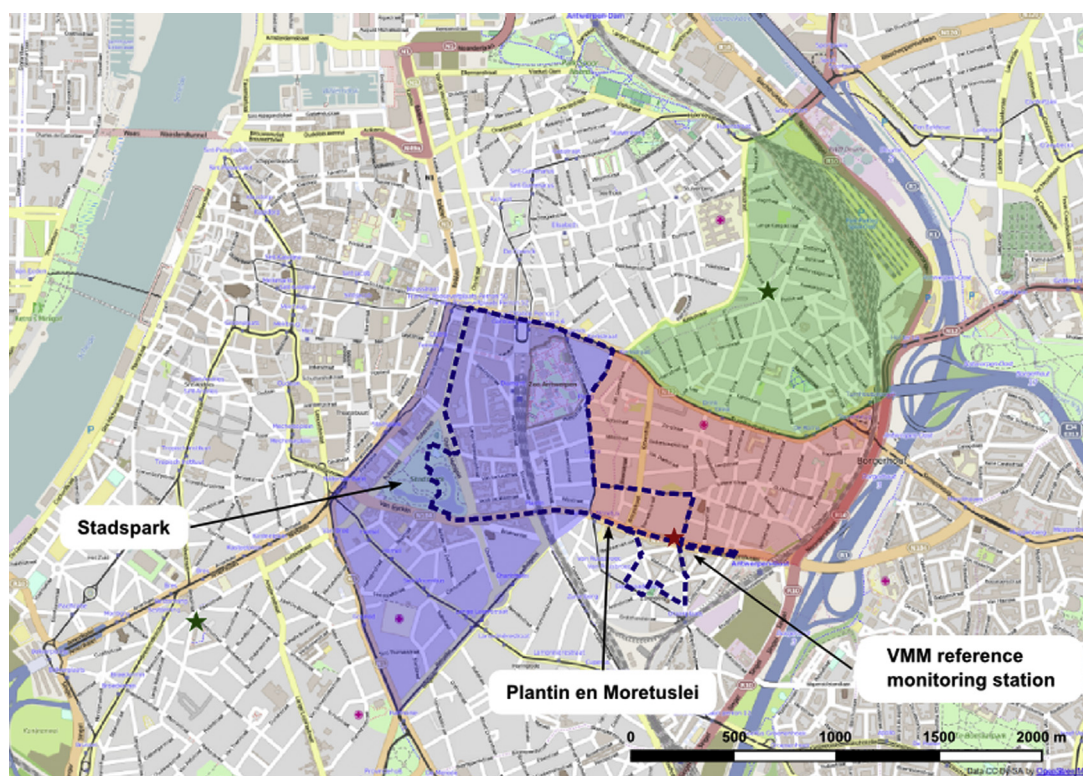

Fig. 2. Overview of the study area. The three coloured areas indicate the zones of the three teams of city wardens and the two green stars indicate their offices. The blue dashed line indicates the route taken in the targeted campaign described in Van den Bossche et al. (2015). The red star indicates the location of the official monitoring station. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

days. Most of the measurements were done between 10 a.m. and 16 p.m. during working days.

Additional stationary measurements from the VMM (Flemish Environmental Agency) official monitoring stations were available. At these stations, BC is measured with Multi-Angle Absorption Photometry instruments (MAAP, Petzold and Schonlinner, 2004; Petzold et al., 2005). One location with two monitoring stations, BETR802 and BETR801 (at the street side and 30 m away from traffic, respectively) is situated in the study area (red star in Fig. 2).

3.2. Quality control and processing of the BC data

The micro-aethalometer measures the concentration of optically absorbing aerosol particles (equivalent black carbon (EBC, in $\mu g\,m^{-3}$) using a mass-specific absorption cross-section (MAC) of 12.47 m^2g^{-1} at 880 nm (Petzold et al., 2013)) by the rate of change in light attenuation of a particle spot on a filter through which air is pumped. The inlet flow rate was set at 100 $mL\,min^{-1}$ and measurements were made at a temporal resolution of one second. The response time is negligible. The optical measurement of the filter is instantaneous and the lag in time due to pumping of the air

through the tube is less than a second for a flow rate of 100 mL min^{-1}. Measuring BC with the micro-aethalometer at one second time resolution is challenging due to the occurrence of signal noise, especially at low concentrations. Therefore, a noise reduction algorithm, the Optimized Noise-reduction Averaging algorithm (ONA algorithm from Hagler et al., 2011), was used with an attenuation threshold of 0.05.

The micro-aethalometer is known to be sensitive to the filter loading, which leads to an underestimation of the BC concentrations at higher ATN values. The filter ticket was changed every 2 measurement days, ensuring that 95% of all measurements have an ATN value below 100. To account for the remaining filter loading effects, the correction algorithm according to Virkkula et al. (2007) was applied. The parameter k of the correction algorithm is estimated based on the comparison to the MAAP instrument from the monitoring station (see Van den Bossche et al. (2015) for details).

Different micro-aethalometers were used throughout the campaign (different instruments for each team, and different instruments were used in the first half and second half of the campaign). For data quality control, the individual micro-aethalometers were all deployed at the VMM monitoring station at the Plantin en Moretuslei (station BETR802) to identify possible bias between instruments. After the filter loading correction and comparing to the mean of all instruments, the micro-aethalometers showed on average a mean absolute error of 0.4 μg m^{-3} at 1 min resolution and 0.1 μg m^{-3} at 30 min resolution (15% and 5% error, respectively). When comparing the micro-aethalometers to the MAAP of the reference station, the overall bias ranged between 0.04 and 0.27 μg m^{-3} (bias of 1—4%), and a mean absolute error of 0.4 μg m^{-3} was found at a 1 h resolution (15% error).

3.3. Quality control and processing of the GPS data

The raw measurements of the GPS device are filtered and processed to ensure data quality. The different processing steps consist of filtering for incorrect or unreliable GPS locations, map matching and spatial aggregation. Firstly, the GPS data are filtered by quality. Of all available GPS data, 71% had a fix (i.e. a minimum of three satellite signals to calculate the position); the remaining data probably consist of indoor periods during the day or outdoor moments with a very bad reception. Starting from these data, some filtering criteria were adopted to remove incorrect or unreliable GPS locations: 1) assuming all data were gathered by bike or on foot, the speed may not be above 25 km h^{-1} and 2) to ensure the GPS quality, the 30 s centred moving mean of the number of satellites should be above 4. After these filtering steps, 86% of the data with a fix remained, corresponding to 393 h of measurements.

Secondly, map matching was applied. In urban environments, the location determined by the GPS is often slightly off track. A commonly used approach to enhance the obtained GPS tracks is map matching, assuming the measurements are always performed on the streets. For each measurement day, those street sections are selected from the database at which measurements have been conducted over the entire section. In a next step, a simple map matching based on shortest distance (max 30 m) is applied for each measurement day separately with the selected street sections. This selection of streets per measurement day is done to exclude map matching to (side) streets that were not actually visited. During map matching, an additional 4% of the filtered GPS data is removed as they are not within 30 m of a street section.

The accuracy of the GPS location of the measurements is related to the quality of the GPS signal (GPS accuracy and projection accuracy) and the synchronization of the measurement device. The uncertainty of the GPS location is difficult to assess, as there is no reference to compare with. The GPS device (Locosys Genie GT-31

GPS) specifies an accuracy of the horizontal position of 10 m. At certain locations in urban environments, this accuracy can be worse due to reflection of signals by high buildings (observed from plotting the geographic data). The accuracy of the perpendicular projection on the streets could not be quantified exactly. The overall accuracy is probably around 10 m.

3.4. Data analysis: spatial aggregation, trimming and temporal adjustment

For further analysis of the mobile measurements, the data are spatially aggregated. For this purpose, the street sections are divided into segments of approximately 50 m. All measurements are allocated to such a segment. For each passage through a segment (*passage* = continuous period of time measurements are performed in one segment), an average BC value is calculated for the segment (= *passage mean*). To assure data quality, we require the calculated average to be based on at least 8 one second measurements (which corresponds to covering 50 m at an assumed maximum speed of 25 km h^{-1}). This removes an additional 3.5% of the data and leads to 40,284 recorded passages. The passage means are further averaged to obtain the mean concentration for each segment (= *segment mean*).

To reduce the impact of extreme values on the average concentration of a segment, a trimmed mean is applied as described in Van den Bossche et al. (2015). The α %-trimmed mean is computed by removing the α % largest and α % smallest values and computing the arithmetic mean of the remaining values. This is a statistical measure of central tendency that is less influenced by the effects of the tails of a distribution. As the air quality data in this study are positively skewed, removing the tail measurements will result in a biased estimator of the population mean. Indeed, this estimator will systematically underestimate the population mean. However, due to the reduced tail effects, the result will be a better estimate of the true mean (i.e. the mean including the tail measurements) when dealing with a limited number of repeated measurements (Van den Bossche et al., 2015). In this paper, a value of 0.5 is used for α. For samples of up to 200 passages, this corresponds to the removal of the highest and lowest passage mean in the calculation of the segment mean (the passage mean itself is not calculated using trimming). Brantley et al. (2014) also discussed processing strategies for mobile monitoring data, and a comparison with the above method is discussed in more detail in Van den Bossche et al. (2015).

The statistical comparison of the pollutant concentrations for cycling and walking data is based on the Kruskal-Wallis test at a critical value of 0.05 (Kruskal and Wallis, 1952). The Kruskal-Wallis test is a non-parametric alternative to a one-way ANOVA and investigates whether the measurements come from the same population.

The mobile measurement data are temporally adjusted based on hourly reference values to minimize the influence of both meteorological day-to-day variations and within-day variation in the urban air quality. This adjustment is performed through a combination of the additive and multiplicative method as used in Dons et al. (2012). The adjusted values are calculated as:

$$BC_{adj,i} = BC_i - BC_{ref,i} + \overline{BC_{ref}} \quad \text{(additive)} \tag{1a}$$

$$BC_{adj,i} = BC_i \Big/ BC_{ref,i} \cdot \overline{BC_{ref}} \quad \text{(multiplicative)} \tag{1b}$$

with BC_i the measured BC concentration at time i, $BC_{ref,i}$ the reference concentration measurement at time i, and $\overline{BC_{ref}}$ the mean reference concentration for the full period. The additive method is

applied to high concentrations (measurement is greater than the reference value) and the multiplicative method to lower concentrations (measurement is smaller than the reference value), see Dons et al. (2012) for more details. The adjustment is applied to passage means, and the hourly reference concentrations are obtained from the stationary measurements at VMM stations in the Antwerp region (mean of the stations BELAL01, BETM802, BETN016, BETR801 and BETR802).

Analyses were performed using Python pandas (PyData Development Team, 2015) and QGIS (QGIS Development Team, 2015).

4. Results

4.1. Characterization of sampling bias

4.1.1. Spatial coverage

Due to the nature of their working routines, the city wardens pass by at varying locations in the city. After one year of measurements, this resulted in a high spatial coverage. Fig. 3 shows a map with the number of times they passed by at a certain segment. Almost all street segments were covered at least once within the study area (see Fig. 2), and 74% of the segments were measured at least 5 times. Only those segments are shown in Fig. 3. The number of measurements varies greatly in space. Some segments were measured more than 400 times, typically those locations close to the wardens' office (start and end point of the measurements). Most segments, however, were measured 9 to 27 times (interquartile range).

4.1.2. Temporal coverage

Apart from the varying spatial coverage among street segments, also their temporal coverage varies. When looking at all data together, the temporal coverage (distribution of the number of passages in time over the day and over the year) is not uniform (Fig. 4). More measurements were performed at the beginning and the end of the measurement campaign, resulting in more measurements during the summer compared to the winter months.

Further, during the day, most measurements were performed between 9 and 16 h. Combined with varying typical concentrations in these periods of the year and the day (illustrated with the concentrations of the official monitoring station that lies within the study area, Fig. 5), this will introduce a bias in the results. The winter months typically have higher concentrations, but less measurements are conducted during this period. The same applies to the morning and evening rush hours. The morning rush is not covered by the measurements, resulting in lower concentrations. Furthermore, this temporal coverage also varies between locations. Examples are given for the Stadspark and Provinciestraat (Fig. 6), where in the Provinciestraat relatively much more measurements were collected in the morning hours. This varying temporal coverage between locations will introduce a different bias between segments.

4.1.3. Mode of transport: walking and cycling

Another possible source of bias in the data set is the mode of transport. The data are gathered both by bike and on foot. It is possible that the cyclists are exposed to higher concentration levels compared to the pedestrians (as being in general closer to the traffic). Next to the difference in temporal and/or spatial coverage, this can be an additional source of variation in the measured concentration levels when data of the two modes of transport are mixed.

The data set is divided by mode of transport (on foot or by bike). Based on the speed measurements of the GPS, a simple division is made based on a cut-off of 5 km/h on the 1 min rolling average speed. The data with low speed are classified as walking data, and those with a speed above 5 km/h as cycling data. Short stops during cycling are dealt with by using a rolling average of the speed. Of all second measurements, 25% are classified as cycling data. Both subsets are processed in the same way as described in Section 3.4 (by taking passage means).

The measured BC concentrations while cycling have a clearly higher overall average: 4.1 vs 2.6 $\mu g\,m^{-3}$ for cycling and walking data, respectively. However, the measurements are not necessarily made at the same locations. When using only those segments that

Fig. 3. Number of passages at each segment. Segments with less than 5 passages are excluded, and the colour scale is truncated at 60 passages (some segments near the office of the city wardens have more than 400 passages).
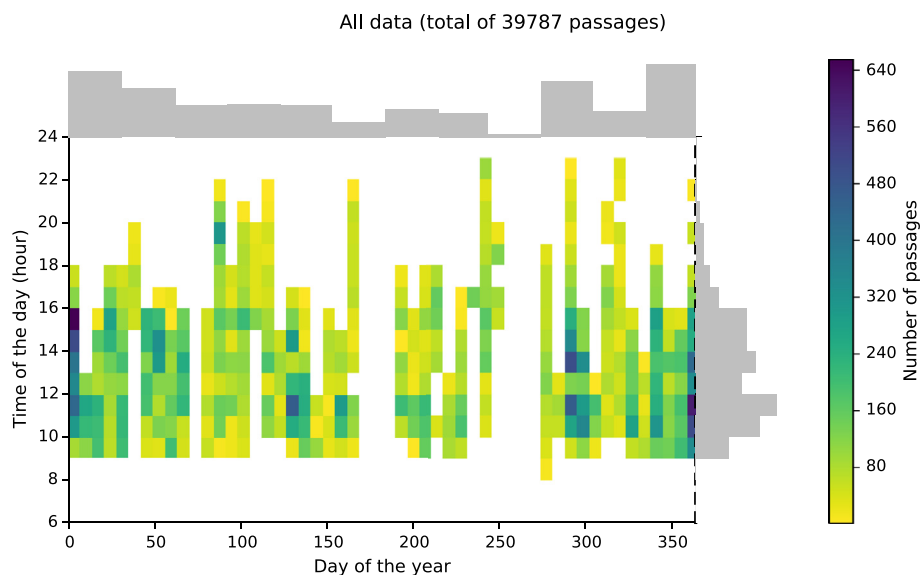
All data (total of 39787 passages)

Fig. 4. Overview of the temporal distribution of the number of passages in time. The horizontal histogram indicates the distribution over the year (per month) and the vertical histogram the distribution over the day (per hour). The internal plot is a combination of both and indicates the number of passages for each hour and each week (the legend for this figure is the colour scale on the right).
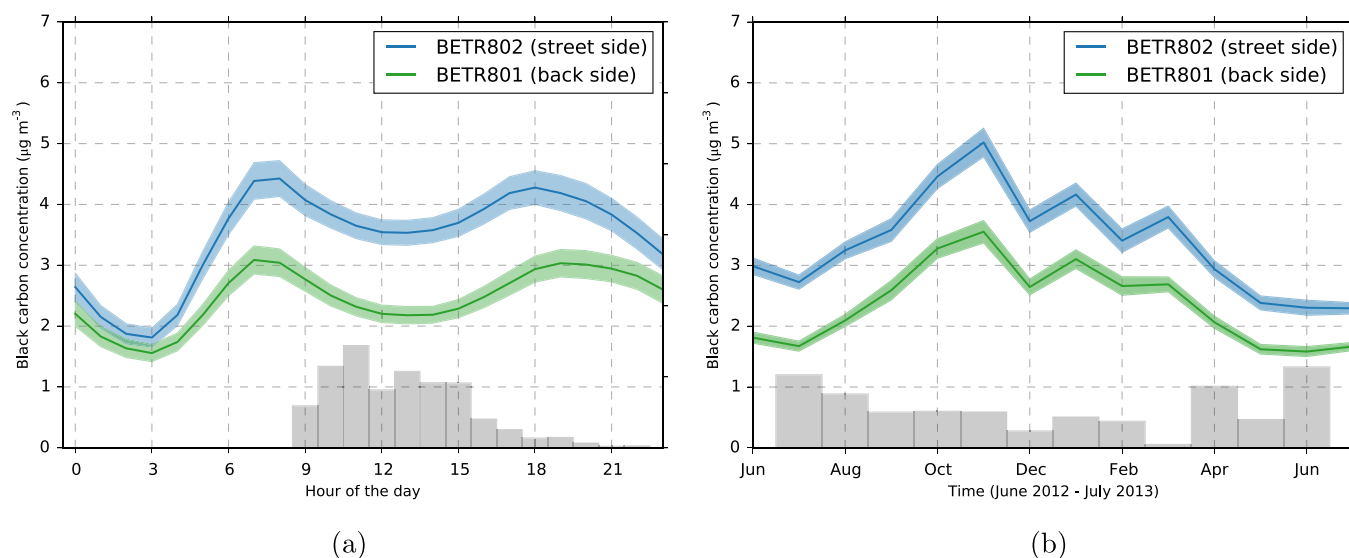

(a)

(b)

Fig. 5. BC concentration at the reference monitoring station near the study area (for two nearby stations, BETR801 and BETR802). Typical daily profile of the concentrations (hourly values averaged over all days during time frame of the campaign) is given in (a) and the monthly averages during the time frame of the campaign in (b). The histogram at the bottom indicates the relative distribution of the mobile measurements.

have data for both modes, the difference decreases (averages of 3.6 vs 3.0 $\mu g\,m^{-3}$, respectively). Fig. 7 also illustrates this difference in a histogram of both data sets. The cycling data are slightly shifted towards higher values. Based on the non-parametric Kruskal-Wallis test, this difference is significant (P < 0.05). Given the distribution of the passage means (Fig. 7), the difference of 0.6 $\mu g\,m^{-3}$ is, however, not large; it is smaller than the standard deviation in passage means (1.8 and 1.5 $\mu g\,m^{-3}$ for cycling and walking data, respectively).

The results above indicate that there is an observable difference in the concentration levels of cycling and walking data. This can be a systematic difference, due to a higher exposure of cyclists. But first some other factors have to be eliminated. Another possible explanation of the difference is the time period of the

measurements. The distribution over the year is not notably different between cycling and walking data, but there are some differences in distribution over the day. The cycling subset has relatively more measurements in the morning (9–12 h), while the concentrations are typically higher in the morning than during the rest of the day (typically 18% higher concentration in the morning compared to the afternoon). This can also contribute to the difference in concentration levels between cycling and walking data. The difference in distribution is, however, small (the morning hours constitute 53% of the data for the cycling subset, and 45% of the data for the walking subset) and does not fully explain the observed difference in concentration levels. Furthermore, when looking at the diurnal pattern, the difference in concentration levels is rather constant during the day. There is also no clear spatial pattern in the
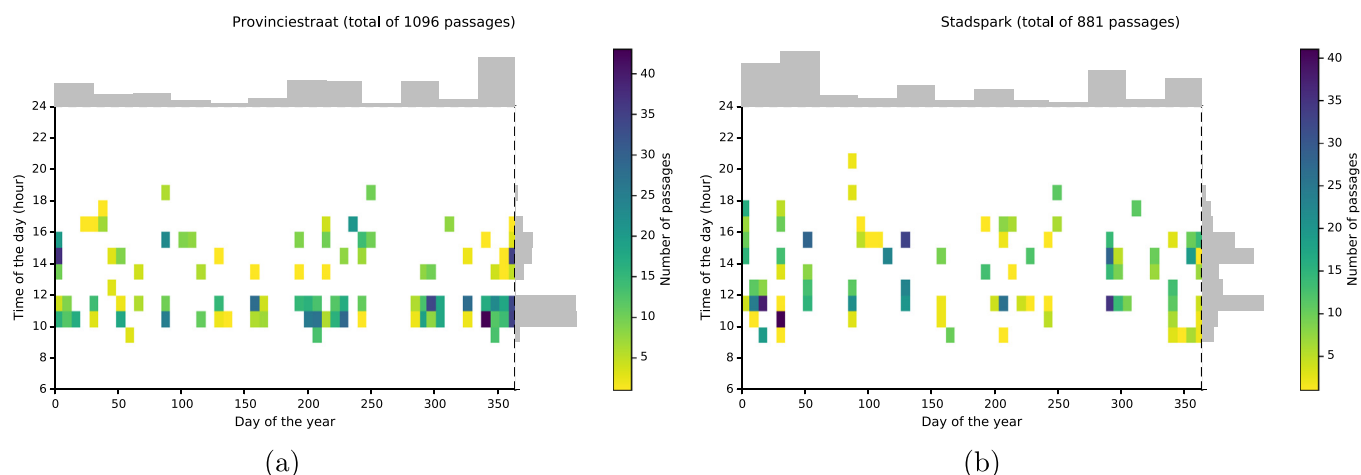
Fig. 6. Overview of the temporal distribution for two different locations: (a) Provinciestraat and (b) Stadspark. The horizontal histogram indicates the distribution over the year (per month) and the vertical histogram the distribution over the day (per hour). The internal plot is a combination of both and indicates the number of passages for each hour and each week (the legend for this figure is the colour scale on the right).
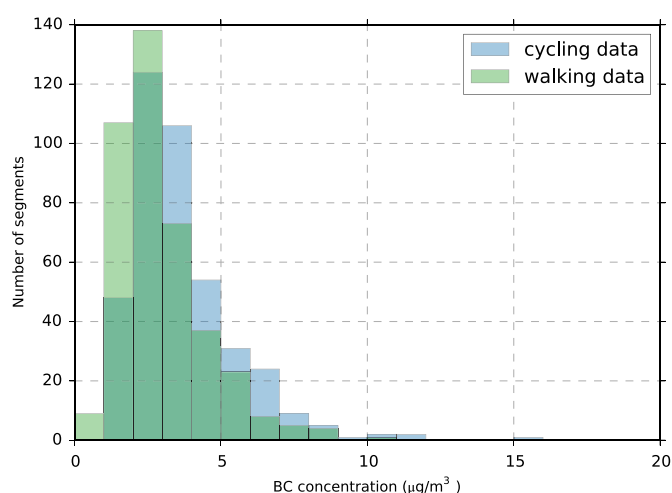

Fig. 7. Histogram of the BC concentration levels (segment means) for cycling and walking data: two partly overlapping histograms indicate the shift towards higher values for the cycling data. Only those segments where data are available for both bike and food mode are included.

difference between cycling and walking data. As a last check, the difference is also consistently present when looking at each of the three teams separately.

We can conclude that there is an observable difference between the exposure during cycling and during walking, however, this difference is not large. For the following analyses in this paper, we will combine the data of both modes of transport (unless explicitly stated).

4.2. Comparison with targeted measurements

A previous case study in the same area was performed and described in Van den Bossche et al. (2015). For this study, an extensive data set of BC measurements was collected by bike in February−March 2012. This targeted campaign consisted of 256 and 96 repeated measurement tracks along two fixed routes that for a large part overlap with the study area of the opportunistic campaign described in this paper (the route is indicated in Fig. 2). Given the high number of repetitions, we can use this data set to

benchmark the results obtained in this paper.

In Figs. 8 and 9, concentration profiles are shown for the Stadspark and surrounding streets and for the Plantin en Moretuslei (figures for the targeted campaign are adapted from Van den Bossche et al. (2015)). Broadly, the patterns in the concentration profiles correspond. For the Stadspark, the concentration increases sharply at the borders of the park towards the surrounding streets (with traffic) in both cases. For the Plantin en Moretuslei, the section between 200 and 600 m corresponds consistently to the section with elevated concentrations, although the exact locations of the highest values do not coincide in both campaigns. In the section between 600 and 1000 m, the Plantin en Moretuslei shows lower concentrations in both cases.

To quantify how good the results of both campaigns correspond, the overlapping 50 m segments are selected where at least 5 passages are recorded. For those segments the average concentration levels are calculated, and based on those values, the correlation (Pearson's correlation coefficient), overall difference (difference in overall mean of all segments) and mean absolute deviation (MAD) are calculated.

First, a basic comparison without adjustments is performed, resulting in a correlation of 0.64, overall difference of 2.7 $\mu g \, m^{-3}$ and MAD of 2.8 $\mu g \, m^{-3}$. This comparison is also illustrated with a scatter plot in Fig. 10. There is a large deviation in absolute value between both data sets: systematically lower values are found by the city wardens compared to the targeted campaign. Some possible explaining factors for this difference in concentration levels are a different mode of transport, period of the year, and period of the day during which the measurements are performed. Correction mechanisms for these factors are applied and a summary of the results of the comparison after correction is given in Table 1. Firstly, the targeted campaign only includes cycling data, while it was demonstrated in Section 4.1.3 that the bike data shows slightly higher concentration levels. When using only the cycling data for the current campaign, the overall difference decreases slightly to 2.5 $\mu g \, m^{-3}$. Secondly, the period of the day can also have an influence as the campaign with the city wardens also includes data in the afternoon, when the concentration levels are typically lower. Selection of the common hours (9−13 h) for both datasets without performing a temporal adjustment decreases the overall difference to 2.4 $\mu g \, m^{-3}$. Thirdly, the period of the year can have a large influence on the measured concentrations. This can be accounted for by performing a temporal adjustment to take the
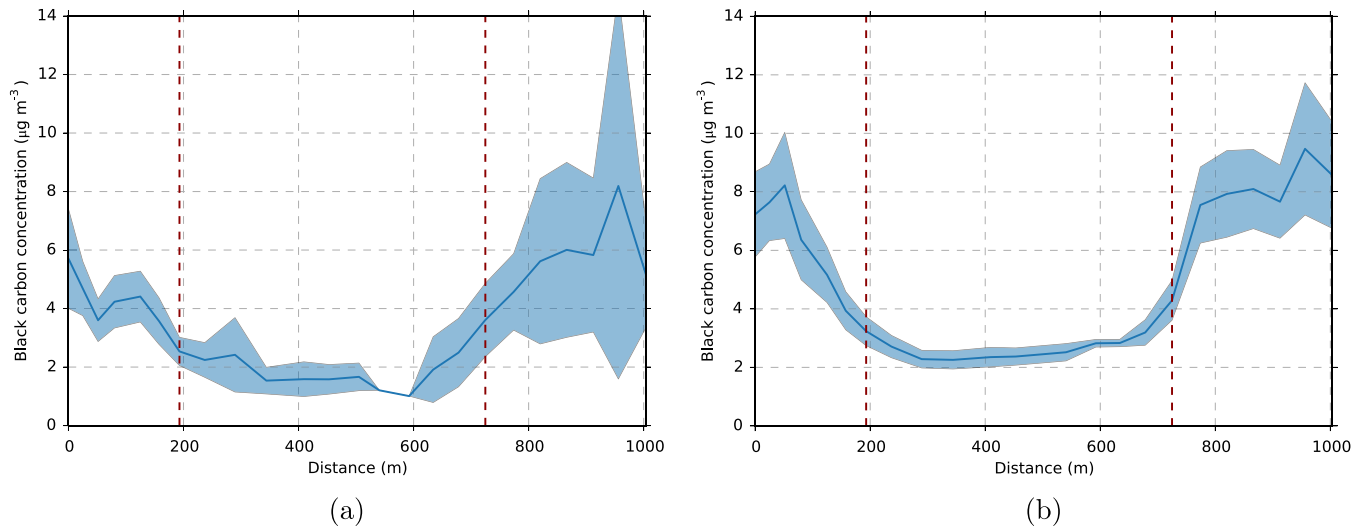
Fig. 8. BC concentration profiles of the Stadspark and surrounding streets for (a) the campaign with city wardens and (b) the targeted campaign (adapted from Van den Bossche et al. (2015)) at 50 m resolution. Shaded area depicts twice the standard error on the mean. The vertical dashed lines indicate the borders of the park.
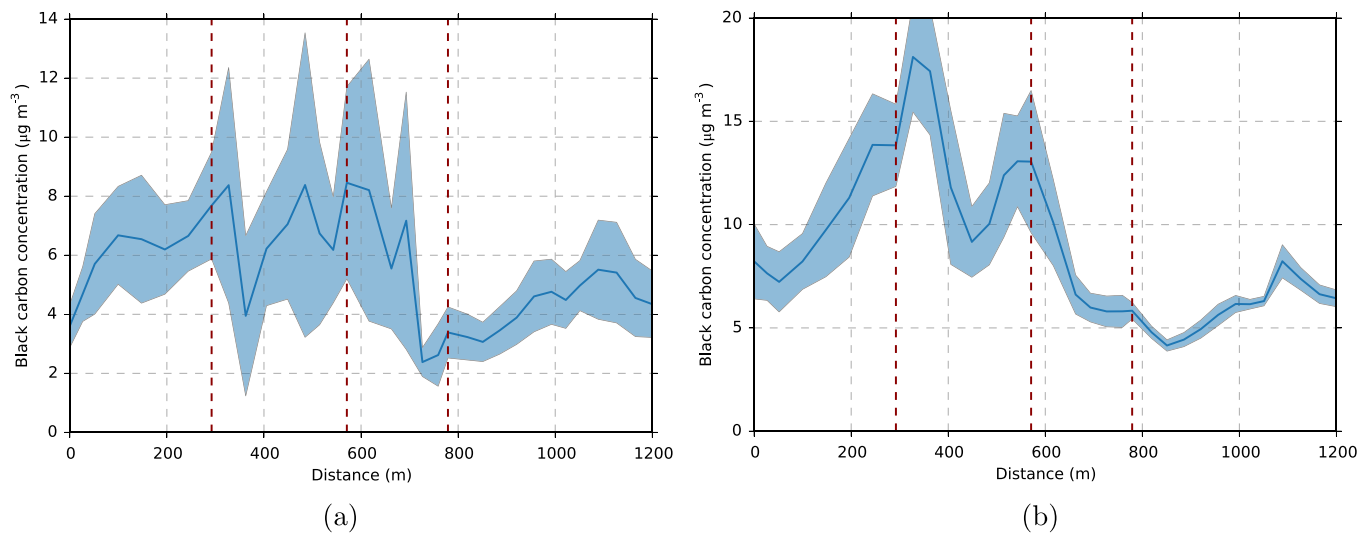

Fig. 9. BC concentration profiles of the Plantin en Moretuslei for (a) the campaign with city wardens and (b) the targeted campaign (adapted from Van den Bossche et al. (2015)) at 50 m resolution. Shaded area depicts twice the standard error on the mean. The vertical dashed lines indicate the three major crossings along the Plantin en Moretuslei.

seasonal and day-to-day variations into account. After performing such adjustment (see Section 3.4 on data analysis for more details), the overall difference decreases to 1.9 $\mu g\,m^{-3}$. However, when using a multiplicative adjustment instead of a combination, the decrease in overall difference is much stronger (down to an overall difference 0.6 $\mu g\,m^{-3}$). When combining these three correction mechanisms, the resulting overall difference is 1.7 $\mu g\,m^{-3}$ and MAD is 2.1 $\mu g\,m^{-3}$.

The same analysis is repeated for another spatial aggregation level, the street level. Instead of 50 m segments, the streets are used as aggregation level, which will result in segments of varying length. The result of the comparison at this level is also included in Table 1.

4.3. High spatial variability in air quality

The comparison in the previous section concerned only the part of the data set where the two campaigns overlapped. However, the spatial coverage of the campaign with city wardens was much larger. In this section, we show some results on the air quality for the entire study area.

First, a map is constructed displaying the BC concentration for each 50 m segment in the study area where at least 5 passages have been recorded (Fig. 11). The main roads clearly stand out, while the green areas and the more residential areas between main roads show much lower concentrations.

Going to a lower spatial resolution (all 50 m segments combined per street), Fig. 12 shows a boxplot of the BC concentration for a selection of streets. Large differences between streets with high BC concentrations and streets with lower concentrations can also be noticed at this level (mean concentration levels ranging from around 1 $\mu g\,m^{-3}$ up to 5 to 6 $\mu g\,m^{-3}$). Typical streets with high concentration levels are busy roads with high traffic intensities (e.g. Turnhoutsebaan, Carnotstraat, Belgiëlei, Noordersingel) and street canyons (e.g. Provinciestraat). The lower concentrations can be observed at green areas (Stadspark,
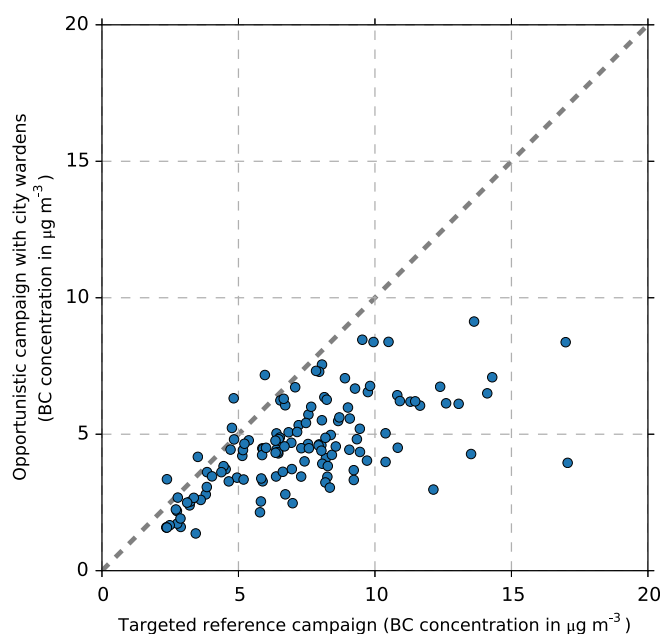
Fig. 10. Comparison of the concentration levels in the overlapping 50 m segments between the campaign with city wardens and the targeted campaign.

Krugerpark) and low-traffic residential streets (e.g. Van Leentstraat, Tuinbouwstraat).

5. Discussion

5.1. Sampling bias

The city wardens operated in a well-confined area in the city mostly during working hours on weekdays, resulting in repeated measurements. The spatial coverage within the study area is very high (Fig. 3), although there is a large variation in the number of passages between different segments. For a lot of the streets and street segments only a small number of repeated measurements were recorded. The temporal coverage (Fig. 4) also shows a non-uniformly spread pattern. Firstly, it shows that not all hours of the day and even not of the working day (9 a.m. - 17 p.m.) are measured proportionally. As a consequence of their work schedule, there are no measurements during rush hours when most people are commuting. Over the year, it can be noted that there are less measurements during winter. This can be caused by a lower motivation in the middle of the campaign, but also to worse weather conditions. Given the different typical concentrations during the day and the year, this will lead to sampling bias when targeting daily average concentrations.

Secondly, the temporal coverage also differs between locations:

Table 1
Comparison of the present opportunistic campaign with city wardens and the targeted campaign at two spatial aggregation levels (50 m segments and street level) and for different scenarios of data treatment. The evaluation statistics are the correlation (Pearson's correlation coefficient), overall difference (difference in overall mean of all segments or streets) and mean absolute deviation (MAD).

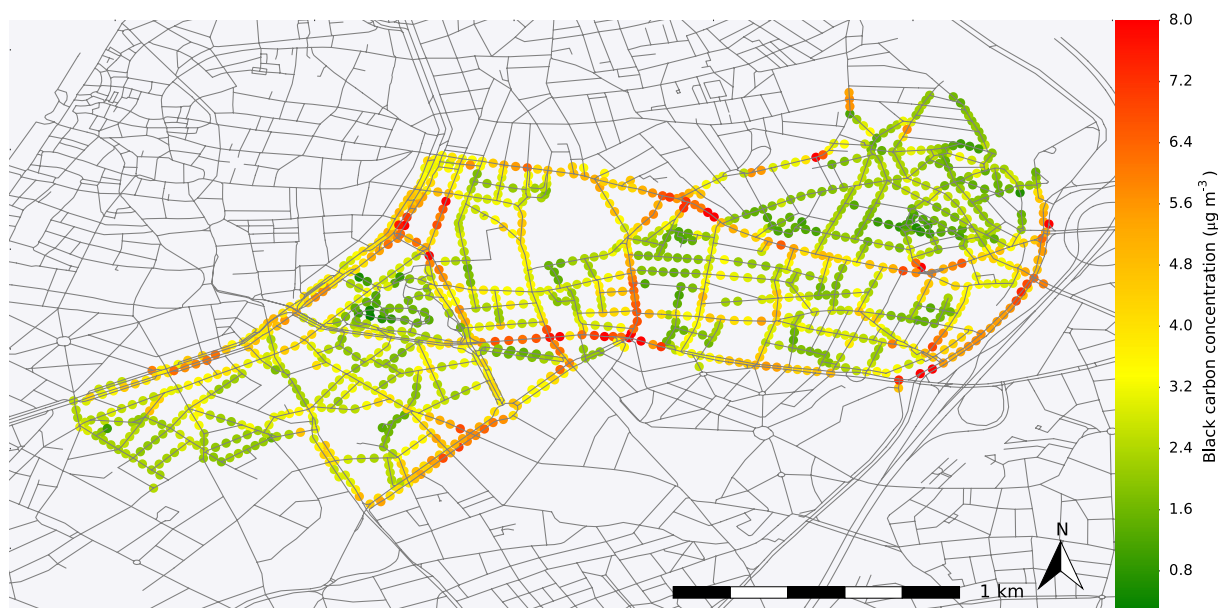|	50 m segments			Street level		
	Correlation	Overall diff.	MAD	Correlation	Overall diff.	MAD
Basic comparison	0.64	2.7	2.8	0.65	2.0	2.1
Only cycling subset	0.65	2.5	2.7	0.60	1.8	2.0
Common hours	0.65	2.4	2.5	0.53	1.7	2.0
Temporally adjusted	0.64	1.9	2.2	0.65	1.2	1.6
Temporally adjusted (multiplicative)	0.56	0.6	1.5	0.75	0.0	1.1
Combination of above	0.65	1.7	2.1	0.59	0.9	1.7

Fig. 11. Map of the BC concentration (50 m resolution): average concentration (trimmed mean) for each 50 m segment within the study area where at least 5 passages have been recorded.
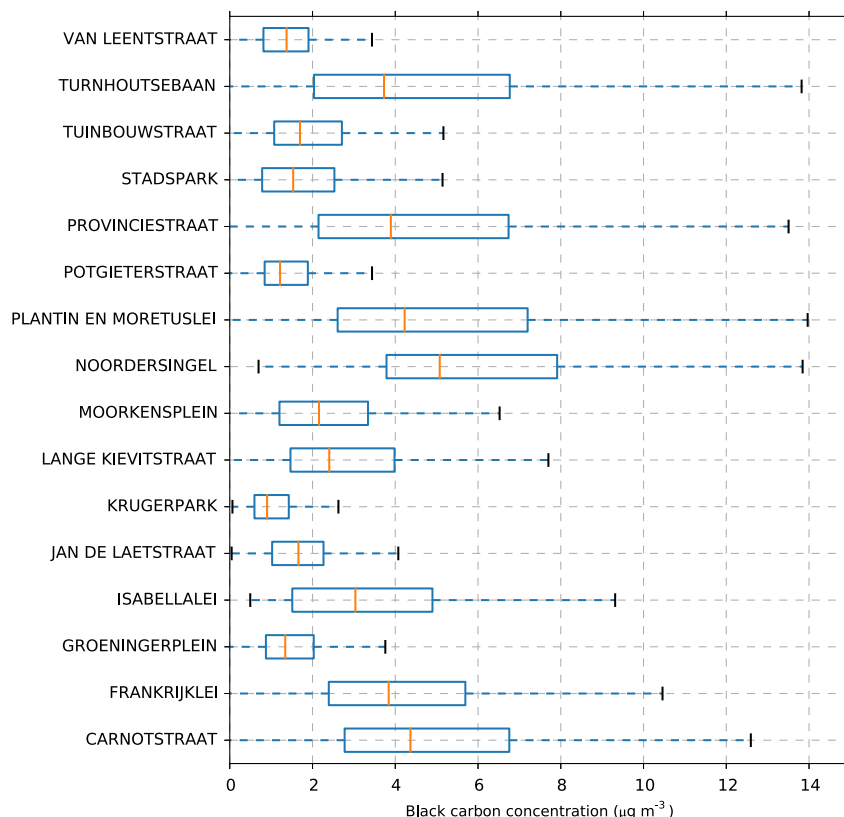
Fig. 12. Boxplots of the BC concentration for a selection of streets (using all passages of the different segments in the specified street).

different locations are measured at different days and/or hours of the day. This leads to a different bias at varying locations, complicating the interpretation of the results. When the concentration levels at two locations are compared, a possible difference can also be partially allocated to sampling bias. To counteract this sampling bias by the time of the measurement, a temporal correction can be used. This will never fully cancel the differences introduced by sampling at other moments in time, but can improve the comparability. Temporal adjustment was applied in the comparison to the targeted campaign, and is discussed there (Section 5.2).

The use of both walking and cycling modes of transport is another factor that complicates the interpretation of the air quality results. When dividing the dataset by mode of transport, there is a clear difference in average levels. After eliminating possible explaining factors (differences in location or time period of the measurements), the difference remains and it can be concluded that this difference is systematic. This can be explained by the different distance to traffic (bike lanes are typically closer to traffic than the pavement) and similar results are reported in the literature (McNabola et al., 2008; Dons et al., 2012; de Nazelle et al., 2012). By combining both data sources, the mixed modes of transport will be an extra factor contributing to the variation in the dataset. Data of both modes combined represent the exposure during active mobility (soft modes of transport).

In general, the data from opportunistic measurement campaigns is biased and this consequently affects the resulting image of the air quality. It is not possible to exactly quantify the overall sampling bias, as the 'true' concentration levels are not known. However, correction mechanisms can be applied to reduce the bias.

5.2. Comparison of the unstructured opportunistic campaign and the targeted campaign

The resulting concentration levels of this unstructured opportunistic campaign were compared to those of the targeted campaign from a previous study (Van den Bossche et al., 2015). This comparison allowed us to investigate whether the approach with the city wardens could lead to comparable results and whether the sampling bias can be reduced or even eliminated. When looking at the concentration profiles (Figs. 8 and 9) and given the correlation of 0.64 (Table 1), the larger spatial patterns (typical places with low or high concentrations, hotspots, …) resulting from both campaigns correspond.

But despite this correspondence, both a large overall difference and a large variability at the individual streets and street segments is present. The large overall difference in concentration levels is caused by the different factors contributing to the sampling bias, i.e. different period of the year, different period of the day and different mode of transport. Of course, other factors could also contribute to the sampling bias (e.g. instrumental bias), but these are not further addressed in this paper. Using only the cycling data slightly improves the comparison (reduction of the overall difference), but also decreases the number of measurements that can be used in the comparison, making the comparison less reliable. The same can be said for selecting the common hours.

The temporal adjustment is used to counteract the bias introduced by measuring on different days, and when using hourly reference values, this can also tackle the bias introduced by measuring at different hours of the day. Applying the temporal adjustment has a higher positive impact on the overall difference, but this depends largely on the used algorithm. In literature, temporal adjustment is a commonly used technique for normalizing

measurements that are not taken simultaneously (e.g. Hoek et al., 2002; Zuurbier et al., 2010; Eeftens et al., 2012; Tan et al., 2014). The adjustment can be additive (difference) or multiplicative (ratio), and it is generally stated that the two methods do not give very different results (Hoek et al., 2002). In the cited studies, the additive method was used. Dons et al. (2012) proposed a combined method to prevent negative values. These findings contrast with the result seen here, were the multiplicative method performs much better than the combined method. The multiplicative method is more suited to correct for temporal variation driven by meteorological conditions, while the additive method corrects for the influence of regional pollution (Tan et al., 2014). BC is mainly coming from local sources (traffic) in the study area, which is a possible reason that the multiplicative method works better in this study. It is clear that, although the correction mechanisms could reduce the difference between both data sets, a considerable difference between the targeted and opportunistic data set remains. Using a sampling experiment, Tan et al. (2014) also showed that correcting measurements from short duration campaigns with data from continuous stationary monitoring sites does not completely remove the uncertainty due to short duration sampling.

Besides the overall difference, there is a large variability on the individual segments (expressed by the mean absolute deviation, MAD). The artefacts are due to the short time period of the individual mobile measurements and the highly variable air quality concentrations at short temporal resolution, combined with a limited number of repeated measurements. The city wardens campaign still has not sufficient repetitions for all segments, and so a higher variability on the individual segments is to be expected. Most segments are measured between 9 and 27 times, while in the previous study on the targeted campaign it was concluded that 24 to 94 repetitions are required to map the BC concentrations at a 50 m resolution with an uncertainty of 25% (5–11 repetitions were required when relaxing this constraint to 50%) (Van den Bossche et al., 2015). We can conclude that not enough data were gathered to assess the air quality at such a high spatial resolution. When looking at a lower resolution (street level), the agreement between both data sets is better.

Despite the systematic difference between both data sets, we can conclude that the larger spatial patterns are captured by the opportunistic campaign, but that there still remains a large uncertainty on the individual segments.

5.3. Evaluation of the monitoring campaign with city wardens

Thanks to the collaboration with the city of Antwerp, we had the opportunity of setting up this campaign with the city wardens. The involvement of the environmental agency of the city and their interest in the results made it easier to engage city personnel for performing the measurements. Working with the city wardens offered some benefits compared to volunteers or other professionals. Because they are city personnel, all instruments were always connected and recharged at their office and the follow-up could be centralized. Further, the city wardens went to varying places in the city as indicated by the high spatial coverage (Fig. 3), e.g. also visited traffic-free and green areas. Other examples of carriers such as postmen or parking wardens can partly have the same advantages (involved through their job), but can have a more limited temporal or spatial coverage. For example, postmen or parking wardens will typically less likely visit traffic-free or green areas.

Some drawbacks or aspects that could be improved in the monitoring approach could be identified. A lot of variability was introduced due to several factors such as varying modes of transport, sampling during different hours of the day and of the year (see Section 4.1). This allowed us to evaluate the potential of a case study that manifests such characteristics, but for the sake of this campaign itself some optimizations could have been performed. A stronger follow-up of the wardens could have kept the level of measurements more constant during the year. The mixed mode of transport of the wardens is an additional complication and it should be avoided if possible (which was not the case in this campaign).

Another problem is the low GPS quality in urban environments, and consequently an intensive processing of these data is needed. In contrast to the targeted monitoring case, no a priori knowledge on the tracks that are monitored is available. Combined with the low GPS quality, this complicated the detailed mapping of the air quality. Due to a possibly large deviation from the true position, some measurements may not have been assigned to the correct street. Also, data are not only collected outdoors. It turned out to be hard to detect, in a reliable way, when a city warden enters a building. We have tried to tackle this by filtering the data on the number of satellites, but this may have led to some excessive data cleaning of good measurements on the one hand or the incorrect use of indoor measurements on the other hand. Possible solutions are the use of better GPS devices (but this can become expensive), better map-matching or transport mode detection algorithms or the use of diaries (this could give more contextual information to interpret the GPS data, but it also puts a much higher burden on the people carrying out the measurements).

A next issue is the instrument used to measure the air quality. The micro-aethalometer signals may sometimes be noisy or show instabilities. This will be an extra factor contributing to the uncertainty on the results. Also, the currently used bag does not protect the micro-aethalometer from the rain. A better system would make it possible to also measure when it is raining.

Collecting data consistently during a one year period by the city wardens is quite a challenge and turned out to be less evident than expected. Good usability of the monitoring equipment is crucial. The measurement set-up needs to be designed to keep the impact on the city wardens limited. Tasks that look simple at first such as making sure the battery of the measurement device is recharged, changing a filter, or turning off measurement devices after completion of the measurement day seem to take more effort than expected. The city wardens seemed to conceive the monitoring as an extra task to their daily job. A stronger follow-up could have kept the data collection at a more constant level throughout the year.

5.4. Relevance for participatory sensing and air quality applications

With the set-up of this campaign, we somehow mimicked an uncoordinated participatory monitoring campaign with actually available sensors. In the future, mobile participatory campaigns could potentially manifest much of the same characteristics of this data set in terms of uneven spatial and temporal coverage and different modes of transport with increasing numbers of data as monitoring equipment becomes cheaper. In this way, this case study emulates the kind of data that would be gathered from participatory campaigns in an achievable way for the time being. We demonstrated that such an approach can indeed be used to identify broad spatial trends and hotspots over a wider area. The results show that the spatial patterns in the urban BC concentrations can be assessed, obtaining a spatial resolution which would not be possible with stationary monitoring stations. This shows that the approach can be used for air quality applications including hotspot identification, personal exposure studies, regression mapping (using the results to build land-use regression models), etc. However, it is clear that repeated measurements and careful processing and interpretation of the results are very important. Even

with a rather large number of measurements (393 h), there is still a large variability on these results. This indicates that an important challenge for unstructured participatory monitoring campaigns will be to collect sufficient data to cover both spatial and temporal variability.

6. Conclusion

We defined *opportunistic mobile monitoring* as data collection making use of existing carriers to move measurement devices around. Based on an opportunistic mobile monitoring campaign with the collaboration of city employees (city wardens) carrying micro-aethalometers during their daily surveillance tours, we were able to map the exposure to black carbon in the urban environment. A high spatial coverage in the selected area was obtained during the one-year data collection. Due to the nature of the wardens' work, there is no control over the route followed. This results in an unstructured data collection both in space and time, leading to sampling bias. A temporal adjustment can partly counteract this bias, but will not fully remove the uncertainty associated with the temporal variation. There is still a rather large uncertainty on the average concentration levels at the individual locations (when looking at a high spatial resolution of 50 m) due to a limited number of measurements and sampling bias. Despite this uncertainty, large spatial patterns within the city are clearly captured. In anticipation of lower-cost sensors, this case study with city wardens demonstrates a promising approach to monitor urban air quality at a high spatial resolution based on opportunistic data collection.

This study illustrates the potential and limitations of participatory monitoring approaches. With the set-up of this campaign, we somehow mimicked an uncoordinated participatory monitoring campaign with the currently available means. If appropriate low cost instruments would be available, citizens can carry it around measuring the air quality in a non-coordinated way. The results in this study demonstrate that such an approach can indeed be used to identify broad spatial trends over a wider area, enabling applications including hotspot identification, personal exposure studies, regression mapping, etc. But they also emphasize the challenges with unstructured opportunistic mobile monitoring. It is clear that repeated measurements and careful processing and interpretation of the results are very important. Unstructured participatory monitoring campaigns will have to collect sufficient data to cover both spatial and temporal variability. Until low-cost air quality sensors are available that can collect data of sufficient quality in an effortless way, continued motivation and follow-up of participants is crucial.

References

Aoki, P.M., Honicky, R.J., Mainwaring, A., Myers, C., Paulos, E., Subramanian, S., Woodruff, A., 2009. A vehicle for research : using street sweepers to explore the landscape of environmental community action. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2009), pp. 375–384.

Brantley, H.L., Hagler, G.S.J., Deshmukh, P., Baldauf, R.W., 2014. Field assessment of the effects of roadside vegetation on near-road black carbon and particulate matter. Sci. Total Environ. 468, 120–129.

Buonocore, J.J., Lee, H.J., Levy, J.I., November 2009. The influence of traffic on air quality in an urban neighborhood: a community-university partnership. Am. J. Public Health 99 (3), S629–S635 (Figure 1).

Burke, J., Estrin, D., Hansen, M., Ramanathan, N., Reddy, S., Srivastava, M.B., 2006. Participatory sensing. In: Proceedings of the Workshop on World-Sensor-Web (WSW'06): Mobile Device Centric Sensor Networks and Applications, pp. 117–134.

Campbell, A.T., Eisenman, S.B., Lane, N.D., Miluzzo, E., Peterson, R.A., Lu, H., Zheng, X., Musolesi, M., Fodor, K., Ahn, G.-S., July 2008. The rise of people-centric sensing. IEEE Internet Comput. 12 (4), 12–21.

D'Hondt, E., Stevens, M., Jacobs, A., October 2013. Participatory noise mapping works! an evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. Pervasive Mob. Comput. 9 (5), 681–694.

de Nazelle, A., Fruin, S.A., Westerdahl, D., Martinez, D., Ripoll, A., Kubesch, N., Nieuwenhuijsen, M., November 2012. A travel mode comparison of commuters' exposures to air pollutants in Barcelona. Atmos. Environ. 59, 151–159.

Dons, E., Int Panis, L., Van Poppel, M., Theunis, J., Wets, G., 2012. Personal exposure to Black Carbon in transport microenvironments. Atmos. Environ. 55, 392–398.

Dons, E., Int Panis, L., Van Poppel, M., Theunis, J., Willems, H., Torfs, R., Wets, G., July 2011. Impact of time-activity patterns on personal exposure to black carbon. Atmos. Environ. 45 (21), 3594–3602.

Dutta, P., Aoki, P., Kumar, A., Mainwaring, A., Myers, C., Willet, W., Woodruff, A., 2009. Common Sense: participatory urban sensing using a network of handheld air quality monitors. In: Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, pp. 349–350.

Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dèdelè, A., Eriksen, K., Galassi, C., Gražulevičienė, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Katsouyanni, K., Korek, M., Krämer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskövi, E., Weinmayr, G., Brunekreef, B., Hoek, G., December 2012. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2-Results of the ESCAPE project. Atmos. Environ. 62, 303–317.

Fruin, S., Urman, R., Lurmann, F., McConnell, R., Gauderman, J., Rappaport, E., Franklin, M., Gilliland, F.D., Shafer, M., Gorski, P., Avol, E., 2014. Spatial variation in particulate matter components over a large urban area. Atmos. Environ. 83, 211–219.

Hagemann, R., Corsmeier, U., Kottmeier, C., Rinke, R., Wieser, A., Vogel, B., September 2014. Spatial variability of particle number concentrations and NOx in the Karlsruhe (Germany) area obtained with the mobile laboratory 'AERO-TRAM'. Atmos. Environ. 94, 341–352.

Hagler, G.S., Yelverton, T.L., Vedantham, R., Hansen, A.D., Turner, J.R., 2011. Post-processing method to reduce noise while preserving high time resolution in Aethalometer real-time black carbon data. Aerosol Air Qual. Res. 11 (5), 539–546.

Hasenfratz, D., Saukh, O., Walser, C., Hueglin, C., Fierz, M., Arn, T., Beutel, J., Thiele, L., January 2015. Deriving high-resolution urban air pollution maps using mobile sensor nodes. Pervasive Mob. Comput. 16, 268–285.

Hoek, G., Meliefste, K., Cyrys, J., Lewné, M., Bellander, T., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., Van Vliet, P., Brunekreef, B., 2002. Spatial variability of fine particle concentrations in three European areas. Atmos. Environ. 36 (25), 4077–4088.

Kapadia, A., Kotz, D., Triandopoulos, N., January 2009. Opportunistic sensing: security challenges for the new paradigm. In: 2009 First International Communication Systems and Networks and Workshops. IEEE, pp. 1–10.

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47 (260), 583–621.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., Britter, R., March 2015. The rise of low-cost sensing for managing air pollution in cities. Environ. Int. 75, 199–205.

Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H.R., Andrews, K.G., Aryee, M., Atkinson, C., Bacchus, L.J., Bahalim, A.N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M.L., Blore, J.D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N.G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R.T., Byers, T.E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J.S., Cheng, A.T.-A., Child, J.C., Cohen, A., Colson, K.E., Cowie, B.C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D.C., Devries, K., Dherani, M., Ding, E.L., Dorsey, E.R., Driscoll, T., Edmond, K., Ali, S.E., Engell, R.E., Erwin, P.J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M.M., Flaxman, S., Fowkes, F.G.R., Freedman, G., Freeman, M.K., Gakidou, E., Ghosh, S., Gio, December 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380 (9859), 2224–2260.

McNabola, A., Broderick, B., Gill, L., August 2008. Relative exposure to fine particulate matter and VOCs between transport microenvironments in Dublin: personal exposure and uptake. Atmos. Environ. 42 (26), 6496–6512.

Peters, J., Theunis, J., Van Poppel, M., Berghmans, P., 2013. Monitoring PM10 and ultrafine particles in urban environments using mobile measurements. Aerosol Air Qual. Res. 13, 509—522.

Peters, J., Van den Bossche, J., Reggente, M., Van Poppel, M., De Baets, B., Theunis, J., 2014. Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. Atmos. Environ. 92, 31—43.

Petzold, A., Ogren, J.A., Fiebig, M., Laj, P., Li, S.-M., Baltensperger, U., Holzer-Popp, T., Kinne, S., Pappalardo, G., Sugimoto, N., Wehrli, C., Wiedensohler, A., Zhang, X.-Y., August 2013. Recommendations for reporting "black carbon" measurements. Atmos. Chem. Phys. 13 (16), 8365—8379.

Petzold, A., Schloesser, H., Sheridan, P.J., Arnott, W.P., Ogren, J.A., Virkkula, A., January 2005. Evaluation of multiangle absorption photometry for measuring aerosol light absorption. Aerosol Sci. Technol. 39 (1), 40—51.

Petzold, A., Schonlinner, M., April 2004. Multi-angle absorption photometry - a new method for the measurement of aerosol light absorption and atmospheric black carbon. J. Aerosol Sci. 35 (4), 421—441.

Pope, C.A., Ezzati, M., Dockery, D.W., January 2009. Fine-particulate air pollution and life expectancy in the United States. N. Engl. J. Med. 360 (4), 376—386.

PyData Development Team, 2015. Pandas: Python Data Analysis Library.

QGIS Development Team, 2015. QGIS Geographic Information System. Open Source Geospatial Foundation.

Setton, E., Marshall, J.D., Brauer, M., Lundquist, K.R., Hystad, P., Keller, P., Cloutier-Fisher, D., January 2011. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. J. Expo. Sci. Environ. Epi-demiol. 21 (1), 42—48.

Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., October 2013. The changing paradigm of air pollution monitoring. Environ. Sci. Technol. 47

(20), 11369—11377.

Tan, Y., Robinson, A.L., Presto, A.A., December 2014. Quantifying uncertainties in pollutant mapping studies using the Monte Carlo method. Atmos. Environ. 99, 333—340.

Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., De Baets, B., January 2015. Mobile monitoring for mapping spatial variation in urban air quality: development and validation of a methodology based on an extensive dataset. Atmos. Environ. 105, 148—161.

Vardoulakis, S., Solazzo, E., Lumbreras, J., September 2011. Intra-urban and street scale variability of BTEX, NO2 and O3 in Birmingham, UK: implications for exposure assessment. Atmos. Environ. 45 (29), 5069—5078.

Virkkula, A., Makela, T., Hillamo, R., Yli-Tuomi, T., Hirsikko, A., Hameri, K., Koponen, I., October 2007. A simple procedure for correcting loading effects of aethalometer data. J. Air & Waste Manag. Assoc. 57 (10), 1214—1222.

Weichenthal, S., Dufresne, A., Infante-Rivard, C., Joseph, L., November 2008. De-terminants of ultrafine particle exposures in transportation environments: findings of an 8-month survey conducted in Montréal, Canada. J. Expo. Sci. Environ. Epidemiol. 18 (6), 551—563.

Wilson, J.G., Kingham, S., Pearce, J., Sturman, A.P., November 2005. A review of intraurban variations in particulate air pollution: implications for epidemio-logical research. Atmos. Environ. 39 (34), 6444—6462.

Zappi, P., Bales, E., Park, J.H., Griswold, W., Rosing, T. v., 2012. The citisense air quality monitoring mobile sensor node. In: Proceedings of the 11th ACM/IEEE Conference on Information Processing in Sensor Networks. China, Beijing.

Zuurbier, M., Hoek, G., Oldenwening, M., Lenters, V., Meliefste, K., van den Hazel, P., Brunekreef, B., June 2010. Commuters' exposure to particulate matter air pollution is affected by mode of transport, fuel type, and route. Environ. Health Perspect. 118 (6), 783—789.