Effect of GPS errors on Emission model

Anders Lehmann*, Allan Gross†
*Aarhus School of Engineering University of Aarhus
anders@ase.au.dk
†Department of Business Development and Technology
University of Aarhus
agr@btech.au.dk

***Abstract*—In this paper we will show how Global Positioning Services (GPS) data obtained from smartphones can be used to model air quality in urban settings. The paper examines the uncertainty of smartphone location utilising GPS, and ties this location uncertainty to air quality models. The results presented in this paper indicates that the location error from using smartphones is within the accuracy needed to use the location data in air quality modelling. The nature of smartphone location data enables more accurate and near real time air quality modelling and monitoring. The location data is harvested from user data gathered in the wild.**

I. Introduction

Air quality can have serious health effects, both immediate and longterm. In many countries limit values for different pollutants are imposed, to keep the health consequences to a minimum. To monitor that the air pollution is well below these limit values, and to be able to alert the affected communities when concentrations are above the limits, extensive air quality models have been developed. These models take different emission sources like long range transport of pollution from regional or transnational sources, urban pollution sources, and emission from traffic into account. All these pollution sources are considered to estimate the influence on the street level air quality. The effects of the weather is also factored in as wind patterns in street canyons will increase the pollution on the lee side of the street and temperature and precipitation will change the rate of chemical reactions and deposition of pollutants.

Air quality in cities is dependent on a number of factors. Pollution can travel for long distances, and thus affect the air quality in a distant city, but often more important are sources local to the city. The local sources of pollution affecting the city air quality are for instance local residential and district heating, local processing industry and traffic. We will in this paper focus on how we can improve the modelling of traffic induced emissions, and thus the traffic related effects on air quality, through data obtained from smartphones used by users of the urban traffic system. Figure 1 shows the different levels of the pollution model.

To model the impact of traffic on air quality the existing models uses traffic counts to estimate the traffic. The traffic counts are typically points measures done rather seldom and for intervals of a few weeks or months. Ordinarily traffic counts only gives information on the number of vehicles passing through the measuring gates, and thus the speeds of vehicles are not measured. Only a few streets in a city is exposed to traffic counts, and the traffic in the remaining streets are estimated from traffic models.

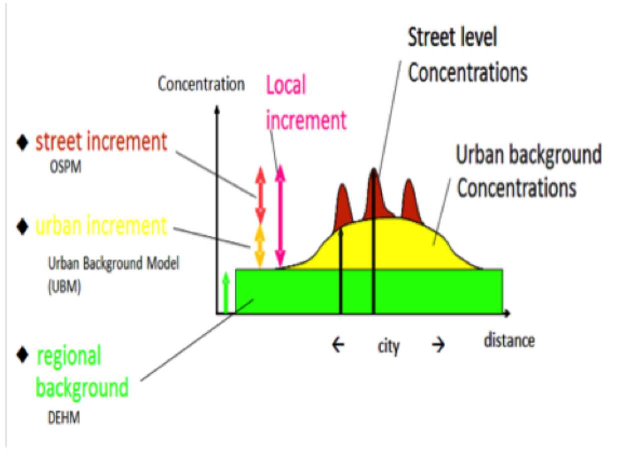

Fig. 1. The pollution in a street is a combined from regional, urban and local sources

In this paper we consider the quality of data collected from smartphones of people moving in an urban transport system, and if this data can be used as a dynamic supplement to the more static traffic counts. We identify two issues to investigate. First the uncertainty of location measurements based on Global Positioning Services (GPS), which will influence how to count vehicles on the streets. The second issue is the uncertainty on the speed measurement from smartphones, which will influence the outcome of the level of emissions estimated from the model.

The data used in this paper is collected in deployments of the project EcoSense, which have provided smartphone apps for different projects promoting sustainable energy, energy savings, sustainable transport and climate change mitigation efforts.

The paper is first considering the available literature on the subjects of air quality and location accuracy in section II. The experiments are presented in section III. Section IV presents the results of the experiments. The paper ends with a discussion of the contributions and results, as well as an outlook towards future work.

II. Related work

Air quality modelling in urban areas has been a scientific field since the privatisation of urban transport became abundant. The large concentration of cars in city centres, leading to serious health issues, has spurred the evolution of methods to monitor the level of pollution. The active measurements of pollution led to an understanding of the sources of the pollution, but was too costly to be deployed in more than a few places in each large city. Modelling approaches were considered to be put in place of active measurements for the rest of the urban areas. One of the more thorough models for urban air quality is the Operational Street Pollution Model [1], [2]. OSPM combines results from Long Range Transport models (i.e. Danish Eulerian Hemisphere model [3]) for regional pollution sources, with Urban background models for urban pollution sources and a street canyon model for local sources and variations due to turbulent winds, to calculate the pollution concentration at the street level. The model is used to monitor the adherence to the limit values set up by national and international regulating bodies [4].

The OSPM model uses an extensive amount of data from diverse sources in order to estimate the emission of pollutants, the mixture and reactions of pollutants, and the transport of pollutants. The data sources stems from traffic counts, traffic surveys, traffic models, fleet information, and meteorological data.

Since the advent and ubiquity of smartphones, many projects have utilised the new opportunities for data collection in the wild. The big problem, from an academic point of view, is that it can be hard to persuade users to submit their data to a research community. Some research projects have used the smartphone location data indirectly, by using location data included in status updates for social media [5], [6]. Other projects have used cell tower data from network operators, to coarsely trace movements of individual mobile phone users [7]. Others, like in this project, have succeeded in using data directly from smartphones to perform their research studies [8]–[10].

The GPS system is well documented and very useful [11]. The location and speed accuracy of non-military commercial receivers has been investigated in different settings. In Witte and Wilson [12] GPS accuracy is measured by a bike driving on an athletics running track. The GPS receiver used had an external antenna placed on the bikers helmet, which is different from a smartphone with a builtin antenna placed in a pocket or bag. The paper uses the placement of the antenna to explain the drop in location accuracy during curves, since the bikers head will be at another place than the bike, due to the leaning necessary for keeping the balance on the bicycle while turning.

In Modsching et al. [13] the GPS location accuracy from four different GPS receiving devices is investigated in an urban environment. The authors report high error rates due to buildings shadowing the radio signals from the GPS satellites, and provide map-matching as a tool to improve the location accuracy. Furthermore the paper finds that there is no correlation between actual location error and device reported error level.

A thorough investigation of error statistics of GPS location for GIS applications, is given in Zandbergen et al. [14]. The error statistics are measured by a mapping grade GPS receiver in a fixed outdoor position over eight hours. The mean error is reported to be 2.4 meters.

To improve the accuracy of GPS location and speed estimation from smartphone data, the paper by Nitsche et al. [15] employ a Kalman filter on the raw GPS data to reduce the location uncertainty. The presented experiments show how the Kalman filter removes outliers, but also shows residual inaccuracy when the filtered GPS data is drawn on a map.

III. Experiments

A series of experiments was conducted, to investigate the error behaviour of GPS in smartphones.

The first experiment concerns stationary indoor accuracy to test experienced system accuracy, without introducing possible distortions from mobility.

The second experiment is about experienced accuracy of moving GPS receivers in smartphones. Data has been collected from smartphones of different users and the reported uncertainty of the GPS location were evaluated.

Lastly an experiment was made to investigate the sensitivity of the Air Quality model to the average speed of the modelled traffic.

A. *Stationary GPS accuracy experiment*

The GPS system involves 27 satellites (24 is the designed number of satellites, the rest are in reserve but operational) in six different medium earth orbits at 20.000 km above earth. The orbit is chosen in such a way that each satellite will trace the same area of the earth two times each day. The recurring events of satellites appearing and disappearing over the horizon, could lead to recurring errors in the localisation of a GPS receiver.

To test this hypothesis we recorded the location and error estimate of a smartphone placed in a fixed indoor location, over a period of 5 weeks. The fixed location was chosen to minimise the effects of changing environments, which could cause location errors from multi-path radio propagation, shadowing or other degradation of the received radio signals. The indoor location was chosen to ensure the integrity of the smartphone and to keep it powered during the experiment. The indoor location may also cause some extra damping of the radio signals from the satellites, which can influence the accuracy of the location estimation. The location was chosen to be a single story home in a residential area. The location and horizontal and vertical error were sampled every second.

The experiment was made in two variants - one with Wifi connected to the local Wifi network, and one with the wifi radio turned off. As the smartphone reported location accuracy is determined from the combined localisation the GSM network, Wifi networks and GPS localisation, the two variants were made to test if the smartphone connected to a

TABLE I
STATISTICS FOR THE "HERNING CYKLER" DATA SET

	Number of trips	Number of location points
Herning Cykler data	900 thousand	250 million
In Herning proper	150 thousand	3 million

local wifi network would influence the location accuracy. The GSM network was disabled in both variants, by removing the SIM card.

B. Mobile GPS accuracy experiment

To assess the accuracy of mobile smartphone localisation data from EcoSense deployments were used. Data on reported GPS accuracy and localisation was extracted from the "Herning cykler" data set. As the data set is collected from users who have downloaded an app from an appstore, the app has been used in many different places, and for different purposes (including boating). As we wish to focus on accuracy in urban environments, points located outside a bounding box containing the city of Herning were removed. The resulting data contains 250 million location points from 900 thousand different trips.

C. Emission modelling speed sensitivity

The OSPM [1] modelling program was used to investigate the potential influence of speed inaccuracy on the results for the modelled air quality. OSPM models the air quality at the street level by combining long range pollution transport models, with urban background pollution model and traffic based street level pollution contributions. In our experiment we choose a street with heavy traffic in a built up area of a city, where the surrounding buildings create a street canyon. In the street canyon, turbulence from cross winds will concentrate the pollution on the lee-side of the street, and thus exacerbate the effects of the traffic emissions. The traffic load was chosen from information on traffic counts performed by the municipality. The builtin meteorological data for the city and fleet data for Denmark was used. A number of simulations was performed where the average speed of the traffic was varied. Numbers for the air quality measures (the concentration of the different pollutants at the height of an virtual sensor) was retrieved from the resulting reports to show the impact of different average speeds.

IV. RESULTS

We present in this section a number of results from our investigation. As the experiments span from single phone stationary experiments, over multiple mobile phones to computer simulations, there are quite diverse results.

A. Stationary GPS accuracy experiment results

The experiment concerning indoor stationary GPS accuracy, was conducted by having an iPhone 4s, without a SIM card, running the app "SensorLog" for five weeks. In the first three weeks the wifi was enabled and connected to a residential wifi network. In the last two weeks the experiment was repeated with the wifi radio disabled.

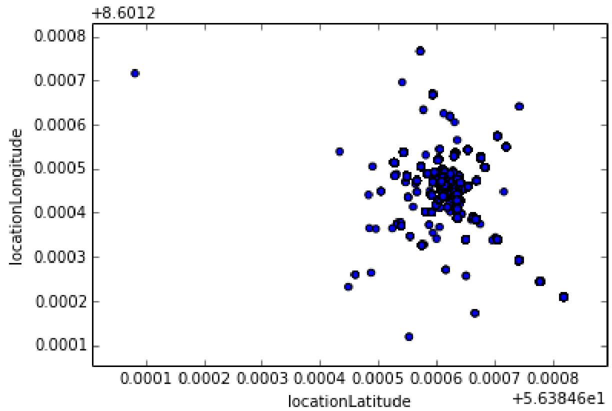

Fig. 2. Reported locations in both stationary experiments

In Figure 2 a scatter plot for 1.2 million position measurements. To be able to see if there are structures apparent only points with a reported location accuracy below 200 meters has been shown. The figure shows that most points are grouped close to each other. The largest distance between points are 67 m. The central grouping shows some signs of preferred directions, and do not resemble a gaussian distribution of points from a fixed centre.

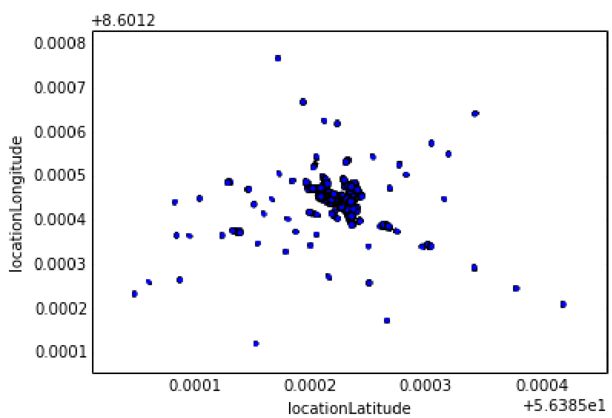

Fig. 3. reported locations in stationary experiment Wifi enabled.

In figure 3 locations from the first three weeks, where the smartphone was connected to the wifi network, and where the wifi radio could be used to aid the location system, by triangulation between neighbouring wifi networks. The single outlier from the previous figure is gone, but the preferred directions are more clearly discernible. When looking at a map of the residential area of the experiment, it is a quite suggestive to match the preferred directions to directions from the smartphone toward the neighbouring wifi access points. There app. 800.000 location points in this data set. The largest difference between the locations is 54 meters.

Figure 4 shows the location data from the second experiment

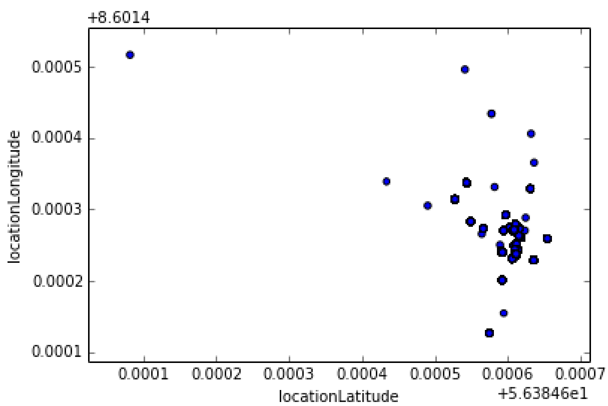

Fig. 4. reported locations in stationary experiment Wifi disabled

where the wifi has been disabled. The figure shows the single outlier from figure 2 and a very tight center with a few strays around the center. The preferred directions are less pronounced in this figure and the largest difference between points is 68 meters, and if the outlier is removed the largest distance between points is decreased to 33 meters. There are almost 430000 location points in this experiment.

From the examination of these two experiments we can conclude that using wifi networks as part of a localisation solution, does have an effect, but maybe it is better to solely rely on GPS for stationary use.

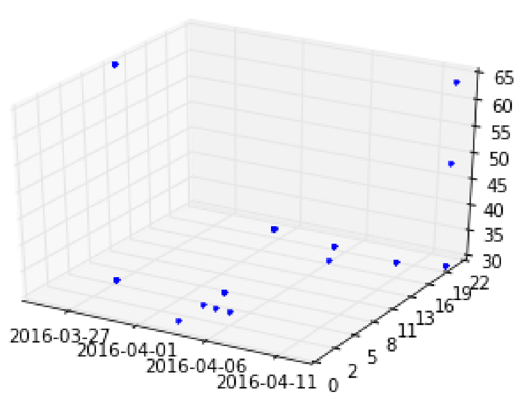

Fig. 5. Large reported errors > 10 meters as a function of time. GPS and Wifi.

To see if there are any recurring accuracy events, which could relate to the recurring constellations of GPS satellites, we show in figure 5 the reported accuracy and the date (x-axis) and time of day (y-axis). The measurements which shows low location uncertainty (less than ten meters) has been removed to not clutter the diagram. Figure 5 shows the 30 location measurements for the 3 weeks where both GPS and wifi was used for location estimation, where the reported accuracy was above 10 meters. There are no apparent recurring daily events so we cannot conclude that satellite constellations seriously affect the accuracy of stationary location measurements. There seems to a small tendency towards larger uncertainty at five in the morning and at ten in the evening.

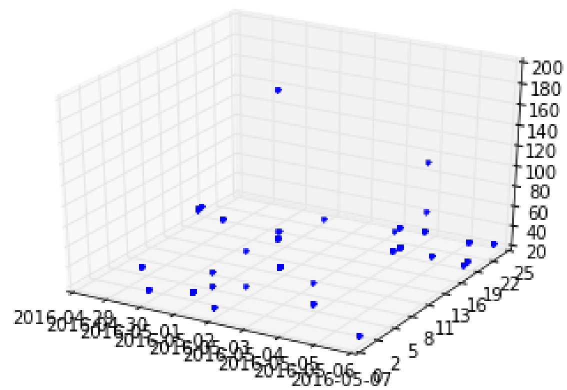

Fig. 6. Uncertainty histogram

In figure 6 we show the same diagram as in figure 5, but for the two weeks with only GPS as localisation device. In these two weeks there were 330 incidents of reported errors larger than ten meters, but there are no apparent recurring events of large localisation uncertainty, thus we cannot find support for the hypothesis that satellites constellations influence GPS accuracy. In this GPS-only experiment it seem that not relying on wifi for localisation introduces more incidents of large accuracy uncertainty.

From the above analysis a few observations can be made. First it can be seen from the figures 2, 3 and 4 that the accuracy of GPS location without wifi based location is better than GPS with wifi based location, except for a single point. From the figures 5 and 6 it seems that the variability of the reported errors is higher when the wifi is switched off, even though that the scatter plots suggested otherwise. It seems that the self reported error is overly relying on the access to known wifi access points.

B. Mobile GPS accuracy experiment results

In this section we report the results of our investigation into reported accuracy from smartphones as used in the wild. The data used is collected through the "Herning cykler" deployment. Even though the app was only promoted in Herning, Denmark, the deployment has generated data from other places in Denmark, and outside Denmark, the last part probably because of some of the informants going on holiday, and bringing their smartphones with them. The dataset contains data from 900 thousand trips, of varying lengths, with 250 million location measurements.

In figure 7 a histogram for the reported uncertainties. The histogram is abbreviated to uncertainties below 200 meters, and the bins are ten meters wide. It is clear to see than the most of the reported uncertainties are ten meters and below. But there is still almost half of the measurements with larger uncertainties.

The figure 8 shows some of the long tail of the uncertainty distribution. The small uncertainties (below 20 meters), has

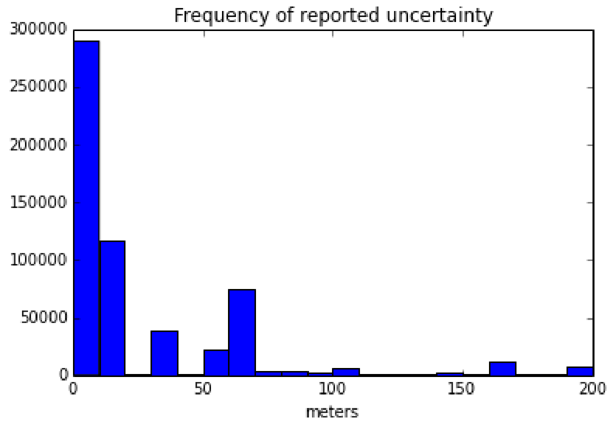

Fig. 7. Uncertainty histogram

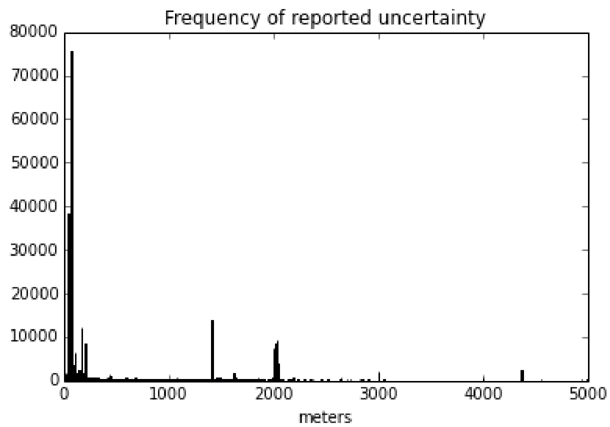

Fig. 8. Uncertainty displays a wide spread. The counts of low uncertainty (< 10) has been removed

ben removed to enable the view of the smaller counts of the larger uncertainties. In the figure 8 one can see that even though most of the location measurements have uncertainties below 100 meters, there is still a considerable amount of measurements with quite large uncertainties above 500 meters.

C. Emission modelling speed sensitivity results

The results of the experiment to quantise the sensitivity of the air quality model towards the speed of the traffic.

The figure 9 shows the concentration the gas NO_2 at a simulated measurement point in a street canyon. The figure shows that the concentration of NO_2 drops as the average speed increases. The drop in concentration is congruent with the speed dependency of emission factors [16], for low speeds the emission per meter is high due to idling or low utilisation of the motor. At higher speeds the motor output is used more effectively, thus the emission factor is lower. As the speed increases towards highway speeds the emission factor grows as the motor has to overcome increasing wind forces. We have only made the simulations for speeds less and equal to 120 km/t, as higher speeds are not deemed relevant for traffic in an urban street canyon.

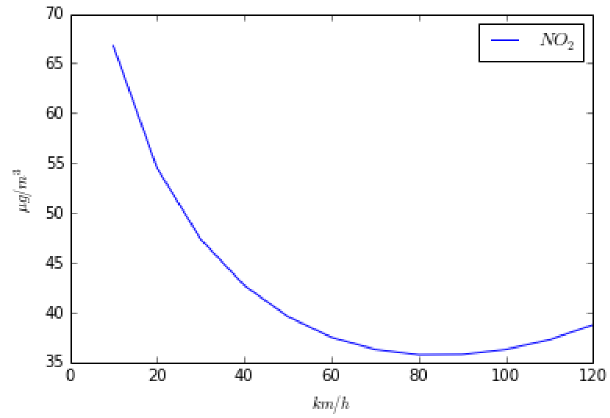

Fig. 9. Modelled NO_2 concentration for different average speeds

From the figure 9 it can be seen that the speed sensitivity of the emission model is largest for low vehicle speeds. To accurately model the street level pollution it is important that the traffic data is accurate for low speeds.

V. Application of GPS location data in street level air quality model

To improve street air quality models like OSPM, with data obtained from smartphones of urban travellers, two essential pieces of information can be supplied. The location of travellers can be used to account for emissions from a trip to the locations traveled through. The speed at different locations is the second important information that the model needs in order to estimate the emission amount.

The smartphones could also supply information on the type and motor size of the vehicle, which could further improve the model estimate, but currently this information has to be gained by asking the user of the smartphone. Research has been done to automatically detect transportation mode [17], but currently it seems impossible to detect motor size and fuel type. One exception is to detect electric vehicles versus internal combustion engine vehicles [18].

VI. Discussion

From the experiment with the OSPM model, it can be seen that the speed of vehicles are very important to the air quality at the street level. But it is also important to be able to attribute the emissions from vehicles surveilled by applications like ours to the right streets. That means that the accuracy of the location of vehicles monitored by smartphone apps must be high enough to assign the location to a single street.

To evaluate the influence of GPS accuracy an air quality modelling we have to determine the connection between location accuracy and speed accuracy and possible methods to overcome inaccuracy problems.

From the GPS design documents the best case error estimate is below six meters, which should be achievable in 95% of the time [11]. The speed is (falsely) reported to be larger than zero in 29 cases in the stationary experiment. As the reported locations as seen in figure 2,3,4 should result in more reported nonzero speed events, if only the raw GPS locations were used to calculate the reported speed. The largest reported speed in the 29 cases of non zero speed is 2 km/h.

Map matching algorithms can reduce the inaccuracy from the GPS data, as seen in [10]. A median filter might also provide a stable location trace.

The errors as seen in the real life data set from "Herning Cykler", points toward the conclusion that we can use smartphone data to supplement our other data sources, for air quality modelling at the street level. This possibility enables more dynamic and up to date air quality monitoring, and possibly also new kinds of health alerts and urban transport planning overview.

Since the projects that have enabled the data collection has been targeted at special audiences, it is likely that there is a bias towards people who cares about sustainability, environment and man made climate change. The validity of the analysis of GPS accuracy does not suffer from this bias.

VII. Future work

In this paper we have only considered the collected data to be from person transport in cars. To further the quality of the collected data, methods to discern between different modes of transportation, to correctly adjust for the emissions from the vehicles, is needed.

VIII. Conclusion

We have measured the accuracy of GPS data from smartphones, and investigated if the data is As traffic models become more ubiquitous, and are able to model the largest cities [19], urban transport planning get more tools for simulation of urban policy initiatives and new infrastructure initiatives, we hope that the work in this paper will lead to urban planning to take air quality into account.

Acknowledgment

We want to thank Henrik Blunck, and Niels Olof Bouvin for their generous support.

This work has been supported by The Danish Council for Strategic Research as part of the EcoSense project (11-115331) ,the TUD COST Action TU1302 Satellite Positioning Performance Assessment for Road Transport (SaPPART), and EU Interreg project Biogas2020.

References

[1] R. Berkowicz, "OSPMa parameterised street pollution model," in *Urban Air Quality: Measurement, Modelling and Management*. Springer, 2000, pp. 323–331.

[2] K. E. Kakosimos, O. Hertel, M. Ketzel, and R. Berkowicz, "Operational Street Pollution Model (OSPM) - A review of performed application and validation studies, and future prospects," *Environmental Chemistry*, vol. 7, no. 6, pp. 485–503, 2010.

[3] J. Christensen, "The Danish Eulerian hemispheric modelA three-dimensional air pollution model used for the Arctic," *Atmospheric Environment*, vol. 31, no. 24, pp. 4169–4191, 1997.

[4] C. Directive, "Council Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air," *Journal of the European Communities*, vol. 50, no. 163, pp. 41–60, 1999.

[5] J.-m. Kang and B.-h. Hong, "A Study on the SNS (Social Network Service) Based on Location Model Combining Mobile Context-Awareness and Real-Time AR (Augmented Reality) via Smartphone," in *Future Information Technology*. Springer, 2011, pp. 299–307.

[6] C. M. Walton, "Urban Travel Demand Analysis for Austin TX USA using Location-based Social Networking Data," *TRB Annual Meeting*, pp. 1–17, 2013.

[7] Y. Zhang, X. Qin, S. Dong, and B. Ran, "Daily O-D Matrix Estimation Using Cellular Probe Data," *Transportation Research Board 89th Annual Meeting*, 2010.

[8] J. Gonc and R. J. F. Rossetti, "Smartphone Sensor Platform to Study Traffic Conditions and Assess Driving Performance," in *IEEE 17th International Conference on Intelligent Transportation Systems*, 2014, pp. 2596–2601.

[9] S. Hemminki, P. Nurmi, and S. Tarkoma, "Accelerometer-based transportation mode detection on smartphones," *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems - SenSys '13*, pp. 1–14, 2013.

[10] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda, "{PEIR}, the Personal Environmental Impact Report, As a Platform for Participatory Sensing Systems Research," in *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services*, ser. {MobiSys} '09. New York, {NY}, {USA}: ACM, 2009, pp. 55–68.

[11] DoD, "GLOBAL POSITIONING SYSTEM STANDARD POSITIONING SERVICE," 2008.

[12] T. H. Witte and A. M. Wilson, "Accuracy of non-differential GPS for the determination of speed over ground," *Journal of Biomechanics*, vol. 37, pp. 1891–1898, 2004.

[13] M. Modsching, R. Kramer, and K. ten Hagen, "Field trial on GPS Accuracy in a medium size city: the influence of built-up," *3rd Workshop on Positioning, Navigation and Communication 2006*, vol. 2006, pp. 1–10, 2006.

[14] A. Zandbergen, "Positional Accuracy of Spatial Data : Non-Normal Distributions and a Critique of the National Standard for Spatial Data Accuracy," *Transaction in GIS*, vol. 12, no. 1, pp. 103–130, 2008.

[15] P. Nitsche, P. Widhalm, S. Breuss, N. Brändle, and P. Maurer, "Supporting large-scale travel surveys with smartphones A practical approach," *Transportation Research Part C*, vol. 43, pp. 212–221, 2014.

[16] O.-K. Nielsen, M. Plejdrup, M. Winther, M. Nielsen, S. Gyldenkærne, M. Mikkelsen, R. Albrektsen, M. Thomsen, K. Hjelgaard, L. Hoffmann, P. Fauser, H. Bruun, V. Johannsen, T. Nord-Larsen, L. Vesterdal, I. Møller, O. Caspersen, E. Rasmussen, S. Petersen, L. Baunbæk, and M. Hansen, "Denmark's National Inventory Report 2014," University of Aarhus, Tech. Rep. 101, 2014.

[17] Z. Sun and X. J. Ban, "Vehicle classification using {GPS} data," *Transportation Research Part C: Emerging Technologies*, vol. 37, pp. 102–117, dec 2013.

[18] M. Wüstenberg, H. Blunck, K. Grønbæk, and M. B. Kjærgård, "Distinguishing Electric Vehicles from Fossil-Fueled Vehicles with Mobile Sensing," in *2014 IEEE 15th International Conference on Mobile Data Management*. IEEE, jul 2014, pp. 211–220.

[19] A. Lehmann and A. Gross, "Using Crowd Sensed Data as Input to Congestion Model," in *IEEE International Conference on Pervasive Computing and Communication Workshops*. IEEE, 2016, p. 6.