On the use of mobile sensors for estimating city-wide pollution levels

Carlos T. Calafate, Bertrand Ducourthial

On the Use of Mobile Sensors for Estimating City-Wide Pollution Levels

Carlos T. Calafate
Department of Computer Engineering (DISCA)
Universitat Politècnica de València
Camino de Vera, S/N - 46022 Valencia, Spain
Email: calafate@disca.upv.es

Bertrand Ducourthial
Heudiasyc Laboratory
Université de Technologie de Compiègne
60205 Compiègne Cedex, France
Email: Bertrand.Ducourthial@utc.fr

Abstract—Obtaining detailed pollution maps for urban environments is an effort that is gathering much interest by allowing to better regulate traffic and protect citizens from hazardous conditions. However, the scarcity of pollution sensors prevents obtaining the desired degree of detail, requiring alternative solutions to be deployed. In this paper we explore the concept of mobile pollution sensing by studying the feasibility of equipping buses with ozone measurement hardware to estimate ozone patterns for the city of Compiègne. Overall, we achieve accurate estimations, with error values typically ranging from 2% to 10%. Compared to solutions based on deploying static sensors on the different bus stops available, we find that the proposed mobile sensing approach is able to provide a degree of accuracy comparable to deploying tens of static sensors, substantially reducing costs and management.

Index Terms—mobile sensing; pollution; ozone; geostatistics.

I. Introduction

Urban air pollution is considered as a growing problem from both citizen and governmental perspectives. Such concerns are supported by different toxicological studies [1], [2] showing that air pollutants such as sulphur dioxide, nitrogen compounds, carbon monoxide and ozone have a serious impact on health when their concentration surpasses certain thresholds.

Depending on the particular atmospheric and orographic conditions, an excessive accumulation of the pollutants can occur in specific parts of our cities. For that reason, having a detailed map of pollutants can help at notifying people living in a specific area about inadequate air conditions, and repeating patterns can be detected by authorities to allow taking action to mitigate such problems.

Currently, pollution mapping relies on very precise sensors placed at strategic locations. However, since such sensors have a very high cost, their number is usually quite limited. Such limitation prevents obtaining detailed pollution maps, being the most accurate maps at a spatial scale of 1km per 1km [3]. From the perspective of small and mid-sized cities, though, even such granularity is too coarse to allow determining the city-wide distribution of pollutants, and certainly does not help at detecting pollution hot-spots within the city.

In this paper we proposed using low-cost sensors to obtain a detailed map of urban pollutants. In particular, we will focus on ozone as a pollutant, and the city of Compiègne, France, will be the target area for our analysis. Ozone is created near the Earth's surface by the action of daylight UV rays on different pollutants, predominantly those emitted during the combustion of fossil fuels. Breathing ozone can trigger chest pain, coughing, throat irritation, and reduce lung function by inflaming the linings of the lungs, among many other problems. Thus, we consider it a relevant pollutant for our analysis.

Since the city of Compiègne has a single ozone measurement station, no detailed ozone maps are available. Thus, for our study, we relied on spatial ozone patterns available for other cities, which were modeled and adapted to the known traffic congestion points in Compiègne to create a realistic ozone map for the city. Using the created ozone map as reference, we then compared the effectiveness of both static and mobile sensing to estimate the ozone levels in a pre-defined Region of Interest (RoI) when varying the number of samples and taking sensor errors into account.

Experimental results show that, even when relying on a single mobile measurement station, and despite of sensor inaccuracy, we are able to estimate the ozone values within the RoI with a good degree of accuracy, being comparable to deploying tens of sensors throughout the city.

The remainder of this paper is organized as follows: in the next section we review some related works in the field. In Section III we describe the problem and detail the methodology followed to solve it. In Section IV we analyze the ozone estimation accuracy when relying on either static or mobile stations. Finally, in Section V, we present the main conclusions of the paper, along with future work.

II. Related works

Estimating air pollution levels in urban areas is a subject that has attracted much research interest in the last decades [4], [5]. In general, the different authors rely on traditional ground sensor stations to obtain measurements, although acknowledging that their reduced number limits estimation accuracy [6], [7].

Concerning the pollution estimation process, Beelen et al. [3] study different techniques to extrapolate from monitoring stations to unsampled locations. Specifically, they compared

the validity of ordinary kriging, universal kriging and regression mapping in developing EU-wide maps of air pollution on a 1×1 km resolution. They focused on different pollutants such as nitrogen dioxide (NO_2, fine particles <10 μm (PM10), ozone (O_3), sulphur dioxide (SO_2), and carbon monoxide (CO). Singh et al. [8] focus on the same problem, and they develop a spatial interpolation system (cokriging) to estimate the 8h mean daily maximum ozone concentrations and daily mean PM10 concentrations over a domain, starting from measured concentration values.

All the aforementioned solutions can be considered standard in the sense that they rely mostly on the typical stationary measurement stations. Participative sensing solutions have emerged as an alternative to traditional sensing solutions by acknowledging that the extensive cost of acquiring and operating official, high-accuracy stations severely limits the number of installations, and thus results in a limited spatial resolution of the published pollution maps. Yajie et al. [9] proposed a solution in this direction by developing low-cost and ubiquitous sensor networks to collect real-time, large scale and comprehensive environmental data from road traffic emissions for air pollution monitoring in urban environments. They rely on a distributed infrastructure based on wireless sensor networks and grid computing technology, and they target different pollutants (SO_2, NO, NO_2, O_3, NH_3, and Benzene). More recently, Hasenfratz et al. [10] addressed this problem by acquiring spatially fine-grained air pollution data with a community-driven sensing infrastructure obtained by connecting a small-sized, low-cost ozone sensor to an Android smartphone.

In this work we study the feasibility of having very few mobile sensors to replace a great number of static sensors when attempting to have an accurate pollution map for a certain RoI. We rely on vehicles, specifically buses following regular routes equipped with ozone sensors, to achieve this goal.

III. PROBLEM STATEMENT AND METHODOLOGY

In this section we detail the procedure followed to create realistic ozone maps for a specific region based on known pollution hotspots. Then we present the ozone map generated, along with the RoI for analysis. Lastly, we present some details about the geostatistical procedure followed to estimate ozone values based on different sampling strategies.

A. Ozone map creation

As stated in the introduction, the limited number of measurement stations prevents obtaining detailed pollution maps for any target urban area. Nevertheless, the few detailed ozone maps available show that spatial ozone distribution is clearly related to the weather, being ozone levels and gradients much higher during summer compared to cold months. The variability of these spatial patterns is specially relevant since the number of samples required, as well as the accuracy of estimation processes, is directly related to the gradients detected.

Table I
RESULTS OF THE CURVE FITTING PROCESS.

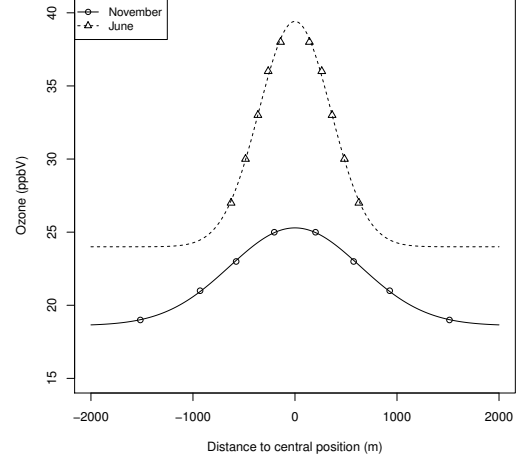| Parameter	June	November
a	15.412	6.684
σ	351.46	640.46
d	24.00	18.61
rms	0.2166	0.0623

Figure 1. Curve fitting using normal distribution for the ozone variations typical of June and November.

In this paper we have taken as reference the results presented by Moral-García et al. [11], who provide detailed urban ozone distribution patterns for the city of Badajoz, a mid-sized city in the southwest region of Spain. Based on the spatial ozone maps for the months of June and November, we took several samples based on which we analyzed the ozone gradients achieved.

Combining data from the different samples we are able to obtain the expected gradients for the months of November and June. Then, by following a curve fitting procedure, we were able to obtain good fits using a Gaussian distribution of type:

$$f(x) = a \cdot e^{-\frac{1}{2}(\frac{x-x_0}{\sigma})^2} + d \tag{1}$$

The results of the fitting process are presented in table I.

Figure 1 represents both the sampled values and the Gaussian curves after fitting. We can clearly observe that, besides a different range of values, the distributions corresponding to the months of June and November have a quite different kurtosis. In particular, the high gradients associated with the distribution for June makes ozone estimations to become quite harder compared to the values for November. For this reason, in the study that follows, we will generate an ozone map based on the pattern associated to June.

In order to create the ozone map for the city of Compiègne we first obtained a two-dimensional equation based on Eq. 1 as follows:

$$f(x,y) = a \cdot e^{-\frac{1}{2}[(\frac{x-x_0}{\sigma})^2 + (\frac{y-y_0}{\sigma})^2]} + d \tag{2}$$

The second step was to mark in the map of Compiègne the traffic hotstops based on knowledge about traffic circulation in the city. These are represented in Figure 2. We can clearly
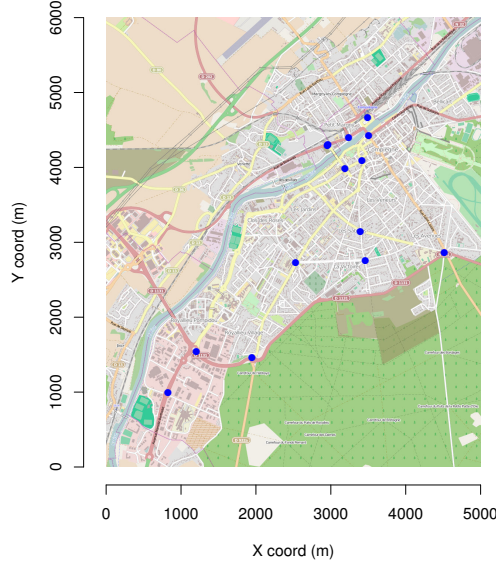
Figure 2. Ozone hotspots defined for the city of Compiègne.

Figure 3. Generated ozone map for the city of Compiègne.

differentiate 3 different regions: (i) 6 hotspots in the north part of the city near the train station with a high traffic load, (ii) 4 hotspots in the mid-part of the city, with moderate traffic load, and (iii) 3 hotspots in the south part of the city, with a high traffic load. By making the coordinates of each of these points match the (x_0, y_0) parameters in Eq. 2, we were able to generate the ozone map shown in Figure 3.

The city is surrounded by forest areas in the east and south areas, and other smaller urban areas followed by rural areas in the western and north parts of the city; in this paper we are not concerned about accurately generating or estimating ozone maps in these areas, focusing solely on the city of Compiègne.

B. Definition of sampling procedures and target geographical area

We study different approaches in order to obtain ozone measurements for the city of Compiègne; specifically we define a target area (RoI) with a parallelogram shape (see Figure 4) that includes the most representative part of the city.

Our first approach to map the ozone in this city is to equip the different bus stops available with ozone sensors. However, as shown in Figure 4 (top), there are more than 100 bus stops, meaning that sensor deployment and maintenance costs could become excessive. So, the first alternative that we study is to create random subsets of different sizes based on the set of all bus stops, to study the impact of subsampling on estimation error.

The second alternative we study is to check which of the five bus lines available has the best match with the RoI defined, and only use that line for ozone measurements. In this case the best match was bus Line #5, depicted in Figure 4 (bottom). Since this bus line includes only 24 stops, the number of sensors to deploy is conveniently reduced.

Finally, as a third alternative, we study the possibility of equipping buses covering Line #5 with ozone sensors, and
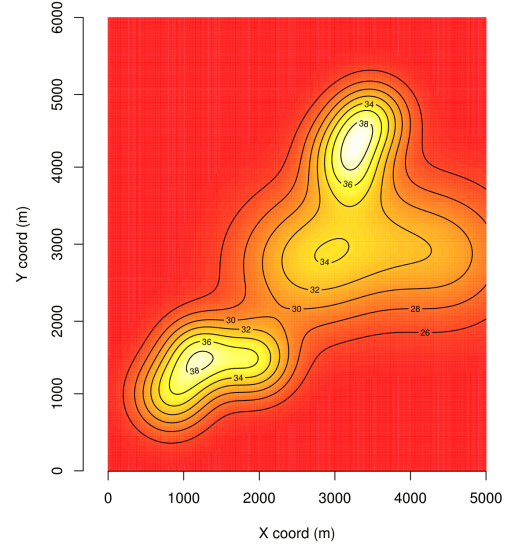
assess the effectiveness of this cheaper solution against the previous ones. This solution is quite more ambitious than the former ones, since we intend to use a single sensor to obtain a similar degree of accuracy as obtained with many distributed sensor units.

For the tests undertaken in Section IV we will analyze the results achieved using the three alternatives referred above. We assume that a representative sample is taken at each bus stop equipped with ozone sensors. In the case of mobile sampling, we have obtained the GPS trace of a real bus following Line #5 to grasp the actual behavior in terms of position, speed, timing, etc. Based on these input data, the sampling process is activated whenever the bus speed is low (less than 20 km/h) to avoid high air flow levels that could affect the measurements, and when the distance to the previous sampling position is of at least 100 meters (at lower distances no significant differences are detected, and the excess of inputs becomes a burden to the kriging process). Based on the GPS trace obtained we find that, despite applying these restrictive criteria, the number of samples obtained during the bus route is of 42, almost twice compared to the number of bus stops in Line #5.

C. Geostatistical procedure for ozone estimation

In our study, after obtaining ozone samples, the next step is to estimate the ozone values corresponding to the whole RoI defined. For this endeavor we rely on geostatistical techniques to perform the desired estimation. Specifically, the technique known as kriging is the most adequate solution in this context, being widely used in this field [6], [11], [8].

When performing kriging, the first step is to generate a semivariogram by combining information from the different samples. The semivariogram is used to describe covariance, which is basically the mean ozone variation as the distance from a sample increases. Based on the shape and characteristics of the semivariogram, we can then perform curve fitting to determine which are the most adequate input parameters
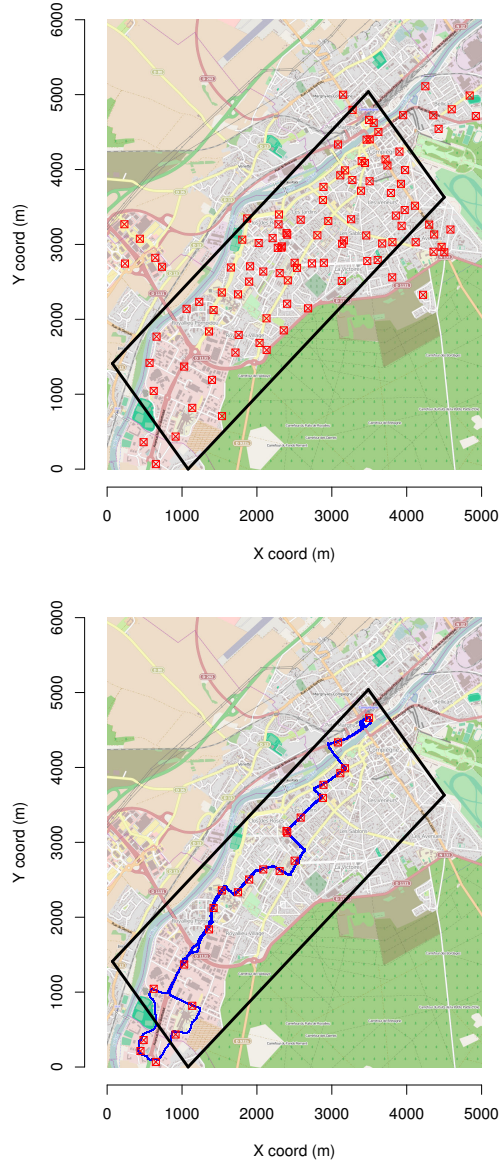
Figure 4. RoI definition and measurement points when using all bus stops (top) or Line 5 bus/stops (bottom).

Figure 5. Semivariogram and fitting output when combining the samples of all bus stops (top) and samples along Line #5 (bottom).

for the kriging process.

In Figure 5 we show the semivariograms for the case of sampling all bus stops, or only the stops along Line #5. We can see that the semivariograms have different shapes and characteristics, being exponential fitting more adequate for the *allstops* case, and Gaussian fitting more adequate when only samples along Line #5 are available.

The output of the fitting process consists of three parameters known as range, partial sill and nugget. They are used as input to the kriging process, meaning that a correct estimation of these three parameters is crucial to obtain a meaningful estimation through kriging.

The kriging process requires as input the ozone samples, along with covariance model and parameters, as stated before. In addition, the target (estimated) positions must also be
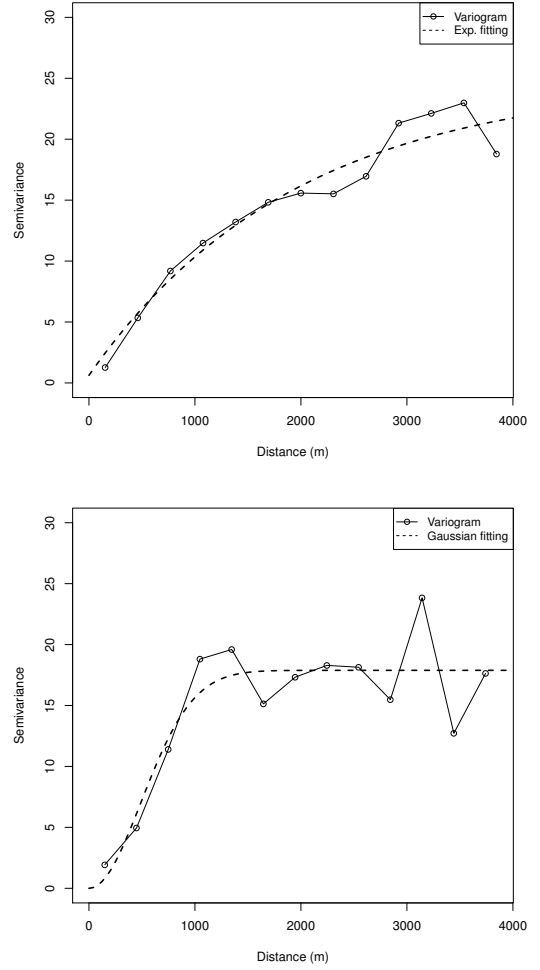
defined. In our case we applied a grid to the map of Compiègne with a 10×10 meters resolution, and selected as target those positions falling within the RoI defined earlier (see Figure 4).

IV. ANALYSIS OF RESULTS

In this section we present and discuss the results of the kriging process when focusing on both traditional (static) and mobile approaches. We assume that both static and mobile nodes are connected to a datacenter where data is collected periodically and processed as detailed above, thereby providing real-time updates of the ozone map for the target city. For the static approach we analyze the case where we can equip all bus stations (or a subset) with ozone sensors, as well as the case where only bus stops along Line #5 are equipped. For the mobile approach we analyze the case where a single bus moving along Line #5 is equipped with ozone sensors along with a GPS module. Both static and mobile devices are also equipped with a 4G communications unit.

In our tests we take sensing error into account. Specifically, we consider that cheap devices such as the MiCS-OZ-47
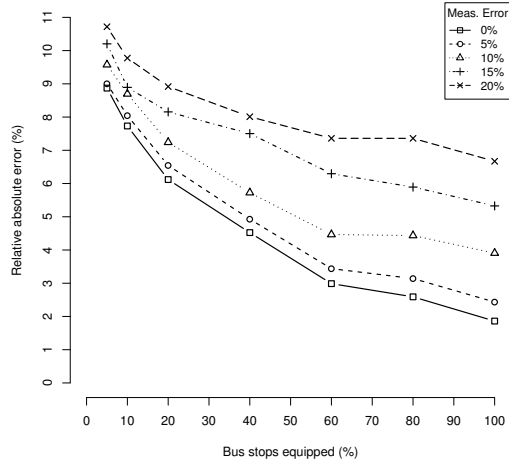
Figure 7. Relative measurement error for the estimated ozone values when varying the number of bus stops equipped with ozone sensors, for different levels of measurement error.

Figure 8. Relative measurement error for the estimated ozone values when varying the sensor measurement error for line 5 stops and moving bus 5.

sensor [12] or similar ones are used to sense the ozone concentration. Since this type of sensors can introduce measurement errors up to 10%, in our tests we also model measurement errors (up to 20%) to assess the impact of sensor inaccuracy on the overall estimation process. We model a measurement error of $\pm\alpha\%$ as uniformly distributed random values, so that measurement s_i will take a random value in the interval: $[(1-\alpha) \cdot s_i, (1+\alpha) \cdot s_i]$.

A. Static vs. mobile approach under ideal conditions

We start our analysis by studying the accuracy of the estimation process under ideal conditions, i.e., when all measurements are accurate. The goal is to determine the upper bound on estimation accuracy.

Figure 6 shows the estimation output for the whole map, including the sampling points (circles) and the parallelogram delimiting the RoI.

Compared to the original ozone map previously presented in Figure 3, we can see that the ozone map estimated using all bus stops available is able to achieve great resemblance with the original one, basically due to the very high number of samples used, and their distribution throughout the map. For the other two cases (Line #5 stops and mobile measurements), the accuracy is smaller if focusing on the whole map. Nevertheless, when focusing on the RoI alone, we find that the differences are rather limited. In particular, when using all bus stops the estimation error (measured as the average value of the relative absolute error) is of only *1.87%*, while, when using Line #5 stops alone, the error is increased to *5.33%*. Using a mobile measurement station allows obtaining quite good results, being the error of only *2.27%*, a value comparable to the all bus stops situation. So, these preliminary values are promising concerning the mobile sensing alternative proposed in this paper.
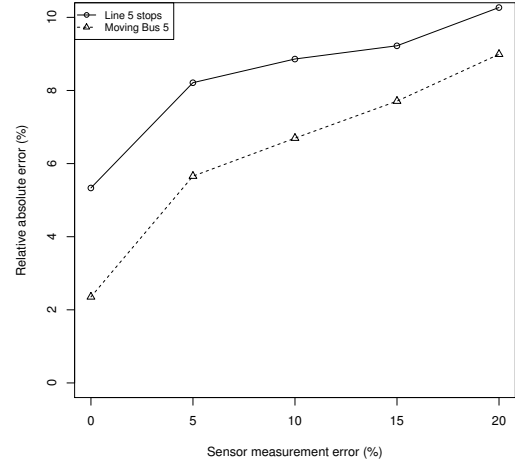
B. Performance under sampling error assumptions

Our analysis now focuses on the impact of sampling errors on the estimation accuracy, being the latter measured using the relative absolute error between the original ozone values and the estimated ones. All the results presented are the average over 20 different simulation runs.

In Figure 7 we show the estimation accuracy results when varying both the sensor accuracy (between 0% and 20%) and the number of bus stops equipped with ozone sensors (random subset between 5% and 100% of all stations). As expected, estimation accuracy improves as we increase the number of equipped bus stops, and as sensing error decreases. It is worth highlighting that, for measurement errors up to 5%, the impact on estimation is minimal. Also, we find that the difference between having all bus stops equipped, or only 60% of them, is not high.

If we now focus on bus Line #5 alone, the results presented in Figure 8 show that sensor measurement errors have a greater impact on estimation error compared to the previous case (all bus stops). This is expectable due to the spatial distribution of samples: while for the "all bus stops" case measurement stations are spread throughout the map, the measurements made in Line #5 (either static or mobile) are mostly along a line and fail to cover the RoI fully, thereby complicating the kriging process and making its output less accurate. Nevertheless, the estimation error remains at acceptable levels, especially for the mobile measurements using bus #5, where a sensing error of 10% results in an estimation error of only 6.7%.

V. CONCLUSIONS AND FUTURE WORK

To mitigate the limitations of current pollution sensors, novel techniques relying on a high number of cheap sensors are envisioned. While crowdsourcing solutions could be a possibility in the future, current mobile devices are not yet equipped with pollution detectors, complicating this approach.
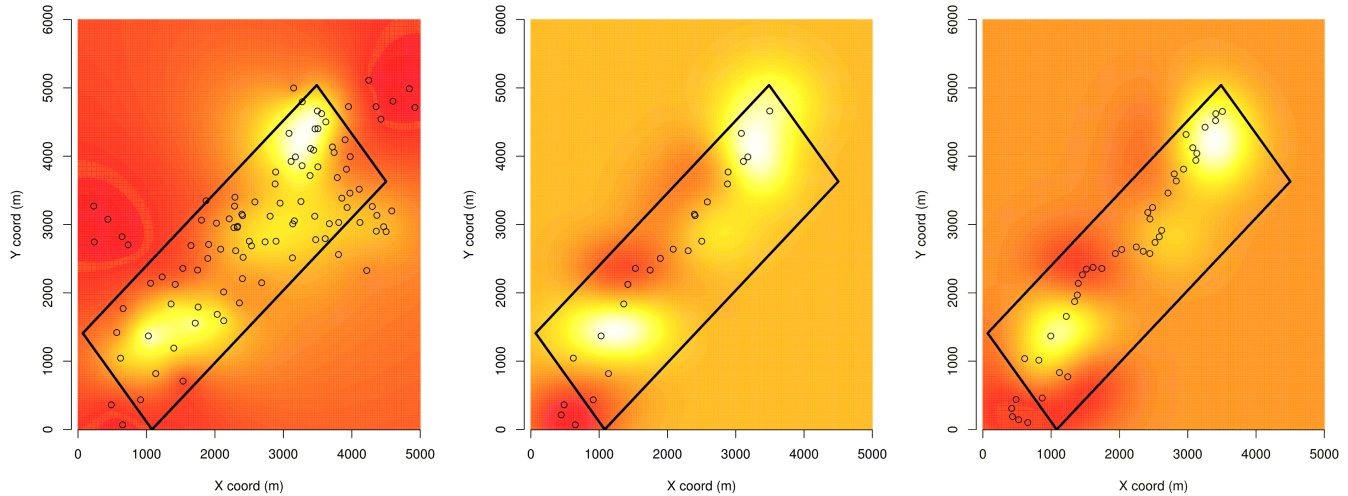
Figure 6. Estimated ozone map for all bus stops (left), Line #5 bus stops (middle) and mobile measurements made on bus #5 (right). The reference ozone map is provided in Figure 3.

In this paper we proposed using cheap ozone sensors to obtain a detailed ozone map for the city of Compiègne, which is currently equipped with a single high-quality ozone sensor. By creating a realistic ozone map based on the known traffic hotspots in the city, we were able to study the effectiveness of both static and mobile sampling strategies in accurately reproducing the spatial ozone variations taking place throughout the city.

Experiments based on the models created showed how the information from these cheaper sensors can be efficiently integrated through geostatistical techniques to allow estimating ozone values for a Region of Interest (RoI) with low estimation errors. Specifically, we showed that a strategy based on equipping buses in one of the bus lines (#5) with ozone sensors is able to achieve results similar, or even better than, equipping 40% of all the bus stations in the city. The mobile sensing solution also outperformed by two percent points an alternative deployment strategy where all the 24 stops available along bus Line #5 were endowed with ozone sensors.

As future work we plan to obtain real measurements by equipping buses covering Line #5 in Compiègne with mobile ozone sensors, and obtain detailed ozone maps for this city for the first time.

REFERENCES

[1] K. Katsouyanni, G. Touloumi, C. Spix, J. Schwartz, F. Balducci, S. Medina, G. Rossi, B. Wojtyniak, J. Sunyer, L. Bacharova, J. P. Schouten, A. Ponka, and H. R. Anderson, "Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 european cities: results from time series data from the aphea project," *BMJ*, vol. 314, no. 7095, p. 1658, 1997.

[2] W. H. Organization, "WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide," tech. rep., 2006.

[3] R. Beelen, G. Hoek, E. Pebesma, D. Vienneau, K. de Hoogh, and D. J. Briggs, "Mapping of background air pollution at a fine spatial scale across the european union," *Science of The Total Environment*, vol. 407, no. 6, pp. 1852 – 1867, 2009.

[4] D. J. Briggs, S. Collins, P. Elliott, P. Fischer, S. Kingham, E. Lebret, K. Pryl, H. Van Reeuwijk, K. Smallbone, and A. Van Der Veen, "Mapping urban air pollution using gis: a regression-based approach," *International Journal of Geographical Information Science*, vol. 11, no. 7, pp. 699–718, 1997.

[5] G. Ibarra-Berastegi, A. Elias, A. Barona, J. Saenz, A. Ezcurra, and J. D. de Argandoña, "From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in bilbao," *Environmental Modelling & Software*, vol. 23, no. 5, pp. 622 – 637, 2008.

[6] P. Kanaroglou, N. Soulakellis, and N. Sifakis, "Improvement of satellite derived pollution maps with the use of a geostatistical interpolation method," *Journal of Geographical Systems*, vol. 4, no. 2, pp. 193–208, 2002.

[7] G. Christakos, A. Kolovos, M. Serre, and F. Vukovich, "Total ozone mapping by integrating databases from remote sensing instruments and empirical models," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, pp. 991–1008, May 2004.

[8] V. Singh, C. Carnevale, G. Finzi, E. Pisoni, and M. Volta, "A cokriging based approach to reconstruct air pollution maps, processing measurement station concentrations and deterministic model simulations," *Environmental Modelling & Software*, vol. 26, no. 6, pp. 778 – 786, 2011.

[9] Y. Ma, M. Richards, M. Ghanem, Y. Guo, and J. Hassard, "Air pollution monitoring and mining based on sensor grid in london," *Sensors*, vol. 8, no. 6, pp. 3601–3623, 2008.

[10] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory Air Pollution Monitoring Using Smartphones," in *Proc. 2nd International Workshop on Mobile Sensing*, 2012.

[11] F.J. Moral Garcia, P. Valiente Gonzalez, and F. Lopez Rodriguez, "Geostatistical analysis and mapping of ground-level ozone in a medium sized urban area," *World Academy of Science, Engineering and Technology*, vol. 4, no. 1, pp. 1244 – 1255, 2010.

[12] "Mics-oz-47 ozone sensing head with transmitter board." Available at: http://www.e2v.com/e2v/assets/File/sensors_datasheets/Metal_Oxide/mics-oz-47.pdf, 2014.