Calibrating Low-End Sensors for Ozone Monitoring

Óscar Alvear(✉), Carlos Tavares Calafate, Juan-Carlos Cano, and Pietro Manzoni

Universitat Politècnica de València, 46022 Valencia, Spain
oscar.alvear@alttics.com

Abstract. Performing pollution measurements is a difficult and costly process. On the one hand, specialized laboratories are needed to calibrate sensors and adjust their readings to units that indicate the level of contaminants in the environment, and, on the other hand, measurements depend on the type of sensor. High-end sensors are very accurate but quite expensive, while low-end sensors are more affordable but have less precision and introduce considerable oscillations between readings. This paper presents a methodology to measure ozone pollution data with low-end mobile sensors, focusing on sensor calibration through historical data and the existing environmental monitoring infrastructure. The proposed methodology is developed in three phases: (i) reduction of data measurements variability, (ii) calculation of calibration equations, (iii) and analysis of the spatial-temporal behavior to reduce variations in time produced when data are captured using mobile sensors.

Keywords: Low-end sensor · Sensor calibration · Ozone sensing

1 Introduction

In recent decades, the monitoring of environmental pollutants has become of great importance for governmental institutions and environmental organizations due to the influence of pollution on our lives. There are several institutions worldwide that monitor environmental pollution. In Europe, the European Topic Center on Air Pollution and Climate Change Mitigation (ETC/ACM) brings together 14 European organizations for the analysis and monitoring of climate change. In the United States, the Environmental Protection Agency (EPA) also tracks the evolution of environmental pollution.

There are many studies that analyze ozone levels in cities like Quebec [1] or Toronto [2]. In addition, projects like [3] rely on Waspmote sensors installed in the public transport system of Belgrade, Serbia, to measure the environmental pollution in the city.

In [4], a vehicular ad-hoc network is proposed to monitor different environmental parameters, focusing on the analysis of the data sending rate and the transmission mechanism to minimize the resources consumed.

B. Mandler et al. (Eds.) IoT 360° 2015, Part I, LNICST 169, pp. 251–256, 2016.
DOI: 10.1007/978-3-319-47063-4_24

In [5], authors propose a system to monitor the concentrations of PM2.5 (particulate matter smaller than 2.5 microns) using crowd-sourcing. This work focuses on the analysis of the mechanical sensor design to optimize the air reception, as well as on data fusion techniques to analyze the data. The calibration of the sensors is achieved by analyzing data produced in the laboratory using neural networks. However, authors do not analyze the variability of the data obtained.

In [6], authors analyze the data obtained from different sources, such as traffic levels, weather conditions and pollution, using different Big Data techniques to infer environmental pollution levels with a better granularity.

Previously described works do quite different types of analysis with the data obtained by the sensors, but they do not analyze the data capture process, neither do they focus on the sensor calibration problem. Hence, in this paper we will address both of these problems, in addition to the time variability of measurements associated to sensor mobility.

In the following sections, we will describe the methodology we propose to solve the aforementioned problems. In Sect. 2, we will discuss the most relevant air pollution monitoring issues, detailing the steps taken in the data capture process. In Sect. 3, we will show captured data and make a comparative analysis against available historical data. Finally, in Sect. 4 we conclude the paper.

2 Pollution Monitoring

The pollution monitoring processes seek to measure pollution levels for a particular contaminant in a specific area by relying on special-purpose sensors. These sensors react by varying their properties when in contact with the element to be monitored, but not other elements.

In general, sensors must have certain characteristics to be considered suitable: (i) being only sensitive to the measured property, (ii) not influencing the measured property, and (iii) having a direct relationship with the measured property. In this regard, the main problems of low-range sensors are that they have a large fluctuation between measurements, and some (e.g. ozone sensors) do not meet all the properties previously described because their measurements are influenced by weather conditions.

In this paper, we have used low-end sensors that can be easily obtained in the market to measure ozone levels. In particular, we used a Waspmote Smart Sense Plug And Environment device, which provides a relatively easy way to measure various environmental parameters. The sensor used is the *Ozone Probe Sensor (MIC-2610)*, which can measure ozone variations ranging between 10*ppb* (parts per billion) to 1000*ppb*. The resistance varies between $11\,\mathrm{k}\Omega$ and $2\,\mathrm{M}\Omega$, and the input voltage for this sensor is 2.5 V.

2.1 Monitoring Process

The first step to monitoring environmental parameters is to capture data through sensors. However, this is not a simple process since many problems must be solved in order to obtain reliable data about existing pollution levels. In particular, the following issues should be taken into account: (i) sensor output data measurements are highly variable in ranges close to the real values, and so such variability should be reduced; (ii) the sensor outputs should be transformed into the respective units for each pollutant. In our case, the measured resistance value must be converted into particles per billion (ppb); (iii) if mobile elements are used, time-dependent variability must be removed.

Below we detail how each of these issues has been addressed.

Data Reading: Data retrieval processes should eliminate the oscillations associated to sensor readings, and for this purpose we performed the following steps. First, we calculated the average value of 25 samples (n = 25), with an interval of 10 ms between each consecutive sample.

Afterward, since the variability was still very high, we used a low-pass filter for the process of data analysis with α equal to 0.95 to reduce variability.

$$O_i = O_r + \alpha \cdot (O_{i-1} - O_r) \tag{1}$$

In this equation, O_i represents the current ozone level, O_{i-1} represents the ozone level in the previous measurement, O_r represents the filtered ozone value, and α represents the filter coefficient.

At the end of this process, we have measurements without the oscillations introduced at measurement time, reducing the standard deviation by 66 % (from 5.37 to 1.82).

Unit Conversion: For calibrating the sensor we have done several measurements on different days, and under different weather conditions, to get a broad range of values. These data have been linked to the data obtained from the monitoring station located at the Technical University of Valencia (UPV), Spain.

Considering that the measurements have a dependency on ozone levels and temperature, we developed a second degree polynomial regression influenced by the temperature and the resistance obtained by the sensor:

$$O = \alpha + \beta_1 t + \beta_2 r + \beta_3 r^2 \tag{2}$$

In this equation, α is a regression coefficient, β_1 is a temperature coefficient, β_2 is a sensor reading coefficient, β_3 is the reading coefficient squared, t is the measured temperature, and r is the sensor reading (Resistance). The output O is the ozone level measured. Final regression output is shown in Eq. 3.

$$O = -156.27 + 2.84t + 10.2r - 0.14r^2 \tag{3}$$

The adjustment obtained for this regression was $|\overline{R^2}| = 0.63$ and, comparatively with historical data, the values are very similar.

Time Variability Reduction: To cover large areas of land with a fine spatial granularity we use mobile sensors, which can capture data at various points although at different time instants. So, the difference between measurements O have both time $\triangle O_t$, and spatial $\triangle O_s$ dependencies. Since our main goal is to determine differences between ozones levels in a particular area, it is necessary to eliminate the time variation as follows.

$$\triangle O = \triangle O_t + \triangle O_s \tag{4}$$

$$\triangle O_s = \triangle O - \triangle O_t \tag{5}$$

For the calculation of the ozone time variations we analyzed data from a monitoring station located at the Technical University of Valencia, focusing on historical data between 2008 and 2014. In the historical data analysis, we analyzed ozone evolution focusing on average monthly measurements between 2008 and 2014. It is noted that the values are higher from April to September, and lower for the remaining months. Figure 1 (right) shows the mean values combined with the standard deviation (shaded area) and the maximum values (line).

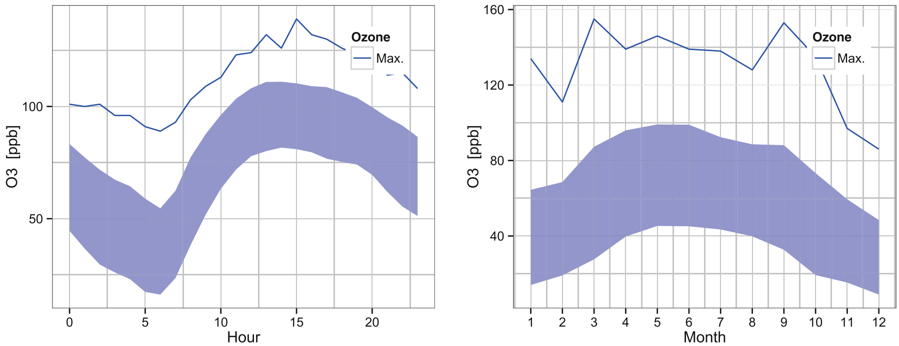

Fig. 1. Ozone evolution in June (left) and the throughout the year (right).

Also, the variation for ozone levels during a representative day of June was analyzed. As shown in Fig. 1 (left), ozone levels reach their lowest value at about 6am, and rise to reach maximum values at 2 or 3 pm, beginning to decline gradually afterward. The behavior for the other months of the year is analogous to the month shown. As a result of the analysis of these data, we can see that ozone has a different behavior during hot periods (from April to September in the northern hemisphere) compared to the other months. During the day, the behavior is very similar to the square logarithmic distribution, with an onset of rapid growth followed by a less pronounced decline. Based on the previous data regarding monthly average values of data between 2008 and 2014, all taken at the monitoring station of the Technical University of Valencia ozone prediction was performed using least-squares logarithmic regression influenced by temperature and season of the year, one for summer, and one for winter. The expression used was:

$$\ln(O_t) = \alpha + \beta_1 s + \beta_2 t + \beta_3 \ln(h) + \beta_4 \ln(h)^2 \tag{6}$$

where h is time of day, s is the season coefficient (3 for winter, 4 for autumn, 7 for sprint and 8 for summer; these values were calculated from the relationship between the means values of ozone), t is the temperature, and the remaining α and β_i values are regression coefficients (β_1 is the season coefficient, β_2 is the temperature coefficient, β_3 is the logarithm of the time of day coefficient, and β_4 is the logarithm of the time of day to the square coefficient).

The values of $|\overline{R^2}|$ range between 0.82 (winter) and 0.91 (summer), showing a behavior very similar to the actual one.

Concerning the procedure followed to correct time-dependent variability, it was: (i) ozone values are calculated at two time instants using Eq. 6; (ii) the difference between the values is obtained; (iii) the calculated variation is reduced from the captured data, according to Eq. 5.

3 Validation

To check the correctness of the proposed methodology, several data collection events took place in different areas of the city of Valencia using the mobile sensor. Different cities areas have been covered, and the data captured was compared against data from the existing public infrastructure.

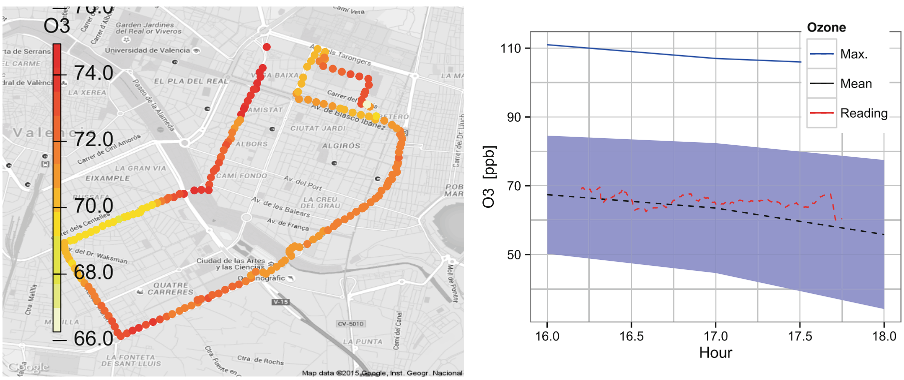

Fig. 2. Data captured (left) and validation expected values for that period (right).

For each route, we have applied the methodology proposed: first, we reduce data oscillation using the low-pass filter (Eq. 1). Next, the readings are adjusted through Eq. 2. Finally, the temporal variation of data is reduced using Eq. 6.

Figure 2 shows data for a particular route and the common values at this time. We can see that ozone levels (reading) are within the range of historical values for the monitored time and close to expected value (mean), which indicates that, with our methodology, we obtain reliable data, allowing to focus our analysis on the spatial variations of pollutants.

4 Conclusions

Environmental pollution monitoring is essential nowadays and, although there are many studies on this topic, few analyze the problems involved in the process of data collection, especially when low-cost mobile sensors are used. In this paper, we have developed a methodology to measure such levels using off-the-shelf sensors to achieve a high spatial granularity compared to that achievable using existing infrastructure.

The proposed process allows measuring and calibrating ozone sensors in a simple and straightforward manner without the need for a specialized laboratory. The data obtained though our method is adjusted to reality using historical data for the target location, and allows analyzing the spatial variability of pollution levels with a small error.

The next steps to be performed include the calculation of the sampling frequency and the spatial granularity of measurements to maintain the evolution of pollution levels in a city under control.

Acknowledgement. This work was partially supported by the *"Programa Estatal de Investigación, Desarrollo e Innovación Orientada a Retos de la Sociedad, Proyecto I+D+I TEC2014-52690-R"*.

References

1. Adam-Poupart, A., Brand, A., Fournier, M., Jerrett, M., Smargiassi, A.: Spatiotemporal modeling of ozone levels in Quebec (Canada): a comparison of kriging, land-use regression (LUR), and combined Bayesian maximum entropy-LUR approaches. Environ. Health Perspect. **122**(9), 970 (2014)
2. Liu, L.-J.S., Rossini, A.J.: Use of kriging models to predict 12-hour mean ozone concentrations in metropolitan toronto-a pilot study. Environ. Int. **22**(6), 677–692 (1996)
3. Brković, M., Sretović, V.: Urban sensing-smart solutions for monitoring environmental quality: case studies from serbia. In: 48th ISOCARP International Congress: Fast Forward: Planning in a (Hyper) Dynamic Urban Context, Perm (2012)
4. Hu, S.-C., Wang, Y.-C., Huang, C.-Y., Tseng, Y.-C.: Measuring air quality in city areas by vehicular wireless sensor networks. J. Syst. Softw. **84**(11), 2005–2012 (2011)
5. Cheng, Y., Li, X., Li, Z., Jiang, S., Li, Y., Jia, J., Jiang, X.: AirCloud: a cloud-based air-quality monitoring system for everyone. In: Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems, pp. 251–265. ACM (2014)
6. Zheng, Y., Liu, F., Hsieh, H.-P.: U-air: when urban air quality inference meets big data. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1436–1444. ACM (2013)