Evaluating air quality by combining stationary, smart mobile pollution monitoring and data-driven modelling

Adriana Simona Mihăiţă[a,b,*], Laurent Dupont[c], Olivier Chery[c], Mauricio Camargo[c], Chen Cai[b]

[a] *University of Technology in Sydney, 81 Broadway Str., Ultimo, NSW*
[b] *Data61/CSIRO Australia, 13 Garden St., Eveleigh, 2015, NSW*
[c] *ERPI Laboratory, Lorraine University, 8 Rue Bastien Lepage-BP647, 54010 Nancy*

Abstract

Air pollution impact assessment is a major objective for various community councils in large cities, which have lately redirected their attention towards using more low-cost sensing units supported by citizen involvement. However, there is a lack of research studies investigating real-time mobile air-quality measurement through smart sensing units and even more of any data-driven modelling techniques that could be deployed to predict air quality accurately from the generated data-sets. This paper addresses these challenges by: a) proposing a comparative and detailed investigation of various air quality monitoring devices (both fixed and mobile), tested through field measurements and citizen sensing in an eco-neighbourhood from Lorraine, France, and by b) proposing a machine learning approach to evaluate the accuracy and potential of such mobile generated data for air quality prediction. The air quality evaluation consists of three experimenting protocols: a) first, we installed fixed passive tubes for monitoring the nitrogen dioxide concentrations placed in strategic locations highly affected by traffic circulation in an eco-neighbourhood, b) second, we monitored the nitrogen dioxide registered by citizens using smart and mobile pollution units carried at breathing level; results revealed that mobile-captured concentrations were 3 to 5 times higher than the ones registered by passive-static monitoring tubes and c) third, we compared different mobile pollution stations working simultaneously, which revealed noticeable differences in terms of result variability and sensitivity. Finally, we applied a machine learning modelling by using decision trees and neural networks on the mobile-generated data and show that humidity and noise are the most important factors influencing the prediction of nitrogen dioxide concentrations of mobile stations.

Keywords: air pollution, eco-neighbourhood, mobile sensing, decision trees, neural networks.

*adriana-simona.mihaita@uts.edu.au, Words in text: 15762.

1. Introduction

Addressing air pollution problems in growing urban cities has become a major problem due to ever increasing traffic in dense populated urban areas, extended industrialisation, high energy consumption, insufficient resources for monitoring and various issues in defining adapted policies (Krupnick, 2008; Kumar et al., 2013). The challenge of managing air pollution becomes more difficult due to its dangerous effects on public health and the multitude of air pollution triggering factors. As a consequence, various studies in recent years have been concentrating on evaluating the impact of bad air quality on citizens, by moving away from traditional monitoring stations which are normally placed in high-altitude locations across cities, towards outdoor and easy-deployable air quality monitoring units, such as mobile sensors installed on cars, bikes or even carried by hand during daily travelling. This new type of collective approach for monitoring air quality brings numerous advantages in terms of real-time pollution measurement and hot-spot identification, but also comes with various challenges due to the amount of data generated and its accuracy. Therefore, there is a true challenge of not only switching towards a mobile air monitoring paradigm (and choosing the best adapted sensing units) but also in modelling efficiently the data generated by all these mobile sensing units. Data-driven modelling is an efficient way of extracting valuable information from generated data-sets, but is less effective when the data is sparse, incomplete or contains many inaccuracies.

There is currently a lack of research studies integrating both real-time air quality measurements through efficient mobile sensing units and afferent data-driven modelling. Appendix A presents a comparative state-of-art table in which we have gathered recent studies adopting new sensing technology available for outdoor air quality monitoring and/or efficient methods for data-driven air quality prediction. The findings have been analysed based on: a) the sensing technology which was used, b) whether the pollution stations were mobile or not, c) whether the units have been used near breathing level (close to human level), d) the time period of the experimentation and most importantly, e) the data analysis procedure which has been applied in order to obtain accurate predictions or meaningful insights. Our initial observation revealed that majority of studies have been using outdoor air quality sensors, but only half of them were used in a mobile setting, and were either fixed, attached to street lamps or inside monitoring boxes. For example, (Castell et al., 2018) installed 17 nodes monitoring NO_2 in kindergartens in Oslo at heights of 2.5-3 meters altitude, (Ercilla-Montserrat et al., 2018) installed outdoors sensors on rooftops of greenhouses or urban houses for detecting heavy mental contaminations due to atmospheric pollution, (Heimann et al., 2015) used 32 low-cost electrochemical sensor nodes in and around the city of Cambridge, UK, to extract underlying pollution levels (baselines) from air measurements, while (Popoola et al., 2018) used 17 low cost portable air quality devices which were installed at the London Heathrow airport in UK on lampposts at 3-meters hight altitude.

While outdoor fixed monitoring can provide a better information about the air quality, their high level of installation (often around 3 meters altitude) does not reflect a true impact of the pollution at human level. More recent studies have started to explore not only fixed outdoors sensing units, but also mobile sensors which can be carried by hand, installed on cars/etc. For example, a recent study published in (SM et al., 2019) used a smart personal air quality system carried by pedestrians walking on predefined paths or by bus in different locations in India; the analysis revealed non-linear relations between the gaseous pollutant concentrations versus the resistance offered by different sensors; there wasn't however any data-driven modelling or predictions of the pollutants being measured. Similar direction was taken by (Minet et al., 2017) which used portable sensors for measuring NO_2 concentrations in Montreal carried by pedestrians walking or biking; the authors applied statistical and land-use regressions for data modelling which revealed sensitivity to the number of road segments and the number of visits per segment. Other studies presented in (Suriano et al., 2015) and (Zappi et al., 2012) deployed similar mobile sensors near to the breathing level, but there is no data-driven modelling attached to the results, except from visual data representation.

Recent advancements in machine learning showcased the power of using such models for any type of investigation where large volumes of data have been generated. For example, a recent study published in (Zhou et al., 2019) presents an artificial intelligence approach based on a Deep Multi-output LSTM (DM-LSTM) neural network model for predicting the air quality in the city of Taipei, Taiwan. The study is very promising but the data being used is generated by five fixed air quality monitoring stations in the city; these data-sets, once again, do not reflect pollution concentrations felt at human levels. Similarly, other powerful data-driven methods have been used for air quality prediction such as DEA (data envelopment analysis) in (Zhou et al., 2018), or Feed-Forward Neural Network and Random Forests in Borrego et al. (2018). But once again, the information being used comes from fixed data sources, not mobile sensing units carried by citizens for daily travel.

In an attempt to address these issues and research gap, this paper presents a research approach for: a) choosing the best mobile sensing units for outdoors air quality monitoring in an urban neighbourhood, b) constructing reliable experimenting protocols for evaluating the accuracy of such units with regards to fixed air pollution stations, and c) constructing a data-driven approach using decision trees and neural networks for predicting the air quality from mobile-generated sensing data. We strongly believe that building such a data-driven modelling over mobile-generated pollution data would help to learn and detect patterns of air monitoring.

The paper is organized as follows. Section 2 presents the challenges faced by air quality monitoring in an international context and the need to use low-cost mobile pollution sensing units in an opportunistic citizen sensing. Section 3 introduces the case study of the Nancy Grand Cœur eco-neighbourhood and the air quality evaluation methods currently deployed in the city. Section 3.1 presents the stationary and mobile smart sensing units chosen to conduct the current study, followed by a description in Section 3.2 of three air quality experimenting protocols which we deployed in the most congested intersections.

This paper is a further extension of our first two air pollution protocols published in (Mihăiţă et al., 2018), by adding a third experimenting protocol and a data-driven prediction modelling on the mobile sensing generated data. The results of the three protocols are provided in Section 4 and concentrate around nitrogen dioxide and noise level evaluations, followed by conclusions and future perspectives of the current work.

2. Challenges in air quality monitoring

2.1. International Context

Currently air pollution is monitored at a regional level by networks of static and sparse stationary Air Quality Monitoring (AQM) stations, equipped with instruments for measuring various pollutants such as: carbon monoxide (CO), nitrogen oxides (NO_x), sulphur dioxide (SO_2), ozone (O_3) and particulate matter (PM). The risk information is often provided as a concentration of pollutants or as an index of air quality (AQI) at a scale which can be easily interpreted by the public. These AQIs can vary in their approach for determining pollutant concentrations, as they follow different regional policies (Plaia and Ruggieri, 2011) which can differ from one country to another. For example, Canada has adopted an Air Quality Health Index on an 11-point scale obtained from a non-linear combination of particulate matter 2.5 microns ($PM_{2.5}$), nitrogen dioxide (NO_2) and ozone (O_3) (Stieb et al., 2008). On the other hand, in Europe all countries are required to comply with EU directives such as the Council Directive 96/62/EC on ambient air quality assessment and management, commonly referred as the Air Quality Framework Directive. Therefore the hourly and daily AQIs are calculated on a scale from 0 to 100 by taking into consideration PM_{10}, NO_2, O_3, and where accessible, $PM_{2.5}$, SO_2 and CO (van den Elshout et al., 2012). The EU directives recommend as well to install a specific number of monitoring stations for individual pollutant monitoring, based on the number of inhabitants and the geographic partitioning of that area/city (EU Air Quality Directive 2008/50/EC).

Although they offer high-precision results, the AQM stations are often high-priced and need a significant amount of resources to be routinely maintained and calibrated (Chong and Kumar, 2003). Often, the temporal and spatial resolution of a network of fixed AQM stations is far too sparse to incorporate the contribution of different sources of pollution without significant constraints and assumptions. The AQM stations would offer a global insight over large urban areas but they can not identify pollution hotspots inside the city centre or around large industrial areas for example. Often there are no real-time pollutant concentration maps available at high-resolution ($< 1m$) for large urban areas because they require a large amount of data, computing facilities and input details which are not available for many cities (Kumar et al., 2015). These aspects led to the emergent idea of using more fine-grained monitoring units in an outdoor setting, which would come at a lower cost and maintenance. The motivation and benefits of using such units are described in the next section.

2.2. The need for low-cost sensing units

Recent improvement of low-cost sensor technology has lead to the development of a multitude of robust micro-sensing units (MSUs) with a lower power consumption which can be used for detailed air quality surveillance. These MSUs can be used either as individual nodes or in an interconnected distributed network, and would collect high-resolution spatial an temporal data when being mounted on cars, bicycles or carried daily by pedestrians. One of the main advantages of using low-cost sensing units is that they provide more input conditions, especially if they are used in significant numbers for detecting pollution hotspots. Their real-time information would allow a rapid assessment of the pollution problem and would lead to more efficient prevention strategies.

Due to a higher granularity provided, easy to handle functionalities and rapid access to real-time pollution concentrations, various research programmes have started to test both fixed and mobile monitoring sensors (Mead et al., 2013). As well, including citizens in the testing and exploration of urban air pollution opens new opportunities for direct environmental awareness, debate and future prevention strategies. Some examples of such projects are: Air Quality Egg (AQE, 2016), Citizen Sense (Gabrys, 2016) and the Smart Citizen Kit (SCK, 2016), which offer a centralized collection of data, processing and real-time map visualisation through on-line platforms and mobile applications. These "citizen sensing" projects intend to expand citizen engagement in environmental issues, and help them making changes in their daily journey-to-work trips in order to avoid polluted urban areas. Currently there are various low-cost air quality sensors which are commercially available (Alphasense, 2016) or prototype sensor networks (MIT, 2016). For a detailed state of the art regarding low-cost pollution sensors the user can refer to the works of (Kumar et al., 2015). For our current study we have also conducted a state-of-art investigation on low-cost sensing units which will be further discussed in Section 3.1.

While this radical change in the air monitoring mentality promises a flexible pollution surveillance solution, the question around the accuracy of the generated data still remains an open research question. The main downfall of low-cost sensors remains their relatively low accuracy compared to official fixed AQM stations or other benchmark devices (Williams et al., 2013). For example, (Moltchanov et al., 2015) has observed a sensor-specific temporal variation of the calibration parameters, and proposed a periodical calibration of wireless sensors based on the nearby AQM stations which would capture the fine and dynamic spatial variability of pollutants at a high temporal resolution. Questions related to the battery power of the sensors and the life-expectancy of low-cost sensors can also be seen as a drawback for adopting MSUs, but their flexibility and remote-control possibility for data transmission and collection increased their popularity. Together with meteorological sensors for measuring humidity, temperature, wind speed and direction, they can form the basis for assessing pollution levels and can induce substantial behavioural changes at a larger scale amongst citizens. But these small monitoring units need to be used at a larger scale in order to provide a global and complete picture of the urban air pollution; therefore the idea of citizens using them on a daily basis brought a higher popularity to this modular and dynamic approach for air pollution monitoring.

2.3. Towards opportunistic citizen sensing

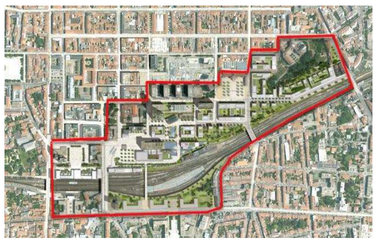
The idea of using low-cost sensing for monitoring air quality has led to a shift in the air quality data collection, generating the notion of opportunistic citizen sensing, which implies that data collected for one specific purpose can be used for other purposes as well (Campbell et al., 2008). Involving citizens in the data collection gave birth to the notion of anthropocentric opportunistic sensing, in which large volumes of sensing data are collected, stored and fused for further analysis and interpretation (Kapadia et al., 2009). Using data analytics for extracting meaningful insights from daily air pollution and noise exposure will provide unparalleled feedback to the citizens regarding their daily trips and route choice behaviour. By following the opportunistic aspect, the air pollution analytics can be coupled together with clinical research studies for analysing correlations between citizen movement and biological exposure (NIEHS, 2014). Emergency alerts could then be triggered when unusual air quality levels are signalised in specific areas of the city or when a significant number of citizens present clinical side-effects of air pollution exposure.

Our current work is driven by the idea of deploying an opportunistic citizen sensing coupled with an intelligent data-driven modelling in order to build future pollution pattern recognition and real-time anomaly detection. This would lead to the creation of a real-time situation awareness for pedestrian and travellers, helping them to customize their journey in order to improve their health condition. The air pollution data generated from such mobile and opportunistic sensing will be further integrated with existing traffic simulation models and improve the air quality prediction. The next section presents our case study and the steps we took for building a real-life mobile air pollution monitoring together with data-driven modelling techniques.

3. Case Study

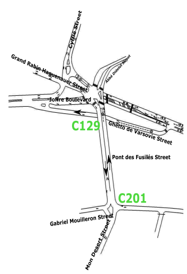
The concept of *eco-neighbourhood* has emerged from a need to build an innovative place for technical, economic and social experimentation. Their role is quite complex as they must meet several principles of sustainable development (MCT, 2016): 1) involve all the city actors, 2) contribute to improving the daily life by developing a healthy and safe living environment for all residents, 3) participate in the economical and local dynamics, 4) promote a responsible resource management and adaptation to the climate change. The eco-neighbourhoods offer the opportunity to experience and anticipate the evolution of cities by guiding the decision makers. The latest changes in the development of digital tools and design practices (collaborative approach, usage integration directly from the design phase, citizen involvement in experimental smart city projects), offer new perspectives for quantifying the impact of the urban changes (Dupont et al., 2015).

Motivation: With the urban project Nancy Grand Cœur (GN, 2012) the Grand Nancy Metropolis in France, wants to rehabilitate the 15-hectares area around the historical train station including its railway and industrial brown field. A visual representation of the train station hosting almost 9 million passengers each year is provided in Figure 1a).



(a) Urban project for NGC by 2020, Source: Arep ville - J.M. Duthilleul.

(b) Urban area in 2013.



(c) Focus on most circulated intersections.

Figure 1: Case study of the eco-neighbourhood Nancy Grand Cœur.

This ecological urban project is intended to be delivered by 2025, and the objectives for this central area are manifold: new fluid mobility, better traffic regulation, reconciliation between historical and modern neighbourhoods of the city, improved air quality, extended green spaces, reduced energy consumption, comfortable homes and offices. An important step to respond to this wide variety of problems is to analyse the air quality inside the neighbourhood, especially at a human level, when passing through the most circulated intersections of the Nancy Grand Cœur (NGC) neighbourhood: $C129$ and $C201$ (see selected area in sub-plots b) and c) from Figure 1)). Understanding how highly circulated streets impact the citizens on their daily journey-to-work trips is a true challenge which can give a clear insight on how the eco-neighbourhood needs to be reconfigured in order to protect its inhabitants.

Solution: The work presented in this paper is a continuation of our previous studies (Mihăiţă et al., 2014, 2016) in which we proposed an integrated air pollution and traffic simulation model for building a simplified NO_2 estimation model which helped predicting and visualizing various environmental changes inside the NGC eco-neighbourhood. Our previous study has used reliable data sets provided by the Air Quality Monitoring Stations (AQM) belonging to the local air-quality management centre, as well as meteorological data. While these data sets are of high accuracy, they only represent global concentrations computed by the AQM station installed at high elevation from the ground (more than 10 meters altitude) in a single location in the city. The real and direct impact that pollution can have at the human level could be completely different than higher dispersed pollution concentrations. Therefore, our major objective is to build a research monitoring framework using mobile smart sensing units transmitting air quality information in real-time, coupled with a prediction engine and situation awareness for citizens travelling daily in NGC. Providing health risk information caused by air pollution is an important step for raising citizen awareness and triggering changes in their daily travelling behaviour.

The NGC project has the initiative to change the structural configuration of the $C129$ and $C201$ intersections represented in Figure 1c), in order to allow a higher inflow of vehicles to cross the neighbourhood every day. A large amount of vehicles in densely populated areas will contribute to an increasing deterioration of the air quality due to higher motor vehicle emissions. In 2012, the U.S. Environmental Protection Agency (US EPA) has shown that 61% of the total emissions of carbon monoxide (CO) and 35% of total emissions of nitrogen oxide were produced by highway vehicles (US-EPA, 2016). The complexity of the air pollution lies in its extent and the large amount of factors changing its behaviour, making it even more difficult to implement measures for protecting the citizens. According to the 2012 air quality assessment (MEDE, 2012), air pollution is caused by various industrial, commercial, domestic, agricultural activities, but the traffic congestion is the major cause. As 56% of the nitrogen dioxide in the air is caused by road transportation (MEDE, 2012), for this initial case study we mostly focus on NO_2 concentrations. The objectives of our study are manifold: 1) measuring air quality at a granular level in the city by using smart pollution sensors, 2) prepare the field for integrating citizens in a daily and global air quality data collection, 3) provide insights by comparing outputs of stationary and mobile smart pollution sensors, 4) derive data-driven predictions by using latest generated data from mobile units.

3.1. Choice of sensing units

In France, the State entrusts the monitoring of air quality to twenty approved AASQA associations (1901 Act) led by the ATMO Federation (ATMO, 2016). Air Lorraine (Lorraine, 2018) is one of the selected air monitoring associations which is responsible for continuously monitoring the air quality inside NGC and which has been our main reference source for testing the accuracy of the mobile sensor units deployed for this study.

As we are currently interested in analysing more granular air quality information, a series of mobile sensing units have been considered; a selection of these MSUs is provided in Appendix B, Table B.5. The choices have been selected after a thorough analysis of existent sensing units on the market at the time of the current project, their accuracy, the feasibility of being used outdoors on a daily basis, their costs and daily maintenance. For this paper, we only present the results obtained when investigating three units which are detailed in the following. Other comparative studies of smart pollution units are currently under testing and evaluation.

Figure 2: a) Passive NO_2 tubes b) Azimut station c) Smart Citizen Kit.

In the following we give a brief discussion about the three MSUs which have been chosen for our experimenting protocols, the reason, scope as well as their advantages and disadvantages for the current study.

1. **Passive tubes** (Figure 2a)): The technique (passive sampling) is based on the passive transfer of pollutants by simple molecular diffusion of ambient air to an adsorbent which is specific to the targeted pollutants. The sampling module is in the form of a porous tube, called "passive tube" which is filled with adsorbent. The passive tubes are fixed in a protection box attached to a support near congested traffic areas. After the exposure time has elapsed, the tubes are sent to the Air Lorraine laboratory for analysis. The concentrations of pollutants obtained by this technique are concentrations averaged over the entire sampling period. This technique has been used for sampling of nitrogen dioxide (NO_2) and has the main advantage of being low cost and not requiring electrical recharge. The passive tubes have been successfully deployed in various project such as the (LigAir, 2009) project for characterizing ambient levels of formaldehyde around industrial sites, or for modelling air quality in the eco-neighbourhood Danube from Strasbourg (ATMO-Alsace, 2012), which was highly impacted by intense circulated areas. The main disadvantage of using the passive tubes is related to the fact that the results are analysed at the end of the experimentation period and cannot detect peaks of localized pollution concentrations during congested traffic hours. These tubes have been used for the implementation of the first experimenting protocol, which is detailed in Section 3.2.1. Nevertheless, they represent an accurate base for comparing NO_2 concentrations with official reported pollution levels from the AQM station during our testing period.
2. **Azimut Station**(Figure 2b)): is a product of Azimut Monitoring (Azimut, 2017) which uses electrochemical gas sensors for measuring the NO_2 emissions. Through a portable emission analyser it can provide continuous real-time monitoring of NO_2, O_3, noise, temperature and humidity. The station can be mounted on cars, bicycles and can be carried by hand while its data is transmitted through GPRS, having a 48-hour autonomy.

The main advantages of this mobile sensing unit relies in its easy installation and utilisation, a two day autonomy and real-time data visualisation. The station has been successfully used for building the open data portal MyGreenServices by INRIA (Trousse et al., 2014) which offers real-time visualisation of environmental data collected by citizens, generates alert services and has a forum for sharing ideas and best practices in terms of eco-responsible behaviour. In an attempt to promote citizen awareness and trigger changes in the daily travelling behaviour of citizens, INRIA has provided for our project one Azimut station for testing, evaluation and comparison. The data analytics provided through the platform have been used for carrying out the second and third experimenting protocols, which are detailed in Sections 3.2.2-3.2.3.

3. **Smart Citizen Kit (SCK)**(Figure 2c)): is a crowd-funded product developed by Fab Lab Barcelona at the Institute for Advanced Architecture of Catalonia (SCK, 2016). This low cost mobile sensing unit can provide real-time data measuring for NO_2 and CO concentrations, noise, temperature, humidity and light. Its solar-panel and low power consumption, together with an ergonomic design make it attractive for daily usage. The device streams the data measured by its integrated sensors over Wi-Fi, using the FCC-certified wireless module on the data-processing board. Results can be visualised through the on-line interface or through a dedicated mobile app. With disregard to is various advantages, the SCK is only produced on order and needs initial settings from the user. In Section 3.2.3 we present the results we have obtained by comparing the NO_2 concentrations provided by the Smart Citizen Kit and the Azimut station.

3.2. Experimenting protocols

This section describes the three experimenting protocols we have deployed during two weeks time period (29^{th} of April 2015 to 13^{th} of May 2015). The length of the experiments has been tied to project constraints for council approval, unit installation, data measuring and processing. For each experimenting protocol we provide insights regarding the purpose, the materials which have been used, the constrains as well as the data acquisition for interpretation.

3.2.1. First Experimenting Protocol

The first experimenting protocol aimed at determining a reliable data source for further comparison of NO_2 concentration with the regional AQM station which happens to be placed in the centre of NGC. For this study, 10 passive tubes provided by Air Lorraine have been installed at 3 meters altitude on street pillars inside two most circulated intersections of NGC ($C129$, $C201$), by using protection cases and fixing clamps.

The placement of the tubes has been chosen to be near some of the most congested streets, as represented in Figure 3. Seven tubes have been used to actively monitor these streets (Figure 3a)), two tubes have been placed near the location of the regional AQM station and one tube has been kept as a

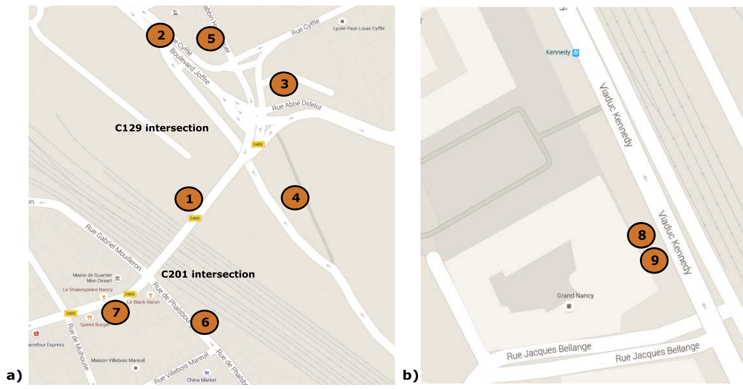


Figure 3: Locations of passive NO_2 tubes in NGC

duplicate reference (Figure 3b)). While the tubes have been used successfully in other eco-neighbourhoods as previously detailed in (ATMO-Alsace, 2012), the placement of the two tubes near the AQM station has the role of double testing and verifying their accuracy during the experiment. In order for the measures to be accurate and non-saturated, certain constraints had to be addressed; tubes needed to be located: a) far from stopping areas such as traffic stops or parking slots in order to avoid over-saturation of the pollutant concentration, b) far away from blooming trees or high-ventilation areas c) at 2-3 meters altitude and d) far away from covering structures which would block air circulation.

At the end of the experimentation period, the tubes have been analysed by Air Lorraine and the results are presented in Section 4.1 of this article.

3.2.2. Second experimenting protocol

The second experimenting protocol used the Azimut mobile station which was carried in hand (at a human level, around 1.5 meters attitude) during the two weeks experimenting period by volunteers walking inside the NGC neighbourhood. In the first experimenting protocol the tubes have been installed at 3-meters altitude due to installation constrains, a height where the pollutant concentrations are starting to disperse in the air. Therefore, this installation altitude is not favourable for a direct evaluation of the "perceived" air pollution impact at a human level, which represents a major challenge and objective of this second experimenting protocol. The daily trajectory of the volunteers would pass near each of the 9 passive tube locations presented in the previous section, where the subject would wait for 5 minutes near each tube. The daily circuit is represented in Figure 4 a) and b).

The advantage of using the Azimut station relies in its high flexibility, mobility and real-time transmission of results through the MyGreenServices platform represented in Figure 5. The platform offers centralised results, personalised filtering, instant evaluation of concentrations based on European air quality monitoring indexes, as well as predefined alerts for raising real-time situation awareness. Having immediate access to results provides a higher awareness re-

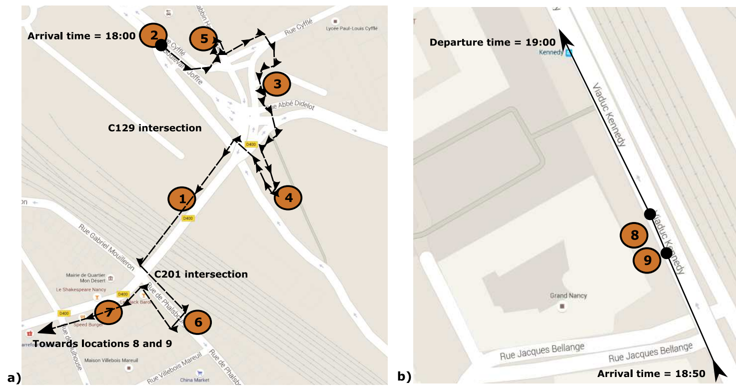


Figure 4: Daily trajectory using Azimut Station in NGC.

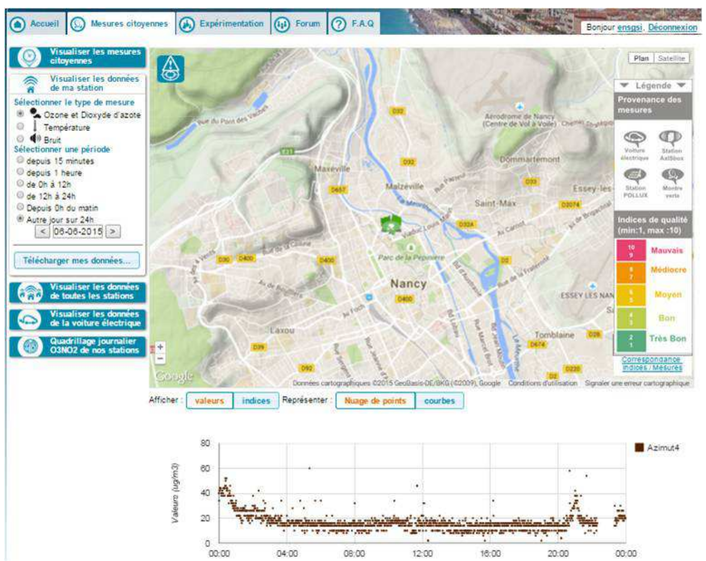


Figure 5: MyGreenServices platform for visualising data collected by the Azimut station.

garding the exposure to pollutants, for both specialists and citizens travelling in the neighbourhood on a daily basis. Using the mobile Azimut station on this predefined trajectory allowed a consistent check of data transmission and quality, which was then compared and matched to the stationary units from the previous experimenting protocol. Despite the above advantages, the main limitations for applying this experimenting protocol were: a) the daily recharge of the Azimut station in order to prevent a discontinuity in the data collection and b) the lack of multiple Azimut stations which would have been tested in parallel on the same trajectory. The data profiling and results obtained during this protocol are further discussed in Section 4.2.

3.2.3. Third experimenting protocol

The third experimenting protocol used both the Azimut mobile station and the Smart Citizen Kit, in order to compare the accuracy and behaviour of both mobile sensing units in the same testing environment. The main purpose of this experiment was to investigate the usage of various mobile smart sensing units which had to monitor the same air pollutants. During 80 hours ($13^{th} - 15^{th}$ of May 2015), both units have been placed outside on an open private balcony facing one of the most circulated streets in $C129$, near the location of passive tube 5.

The main constraint for this experiment was the impossibility to use the SCK in the mobile circuit (previously presented in the second protocol), mostly due to the lack of Wi-Fi availability for transmitting the data to the online platform when walking in NGC. The SCK needs local Wi-Fi configuration for data transmission which was not available through mobile tethering, while local data storage was not possible with the received unit; therefore, we have configured the SCK for Wi-Fi usage on a private balcony near the closest passive tube. The data collected from the Azimut station followed the procedure explained in the previous protocol, while SCK provided real-time access to the collected data trough the Smart Citizen platform (SCK, 2016), as represented in Figure 6. The results and analysis of this third experimenting protocol are further discussed in Section 4.3.

Figure 6: Smartcitizen.me platform for visualising data collected by the Azimut station.

4. Data profiling and results

4.1. First Protocol results

As previously mentioned in Section 3.2.1 the purpose of the first experimentation protocol was to establish an accurate and reliable source of information regarding the air pollution inside the eco-neighbourhood around main circulated areas (hotspots). The NO_2 levels collected from the passive tubes have been investigated in the air quality laboratory of Air Lorraine (Lorraine, 2018), and evaluated according to the ATMO indexes defined nationally by the French government: see decree on 22^{nd} of July 2004 related to air quality indices (MEDD, 2004).

Value($\mu g/m^3$)	Index	Qualifier
$0 - 29$	1	Very Good
$30 - 54$	2	Very Good
$55 - 84$	3	Good
$85 - 109$	4	Good
$110 - 134$	5	Medium
$135 - 164$	6	Medium
$165 - 199$	7	Poor
$200 - 274$	8	Poor
$275 - 399$	9	Bad
≥ 400	10	Very Bad

Table 1: ATMO French National Index scale for NO_2. (Lorraine, 2018).

The evaluation scale is provided in Table 1 and uses indexes from 0 to 10 for different NO_2 concentration levels depending on their severity (0 and 10 standing for a very good, respectively very bad air quality index). The results for each tube are provided in Table 2 and are coloured accordingly to these standard indexes. We make the observation that Tube 10 has been kept as a duplicate for verification purposes, whereupon the low scored value. The results have been obtained during a time period which registered a mean temperature of 13.6°C and a mean pressure of 1013.0 hPA. The investigation results indicate that overall the tubes have registered very good NO_2 concentrations corresponding to index 1 or 2 (according to Table 1). Tube 7 presented higher NO_2 levels which is explained by its position near a narrow but highly circulated road in the neighbourhood.

Tube Number	Registered Value($\mu g/m^3$)	Qualifier
1	22.2	Very Good
2	27.0	Very Good
3	24.7	Very Good
4	14.9	Very Good
5	29.2	Very Good
6	22.2	Very Good
7	53.6	Very Good
8	23.2	Very Good
9	23.1	Very Good
10	0.3	Very Good

Table 2: NO_2 results of the passive tubes investigation

According to Table 2, Tubes 8 and $\dot{9}$ which have been placed near the AQM fixed station of Air Lorraine (located near the train station) presented an average NO_2 level or $23.15(\mu g/m^3)$. The official NO_2 concentration registered by the

AQM station during the same period of time indicated a level of $24.11(\mu g/m^3)$ which translates in a 3.9% error between the tubes and the AQM station. The location of these tubes near the AQM station (see Figure 4b)) has been intentionally chosen for re-verifying the accuracy of passive tubes against the official reported NO_2 levels at the whole regional level.

Findings: The findings of the first protocol confirm a high accuracy of the fixed tubes which have been later used for the comparison analysis with the smart mobile stations. This protocol was the base set-up for comparing mobile sensing results and making sure their accuracy is validated.

4.2. Second Protocol results

4.2.1. NO_2 findings

The second experimentation protocol aimed at investigating the air pollution and noise levels as reported by the smart mobile unit Azimut. As previously detailed in Section 3.2.2, the experiments took place between the same time period when the fixed pollution tubes have been tested. For easing the experimental result interpretation, in this section we present the analysis results obtained during the evening peak hour (6pm-7pm). Figure 7 presents the NO_2 concentration levels registered for every day of the study period during PM traffic peak with the average values ranging from a minimum of $41.48(\mu g/m^3)$ at 7pm up to a maximum of $91.3(\mu g/m^3)$ at 6:45pm. Figure 7 contains as well markers on the X-axis of the time period that corresponds to the waiting time near each tube location along the trajectory shown in Figure 4 (for example between 6:00pm and 6:05pm the Azimut carers would be stopping near Tube 2 in order to record the concentration in this hotspot of the neighbourhood).

According to Figure 7, the lowest pollution scores have been obtained during Sunday 03/05/15 as traffic activity in the city center was low. The highest NO_2 levels reached $152(\mu g/m^3)$ during Monday 11/05/15 which corresponds to a "medium" towards "poor" pollution level according to Table 1. This can be explained by an increased traffic demand on Mondays when citizens returned to work after long weekend (8/05/15 was public holiday). An important observation is that the average peak of NO_2 concentrations over the whole study period has been registered between 6:45pm and 6:50pm, which corresponds to the waiting time near Tube 7. A possible explanation comes from the narrow street configuration and dense traffic that circulates in this area compared to other locations which have a wider exposure to air flow and multiple circulation lanes.

Once again, the finding confirms that pollution levels near Tube 7 (as registered using the mobile pollution sensor Azimut) are higher than those of other tube locations. Overall, the mobile station seems to register lower NO_2 concentration levels towards 7pm, at the location of Tube 8 and 9, which are placed near the AQM station of Air Lorraine. The findings confirm a similar trend between the NO_2 values registered around the tube locations by both stationary and mobile sensors with differences which will be discussed in the following.

To summarise, the real-life mobile sensing we have deployed revealed its high capability to detect correctly the hot-spots for NO_2 concentrations under the influence of severe traffic congestion.

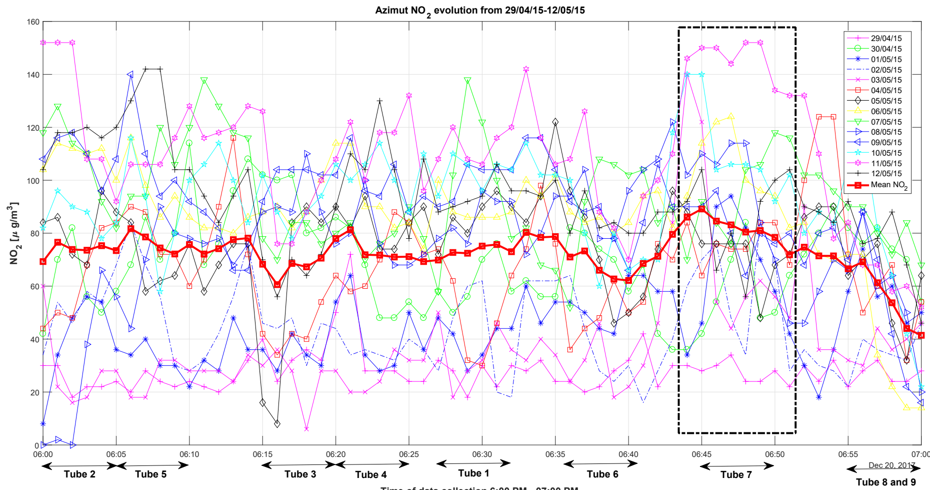


Figure 7: Daily NO_2 concentrations registered by Azimut mobile station.

4.2.2. Weather impact on NO_2 concentrations

An important aspect that we considered for this study is the weather impact on the NO_2 evolution. Our previous findings (Mihăiţă et al., 2016) revealed that temperature, wind and humidity play a very important role in influencing the pollution dispersion or accumulation in the city and that predicting the air pollution levels is not only impacted by a specific day profile (day of the week/weekend/public holiday) or mobility patterns (peak/non-peak hours), but also by various exogenous factors which can highly affect the pollutant evolution in time. In order to understand the current emission levels, one needs to analyse not only the daily traffic patterns, but also previous weather conditions that have led to the current concentration levels. In the following we conduct a comparative analysis of different day profiles, weather conditions and traffic counts registered during the study period in order to understand how these factors can influence air pollution monitoring at both human level and stationary monitoring stations placed at higher heights.

Figure 8 presents the NO_2 evolution for Mondays, Tuesdays, Thursdays and Fridays, while Table 3 summarises the $Temperature(°C)$, $Humidity(\%)$, $Precipitations(mm/h)$ and $Wind(km/h)$ registered during the study period. The NO_2 concentration levels for two typical Mondays are shown in Figure 8a) and although both days presented an average of 550 cars per hour passing the predefined trajectory shown in Figure 4, one can easily observe that the NO_2 concentration on 4/05/2015 was significantly lower than that of 11/05/2015. Although the weather parameters during these two days are almost similar according to Table 3 (temperature was around $22-25°C$, wind around $11-13km/h$), by analysing the three previous days to our chosen dates, one can notice different weather conditions: prior to 4/05/2015 the humidity was higher, temperature lower and there was less sunshine, while prior to 11/05/2015 there were lower precipitations, a higher temperature and more sunshine. This aspect indicates that high humidity, low temperature and high precipitations can reduce the NO_2 accumulation in the city. The finding is also supported by the comparison between Thursdays as presented in Figure 8c); similarly, the highest NO_2 levels were registered during 07/05/2015, a day with higher temperature, lower humidity and low wind levels, when compared to 30/04/2015 which registered much lower temperatures and higher precipitations.

Day	$T(°C)$		H(%)		Pr (mm/h)		W (km/h)	
	6pm	7pm	6pm	7pm	6pm	7pm	6pm	7pm
29/04/15	9	8	80	82	0.7	0.68	10	9
30/04/15	10	9.8	85	86	0.8	0.6	11	9
01/05/15	9.8	8.5	90	92	0.8	0.8	13	13
02/05/15	13	12.8	89	90	0.2	0.2	7	9
03/05/15	18.6	17.6	81	88	0	3.5	11	9
04/05/15	22.1	21.2	62	67	0	0	11	7
05/05/15	20.5	20	46	46	0	0	20	24
06/05/15	17.1	16.1	40	44	0	0	24	22
07/05/15	17.4	16.9	41	43	0	0	9	7
08/05/15	20.0	18.4	46	60	0	0	15	17
09/05/15	19.9	19	43	46	0	0	19	19
10/05/15	21.6	20.9	49	51	0	0	7	11
11/05/15	25.8	25.1	43	48	0	0	13	11
12/05/15	24.6	21.8	60	63	0	0	26	24

Table 3: Weather conditions during the study period.

A special case is the comparison between Tuesdays (see Figure 8b) when humidity and wind conditions were almost similar during the observed time period; this translated in similar NO_2 levels, except from 6:00pm until 6:18pm on 12/05/2015 when the higher temperature (24.6°C) registered at 6:00pm induced higher NO_2 levels. After the temperature decreased to 21°C around 6:18pm, the NO_2 level presented similar evolutionary patterns as one week before. The comparison for Friday is shown in Figure 8d) and strengthens even more our previous findings, with the observation that the lower concentration levels registered on 01/05/2015 were also influenced by the reduced traffic flow as this day was a national public holiday. The public holiday on Friday 1/05/2015 has influenced as well the NO_2 levels on the next Saturday 2/05/2015 and Sunday 3/05/2015, as represented in Figure 9. The mean number of cars during the chosen study period averaged around 234 cars, which is almost half than during a normal week day. Moreover, Figures 9a) and b) indicate a clear difference between the average NO_2 concentrations registered in a weekend preceded by public holiday [44.68($\mu g/m^3$) for Saturday 2/05/2015 and 38.09($\mu g/m^3$) for Sunday 3/05/2015] and the next regular weekend with no public holiday [90.49($\mu g/m^3$) for Saturday 9/05/2015 and 93.47($\mu g/m^3$) for Sunday 10/05/2015]. The dif-

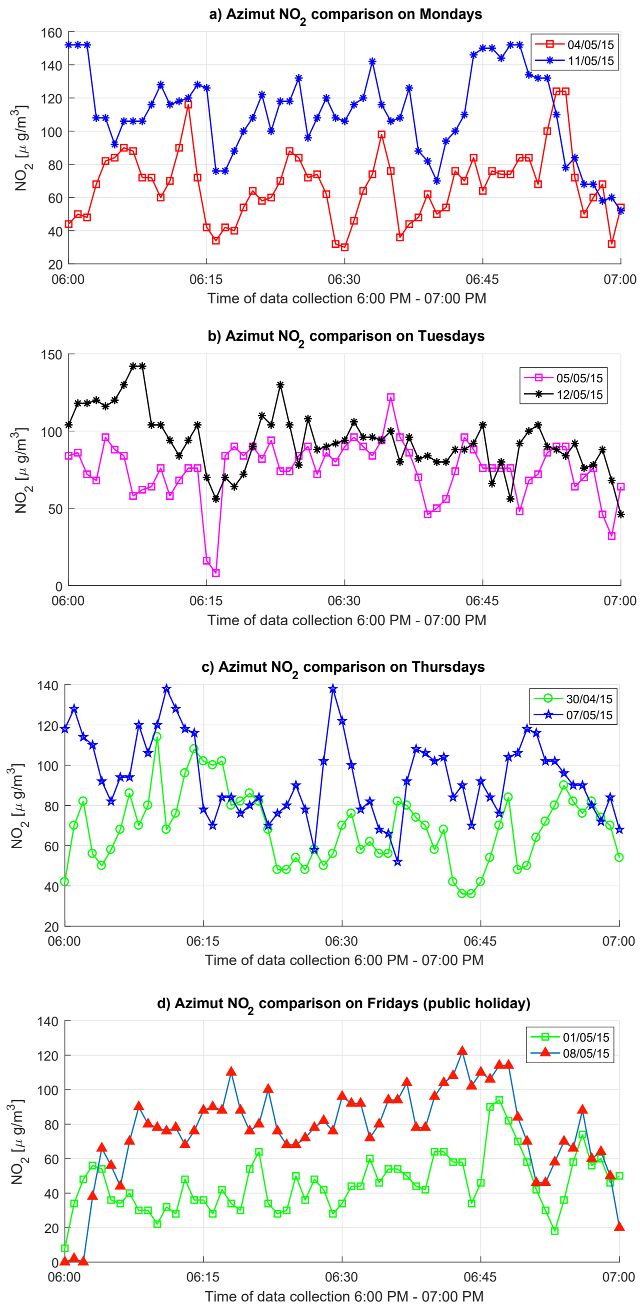


Figure 8: Daily comparison of NO_2 concentrations.

ference is not only caused by the increased number of cars during a regular weekend, but also by higher temperatures and lower air humidity.

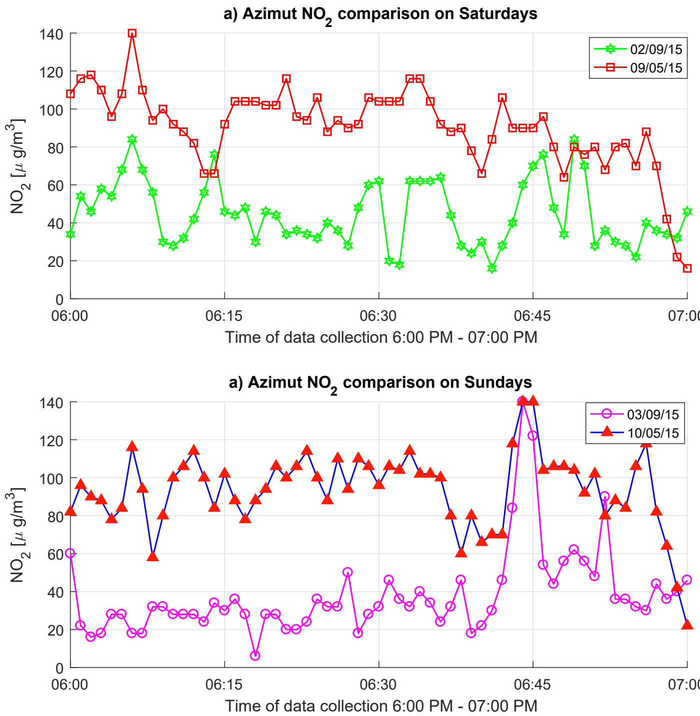


Figure 9: Weekend comparison of NO_2 concentrations.

Findings: To resume, the weather impact investigation revealed a higher sensitivity of the mobile sensing units to external factors such as temperature, wind, humidity and atmospheric pressure. This could be caused by direct exposures of the units to pollution concentrations close to ground levels, but also to the influence of traffic congestion. The daily profiles of citizen trips coupled with weather information across extended time periods can be used to establish accurate daily patterns of pollution along specific urban routes in the city.

4.2.3. Noise pollution from mobile sensing

Besides NO_2 levels, the Azimut station continuously registered noise levels at the human level while following the proposed daily circuit. Figure 10 presents the mean and daily noise evolution registered during the study period with the associated European noise scale. The measurements indicate that noise levels ranged between 53.48 dB(A) and 89.76 dB(A), with an average reaching often 72.77 dB(A) which indicates a highly noisy/hazardous environment. In comparison to the NO_2 levels which have a dispersed behaviour and are harder to be analysed in time, noise levels seem to have a homogeneous evolution and follow almost similar trends from one day to another.

By undertaking a daily noise comparison similarly to the previous NO_2 analysis, one can easily identify almost similar evolutionary patterns of noise levels during a normal week day (as seen in Figure 11a) and b); lower noise levels

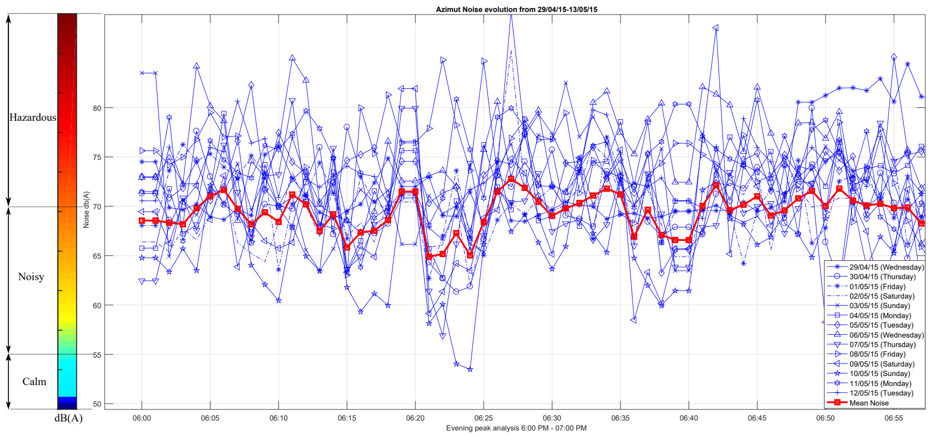


Figure 10: Daily and mean noise levels registered by the Azimut mobile station.

were registered during public holidays, when traffic is heavily reduced in the city centre (see Figure 11c)). Overall, the current analysis revealed unexpected high noise levels inside the eco-neighbourhood NGC, which is the contrary objective of Grand Nancy Metropolis who wants to increase the liveability for its citizens, not only by offering good public transport services and multi-modal interconnection, but also good levels of air quality, reduced traffic jams and implicitly, noise levels.

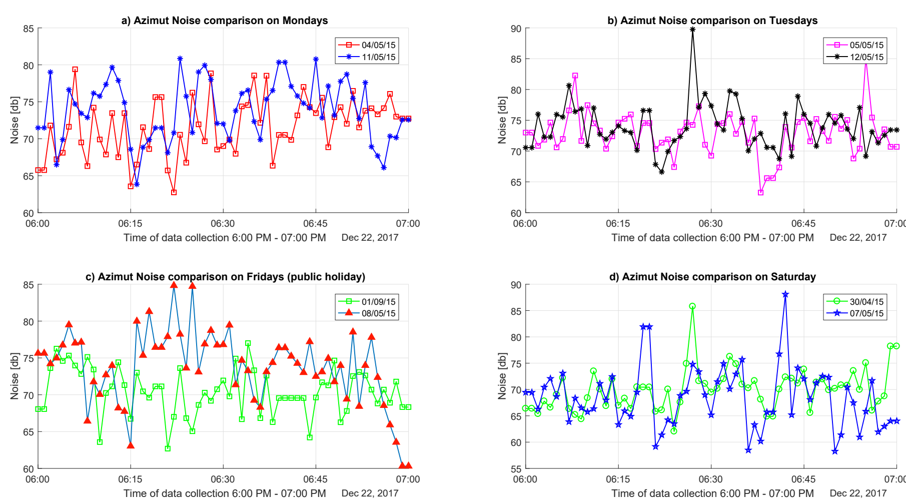


Figure 11: Daily comparison of noise levels registered by the Azimut station.

Findings: To resume, the second protocol revealed that:

- mobile sensing units are capable of accurately detecting hot-spots of NO_2 concentrations, especially under the influence of traffic congestion,
- mobiles sensing units are highly sensitive to the weather conditions but

reflect an accurate impact of pollution at breathing level,

- mobile generated data can be accurately used for building daily patterns of pollution along specific urban routes; this could be further expanded to any data-driven modelling which could learn from historical patterns and generate situation awareness alerts whenever a real-time pollution concentration falls out of historical expected trends.

- noise can also be used as a data feature for any pollution prediction modelling due to their high relevance towards increased traffic congestion in affected areas.

4.3. Third protocol results

Adopting a smart mobile station for measuring the air quality at the human level can bring numerous benefits and additional insights to citizens, but one needs to verify the accuracy of the mobile station for calibration and validation purposes. As previously mentioned, the SCK was adopted not only for testing against fixed pollution units or large AQMs, but also to make the comparison between the robustness of various mobile devices for air quality monitoring. Although the SCK provided a bigger autonomy due to its integrated solar panel, the collected data couldn't be transmitted in real-time towards the on-line platform unless there was a continuous Wi-Fi availability in the area. The SCK used in this experiment offered a 80-hour continuous monitoring together with the Azimut Station, while being placed outside a balcony near the location of tube 5.

Figure 12a) presents the comparison between NO_2 levels registered by both Azimut and SCK; overall one could observe the higher variability of the Azimut station compared to the SCK, but also a long-term steadiness of results especially towards the end of the experimentation protocol when the SCK station registered an unexplained rise in NO_2 levels. Figure 12b) showcases the noise comparison between the two mobile devices, with a more steadier but higher noise levels registered by SCK when compared to Azimut. Although Azimut registered a lower noise range, one could observe its sensitivity to short-term variations when compared to SCK, which registered an overall noise value of 50dB with seldom higher peaks that have reached a maximum of 65dB. Overall, these results translate as medium to good noise levels when compared to standard scales, but the lack of noise fluctuations in the SCK behaviour on the long term can indicate less sensitivity to smaller noise variations in the surrounding environment.

Findings: The main and surprising finding of the 3rd protocol revealed that different mobile sensing units can report different concentrations of the same pollutant at the same moment in time, and can be very sensitive to different external factors. As previously mentioned, the main limitation of this protocol was mainly related to the impossibility of a long-term usage of the SCK in mobile outdoor circuits as the one in Figure 4. The complete different behaviour of the two stations indicate a further need to conduct long-term data

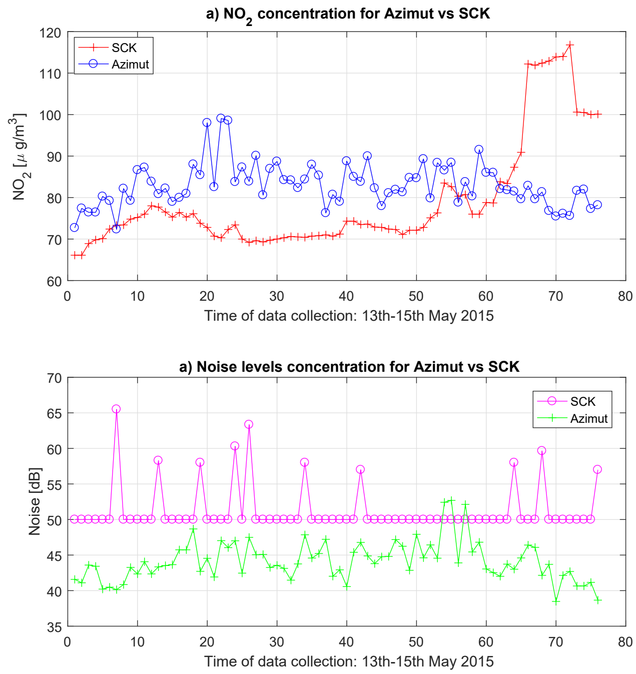


Figure 12: Noise and NO_2 comparison between Azimut and SCK.

collection, analysis and interpretation in order to fully understand the cause in different behaviour and data variability. One would need to conduct several experimenting protocols, in different settings and on longer time periods before adopting one mobile unit over the other for outdoor and mobile air pollution monitoring.

4.4. Data driven models for air quality prediction

We have further performed a closer data-driven investigation over the factors that could influence the NO_2 concentration levels registered by mobile pollution units and the accuracy of the generated data sets for further air quality predictions. As Azimut was the sensing unit which provided continuous data collection (every minute) over the entire experimenting protocol, we have used this data to build the following features (variables): latitude, longitude, temperature, humidity and noise. We store these features in a matrix $X_t = [X(i,j)]_{i=1,..N_p}^{j=1..5}$ and consider the corresponding NO_2 vector as $P_t = [N_i]_{i=1,..N_p}$, where N_p is the total number of data record transmitted by Azimut during the second experimenting protocol which summed around 20160 records. We then consider the regression problem of predicting P_t from X_t, so as to determine the highly predictive features which influence the NO_2 concentration levels.

The data has been separated into a training set comprised of 75% of all records, a test data set of 15% and a validation data set of also 15%. We then fit a regression model on the training set and evaluate the model performance using the mean squared error (MSE). As a baseline we use a trivial model which predicts the mean NO_2 concentrations in the training set; any model that performs worse then this is typically useless.

The first underlying model is a decision tree using the CART algorithm. This intuitive model can fit the non-linearity in the data well, as it involves making splits in the data on some simple thresholds. Figure 13 shows the output of a decision tree with 3 levels of depth which indicates that the most predictive features are noise, humidity and location. The MSE of this model is 303.32 which is almost a 50.6% improvement from the baseline MSE of 598.62. The leafs of the tree represent a hard prediction of the target variable P_t (in our case NO_2 concentrations), which is typically done by averaging the points that fall into a particular leaf node. The models seems to treat separately the cases when the registered noise is lower than $65dB$ or higher, but after a location investigation in both cases, humidity seems to be the most predictive feature of the model.

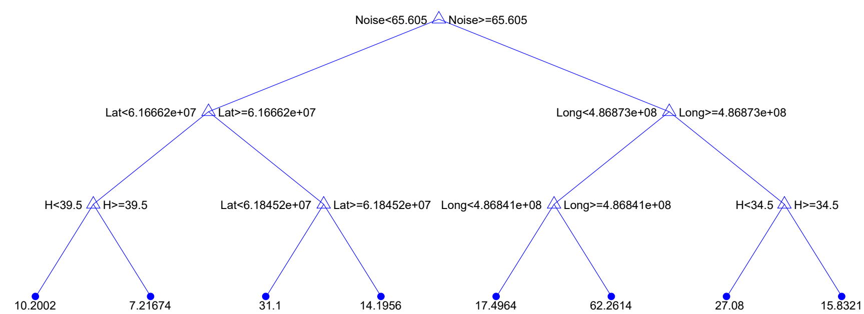


Figure 13: Decision Tree outcome for P_t.

Further on, we believe that data collected by mobile sensing unit can be used learn patterns of air pollution evolution, especially when being used in particular urban locations. When passing through a polluted area, if the pattern analysis detects anomalies and historical high pollution levels, the mobile unit could release alarms to the user to avoid the specific area. In order for this to happen, the data collected by the mobile unit needs to be accurate enough and has to contain sufficient information that could be used for predicting air pollution depending on location temperature, humidity, etc.

We have further used the above divided datasets from the Azimut mobile unit to train the second model: a neuronal network which is generally good at fitting specific practical functions. The neuronal network used for our predicting problem contains a hidden layer of 10 neurons and was trained first by using a Levenberg-Marquardt algorithm which typically requires more memory but less computational time; training automatically stops when the generalization stops improving, as indicated by an increase in the mean square error of the validation samples. The obtained MSE revolved around 150.59 which is a further 25% improvement from the decision tree algorithm and the overall R^2 value was 0.71. Despite fast running time (1 second) and 97 iterations, we have found that by using a Bayesian regularisation algorithm on the training set, the results improved significantly. Although the Bayesian regularisation typically requires more time for training on the dataset, it usually results in good generalisation for difficult, small or noisy datasets. The training stops according to adaptive weight minimization (regularization). The obtained MSE rounded up to 126.09 which is a further 4.09% in the prediction accuracy.

Figure 14 shows the R^2 values obtained for all datasets: training, testing and validation which reached an overall score of 0.81 when using the Bayesian Regularisation. While there is still place for improvement, the results are promising for further training on larger datasets when available. This is the best results obtained on the air pollution dataset generated by the Azimut which indicated that with a confidence of 81% one could predict in the future the NO_2 concentrations when using a mobile sensing unit in outdoors environment. We also make the observation that the current results apply to the specific mobile sensing unit and further tests and analysis could be imagined for a better performance evaluation of both air quality monitoring and prediction accuracy.

Findings: Machine learning modelling is an efficient tool for predicting accurately the mobile air quality. Decision-trees and neuronal network showed good capabilities for air quality prediction which could be further improved if more data would be available. A true challenge to extend this modelling would be to build collective data-driven predictions and anomaly detection algorithms for insuring a continuous real-time situation awareness.

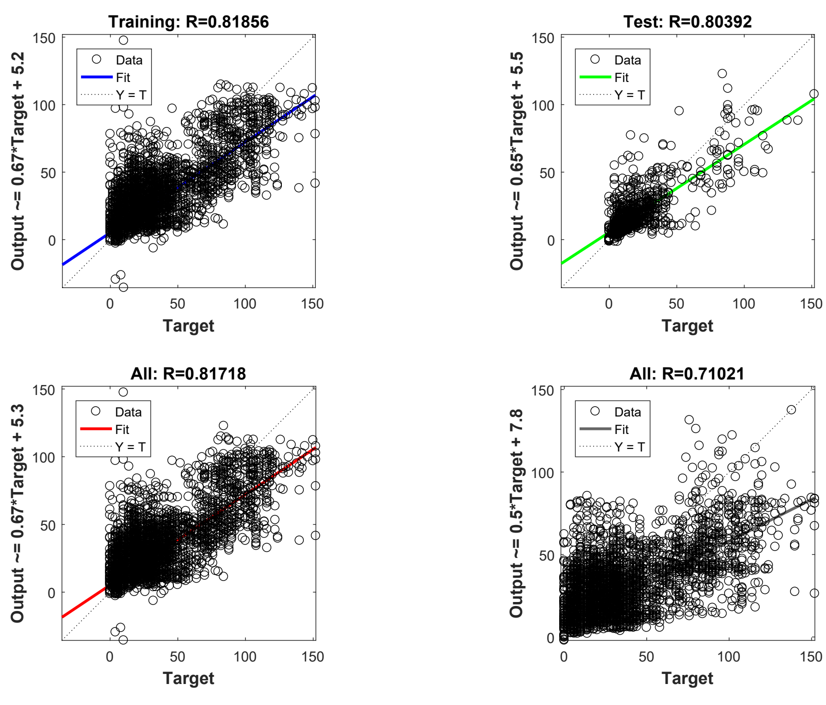


Figure 14: Results for Neuronal Network training using Bayesian Inference.

4.5. Impact evaluation of fixed versus air mobile sensing

While the original purpose for this study was to propose a mobile sensing investigation coupled with data-driven modelling for air quality prediction, the biggest finding is mostly related to the difference and high impact of pollution concentrations registered at the human breathing level when compared to those reported by official stationary monitoring units. Figure 15 shows the summary of mean NO_2 levels registered during the study period by both the Azimut station and the passive tubes (experiments 1 and 2). Although the overall average concentration levels are in good evaluation scales (less than $90(\mu g/m^3)$, the difference between the two experimentation protocols reveal significant differences between monitoring techniques and an alarming direct citizen impact. From Figure 15 one can identify that most of the pollution levels registered by the mobile station carried at the human level near the locations of the passive tubes are almost three times higher than the stationary levels monitored at higher levels: tubes 8 and 9 registered almost $23.3(\mu g/m^3)$ from the passive tubes (placed at 3 meters altitude) and validated by the AQM station (placed at around 10 meters altitude) in comparison to $61.2(\mu g/m^3)$ recorded by the Azimut station carried at human level (1.5 meters altitude) near these tubes. The biggest difference between fixed and mobile air pollution monitoring is showcased by the passive tube 4, which recorded an NO_2 concentration of $14.9(\mu g/m^3)$ in comparison to $74.17(\mu g/m^3)$ registered by the mobile station Azimut; this translated in a human-level pollution score which is almost 5 times higher than official reported scores by the stationary monitoring devices.

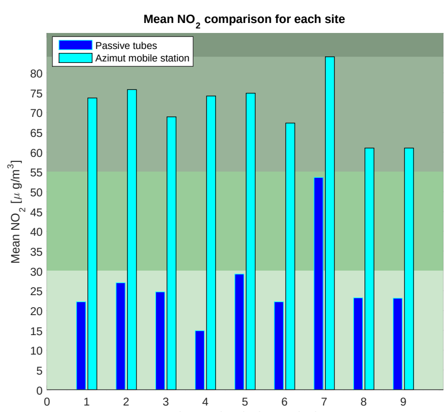


Figure 15: NO_2 pollution level registered by passive tubes places at 3 meters altitude versus the mobile station Azimuth carried at the human level.

While various reasons and factors could be further taken into consideration for explaining the significant difference between the investigated emission levels, this finding brings a solid awareness towards the real impact that air and noise pollution can have on human health and the risk that citizens are facing when walking in extremely crowded and congested areas in the city. It is also a proof that mobile sensing units for air pollution monitoring are better at quantifying the exact exposure to air pollutants and bring a solid and real-time situation awareness.

5. Conclusions

This paper proposed a mobile air pollution monitoring framework coupled together with a data-driven modelling approach for predicting the air quality inside urban areas, at human breathing level. Three experimenting protocols have been implemented by using: a) fixed passive tubes for NO_2 monitoring which have been verified against a reliable AQM station (this represented the baseline verification for the accuracy of any future mobile sensing in this area), b) smart and mobile sensors with real-time data transmission and collection deployed through the use of citizens travelling daily in the neighbourhood and c) a static monitoring protocol for comparing the performance of two different mobile sensing units. The proposed experimentation protocols not only showed a significant higher impact of NO_2 concentrations when using smart sensors carried at human level and walking inside highly circulated urban areas, but also poor noise levels registered especially during evening peak hours. Weather conditions are also important factors to be used when analysing the pollution concentration due to their strong correlation and high temporal influence. Evenmore, the data-driven investigation revealed that data generated by the mobile sensing units when used outdoors can be accurately used to predict future NO_2 levels in urban areas.

Limitations: Besides the advantages and disadvantages of using each of the mobile units detailed in Section 3.1, the main limitations for monitoring, investigating and evaluating air quality by using a crowd-sensed initiative consists in the evaluation of the data accuracy. While multiple sensing units are available for testing and usage, one needs to verify the accuracy of the mobile stations for calibration and validation purposes, under different traffic and weather conditions. Despite important advantages of using low-cost sensing units for measuring air pollution at a very granular city scale, various questions about the use and large-scale utilisation of such devices remain challenging and unanswered. Important aspects which are currently under investigation relate to the regulated production and marketing of such units, the use and ownership of generated data, cost of maintenance and installation, etc. A very important question concerns the electronic waste and the impact on public health (Grant et al., 2013), especially as many cities around the world are switching towards a sustainable and ecological paradigm (Bayulken and Huisingh, 2015).

Future applications: To the best of our knowledge there isn't currently any research approach trying to apply advanced machine learning methods on mobile sensing-generated data for improving citizen's health. There is a lack of solutions proposing both real-life air quality monitoring at human level and data-driven prediction approaches for situation awareness and real-time alert generation. Data-driven modelling has a true potential for real-time operations as it can automatically detect non-linear spatial relationships between sensing units and could easily aggregate results for regional investigations as well.

A future real-life application of our study is to: a) extend the usage of mobile sensing units to more citizens driving/walking/cycling every day in the NGC eco-neighbourhood, b) improve the data driven modelling for air quality prediction and incident hot-spot identification (this would require extensive machine learning research investigations for addressing data sparsity and missing features), and c) build a situation awareness module which would learn from previous air pollution episodes and release pollution alerts to citizens in real-time.

A future data analysis and experimental investigation awaits for Metropolis research approval. The duration of the monitoring could be also extended to longer periods which can be a true challenge due to higher costs involving both human resources, material acquisition, data processing and interpretation. Seasonality could also be included in the analysis when more data would become available, as in our previous studies (Mihăiţă et al., 2016).

One of the benefits of predicting air pollution hot-spots is to determine a change in the citizen driving behaviour which would change not only their routes to avoid polluted areas, but also their residential areas/recreational areas, etc. On the long-term this would lead to a reconfiguration of cities based on human health prioritisation.

The NGC project is further developing more studies on how to better integrate accurate air quality information with traffic congestion monitoring (Mihăiţă et al., 2017), but also how to involve citizens in an active crowd-source activity for raising awareness around pollution and traffic behaviour. Offering the correct monitoring tools will trigger more adapted urban actions which will improve on the long-term the life of inhabitants in such complex environments.

Acknowledgements

The authors of this work are grateful for the support provided by various actors in this project: 1) Grand Nancy (GN) through the project Chaire REVES, 2) Brigitte Trousse, the president of France Living Labs for providing the Azimut station, 3) Air Lorraine (Alexandre Ockler and Gerard Pol) for providing the passive tubes, 4) Barcelona Fab Living Lab for providing the Smart Citizen Kit in time for the experiments, and 5) Aline Carier, Thibault Frère, Oussama Okasha and Ophélie Voirin for conducting the daily experimenting protocols. The experiments of this work have been done in the ERPI laboratory from Nancy France. The final writing and correction of this article has been done in Data61/CSIRO, Australia.

Appendix A.

The table A.4 in this Appendix section lists all the studies which have been working towards the analysis and prediction of air quality by either using fixed or mobile sensors, or by developing new data-driven methods. They are compared based on the technology used (what sensors, what data), their mobility, their proximity near the breathing human level, and their innovative data-driven approaches. The results have been generated by using the following key words on major research platforms such as Science Direct, Web of Science, Elsevier and Google Scholar: "mobile sensors air quality pollution" and "air quality prediction". Please note that the list s not exhaustive and other research papers meeting our research criteria might be available.

Table A.4: State-of-art comparison.

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Zhou et al., 2019)	This study presents an artificial intelligence approach based on a Deep Multi-output LSTM (DM-LSTM) neural network model for predicting the air quality in China.	The proposed DM-LSTM model was evaluated by three time series of PM2.5, PM10, and NOx simultaneously at five air quality monitoring stations in Taipei City of Taiwan.	No	No	This study employed hourly data of eight air quality factors (PM2.5, PM10, O3, NOx, NO2, NO, SO2, CO) and five meteorological factors (rainfall, temperature, wind speed, wind direction, and relative humidity).	7 years of data (2010 to 2016).	Results demonstrated that the proposed DM-LSTM model incorporated with three deep learning algorithms could significantly improve the spatio-temporal stability and accuracy of regional multi-step-ahead air quality forecasts.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Wu et al., 2018)	This study deployed grey prediction theory (which can deal with the forecasting problem with limited samples), for predicting the air quality in the Beijing-Tianjin-Hebei region between 2017 and 2020.	The indicators data are from the Air Quality Assessment Report IV which depend on a new standard (GB 3095-2012) published in 2012. The PM2.5 values were mandatorily included in this new standard for the first time. Only a few sets of annual data were available.	No	No	The proposed model predicted the average annual concentrations of PM2.5, PM10, SO2, NO2, 8-h O3, and 24-h O3.	2013-2016.	Results demonstrated that the FGM(1,1) model is significantly higher than that of the traditional GM(1,1) model. The prediction results from 2017 to 2020 indicate that the concentrations of PM2.5, PM10, SO2, and NO2 will decrease, whereas the 8-h O3 and 24-h O3 concentrations will increase.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Zhou et al., 2018)	This work constructed a set of data envelopment analysis (DEA) models for calculating air quality in 31 main cities in China by using the daily Air Quality Index.	All the data is collected from China National Urban air quality real-time publishing platform. The composition of different levels of AQI appeared in a two-years data base for 2014-2015.	No	No	The proposed model calculated the average monthly AQIs for each city.	2014-2015.	The work included integer and zero-sum gain constraints into a DEA approach for evaluating the air quality of 31 main cities in China by using daily AQI data. There is no further inclusion of other pollution data sources from these cities.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Borrego et al., 2018)	This study presents part of the results of the intercomparison campaign in Aveiro and is intended to estimate the uncertainty of the measurements according to the DQO of the European Air Quality Directive and to improve their performance with the aid of a computationally-oriented methodology.	Study used micro-sensor nodes which are low-cost devices with considerable application potential, offset by important limitations when applied to urban air quality monitoring.	No	No	Both units utilised Alphasense electrochemical cell (ECC, model B4/BH) for measuring NO, NO_2, O_3, CO, SO_2 and Total VOC. Measurements of CO_2 (SenseAir K30) and particulate matter (University of Hertfordshire CAIR) were also undertaken. The CAM10 and CAM11 boxes measured wind speed, wind direction, temperature and humidity, allowing comparison of meteorological variables and source apportionment.	Two weeks in October 2014.	Linear regression in addition to the Feed-Forward Neural Networks and Random Forests algorithms were employed, in an effort to improve sensor performance based solely on sensor data and local meteorological data from a reference station.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Castell et al., 2018)	This study installed 17 nodes monitoring NO_2 in the playground kindergartens in Oslo and one node in the top of a reference air quality monitoring station. The results showed that indeed low-cost sensors can provide a good indication of the air pollution levels, being capable to reproduce the trends during a high pollution episode.	The sensor nodes were located outdoors, on the playground of the kindergarten, at heights of around 3 meters. Additionally, 6 sensor nodes were installed in the streets in Oslo and 9 reference stations were used from Oslo. A commercial low-cost platform AQMesh $v3.5$ was used for monitoring results (provided by Environmental Instruments Ltd, UK).	No	N/A	Their battery driven stationary platforms measured four gaseous components (CO, NO, NO_2 and O_3), particulate matter, temperature, relative humidity and atmospheric pressure. The NO_2 sensor employed in the AQMesh v3.5 is an electrochemical sensor provided by Alphasense (NO2-B42F) that incorporates a filter to eliminate cross-sensitivity issues with O_3.	From 1st-31st of January 2016.	The air quality modelling was done using the EPISODE dispersion model which is a 3-D Eulerian/Lagrangian dispersion model that provides urban and regional-scale atmospheric pollutant concentrations such as NO_2. The study mentions an air quality visualisation portal and mapping using data fusion techniques.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Ercilla-Montserrat et al., 2018)	This study focuses on the potential contamination of heavy metals in hydroponic lettuce crops due to atmospheric pollution in high-traffic areas.	To collect air samples, high-volume sensors (MCV CAV-A/mb) were used, working at a volume of $30m3/h$ in 48-h periods, using glass microfibre filters.	No	No	The following metals were analysed: Ni, As, Cd and Pb on 4 sites: a peri-urban integrated rooftop greenhouse, a peri-urban rooftop, an urban courtyard and an urban rooftop.	1st test: 40 days starting from 4th of April 2017; 2nd test: 34 days starting from 22nd of June 2017.	Air and crop sample analyses were carried out by the Agency of Public Health of Barcelona with ENAC accreditation (UNE-EN ISO/ IEC 17025:2005).

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Heimann et al., 2015)	This study shows a purely measurement-based approach to extract underlying pollution levels (baselines) from air measurements by exploiting different relative frequencies of local and background pollution variations.	The study measured the carbon monoxide (CO) by deploying a network of 45 low-cost electrochemical sensors, in and around the city of Cambridge, UK (out of which only 32 were successfully reporting integral data sets).	No	No	CO concentration measurements only were obtained from the Greenhouse Gas Laboratory at Royal Holloway, University of London (RHUL), 75 miles south-west of Cambridge in Egham, Surrey. Every sensor node was referenced to this one station. To remove local influences on the measurements, a meteorological filter was applied.	2 months starting from 28th of March 2010.	A novel and flexible method has been developed to determine sensor baselines, i.e. underlying variation, of measurements (that represent non-local emissions) which is suitable for application to large data sets.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Popoola et al., 2018)	This work demonstrates a technique for source apportionment in complex environments, in this case a major international airport, using a low-cost air quality sensor network.	17 low cost portable air quality devices developed at the Department of Chemistry in the University of Cambridge, UK, were installed at the London Heathrow airport (Alphasense B4, Sensair K33, Alphasense PID-AH, Pt1000, Honeywell HIH4000, Gill WindSonic University of Hertfordshire).	No	No (fixed on lampposts in the airport).	CO, NO, CO_2, $VOCs$ (total), Temperature, Relative Humidity, Wind Speed/Direction, Size Specialised Particulates.	5 weeks starting from 4th of October 2012.	They proposed an extension of the street scale resolution ADMS-Urban dispersion model for complex urban environments with additional capability for the explicit modelling of aircraft jet engine emissions as jet sources.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(SM et al., 2019)	This paper deals with the development of a smart personal air quality monitoring system (SPAMS) for real time air quality monitoring, in order to measure the individual exposure to air pollution. Monitoring campaign was designed for both pedestrian exposure as well as exposure while travelling in bus. Sensors were set-up near to the breathing level throughout the monitoring campaign to represent the actual personal exposure to the air pollutants.	Low-cost and lightweight sensors (SPAMS) were used for measuring CO (TGS 2442), NO_2 (MiCS 2714), O_3 (MiCS 2614), PM, temperature and humidity.	Yes	Yes	CO, NO_2, O_3, PM, Temperature, Humidity. Field measurements were performed by walking on both footpaths and travelling in the bus during various times of a day (morning, afternoon and evening) and different days in a week at selected locations in Chennai city, India.	Pedestrians walking (random 2015); bus travel (November 2015 to January 2016).	Analysis revealed that the correlation between temperature, RH and output voltage of SPAM was linear. The best fit line for each parameter was used to get sensor response function. These response functions were later used in the sensor system to measure the real-time gaseous pollutant concentrations.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Ripoll et al., 2019)	In this work, the performance under field conditions of two custom-made types of ozone sensing devices, based on metal-oxide and electrochemical sensors, was tested.	A large array of 132 metal-oxide (Sensortech MICS 2614) and 11 electrochemical (Alphasense) ozone sensors, built into 44 sensing devices, was co-located at reference stations in Italy (4 stations) and Spain (5).	No	N/A	O_3	Between May and October 2017.	The individual sensor datasets were calibrated applying multi-linear regression (MLR) analysis.
(Pereira et al., 2018)	This is a student project which turned smartphones into dynamic sensor nodes which send data to a centralised platform. Citizens play a significant and empowering role regarding healthcare conditions.	The proposed solution has an out-of-the-box central platform available in the cloud which gathers available IoT sensor data (ThingSpeak) and (2) an online custom developed central platform.	Yes	N/A	NO_2, CO and O_3.	Not an exact set-up experimentation.	N/A.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Hasenfratz et al., 2015)	This work analyses one of the largest spatially resolved UFP data set publicly available which contains over 50 million air measurements but also proposed a mobile measurement system. Node mobility trades off temporal resolution against spatial resolution, enabling a high spatial resolution across large areas without the need for a huge number of fixed sensors.	The mobile measurement system consists of 10 sensor nodes installed on top of public transport vehicles, which are equipped with a semiconductor-based O_3 sensor, electrochemical-based CO and NO_2 sensors, and a compact device to measure UFP concentrations. Additionally, the nodes monitor radio-frequency electromagnetic fields, temperature and humidity.	Yes	No (3 meters high).	NO_2, CO and O_3.	2 years starting from April 2012.	Based on these data, authors develop land-use regression (LUR) models to create pollution maps with a high spatial resolution of $100m$x$100m$. Authors compared the accuracy of the derived models across various time scales and observed a rapid drop in accuracy for maps with sub-weekly temporal resolution.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Minet et al., 2017)	This study highlights the challenges of short-term mobile sampling campaigns in terms of the resulting exposure surfaces. A mobile monitoring campaign was conducted in 2015 in Montreal; NO_2 levels at 1395 road segments were measured under repeated visits.	The study used NO_2 and O_3 Aeroqual Series 500 portable sensors (Aeroqual, 2014a), two types of GPS Garmin Edge 800 as well as MapMyRide, a smartphone application.	Yes	Yes (walk and bike).	NO_2 and O_3 concentrations were recorded by cyclists and pedestrians. Hourly temperature (T), relative humidity (RH) and wind speed from the Montreal Pierre Elliott Trudeau International Airport weather station were recorded and synchronized with the other measurements.	Between May and August 2015.	The authors applied a statistical analysis and land-use regression based on sub-segments, categorized in terms of the number of visits per road segment. Authors observed that LUR models were highly sensitive to the number of road segments and to the number of visits per road segment.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Suriano et al., 2015)	This work is a technical report on preliminary real-world measurements by low-cost gas sensor-systems for air quality monitoring.	Authors used a sensor-system AIRBOX in Bari Italy, which is equipped with low-cost electrochemical gas sensors, optical particle detectors, NDIR infra-red sensors, photo-ionisation detectors, and miniaturized sensors for temperature and relative humidity. Another ENEA prototype of portable sensor system which connects to the smart-phone by Bluetooth has been used inside cars for air pollution monitoring.	Yes-Airbox; No-Enea	Yes	NO_2, CO, O_3, CO_2, SO_2, VOCs, $PM1.0$, $PM2.5$, $PM10$, Temperature, humidity.	between 2015-07-21 and 2015-08-07.	No data-driven analysis was provided, only a graphical representation of data.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Devarakonda et al., 2013)	This work presents a vehicular-based mobile approach for measuring fine-grained air quality in real-time. They proposed two cost effective data farming models: one that can be deployed on public transportation and the second a personal sensing device.	Authors used a) a Mobile Sensing Box which can be mounted on vehicles and contains a micro-controller, dust and carbon monoxide sensors, GPS and a cellular modem, and b) personal sensing devices: a mobile air quality sensor and a smart phone to act as an interface with the central repository hosted on a cloud server.	Yes	Yes	They selected OXA and CLIMA modules to measure carbon monoxide, humidity, temperature, ambient light and barometric pressure.	A car tour for each experiment.	Linear Regression in Pollution Measurements and a heat-map of Carbon Monoxide Concentrations were generated.

Continued on next page

Table A.4 – *State-of-art comparison of mobile sensing and data-driven air pollution modelling - continued from previous page.*

Ref.	Study objective	Sensors technology	Mobile	Near breathing level	Parameters measured	Period	Data analysis
(Al-Ali et al., 2010)	This work proposed an on-line GPRS-Sensors Array for air pollution monitoring which consists of a Mobile Data-Acquisition Unit (Mobile-DAQ) and a fixed Internet-Enabled Pollution Monitoring Server.	The Mobile-DAQ unit integrates a single-chip microcontroller, air pollution sensors array, a General Packet Radio Service Modem (GPRS-Modem), and a Global Positioning System Module (GPS-Module). The Mobile-DAQ was mounted on a University bus that was driven around the campus of the American University of Sharjah (AUS) to collect pollutant data.	Yes	No	CO, NO_2, and SO_2.	The pollutant data was collected for 12 h.	N/A

Appendix B.

In Table B.5 we summarize the main air pollution sensors which have been initially considered for this study, their characteristics, advantages and disadvantages which correspond to our current purpose and needs for this study. While various comparative studies could be further undertaken, our main objective was to choose an appropriate mobile sensing unit which could be easily adopted at larger scale but which could also provide accurate results when compared to highly efficient AQMs.

Table B.5: Air quality sensing unit comparison table.

Sensor	Characteristics	Advantages	Disadvantages
Azimut 	-Measures O_3, NO_2, noise and temperature by using an electrochemical detection method (description provided in Section 3.1).	- The mobile sensing unit can be fixed on cars, bikes, held in hand, etc. - Provides real-time information. - Offers access to a centralised platform for data monitoring and pollution alert.	- Azimut has a two days autonomy. - Higher acquisition price than passive tubes.
Passive tubes 	- Measures NO_2 by passive sampling (passive transfer of pollutants by molecular diffusion of ambient air to an adsorbent specific to the targeted pollutant). - Needs physical and fixed installation at 2-3 meters from the ground (see description in Section 3.1). - The result analysis is done in a dedicated air quality laboratory.	- Easy preparation, installation and result investigation. - Low cost unit (under 10 Euro per analysis). - Possibility to use the units at larger scale. - Does not require any electrical charge.	- Does not provide any real-time information. - At least 1 or 2 weeks of waiting period for obtaining complete results. - Fine-granular pollution peaks which are hourly-based can not be detected.

Continued on next page

Table B.5 – *Air quality sensing unit comparison table - continued from previous page.*

Sensor	Characteristics	Advantages	Disadvantages
Smart Citizen Kit 1 	- Measures NO_2, O_3, CO_2, humidity, luminosity and noise. - Mobile unit with both electrical recharge and integrated solar panel which can be used for both indoor and outdoor air pollution monitoring.	- Low energy consumption and long autonomy. - Affordable cost unit price (averages around 200 Euro). - The unit comes with an online platform access for data-collection and visualisation (smartcitizen.me).	- Delivery time can be longer as each unit is made on-order. - Unit needs permanent Wi-Fi connection for data transmission. - Unit needs a special case for outdoor protection.
Pollux'nz city 	- The mobile unit contains PM_{10} particle pollution sensors, temperature, sound and oxygen sensors. - There are two main components to the system: a base station which collects the data from the array of sensors and a receiving module which measures and transmits the data (CKAB, 2017). - Each sensor runs off of a battery, but features a PV solar panel which keeps the power source topped off.	- Autonomous with a feature of direct measures and lecture from the online platform. - It uses an Arduino to drive the system, and an XBee radio for communications. The base station also uses an XBee radio to poll the network, but it is not driven by an Arduino.	- It does not measure the NO_2 concentrations directly. - It needs high maintenance and configuration before the results can be analysed.

Continued on next page

Table B.5 – *Air quality sensing unit comparison table - continued from previous page.*

Sensor	Characteristics	Advantages	Disadvantages
City Pulse 	- This is a product of the Green Watch Project (CityPulse, 2015) which in 2009 used a prototype equipped with two environmental sensors (for measuring ozone and noise), a GPS chip and a Bluetooth chip. The device has the shape of a watch that his wearer takes with him to the city, capturing and storing measurements that are then published on the network. - The product measurements are sent to both the mobile phone and an open platform which can be used for monitoring, analysis and air quality modelling.	- Autonomous, easy to carry and well connected. - Comes with a platform for data collection and visualisation.	- No NO_2 monitoring. - The unit was hard to purchase and test. The project has been discontinued while waiting for the second round of financing.
Dylos Station 	- Indoor air quality monitoring for fine particles (Coporation, 2017). - Provides small and large particle counts with a direct reading on the unit.	- Has an increased lower sensitivity for detecting particles down to 0.5 micron. - The units have a higher range of unit prices (199-425 Euro).	- No NO_2 or noise monitoring. - The unit cannot be used for outdoor pollutant monitoring. - It requires special software installation for historical data storage or visualisation.

Continued on next page

Table B.5 – *Air quality sensing unit comparison table - continued from previous page.*

Sensor	Characteristics	Advantages	Disadvantages
Geocube 	- The Geocube station measures NO_2, O_3, SO_2, PM_{10}, temperature, wind speed, wind direction and noise (IGN, 2006). - Contains three modules including GPS, data and radio management.	- Outdoor station for monitoring landslides which can communicate together in a global network for transmitting data to a central node connected to the internet. - The units are autonomous for months and have also solar panels integrated. - They have an embedded alert system for raising awareness.	- The unit has a competitive price which makes it less affordable for daily citizen purchase and usage. - It needs renewal after 500 cycles of utilisation.
Footbot 	- Indoor station unit for monitoring particulate matter, temperature, humidity and chemical pollutants (Formaldehyde, Iso-Butane, Toluene, Methane, Ammonia, Benzene) - It uses a ”video stream” approach with affordable sensors for a continuously monitoring. (US, 2018) - The unit has embedded software to send in real-time to the Could the recalibration results via Wi-Fi and rapidly process the data measured.	- The data can be retrieved remotely from the Cloud service, and any updates or investigations can be easily followed. - The platform sends notifications to the users which can also use an API for data display. - Price aligns with other mobile units for air quality monitoring (179 Euro).	- The unit is mainly dedicated for indoor not outdoor air quality monitoring due to its Wi-Fi connectivity needs. - No NO_2 or noise monitoring.

Continued on next page

Table B.5 – *Air quality sensing unit comparison table - continued from previous page.*

Sensor	Characteristics	Advantages	Disadvantages
Air.air! 	- Portable indoor station unit for monitoring particulate matter $PM_{2.5}$ and temperature which are then sent to a smartphone app by using a Low Energy Bluetooth (BLE) (AirAir, 2014). - It has a compact optical dust sensor with a IRED infra-red-emitting diode and a photo transistor, which are diagonally arranged in the device to measure the reflected light off of the airborne dust particles.	- The unit can be used both indoors and outdoors and comes at an attractive price starting from 59$.	- No NO_2 or noise are monitored. - No indication regarding the autonomy is provided.
NetAtmo 	- NetAtmo provides a package of two stations for both indoor and outdoor air quality monitoring which register mainly temperature, humidity, atmospheric pressure, CO_2 and noise (NetAtmo, 2018). - Results are being sent to a cloud storage for later analysis and smart-phone consultation.	- Long lasting autonomy for up to 2 years.	- NO_2 is not monitored. - Needs permanent Wi-Fi connection at no more than 100 meters.

References

AirAir, 2014. Air.air! http://kck.st/2DmnVC5. Accessed: 2018-01-21.

Al-Ali, A.R., Zualkernan, I., Aloul, F., 2010. A mobile gprs-sensors array for air pollution monitoring. IEEE Sensors Journal 10, 1666–1671.

Alphasense, 2016. Alphasense air, sensors for air quality networks. http://bit.ly/2tVY2ca. Accessed: 2018-01-05.

AQE, 2016. Air quality egg. http://airqualityegg.com/. Accessed: 2018-01-05.

ATMO, 2016. La carte des aasqa. http://bit.ly/2cjX6QX. Accessed: 2018-01-05.

ATMO-Alsace, 2012. Modélisation de la qualité de l'air sur le futur éco-quartier Danube. Technical Report. Association pour la Surveillance et l'étude de la Pollution Atmosphérique en Alsace.

Azimut, 2017. Fireflies azimut monitoring. www.azimut-monitoring.com. Accessed: 2017-06-05.

Bayulken, B., Huisingh, D., 2015. Are lessons from eco-towns helping planners make more effective progress in transforming cities into sustainable urban systems: a literature review (part 2 of 2). Journal of Cleaner Production 109, 152 – 165. Special Issue: Toward a Regenerative Sustainability Paradigm for the Built Environment: from vision to reality.

Borrego, C., Ginja, J., Coutinho, M., Ribeiro, C., Karatzas, K., Sioumis, T., Katsifarakis, N., Konstantinidis, K., Vito, S.D., Esposito, E., Salvato, M., Smith, P., André, N., Gérard, P., Francis, L., Castell, N., Schneider, P., Viana, M., Minguillón, M., Reimringer, W., Otjes, R., von Sicard, O., Pohle, R., Elen, B., Suriano, D., Pfister, V., Prato, M., Dipinto, S., Penza, M., 2018. Assessment of air quality microsensors versus reference methods: The eunetair joint exercise – part ii. Atmospheric Environment 193, 127 – 142.

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., 2008. The rise of people-centric sensing. IEEE Internet Computing 12, 12–21.

Castell, N., Schneider, P., Grossberndt, S., Fredriksen, M., Sousa-Santos, G., Vogt, M., Bartonova, A., 2018. Localized real-time information on outdoor air quality at kindergartens in oslo, norway using low-cost sensor nodes. Environmental Research 165, 410 – 419.

Chong, C.Y., Kumar, S., 2003. Sensor networks: Evolution, opportunities, and challenges. Proceedings of the IEEE 91, 1247–1256. Cited By 1735.

CityPulse, 2015. Green watch city pulse. http://bit.ly/2GFZZMt. Accessed: 2018-01-05.

CKAB, 2017. Pollux'nz city. http://bit.ly/2DsQn6i. Accessed: 2018-01-14.

Coporation, D., 2017. Dylos air quality monitor. http://bit.ly/2FUmPPq. Accessed: 2018-01-14.

Devarakonda, S., Sevusu, P., Liu, H., Liu, R., Iftode, L., Nath, B., 2013. Real-time air quality monitoring through mobile sensing in metropolitan areas, in: Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing, ACM, New York, NY, USA. pp. 15:1–15:8.

Dupont, L., Morel, L., Guidat, C., 2015. Innovative public-private partnership to support smart city: the case of {Chaire REVES}. Journal of Strategy and Managment 8, 245–265.

van den Elshout, S., Bartelds, H., Heich, H., Leger, K., 2012. CAQI Air quality index, Comparing Urban Air Qualiy across borders. Technical Report. Citeair.

Ercilla-Montserrat, M., Muñoz, P., Montero, J.I., Gabarrell, X., Rieradevall, J., 2018. A study on air quality and heavy metals content of urban food produced in a mediterranean city (barcelona). Journal of Cleaner Production 195, 385 – 395.

Gabrys, J., 2016. Citizen sense. http://www.citizensense.net/. Accessed: 2018-01-05.

GN, 2012. Nancy grand cœur. http://bit.ly/1RzsrNu. Accessed: 2018-01-05.

Grant, K., Goldizen, F.C., Sly, P.D., Brune, M.N., Neira, M., van den Berg, M., Norman, R.E., 2013. Health consequences of exposure to e-waste: a systematic review. The Lancet Global Health 1, e350–e361.

Hasenfratz, D., Saukh, O., Walser, C., Hueglin, C., Fierz, M., Arn, T., Beutel, J., Thiele, L., 2015. Deriving high-resolution urban air pollution maps using mobile sensor nodes. Pervasive and Mobile Computing 16, 268 – 285. Selected Papers from the Twelfth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom 2014).

Heimann, I., Bright, V., McLeod, M., Mead, M., Popoola, O., Stewart, G., Jones, R., 2015. Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors. Atmospheric Environment 113, 10 – 19.

IGN, 2006. Geocube. http://bit.ly/2FKpbEw. Accessed: 2018-01-05.

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

Krupnick, A.J., 2008. Challenges to managing air pollution. Journal of Toxicology and Environmental Health, Part A 71, 13–23.

Kumar, P., Jain, S., Gurjar, B., Sharma, P., Khare, M., Morawska, L., Britter, R., 2013. New directions: Can a “blue sky” return to indian megacities? Atmospheric Environment 71, 198 – 201.

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

LigAir, 2009. Caractérisation des niveaux ambiants de formaldéhyde autour du site industriel. Technical Report. LigAir, surveillance de la qualité de l’air en région Centre.

Lorraine, A., 2018. Atmo grand est. `http://www.air-lorraine.org/`. Accessed: 2018-01-05.

MCT, M.d.l.c.d.t., 2016. Les éco-quartiers. `http://bit.ly/2xF3kcl`. Accessed: 2018-01-05.

Mead, M., Popoola, O., Stewart, G., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J., McLeod, M., Hodgson, T., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J., Jones, R., 2013. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. Atmospheric Environment 70, 186 – 203.

MEDD, M.d.l.e.d.d.d., 2004. Arrêté du 22 juillet 2004 relatif aux indices de la qualité de l’air. `http://bit.ly/2Av37Gp`. Accessed: 2017-12-20.

MEDE, 2012. Bilan de la qualité de l’air en France en 2012. Technical Report. Paris: Ministére de l’écologie et du développement durable, Direction Générale de l’Energie et du Climat.

Mihăiţă, A., Camargo, M., Lhoste, P., 2014. Evaluating the impact of the traffic reconfiguration of a complex urban intersection. 10th International Conference on Modelling, Optimization and Simulation (MOSIM 2014), Nancy, France .

Mihăiţă, A.S., Benavides Ortiz, M., Camargo, M., 2017. Predicting air quality by integrating a mesoscopic traffic simulation model and simplified air pollutant estimation methods. International Journal of Intelligent Transportation Systems Research (under revision) .

Mihăiţă, A.S., Dupont, L., Cherry, O., Camargo, M., Cai, C., 2018. Air quality monitoring using stationary versus mobile sensing units: a case study from lorraine, france. IET Intelligent Transport Systems journal, Special edition ITS World Congress 2018 (under revision) .

Mihăiţă, A.S., Ortiz, M.B., Camargo, M., 2016. Integrating a mesoscopic traffic simulation model and a simplified no2 estimation model for predicting the impact of air pollution. 23rd World Congress on Intellient Transportation Systems (ITSWC 2016), Melbourne, Australia, 10-14 October 2016 .

Minet, L., Gehr, R., Hatzopoulou, M., 2017. Capturing the sensitivity of land-use regression models to short-term mobile monitoring campaigns using air pollution micro-sensors. Environmental Pollution 230, 280 – 290.

MIT, 2016. Clarity, department of civil and environmental engineering. http://bit.ly/2poIAjo. Accessed: 2018-01-05.

Moltchanov, S., Levy, I., Etzion, Y., Lerner, U., Broday, D.M., Fishbain, B., 2015. On the feasibility of measuring urban air pollution by wireless distributed sensor networks. Science of The Total Environment 502, 537 – 547.

NetAtmo, 2018. Netatmo individual station. http://bit.ly/2z2U3td. Accessed: 2018-01-21.

NIEHS, 2014. Exposure biology and the exposome, u.s. department of health and human services, national institute of environmental health sciences. http://bit.ly/2HJ6mhr. Accessed: 2018-01-04.

Pereira, A., Patrício, B., Fonte, F., Marques, S., Reis, C.I., Maximiano, M., 2018. Collecting information about air quality using smartphones. Procedia Computer Science 138, 33 – 40. CENTERIS 2018 - International Conference on ENTERprise Information Systems / ProjMAN 2018 - International Conference on Project MANagement / HCist 2018 - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 2018.

Plaia, A., Ruggieri, M., 2011. Air quality indices: a review. Reviews in Environmental Science and Bio/Technology 10, 165–179.

Popoola, O.A., Carruthers, D., Lad, C., Bright, V.B., Mead, M.I., Stettler, M.E., Saffell, J.R., Jones, R.L., 2018. Use of networks of low cost air quality sensors to quantify air quality in urban settings. Atmospheric Environment 194, 58 – 70.

Ripoll, A., Viana, M., Padrosa, M., Querol, X., Minutolo, A., Hou, K., Barcelo-Ordinas, J., Garcia-Vidal, J., 2019. Testing the performance of sensors for ozone pollution monitoring in a citizen science approach. Science of The Total Environment 651, 1166 – 1179.

SCK, 2016. Smart citizen kit. https://smartcitizen.me/. Accessed: 2018-01-05.

SM, S.N., Yasa, P.R., MV, N., Khadirnaikar, S., Rani, P., 2019. Mobile monitoring of air pollution using low cost sensors to visualize spatio-temporal variation of pollutants at urban hotspots. Sustainable Cities and Society 44, 520 – 535.

Stieb, D.M., Burnett, R.T., Smith-Doiron, M., Brion, O., Shin, H.H., Economou, V., 2008. A new multipollutant, no-threshold air quality health index based on short-term associations observed in daily time-series analyses. Journal of the Air & Waste Management Association 58, 435–450.

Suriano, D., Prato, M., Pfister, V., Cassano, G., Camporeale, G., Dipinto, S., Penza, M., 2015. Short term forecasting of the level of service on a motorway network, by using a microscopic simulation model. Proceedings of the Fourth Scientific Meeting EuNetAir 2015-06-03 - 2015-06-05 Linkoping University, Linkoping, Sweden .

Trousse, B., Palloth, M., Tiffonh, C., 2014. Co-creating environmental services based on pollution citizen sensing. Interdisciplinary Studies Journal 3, 331–350.

US, A., 2018. Footbot. `https://foobot.io/`. Accessed: 2018-01-15.

US-EPA, 2016. What are the six common air pollutants? `http://bit.ly/1s74zuH`. Accessed: 2018-01-05.

Williams, D.E., Henshaw, G.S., Bart, M., Laing, G., Wagner, J., Naisbitt, S., Salmond, J.A., 2013. Validation of low-cost ozone measurement instruments suitable for use in an air-quality monitoring network. Measurement Science and Technology 24, 065803.

Wu, L., Li, N., Yang, Y., 2018. Prediction of air quality indicators for the beijing-tianjin-hebei region. Journal of Cleaner Production 196, 682 – 687.

Zappi, P., Bales, E., Park, J., Griswold, W., Rosing, T., 2012. The citisense air quality monitoring mobile sensor node. Proc. 11th ACM/IEEE Conf. Inf. Process. Sens. Networks, Beijing, China .

Zhou, Y., Chang, F.J., Chang, L.C., Kao, I.F., Wang, Y.S., 2019. Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. Journal of Cleaner Production 209, 134 – 145.

Zhou, Z., Guo, X., Wu, H., Yu, J., 2018. Evaluating air quality in china based on daily data: Application of integer data envelopment analysis. Journal of Cleaner Production 198, 304 – 311.