

# CITIZEN SENSING FOR ENVIRONMENTAL INFORMATION



QIJUN JIANG

## **Propositions**

1. Citizen environmental sensing will revolutionise environmental information provision.  
(this thesis)
2. Integration of formal and informal environmental sensing data is essential for decision making.  
(this thesis)
3. To extract valuable information from big data, small data analysis is essential.
4. Open hardware accelerates research on natural resource management and recreation.
5. Low Power Wide Area Network (LPWAN) is the solution for sustainable monitoring of precision agriculture.
6. Robots are the new steam engines.
7. Smart citizens create smart cities.

Propositions belonging to the thesis, entitled

Citizen sensing for environmental information

Qijun Jiang

Wageningen, 23 May 2018

## **Citizen sensing for environmental information**

Qijun Jiang

## **Thesis committee**

### **Promotor**

Prof. Dr A.K. Bregt  
Professor of Geo-information Science  
Wageningen University & Research

### **Co-promotor**

Dr L. Kooistra  
Associate professor, Laboratory of Geo-information Science and Remote Sensing  
Wageningen University & Research

### **Other members**

Prof. Dr R. Uijlenhoet, Wageningen University & Research  
Dr L.J. Carton, Radboud University, Nijmegen  
Dr J. van Etten, Bioversity International, Costa Rica  
Dr G.J. Steeneveld, Wageningen University & Research

This research was conducted under the auspices of the C.T. de Wit Graduate School of Production Ecology & Resource Conservation (PE&RC)

# **Citizen sensing for environmental information**

Qijun Jiang

## **Thesis**

submitted in fulfilment of the requirements for the degree of doctor  
at Wageningen University  
by the authority of the Rector Magnificus,  
Prof. Dr A.P.J. Mol,  
in the presence of the  
Thesis Committee appointed by the Academic Board  
to be defended in public  
on Wednesday 23 May 2018  
at 4 p.m. in the Aula.

Qijun Jiang  
Citizen sensing for environmental information,  
156 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2018)  
With references, with summary in English

ISBN 978-94-6343-849-0  
DOI: 10.18174/444457

# Contents

	<b>Page</b>
<b>Chapter 1.</b> Introduction	7
<b>Chapter 2.</b> Citizen sensing for improved urban environmental monitoring	13
<b>Chapter 3.</b> Field calibration of electrochemical NO <sub>2</sub> sensors in a citizen science context	31
<b>Chapter 4.</b> Formal and informal environmental sensing data and integration potential: Perceptions of citizens and experts	55
<b>Chapter 5.</b> Approaches for Formal and Informal Environmental Sensing Data Integration	75
<b>Chapter 6.</b> Synthesis	93
<b>References</b>	107
<b>Appendices</b>	119
<b>Summary</b>	145
<b>Acknowledgements</b>	149
<b>List of publications</b>	150
<b>Short biography</b>	151
<b>PE&amp;RC Training and Education Statement</b>	152





# **Chapter 1**

## **Introduction**

## 1.1 Background

Citizens in Amsterdam want to know the air quality related to their everyday living environment. For instance, citizens may want to know the healthiest route to their work or to school. Subsequently, they check if there is any information available to provide the answer. However, in the current situation they will find out there are only 12 official air quality measurement stations in Amsterdam city. For some citizens, this means, that the closest station from their home is still several kilometres away. Therefore, they cannot get accurate air quality information of their neighbourhood from such measurement stations. These official measurement stations produce high quality data at the measured locations. However, finding the appropriate answers for cleaner routes, is difficult from such a limited number of stations. Of course, air quality of unmeasured locations can be predicted by modelling, but how accurate are these predictions? What if citizens can measure their environment by themselves with low-cost sensors? Then it would be possible to measure air quality at citizen's home, on their way to school or at work more precisely. This offers them the possibility to find healthy routes. What if they share and open their measured data? If so, these data can be used to derive information for better understanding of our environment. Citizens but also policy makers could make better informed decisions.

The case described above is not a specific case for Amsterdam. It is a global issue even though circumstances may differ from city to city. What do we get from this story? First, we see that to make proper decisions for a healthier living environment, data or information produced by environmental monitoring are crucial. This will be further elaborated in the next section. Second, the story tells us that except official measurements (formal data), citizens can also measure the environment to collect data (informal data) for their decisions. These activities are called citizen science, and the development and the relevance for this research will be elaborated in section 1.3. The low-cost sensors which can be used for citizen environmental monitoring will be elaborated in section 1.4. Finally, the story raises questions about how these data produced by citizens should be used. The potential combined use of formal and informal environmental sensing data will be elaborated in section 1.5.

## 1.2 Environmental monitoring

Environmental monitoring is essential for information provision to understand our living environment. Only when this information is provided, policy makers, the public and individuals are able to make proper decisions relevant to environmental issues. Environmental monitoring organizations get an increasing number of requests from the public and policy makers for more detailed environmental information (Minkman et al., 2017). However, traditional environmental monitoring networks use certified sophisticated equipment which normally are expensive, static and need specialists to maintain. Therefore, it is difficult and costly to have a denser network of these environmental monitoring stations. On the other hand, citizens are increasingly interested to measure their environment using low-

---

cost sensing devices. These low-cost devices are mainly used by citizens and not often by official monitoring organizations. Mainly because official monitoring organizations are based on legal requirements and specifications from for example the national government, the European Commission or United Nations. Low-cost sensors are often not used by official organizations due to unknown data accuracy (Lewis and Edwards, 2016). Consequently, community or citizen level environmental data collection using emerging low-cost sensors is rising (Snyder et al., 2013).

### 1.3 Citizen science

Citizens have been involved in environmental monitoring for a long time. In the geo-information science domain, data collected by citizens are called voluntary geography information (Goodchild, 2007). In the environmental domain, it is also called participatory monitoring. Crowdsourcing is another term used in other disciplines. These are just a few terms used to describe citizens' involvement in science. Eitzel et al. (2017) discussed those terms in more detail. In our research, we use the term 'citizen science'. We consider citizen science as individuals and/or communities performing scientific activities from raising questions, collecting data to findings answers with or without the involvement of scientists (Bonney et al., 2016; Haklay, 2013; Lewenstein, 2016; See et al., 2016).

Citizens have been involved in environmental monitoring for data collection in various domains. For example, weather monitoring by amateurs (Bell et al., 2013), invasive species monitoring (Delaney et al., 2007) and water quality monitoring (Jollymore et al., 2017). Actually, citizens have been contributing to science earlier than scientist became a profession in late 19th century (Silvertown, 2009). For instance, Antonie van Leeuwenhoek, who did pioneering work in microscopy and contributed to the establishment of microbiology, was a Dutch businessman (Ruestow, 1996). Nowadays, due to the development of information and communication technology, citizens can participate and initiate citizen science projects easier than ever (Sullivan et al., 2009). First, citizens can access information and contribute data through Internet by computers, laptops, smartphones and other digital devices. Second, citizens can communicate with each other, with scientists and policy makers using social media applications and tools like Facebook, Twitter, Meetup and many more. Third, they can also use digital applications and tools for online and offline citizen science activities, such as organization, collection, processing, sharing, analysing and utilization of data. In short, the technology development has profound impact on citizen science. Specially, the development of sensor technology not only gives way to new forms of environmental data collection but could potentially change environmental monitoring profoundly.

## 1.4 Sensors for environmental monitoring

Sensors can detect our environment in different ways. A sensor is defined as: 'a device that converts a physical measure into a signal that is read by an observer or by an instrument' (Chen et al., 2012). For instance, gas sensors can use an electrochemical mechanism to measure concentrations of certain gasses, but can also detect gasses based on a metal oxide semiconductor mechanism (Mead et al., 2013; Pang et al., 2017; Thompson, 2016). Thanks to technological developments, the sensors are becoming smaller, cheaper, portable and more functional. Some sensors can measure and communicate environmental parameters in near real time. These sensors potentially can help us to collect environmental information cheap, timely, and frequently. Therefore, dense networks of sensors would be possible. This will help us to understand more about our dynamic environment.

Kumar et al. (2015) discussed how the use of low-cost sensing for urban air pollution monitoring is emerging and indicated as main challenge data reliability. Low-cost sensors have also been used in weather monitoring (Bell et al., 2013) and noise mapping (Aumond et al., 2017). Besides an increasing interest in using sensors for environmental monitoring, there are also many challenges related to data quality regarding instrument biases (Bell et al., 2015; Borrego et al., 2016), robustness and measurement repeatability (Castell et al., 2017a), accuracy (D'Hondt et al., 2013; Duvall et al., 2016), interferences (Lewis et al., 2016) calibration (Hasenfratz et al., 2012; Lin et al., 2015; Spinelle et al., 2015b), energy consumption, and networking. Lewis and Edwards (2016) raised the concern of validation for air quality sensors regarding the rising interest of using low-cost sensors to measure air quality. How to use these sensors properly and how the produced data should be used are still questions which should be comprehensively studied.

Another question, raised by (Kumar et al., 2015), is how these informal sensor data from citizen science can be used. Can these data be used to complement formal environmental sensing data? And on the other hand, how to use formal environmental sensing data to enhance informal sensing data?

## 1.5 Formal and informal environmental sensing data

Formal environmental sensing data is defined in this research as data produced by official environmental institutes using standardised equipment and data collection procedures. On the other hand, we define informal sensing data as data produced by citizens using low cost sensors. Formal environmental sensing data are generally considered accurate, consistent and complete since they are based on legal requirements and specifications (Snyder et al., 2013). However, they are often collected at limited locations at a sparse number of stations. Informal environmental sensing data, on the other hand, produce increasing amounts of environmental data with denser networks and increasing interest from citizens and experts in citizen science. But informal environmental sensing data is currently considered as low-quality data. Nevertheless, informal environmental sensing data

---

are increasingly studied and tested in environmental monitoring to study the reliability, challenges and opportunities (Bell et al., 2013; Conrad and Hilchey, 2011; Kumar et al., 2015).

Formal environmental sensing is essential for qualified information provision. However, informal environmental sensing can also be important to provide complementary information. For example, after the Fukushima Daiichi nuclear disaster in Japan in 2011 citizens were eager for radiation information. Their demands were not satisfied by formal radiation information sources. A community of interested citizens then create their own tools to monitor radiation and share this information with others (Hemmi and Graham, 2014). These data collected by citizens provided information not only for themselves but also for scientists and policy makers. In another example, scientists deployed a low-cost ozone sensor network in Auckland, New Zealand, where limited formal ozone monitoring stations were existing. The deployed low-cost sensor network produced complementary information to let scientists better understand the dynamics of ozone in that area (Weissert et al., 2017).

More importantly, if formal and informal environmental sensing data can be integrated, we might gain more information than using either one separately. Mazzoleni et al. (2017) assimilated data derived from crowdsourced observations of water levels in hydrological modelling and concluded that citizen collected data can complement formal data to improve flood forecasts. However, formal and informal sensing data collection are organized and collected in different ways with different tools, standards, metadata, data quality control and so on. As a result, it is still challenging to integrate formal and informal sensing data.

Since informal environmental sensing is an emerging phenomenon, the perception on the usefulness of it by both experts and citizens is an important issue. Perception, defined as “something is regarded, understood, or interpreted” (Oxford Dictionaries, 2017), from citizens and experts on formal and informal environmental sensing data and the potential for integration influences the use of these data. For formal and informal environmental sensing data, citizens and experts may not have the same perception on, for instance, data accuracy. Even though these perceptions do not always reflect reality, they play a key role in the willingness to use these data.

## 1.6 Research objectives

The overall objective of this thesis is to assess opportunities for informal sensor data and their integration with formal sensor data to improve environmental information provision by investigating challenges and obstacles regarding technical and societal aspects.

The thesis was guided by the following main research questions:

1. How does a bottom-up citizen science project develop, and can it contribute to environmental monitoring?

2. How to calibrate low cost air quality sensor data, and how valuable are these data?
3. What are the perceptions of citizens and experts on formal and informal environmental sensing data and their integration potential?
4. How can informal and formal sensing data be integrated to provide enhanced environmental information?

## 1.7 Thesis outline

This thesis consists of six chapters including this introduction chapter. Chapter 2 provides an overview of citizen environmental sensing focusing on an urban air quality monitoring case study in Amsterdam. It examined how a citizen environmental sensing campaign was developed and it discusses the value of citizen environmental sensing. Next, in Chapter 3, I focus on sensor calibration in a citizen science case to examine the value of informal sensing data: lessons learnt and recommendations are presented. In Chapter 4, an analysis is presented on the perceptions from experts and citizens on formal and informal environmental sensing data based on a questionnaire survey. In Chapter 5, a framework for formal and informal sensing data integration is presented. A literature study and an expert validation method are used to evaluate the proposed framework and integration approaches. Chapter 6 provides a synthesis and conclusion on the research results of this thesis regarding the proposed research questions. Recommendations for future research are also presented.

# Chapter 2

## **Citizen sensing for improved urban environmental monitoring**

This chapter is based on:

Jiang, Q., Kresin, F., Bregt, A.K., Kooistra, L., Pareschi, E., van Putten, E., Volten, H., Wesseling, J., 2016. Citizen Sensing for Improved Urban Environmental Monitoring. *Journal of Sensors* 2016, 9.

## **Abstract**

Citizen science is increasingly being used in diverse research domains. With the emergence and rapid development of sensor technologies, citizens potentially have more powerful tools to collect data and generate information to understand their living environment. Although sensor technologies are developing fast, citizen sensing has not been widely implemented yet and published studies are only a few. In this paper, we analyse the practical experiences from an implementation of citizen sensing for urban environment monitoring. A bottom up model in which citizens develop and use sensors for environmental monitoring is described and assessed. The paper focuses on a case study of Amsterdam Smart Citizens Lab using NO<sub>2</sub> sensors for air quality monitoring. We found that the bottom up citizen sensing is still challenging but can be successful with open cooperation and effective use of online and offline facilities. Based on the assessment, suggestions are proposed for further implementations and research.



---

## 2.1 Introduction

Highly dynamic environmental phenomena call for detailed and timely environmental information to support decision making. For instance, in the case of air pollution, it is essential to understand where this pollution comes from and how to reduce it (Kolb et al., 2004). For environmental disasters such as floods (Schnebele et al., 2014a), severe weather (Zhang et al., 2011) and volcanic eruptions (Werner-Allen et al., 2005), sufficient and timely environmental information is essential for risk forecasting and early warning. To derive this information, environmental data needs to be collected.

Traditionally, environmental monitoring is conducted by official authorities which usually spend large amounts of money for high quality but expensive monitoring equipment followed by continuous labour and money investments on maintenance and calibration (Mead et al., 2013) which leads to often low spatial and temporal resolution (Hasenfratz et al., 2012). Therefore, these data sources are often too sparse to meet the information demands from the public and organizations. For example, there are twelve air quality monitoring stations in Amsterdam operated by the Public Health Service of Amsterdam (GGD Amsterdam) and the National Institute for Public Health and the Environment (RIVM) at selected locations and most of these stations measure a limited number of air quality parameters; compared to Amsterdam, other cities in the Netherlands even have less official air quality stations (<https://www.luchtmeetnet.nl>). In addition, in some developing countries, official environmental monitoring systems are completely absent (Mead et al., 2013). For instance, during the 26 December 2004 tsunamis around the Indian Ocean, affected countries could have had enough time to avoid the disaster if they had employed a functional alarm system earlier (Kelman, 2006; Samarajiva, 2005).

There is a general need for flexible and affordable alternatives to complement the official or formal environmental monitoring stations. Recent developments of sensor technologies allow citizens to buy affordable sensors and electronic components like Arduino (<https://www.arduino.cc>) and Raspberry Pi (<https://www.raspberrypi.org>) to create sensor systems by themselves or with help from communities which provide alternative approaches to collecting environmental data (Akyildiz et al., 2002; Zerger et al., 2010). So-called informal sensors operated by citizens are not only raising public awareness of environmental problems from social aspects but also are potentially capable to complement the quantity and spatial-temporal resolution of the formal environmental data sources (Corke et al., 2010). For example, after the Fukushima Daiichi accident on March 11, 2011, local citizens started using sensors, developed by a project called Safecast, to personally detect the radiation levels and shared the data to a website which gathered all these data coming from citizens. Now this has become an international community for radiation monitoring all over the world (Hemmi and Graham, 2014) Citizens contribute to the data source and benefit not only themselves but also the larger public and policy makers with an independent source of environmental data. Similarly, weather stations operated by amateurs in

the UK are rapidly growing and have become crucial data sources supplying timely and high density monitoring data for local weather information (Bell et al., 2013).

Citizen science has existed for a long time. It has been used for bird observations (Sullivan et al., 2009), invasive species monitoring (Delaney et al., 2007) and other domains. Most of them use top down approaches in which scientists design the research project and subsequently citizens are asked to join, mainly collecting data. Due to development of sensor technology, citizens can now use affordable sensors to monitor their living environment by themselves. They can even create their own sensor systems for their interest. However, it is still not clear how this bottom up approach should be organized; how to tackle the calibration and implementation problems; what is the data quality of informal sensor networks compared to formal sources; how is the knowledge dissemination on sensor development, and how to make sense of these data. According to previous studies, citizen science can be classified as: community consulting model that citizens only define the problems, community workers model that citizens are mainly involved in collecting data, and community-based, participatory research model that citizens are involved in all research activities (Wilderman, 2007).

In this chapter, by taking a case study focusing on air quality monitoring in the city of Amsterdam, a bottom-up citizen science approach for informal sensor environment monitoring is developed and evaluated.

## **2.2 A bottom up approach for informal sensing**

The project named Amsterdam Smart Citizens Lab (<http://waag.org/en/project/smart-citizens-lab>) uses a bottom up approach to organize citizen sensing for urban environmental monitoring. This approach can be classified as co-created class and is recognized as a community-based, participatory research model that citizens involved in all steps of the project (Cooper et al., 2007; Morzy, 2015; Wilderman, 2007). Within this approach, citizens are involved in each step (as shown in Figure 2.1) of sensing strategy together with project partners.

The project was initiated by Waag Society, Institute for Art, Science and Technology, a pioneer in the field of digital media at Amsterdam. The approach is completely bottom-up. Waag Society is responsible for community coordination and provides a place for meetings and for sensor making in their Fablab Amsterdam, a place for makers. As organizer, Waag Society invited other partners who have expertise in different aspects to help citizens, such as Netherlands Organisation for Applied Scientific Research (TNO), Amsterdam Smart City (ASC), RIVM, and SenseMakers. Firstly, Waag Society organized a meetup called Topical BarCamp in which citizens raise their urban environment concern (issue mapping in Figure 2.1). According to the issue mapping from citizens and their interests, the community was divided into small groups to develop and test sensor systems for specific urban environment problems. During this period, the Fablab Amsterdam was open every Tuesday for teams to use their facilities for making prototypes. This step is called sensors making and done by an Open Hardware Bootcamp. After the prototypes were

developed and tested, citizens used the sensors they developed themselves to monitor the city environment and collect data (sensing). These data were then interpreted and visualized with the help of experts. For each step, Waag Society invited experts to inspire and share their experience. In the next section, the adopted methods and results for every step of the proposed bottom-up approach are presented in more detail.

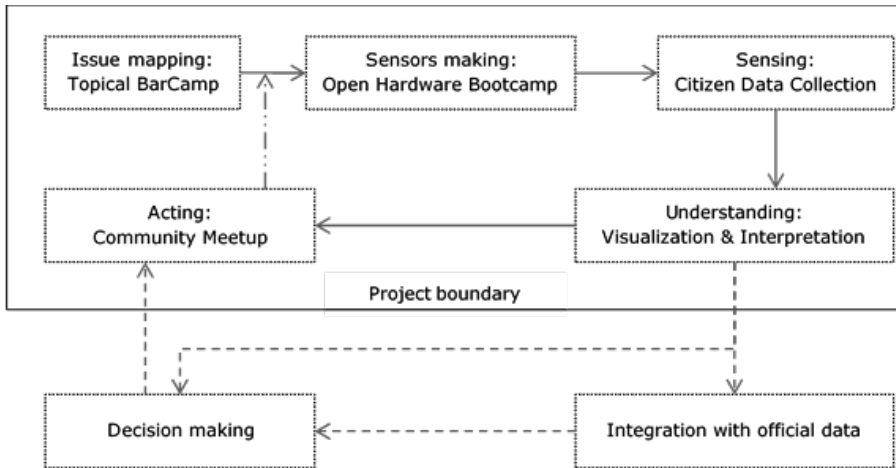


Figure 2.1. The project steps of Amsterdam Smart Citizens Lab for bottom up citizen sensing.

After Issue mapping, four sensing topics (air quality, noise pollution, wind and road bumpiness) were raised by citizens. For this paper the air quality topic was selected as a case study to describe and evaluate the bottom up citizen sensing approach. This case was inspired by previous experience within the Smart Citizen Kit (<https://smartcitizen.me>) pilot in 2014. During this pilot, Amsterdam citizens measured temperature, humidity, light intensity, sound levels, carbon monoxide and nitrogen dioxide in the city using Smart Citizen Kit 1.1 version. One outcome of the Smart Citizen Kit pilot was that the semiconductor air quality sensors (CO and NO<sub>2</sub>) used in Smart Citizen Kit 1.1 version were not suitable for urban air quality monitoring (Horn and Boonstra, 2014). A community was established including citizens and experts during this pilot. Some of them decide to join the Amsterdam Smart Citizens Lab to explore new sensors which can be used for urban air quality monitoring. Compared to other groups, this air quality group is relatively diverse including citizens, hardware developers, air quality experts and university researcher.

## 2.3 Results

In this section, we present the results from each step of the bottom up approach as indicated in Figure 2.1.

### 2.3.1 Issue mapping: Topical BarCamp

At the issue mapping meeting (Figure 2.2), people discussed what kind of environmental problems they were concerned about and would like to solve. The meeting was organized as a Topical BarCamp which means that the content is provided by the participants. Air quality, noise pollution, wind and road bumpiness were raised as main issues of concern. The Meetup environment (<http://waag.org/en/project/smart-citizens-lab>) is used as a community platform that has functions like member registration, events organization, sharing of information, communication and so on. Citizens were informed to join this platform during this meeting.



Figure 2.2. Citizens discuss environmental problems and sensing strategies during the issue mapping phase.

### 2.3.2 Sensors making: Open Hardware Bootcamp

One aim of the air quality group is to find an air quality sensor with the proper requirements for urban air monitoring. The main three requirements are that the sensor should 1) be able to measure a pollutant relevant for urban environments, 2) be sensitive enough to measure typical ambient concentrations, and 3) be affordable. Over recent years, most of the concentration levels in the Netherlands have decreased substantially. Presently, the nitrogen dioxide levels are the most likely ones to lead to exceedances of legal threshold values. As a result, there is much emphasis on nitrogen dioxide, both from official authorities as well as from concerned citizens. Several measuring campaigns using passive NO<sub>2</sub> samples were undertaken during the last five years, both by municipalities and concerned citizens. Combined with the available state of the art in sensors for gasses at ambient concentration levels this has led us to focus on nitrogen dioxide concentrations. Based on these requirements and the experience of the Smart Citizen Kit pilot (<http://waag.org/en/project/smart-citizen-kit>), different sensor options were proposed by group members. With scientific proof (Hasenfratz et al., 2015; Mead et al., 2013), after group discussion and comparison, the electrochemical Alphasense NO<sub>2</sub> sensor (NO<sub>2</sub>-B42F) (<http://www.alphasense.com>) was chosen for this experiment, according to the specification, the measurement range is 20 ppm NO<sub>2</sub> limit of performance warranty. The Arduino Uno or Arduino compatible microcontrollers were chosen to connect the sensor and other components such as the power supply, a real time clock (RTC) and a storage module

(Figure 2.3). The RTC is used for the timestamp, the temporal resolution can be programmed accordingly. For the measurement campaign, we measured every minute. To read the analogue signal from sensors, a high resolution analogue to digital converter (ADC) above 16-bit is required. The Arduino board only has a 10-bit ADC on board, which is not enough to accurately determine the output of the sensor that varied only a few millivolts (mV). Therefore, an external ADC was used. To reduce the noise of the data, a stable power supply is also recommended, which was not implemented. During the Open Hardware Bootcamp, the sensors were made by the citizens with assistance of experts of Fablab Amsterdam, RIVM and Wageningen University. In total, five NO<sub>2</sub> sensor boxes were prepared in this phase of the project.

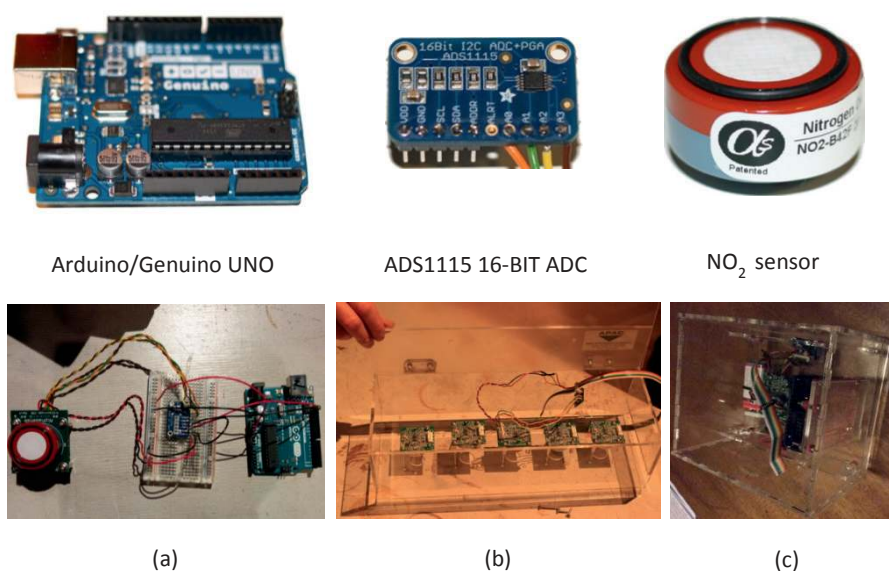


Figure 2.3. Main hardware, sensor and different prototypes: (a) function test, (b) sensors with waterproof enclosure for outdoor test and (c) sensor box for monitoring campaign.

To test the performance of the Alphasense NO<sub>2</sub> sensor, we first located four Alphasense NO<sub>2</sub> sensors together indoors at Waag Society as these sensors had no enclosures designed yet. Besides these Alphasense sensors, there is also a Smart Citizen Kit measuring outdoor of Waag Society. As can be seen in Figure 2.4, all Alphasense NO<sub>2</sub> sensors show good sensitivity compared to the Smart Citizen Kit, which did not show any variation of the 24 hours measurement period. However, although Alphasense NO<sub>2</sub> sensors indicated quite similar trends, the four sensors did not show the same concentration. Clearly, a calibration procedure is needed.

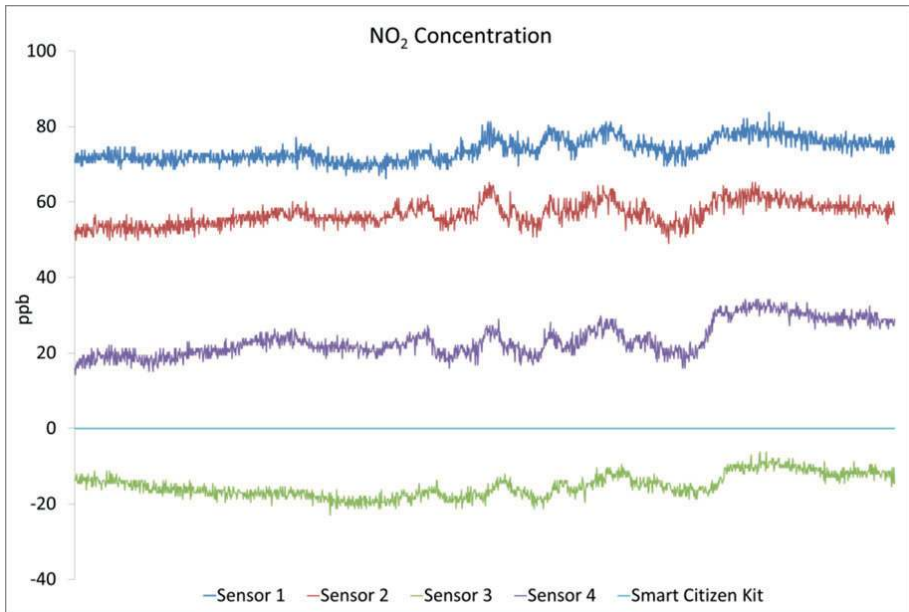


Figure 2.4. The calculated NO<sub>2</sub> concentrations measured by four sensors at the same test location (indoor in the building of Waag Society) and the measurement of a Smart Citizen Kit (outdoor of Waag Society) every minute on September 20th, 2015, for around 24 hours.

Next, four sensors were mounted in a common box and placed outside a window of the Waag Society building. Measurements were taken during five days and afterwards the raw outputs voltages of the sensors were converted to estimated NO<sub>2</sub> concentrations. For the calibration, readings from a nearby located official measuring station (Oude Schans, roughly 300 meters from the Waag Society) and one other station southwest of the centre of Amsterdam (Vondelpark, roughly 2700 meters from the Waag Society) were used. This latter station provides an indication of the city background concentration for NO<sub>2</sub> during the prevalent western winds.

According to “Alphasense 4-Electrode Individual Sensor Board (ISB); User Manual 085-2217” the sensors output two voltages, the “Working Electrode” (WE) and the “Auxiliary Electrode” (AE). Both have to be corrected for a zero-offset of typically 225-245 mV. The value of the corrected AE is subtracted from the value of the corrected WE and the remaining voltage is divided by sensitivity in mV/ppm of typically 0.175-0.185. In order to get a similar set of hourly concentrations for all sensors offsets for WE of 234 mV were combined with AE offsets between 220 and 245 mV, combined with sensitivities between 0.5 and 0.8 mV/ppm. Finally, a conversion from ppb to micrograms per cubic meter of air ( $\mu\text{g}/\text{m}^3$ ) was performed as all Dutch concentrations are reported in  $\mu\text{g}/\text{m}^3$ .

In the beginning of the comparison the Alphasense sensors took several hours to become stable. With the adjustments, the concentrations show a roughly similar

pattern in time as the official measuring stations (Figure 2.5). The concentrations show a daily pattern. The reason for the high concentrations at station “Oude Schans” between the hours 30 and 45 is not clear. The concentration values are averaged and presented in hourly values as shown in Figure 2.5.

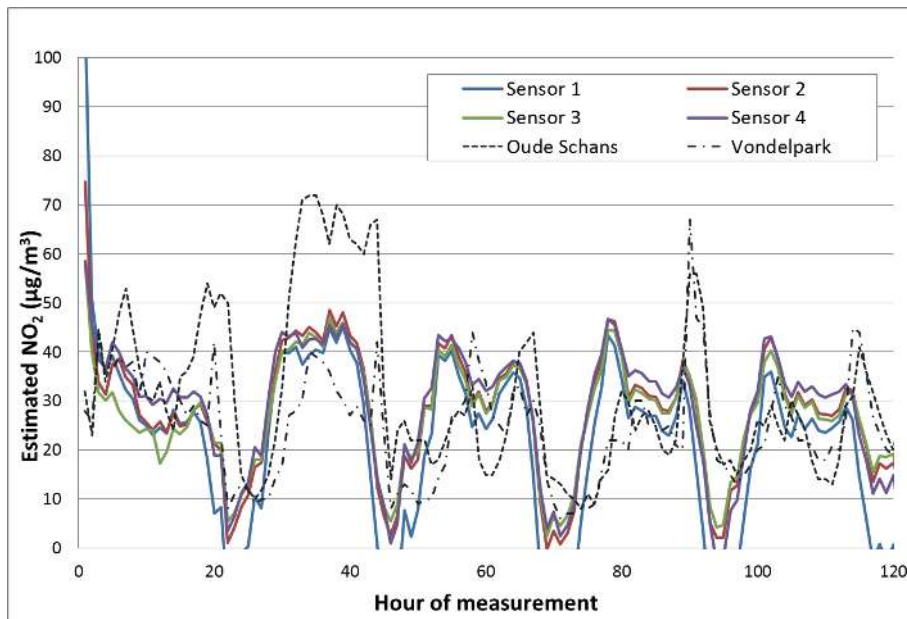


Figure 2.5 Measured NO<sub>2</sub> concentrations outside Waag Society during five days. Oude Schans and Vondelpark are official stations operated by GGD Amsterdam.

In order to obtain a similar behaviour for all sensors, the parameters used to convert output voltages into concentrations varied substantially between sensors. The concentrations show a roughly similar pattern in time as the official measuring stations do.

### 2.3.3 Sensing: Citizen Data Collection

After system testing and improving, an air quality campaign test was conducted by the air quality monitoring group on December 2, 2015. After discussion on a paper map for measurement campaign location planning, the digital map was developed for online access (Figure 2.6). The developed sensor prototypes do not have Global Positioning System (GPS) module, to get the measurement location's latitude and longitude coordinates, a measurement location retrieval tool was developed which can use GPS and Wi-Fi signals on the smartphone to retrieve location coordinates (Figure 2.7). A website was developed to host these prototype tools and visualization map. Twenty-Seven locations were selected including city background, traffic streets and parks. Figure 2.8 shows a monitoring example that is close to an official monitor station (Amsterdam-Stadhouderskade) for

comparison. For each location, the air quality sensor was operated for a minimum of 15 minutes, which enabled the sensors to stabilize.

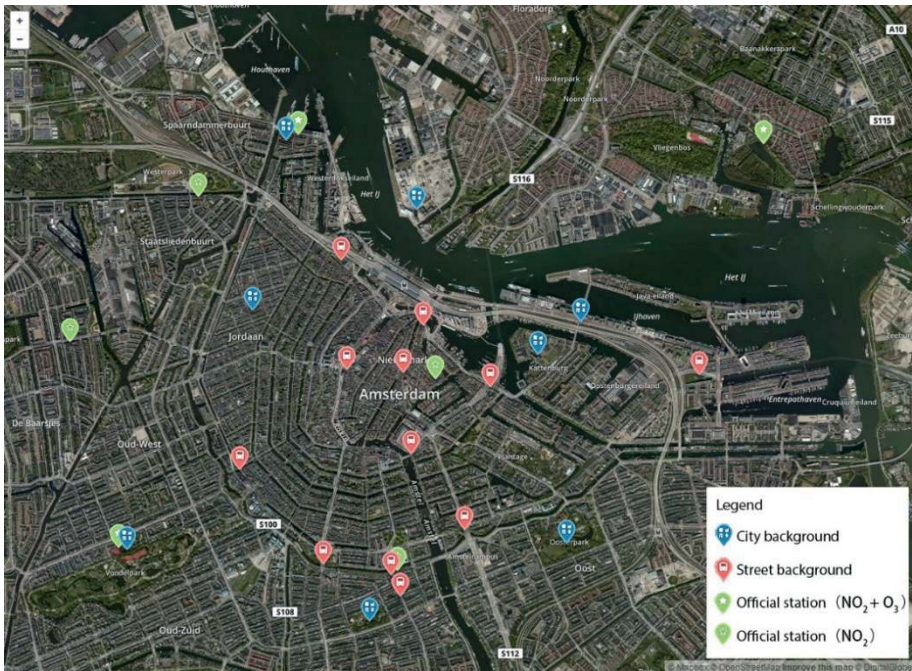


Figure 2.6. Plan for locations in Amsterdam to perform air quality monitoring.

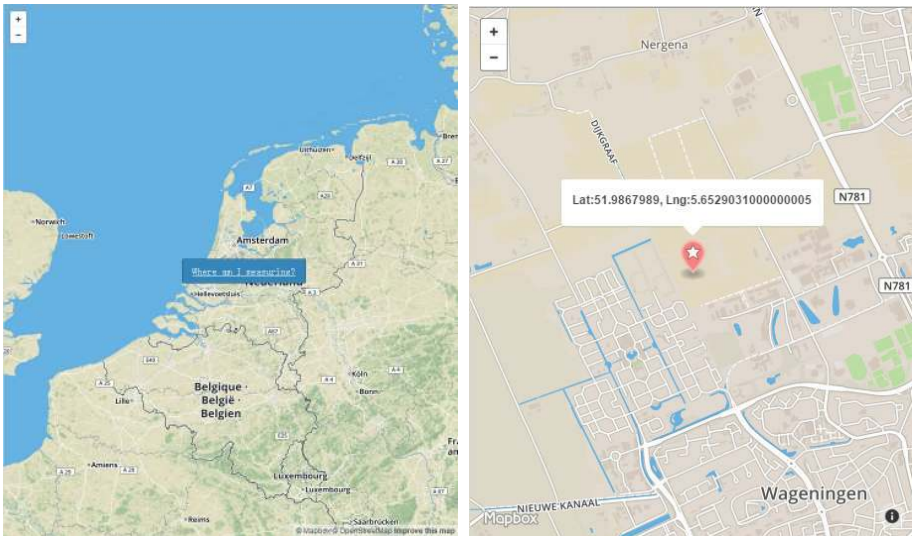


Figure 2.7. Tool to retrieve latitude and longitude of measurement locations.





Figure 2.8. An informal measurement point is close to an official station (Amsterdam-Stadhouderskade) for data comparison.

### 2.3.4 Understanding: Visualization & Interpretation

A typical concentration pattern is shown in Figure 2.9. First a measurement was done outside (traffic), next the sensor was transported to a new location and was put in a bag for some time before taking out and moving to next location, and subsequently a follow-up measurement was done at the new location (park). As can be seen in Figure 2.9, the  $\text{NO}_2$  sensor shows considerable different results between in the bag and out of the bag. No large differences between results in traffic and park locations are observed. In order to evaluate the performance of Alphasense  $\text{NO}_2$  sensors, we installed sensors close to the GGD Amsterdam official air quality monitoring stations (see Figure 2.8 as an example). The official station data can be downloaded from the web portal Luchtmeetnet (<http://www.luchtmeetnet.nl>) and are on hourly basis (Figure 2.10). Compared to the official measurements, the informal Alphasense  $\text{NO}_2$  sensor measured concentrations in minutes and shows a similar trend (Figure 2.10).

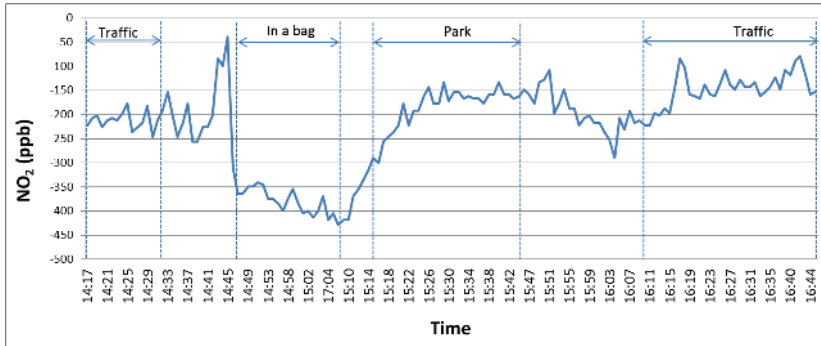


Figure 2.9. Roughly comparison of measurements from one sensor.

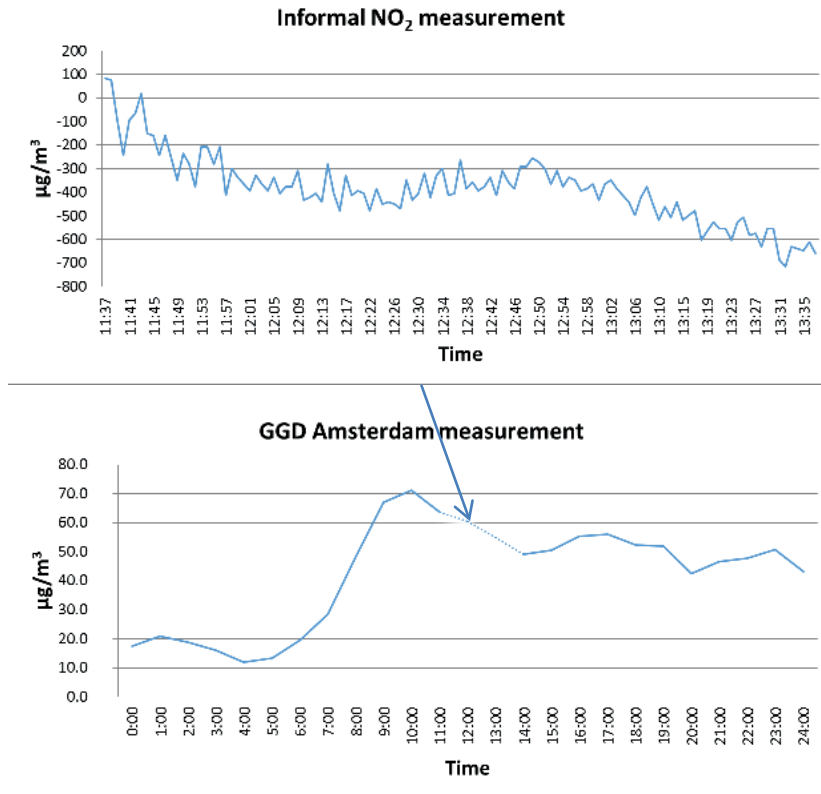


Figure 2.10. Roughly the same period of the informal sensor measurement (without calibration, 1 ppb = 1.91 µg/m<sup>3</sup> is used for the conversion) every minute (11: 37 - 13: 36) and the hourly official measurement in dashed line (11: 00 - 14: 00) at Amsterdam-Stadhouderskade.

To visualize NO<sub>2</sub> concentration data, an online map service (Figure 2.11) was developed using open source Mapbox and GitHub Pages. The concentration is colour coded according to the Dutch national standard.



Figure 2.11. The data visualization map prototype for NO<sub>2</sub> measurement campaign on December 2, 2015.

### 2.3.5 Community analysis

As shown in Figure 2.12, the number of members of the Amsterdam Smart Citizens Lab has been increasing continually which means that this community received continuous attention from the public. However, according to the Figure 2.13, the number of active participants for each community event did not increase substantially.

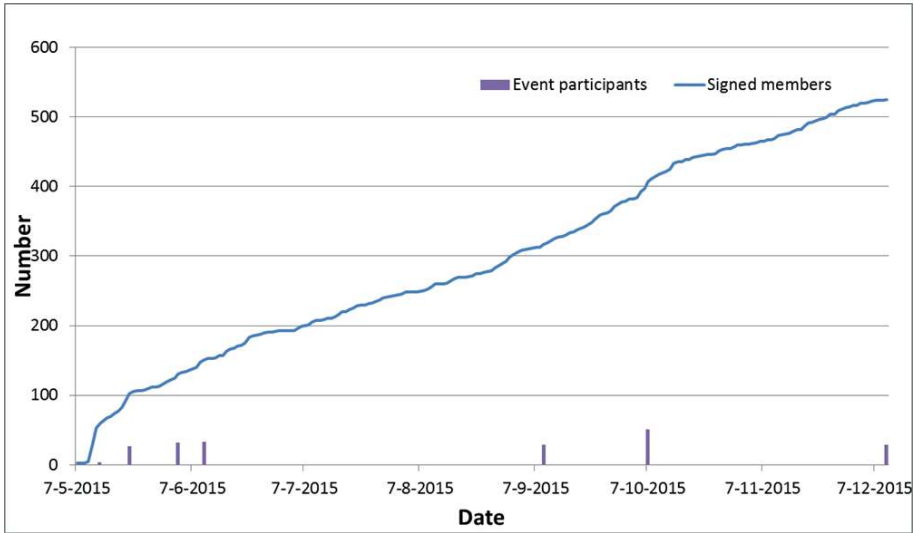


Figure 2.12. The growing number of total signed members (data collected from Meetup) compared to event participants indicated in purple bars.

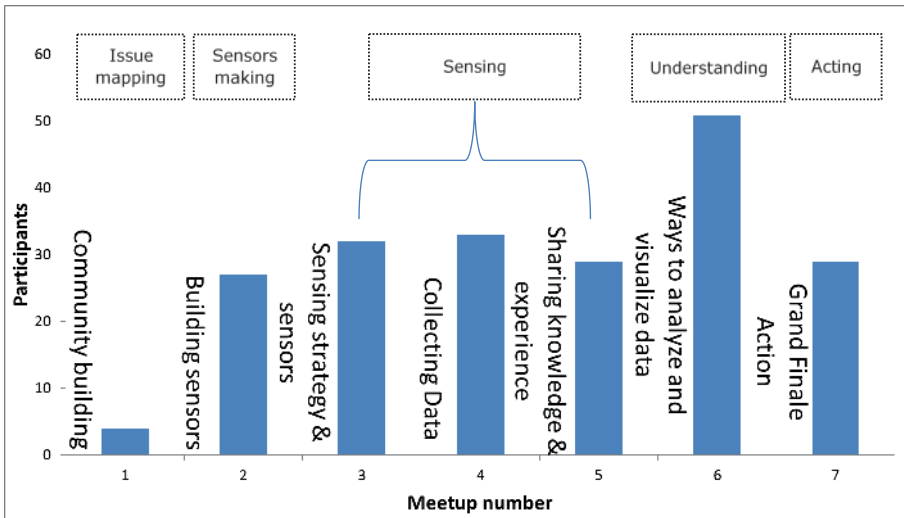


Figure 2.13. The variation of community participation: data collected from Amsterdam Smart Citizens Lab Meetup, the actual participants may be different, especially the first meeting in which most citizens have not registered for the Meetup (see Figure 2.2).

---

## 2.4 Discussion

In this part, we first take air quality monitoring as an example to discuss the technical and operational challenges. Next, we analyse and discuss the bottom up approach for citizen sensing.

### 2.4.1 Citizen sensing is still challenging

The Arduino board is a relatively simple and powerful tool for experimentation. However, to retrieve reliable data, professional knowledge and support are still needed. An important issue for air quality measurement is sensor data calibration and analysis. With different tests, several problems were identified but we did not have sufficient time within this project to find solutions or to test potential solutions. The NO<sub>2</sub>-B42F sensor is cross-sensitive to ozone, temperature and humidity, as described in the specifications. Our tests suggest that other environmental factors may also influence the results of measurements. The indoor test as shown in Figure 2.4 shows that even each sensor package was calibrated individually by the sensor company; the four sensors indicated quite different concentrations. This is an obstacle for citizens, because if further calibration is needed, citizens normally do not have the facilities to do so.

This project was based on the experience of the Smart Citizen Kit pilot (Horn and Boonstra, 2014). The general aim was to encourage citizens to measure their environment by developing sensor platforms together in a community. According to the outcomes and observations of Amsterdam Smart Citizens Lab, this complete bottom up approach is challenging but can be successful. Still an important precondition for success is that expert communities are involved. They provide the required support to build a fully functional sensor system which can collect reliable data. Independent development would still be challenging for citizens. There are quite a lot of sensors and sensor-related electronic products, which are assumed to be “plug and play” products. To create a functional environmental sensing system is an interdisciplinary task which requires knowledge from different domains like electronics, environmental sciences, communication, information technology, design and so on.

### 2.4.2 Bottom up approach calls for broad open cooperation

The continually reduced costs and increased functions of sensors and microcomputers offer citizens possibilities to measure the environment by themselves, which has already made an impact for environment awareness and decision making. There are a lot of open hardware and software resources that can be used by citizens. Furthermore, the fab labs and the maker movement also play an important role in citizen science. The fab labs provide facilities for citizens to get together to communicate, use facilities and get help to make their sensor platforms. From this, we see organizations that support citizens sensing as an important resource. Research conducted by Balestrini et al. (2015) concluded that the supporting organization is important in terms of connecting people, supplying

guidelines and helping each other. From our observation, this is confirmed in this case study. In particular, we observe that the support or involvement from professionals in pushing the community by supplying information, suggestions and technical assistance is very valuable. Since environmental sensing by citizens is an interdisciplinary field, the experts from different domains foster the community and keep the project running.

### **2.4.3 Stimulate citizens for a sustainable and growing community-online plus offline**

In the Amsterdam Smart Citizens Lab, Fablab Amsterdam plays an important role for offline meetups for citizens to learn skills, work with like-minded people, and create and test prototypes. In addition, the online meetup platform is not only used to organize offline meetups but also functioned as a broader virtual community platform for information sharing and attracting more people to join the community. Even though from Figure 2.12, the active participants did not increase substantially compared to the continuous growing online community, it shows that citizens are interested in the project and potentially could become active participants. From the Meetup platform, citizens know what has been achieved by the Amsterdam Smart Citizens Lab community, they can join the community and keep it updated for further activities. This “online plus offline” model is a good approach to organizing citizen sensing. There are 602 fab labs in the world and the number is still growing (<https://www.fablabs.io/labs>). In addition, there are diverse open living labs (<http://www.openlivinglabs.eu>) and hacker spaces. For online tools, besides Meetup (<http://www.meetup.com>), social media like Facebook, Twitter, YouTube and blogs can also be used for community building. Trello and Slack can be used for teamwork. If these tools and open offline physical places can be used together effectively and efficiently, this may help to stimulate citizens to create a sustainable and growing global and local community.

## **2.5 Conclusions**

In the present paper, we present a bottom up approach for citizen science to collect informal urban environmental sensor data. We found that highly sensitive electrochemical sensors potentially have better performance than semiconductor sensors for urban air quality monitoring but need appropriate hardware and software design, careful calibration and post-processing to deliver correct and usable data. This leads to challenges for citizens to build sensor systems from scratch. Therefore, wide cooperation from different aspects such as community building, maker spaces, fab labs and different types of professional support for domain knowledge is essential for citizen sensing. This wide cooperation, together with effective use of online and offline facilities can keep the citizen sensing community sustainable and growing.

## **2.6 Future work**

For future work, the focus will be on how citizens and experts can work together to optimize the citizen environmental sensing model for reliable data provision to benefit citizens, official organizations, and more importantly the whole society. For example, the data quality is a big concern for both official organizations and citizens. Official organizations can, for instance, guide and help citizens to calibrate sensors in order to improve data quality. How local governments react on citizen sensing can be another interesting research topic. In terms of the environmental monitoring, other air pollutants like particulate matter, CO and O<sub>3</sub> need also to be considered.

### **Acknowledgements**

The valuable support of Huib de Bats, Bob Beertema, Mirjam Bekker, Co de Boer, Pieter van Boheemen, Guillem Camprodon Pujol, Tomas Diez, Jimena Gauna, Sybrand Hekking, Christine van den Horn, Wo Meijer, Fokko Visser and other members of Amsterdam Smart Citizens Lab is highly appreciated by the authors.





# Chapter 3

## **Field calibration of electrochemical NO<sub>2</sub> sensors in a citizen science context**

This chapter is based on:

Mijling, B., Jiang, Q., de Jonge, D., and Bocconi, S., 2018. Field calibration of electrochemical NO<sub>2</sub> sensors in a citizen science context. *Atmospheric Measurement Techniques* (accepted).

## Abstract

In many urban areas the population is exposed to elevated levels of air pollution. However, real-time air quality is usually only measured at few locations. These measurements provide a general picture of the state of the air, but they are unable to monitor local differences. New low-cost sensor technology is available for several years now and has the potential to extend the official monitoring network significantly even though the current generation of sensors suffer from various technical issues.

Citizen science experiments based on these sensors must be designed carefully to avoid generation of data which is of poor or even useless quality. This study explores the added value of the 2016 Urban AirQ campaign, which focused on measuring nitrogen dioxide (NO<sub>2</sub>) in Amsterdam, the Netherlands. 16 low-cost air quality sensor devices were built and distributed among volunteers living close to roads with high traffic volume for a two-month measurement period.

Each electrochemical sensor was calibrated in-field next to an air monitoring station during an 8-day period, resulting in  $R^2$  ranging from 0.3 to 0.7. When temperature and relative humidity are included in a multilinear regression approach, the NO<sub>2</sub> accuracy is improved significantly, with  $R^2$  ranging from 0.6 to 0.9. Recalibration after the campaign is crucial, as all sensors show a significant signal drift in the two-month measurement period. The measurement series between the calibration periods can be corrected in hindsight by taking a weighted average of the calibration coefficients.

Validation against an independent air monitoring station shows good agreement. Using our approach, the standard deviation of a typical sensor device for NO<sub>2</sub> measurements was found to be 7  $\mu\text{g m}^{-3}$ , provided that temperatures are below 30°C. Stronger ozone titration at street sides causes an underestimation of NO<sub>2</sub> concentrations, which 75% of the time is less than 2.3  $\mu\text{g m}^{-3}$ .

Our findings show that citizen science campaigns using low-cost sensors based on the current generations of electrochemical NO<sub>2</sub> sensors may provide useful complementary data on local air quality in an urban setting, provided that experiments are properly set up and the data are carefully analysed.

### 3.1 Introduction

Because air pollution is difficult to measure, instrumental and operational costs of official measurement stations are usually high. Air quality networks in cities, if present at all, are therefore usually sparse. Diffusive sampling is a common addition to these real-time measurements and are successfully used to monitor local differences (see e.g. (Cape, 2009)). However, these differences are poorly attributed to an emission source due to the long averaging time of these measurements (usually 4-weekly). Emerging low-cost sensor technology has the potential to extend the official monitoring network significantly, and improve our understanding of local urban air pollution. Miniaturized and affordable sensors potentially enable citizens to measure their environment in more detail in space and time (Kumar et al., 2015). Most commercially available sensors, however, suffer from various technical issues which limit their applicability. Despite their limitations many experiments are done with air quality devices containing these sensors, often by motivated but not necessarily scientifically trained people. Comprehensive calibration and validation of these devices is crucial (see e.g. (Lewis and Edwards, 2016; Lewis et al., 2016)), but often overlooked. The resulting poor data quality is of concern to health authorities, scientists and citizens themselves.

Several studies have been done to explore the performance of low-cost air quality sensors, e.g. Jiao et al. (2016), Duvall et al. (2016); Mead et al. (2013); Moltchanov et al. (2015). For NO<sub>2</sub> monitoring, mostly metal oxide and electrochemical sensors are used (Borrego et al., 2016; Spinelle et al., 2015b; Thompson, 2016). Typical ambient concentrations of NO<sub>2</sub> are at part-per-billion (ppb) level. The main problems encountered in NO<sub>2</sub> sensor evaluations in these real-world environments are low sensitivity, poor selectivity, low precision and accuracy, and drift. Especially metal oxide sensors are not very stable (Spinelle et al., 2015b; Thompson, 2016) and suffer from lower selectivity. Therefore, in this study, we opted for electrochemical sensors to measure NO<sub>2</sub>.

Mead et al. (2013) already noted the strong interference of ozone and other ambient factors in electrochemical NO<sub>2</sub> sensors. The performance can be increased significantly when adding additional measurements of e.g. temperature and humidity in a regression model or neural network, as shown by e.g. Piedrahita et al. (2014), Spinelle et al. (2015b), Masson et al. (2015). Coping with sensor degradation remains a serious issue. Some studies, such as Jiao et al. (2016), include an additional temporal term in their linear regression which improves the predicted NO<sub>2</sub> slightly.

In the following sections we assess the data quality of the 2016 Urban AirQ campaign. As many similar initiatives depending on participating citizens, this campaign was not set up as a strictly controllable scientific experiment such as in the previously mentioned studies. However, we will demonstrate that citizen air quality monitoring using the current generation of electrochemical NO<sub>2</sub> sensors may provide useful data of urban air quality, by using a practical method for field calibration and correcting for sensor degradation in hindsight.

## 3.2 The Urban AirQ project

The Urban AirQ project explores the added value of alternative air quality measurements in the city, by addressing citizens' questions about their local air quality. It focusses on a  $2 \times 1$  km<sup>2</sup> area around Valkenburgerstraat, a primary road in the East-central part of Amsterdam, see Figure 3.1. Its dense traffic causes regular exceedances of the European annual limit value for nitrogen dioxide ( $40 \mu\text{g m}^{-3}$ ).



Figure 3.1 Locations of the sensor devices during the citizen measurement campaign. The green marker indicates the calibration location at GGD Vondelpark. In the circle the location of SD04 and the GGD station at Oude Schans (in red). The location of Valkenburgerstraat is highlighted in yellow.

Two town hall meetings were organized in which residents of this area were invited to raise their concerns about air pollution in their neighborhood and to formulate related research questions. Topics included the relation between traffic density and air pollution, the difference between main roads and side streets, the front side of an apartment compared to its backside, the influence of apartment height, and the influence of cut-through traffic at nighttime. The residents were invited to participate in finding answers to their questions by measuring their outdoor air quality with 16 experimental low-cost sensor devices (labeled SD01 to SD16), built for this purpose by Waag Society.

Measurements were done from June to August 2016. Beforehand, the sensor devices were calibrated using side-by-side measurements next to an official air quality measurement station. With a second calibration period after the campaign, individual sensor drift was assessed and compensated in hindsight.

The Urban AirQ experiment is unique in the sense of the used number of devices, the duration of the experiment, the direct involvement of citizens, and the use of open hardware and generation of open data.

### 3.3 Urban AirQ sensor devices

The concept of the Urban AirQ sensor is building a device with low-cost electronic components which is easy to operate, so citizens can do their own air quality measurements. It builds on the basic design described by Jiang et al. (2016), having an improved power supply, weather resistant housing, WiFi connectivity, and additional sensors for temperature, relative humidity, and particulate matter. The sensor development is part of an open hardware project; detailed technical information can be found at <https://github.com/waagsociety/making-sensor>.

Central is the microcontroller board (Arduino UNO) which handles the reading of the sensors and sends the data to the WiFi module (ESP8266), see Figure 3.2.



Figure 3.2 Hardware modules of a sensor device (left), and the integration in the casing: open (middle) and closed (right).

For  $\text{NO}_2$  measurements, an electrochemical cell is used from Alphasense Ltd (Essex, United Kingdom). The cell contains four electrodes. The target gas,  $\text{NO}_2$ , diffuses through a membrane where it is chemically reduced at the Working Electrode, generating a current signal. This electric current is balanced by an opposite current from the Counter Electrode. The Reference Electrode sets the operating potential of the Working electrode. The sensor also includes an Auxiliary Electrode, which is used to compensate for baseline changes in the sensor. To get full sensor performance, low noise interface electronics is necessary. An individual sensor board with amperometric circuitry, also provided by Alphasense, is used to guarantee a low noise environment and to optimize the sensor resolution at low ppb levels. The sensor signal is read by a 16-bit analog to digital (A/D) converter (ADS1115). Two sensor devices (SD01 and SD02) contain model NO2-B42F for  $\text{NO}_2$  measurements, the other 14 contain the newer NO2-B43F sensor.

12 of the 16 sensor devices are also equipped with a Shinyei PPD42NS sensor in order to measure particulate matter optically. The present paper, however, will focus only on the assessment of the NO<sub>2</sub> measurements. All devices measure internal temperature and relative humidity (RH) with a DHT22 sensor from Aosong Electronics.

The system is supplied with a 7.5V voltage output adapter and a regulator board which generates 5V for the Arduino and the sensors. The microcontroller consumes a 10 mA current (measured). The PM sensor needs up to 80 mA (measured), the NO<sub>2</sub> sensor about 10 mA (measured), and the DHT22 less than 1 mA. The WiFi module peaks periodically to 350 mA when establishing an internet connection.

### 3.3.1 Averaging and filtering

Raw sensor measurements are stored in a central database on a one minute base. However, the calibration analysis is based on hourly averages to enable direct comparison between the ground truth (also provided as hourly values), and to improve the signal to noise ratio.

The NO<sub>2</sub> sensor measurements are done at the Working Electrode ( $S_{WE}$ ) and the Auxiliary Electrode ( $S_{AE}$ ). They are provided as counts from the A/D converter. Sensor readings of temperature and RH are converted according to the indication of the manufacturer to degrees Celsius and percentages respectively.

Raw, hourly averaged, sensor data are shown in Figure 3.3. The spread in temperature and RH displayed in the raw data is partly explained by the sensor-to-sensor variability. By looking at night-time temperatures (to eliminate the effect of local heating by exposure to direct sunlight) we see that the internal sensor temperatures are 2-5°C higher than ambient temperature. The devices are not actively ventilated, which means that the energy dissipation of the electronics influences their internal temperature. The variable position of the temperature sensors with respect to these heat sources further explain the variance in temperature and relative humidity.

Careful filtering is needed before the data can be further processed. We have applied the following rules:

- Raw, minute-based,  $S_{WE}$  and  $S_{AE}$  measurements outside a  $\pm 10\%$  range of their mean value during the entire measuring period are considered outliers. This filters out 0.33% of all measurements. This criterion was used for its simplicity and effectiveness. Note that, due to the large offset in the raw  $S_{WE}$  and  $S_{AE}$  signal, realistic NO<sub>2</sub> peak values are still detectable as the corresponding sensor response is still within a 10% bandwidth.

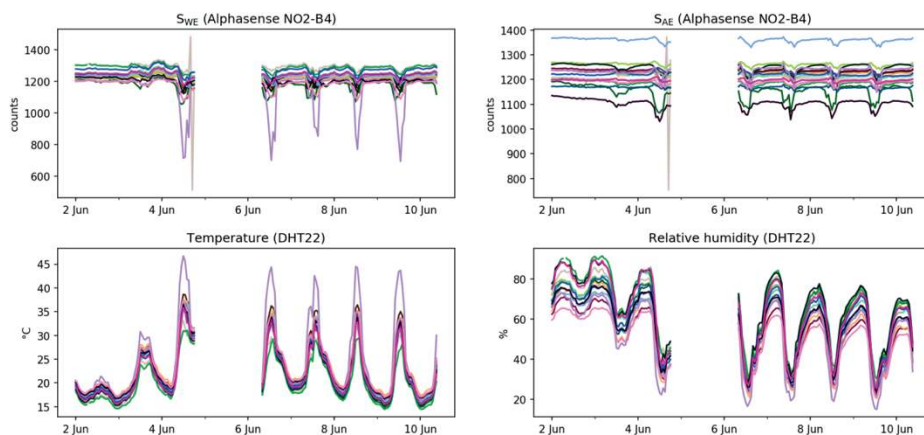


Figure 3.3 Raw sensor data, unfiltered but hourly averaged, from the 16 sensors during the first calibration period, 2-10 June 2016. The data gap around 5 June is due to a connectivity problem to the central database.

- All readings at sensor temperatures above 30°C are discarded to avoid non-linear temperature dependence of the electrochemical NO<sub>2</sub> sensor (see Sect. 4.4). This filters out 4.53% of the measurements during the entire period.
- At least 20 valid minute-based measurements are required to calculate a representative hourly mean. This criterion was found to be a good trade-off between noise reduction by averaging and not losing too many hourly measurements.

During the first calibration period, the sensors were measuring 79% of the time on average. After applying the criteria above, this resulted in 70% valid hourly measurements. During the measurement campaign, the sensors produced 79% valid hourly measurements on average, with the uptime dropping to 50% in places where sensors experienced connectivity problems due to limited range of the participant's WiFi network.

### 3.3.2 Calibration periods

Calibration of the sensors devices have been done by placing the 16 sensors side by side on the rooftop of the air quality station at Vondelpark, operated by the Public Health Service of Amsterdam (GGD). This station is classified as a city background station. It measures nitrogen dioxide, nitrogen monoxide (NO), ozone (O<sub>3</sub>), particulate matter (PM<sub>10</sub>, PM<sub>2.5</sub>, particle number and size distribution), black carbon, and carbon monoxide (CO). For NO and NO<sub>2</sub> measurements, GGD alternates a Teledyne API 200E and a Thermo Electron 42I NO/NO<sub>x</sub> analyser, both based on chemiluminescence. The validated measurements used in this study are considered to be the ground truth. The calibration period spanned several days to be able to test the sensors under a wide range of ambient conditions. To assess the stability of the calibration, the sensors were brought back after the two-month measurement

campaign to the calibration facility for a second calibration period. The Urban AirQ campaign consisted therefore of three phases.

The first field calibration period at GGD Vondelpark station started at 2 June 2016, 00h LT (local time), and ended at 10 June 2016, 10h (8.5 days; 204 hours). Due to connectivity problems sensor data were missing between 4 June 19h and 6 June 9h.

During the following citizen campaign, 15 sensors were distributed among the participants. One sensor (SD03) was kept at the Vondelpark station as a reference. The first sensor was installed and connected at 13 June 2016, 18h, and the last sensor connected at 17 June 2016, 17h. At 15 August 2016, 9h, the first sensor was disconnected, and at 16 August 2016, 18h, the last sensor was disconnected. In this 1537-hour period the devices produced 1204 valid hourly measurements on average.

The second field calibration period at GGD Vondelpark station started at 18 August 2016, 15h, and ended at 29 August 2016, 00h (10.4 days; 249 hours). Due to connectivity problems sensor data were missing between 26 August 12h and 27 August 11h.

Figure 3.4 shows the distribution of temperature, relative humidity, NO<sub>2</sub>, and O<sub>3</sub> during the different periods. Looking at the 75<sup>th</sup> percentile of the distributions, the calibration periods are characterized by higher temperatures and ozone levels than the campaign period. The range of NO<sub>2</sub> concentrations at the Vondelpark station in the calibration periods is larger than in the campaign, reaching more frequently higher NO<sub>2</sub> values. During the campaign the sensors are closer to the GGD station at Oude Schans, where measured NO<sub>2</sub> values are generally a few  $\mu\text{g m}^{-3}$  higher than at Vondelpark. The Oude Schans site does not measure ozone.

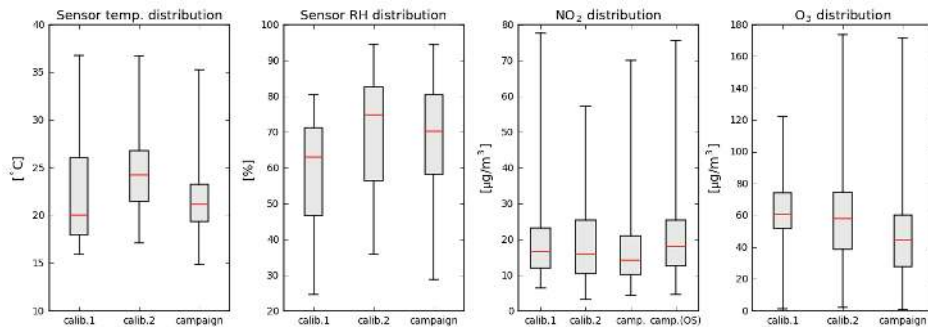


Figure 3.4 Box whisker diagrams of hourly ambient parameters during the two calibration periods and the measurement campaign. The box edges indicate the 25<sup>th</sup> – 75<sup>th</sup> percentile; the whiskers the minimum and maximum values. The median is indicated in red. Temperature and RH are based on the average values of all sensors devices, NO<sub>2</sub> and ozone are taken from the reference station at Vondelpark. For comparison, NO<sub>2</sub> from the reference station at Oude Schans (OS) is also shown.



### 3.4 NO<sub>2</sub> calibration

Electrochemical sensors such as the Alphasense NO<sub>2</sub>-B series are known to be sensitive to interfering species and ambient factors. Especially ozone, temperature, and relative humidity influence the sensor reading (see e.g. Spinelle et al. (2015a)).

#### 3.4.1 Explaining the NO<sub>2</sub> sensor signal

To understand better the behaviour of the NO<sub>2</sub> sensor, we study its sensitivity to different ambient factors. We use the first calibration period to test the correlation of the measured  $S_{WE}$  and  $S_{AE}$  signal with NO<sub>2</sub>, ozone, temperature and humidity by making a best fit through the hourly time series, e.g.

$$S_{WE}(t) = c_0 + c_1 NO_2(t) \quad (1)$$

Temperature and RH were not readily available from the GGD Vondelpark station data. We take temperature and RH from the average readings from the DHT22 sensors instead, which better reflect the internal sensor conditions than ambient air measurements.

Figure 3.5 shows scatter plots for an average performing sensor and the  $R^2$ , the coefficient of determination. The measured  $S_{WE}$  signal can be explained by ambient NO<sub>2</sub> ( $R^2=0.20$ ), but better by its anti-correlation with ozone ( $R^2=0.49$ ). Temperature alone is an even better predictor for the sensor signal ( $R^2=0.73$ ), because of the sensors' direct dependence on temperature, and indirect dependence on temperature (being a reasonable proxy for both NO<sub>2</sub> and O<sub>3</sub> concentrations). Also the correlation with relative humidity is very strong ( $R^2=0.73$ ). The measured  $S_{WE}$  signal can best be explained as a linear combination of NO<sub>2</sub>, O<sub>3</sub>, T, and RH together, resulting in a correlation of 0.98 ( $R^2=0.96$ ).

The  $S_{AE}$  signal is practically insensitive to NO<sub>2</sub>. This suggests that a combination of  $S_{WE}$  and  $S_{AE}$  is more sensitive to NO<sub>2</sub> and less to the other interfering factors, as intended by the manufacturer.

#### 3.4.2 NO<sub>2</sub> calibration models

For NO<sub>2</sub> measurements, the sensor manufacturer suggests to correct both Working Electrode and Auxiliary Electrode for a zero-offset with  $S_{WE,0}$  and  $S_{AE,0}$  respectively. Then a sensitivity constant  $s$  is applied to convert from mV to ppb NO<sub>2</sub>:

$$NO_2[\text{ppb}] = \frac{(S_{WE} - S_{WE,0}) - (S_{AE} - S_{AE,0})}{s} \quad (2)$$

In practice, the factory-supplied constants  $S_{WE,0}$ ,  $S_{AE,0}$ , and  $s$  do not result in realistic values of NO<sub>2</sub>, see e.g. Cross et al. (2017). As an alternative, we propose a linear combination of the signals  $S_{WE}$  and  $S_{AE}$  (calibration model A):

$$NO_2[\mu\text{g m}^{-3}] = c_0 + c_1 S_{WE} + c_2 S_{AE} \quad (3)$$

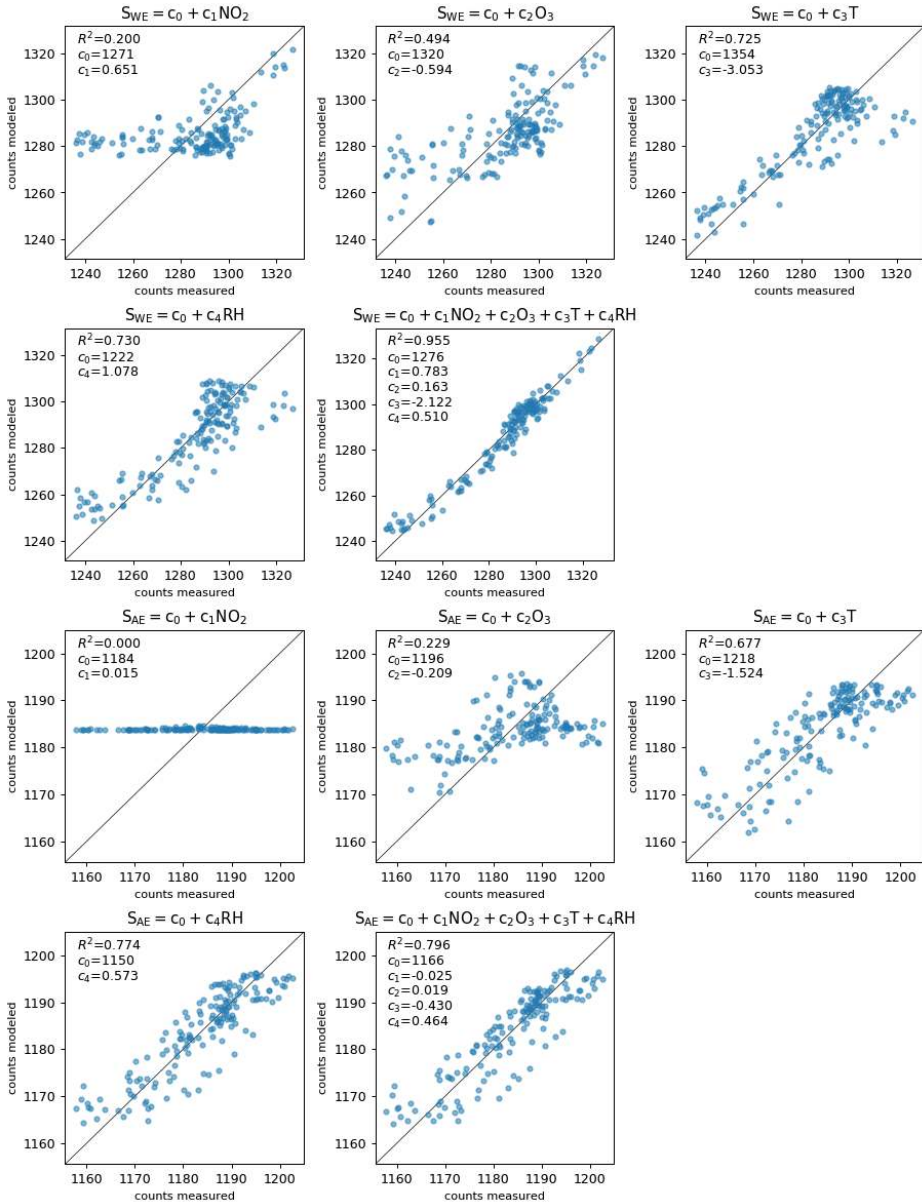


Figure 3.5 The reading of a typical performing NO<sub>2</sub>-B43F sensor (SD10) explained as a linear regression of respectively NO<sub>2</sub>, O<sub>3</sub>, T, RH, and all variables. The top two rows show the results for the Working Electrode; the bottom two rows for the Auxiliary Electrode. The axes represent the A/D converter counts, which are proportional to the currents generated by the sensor at the corresponding electrode.

The coefficients  $c_1$  and  $c_2$  are determined with data from the calibration period using ordinary least squares (OLS). As can be seen from the fit results in Table 3.1, within the batch of sensors there is a large variability of direct sensitivity to ambient  $\text{NO}_2$ .

Table 3.1 Fit results for regression model A.

Sensor ID	$c_0$	$c_1$ ( $S_{WE}$ )	$c_2$ ( $S_{AE}$ )	$R^2$
SD01 <sup>1</sup>	455.4	0.6977	-1.0835	0.47
SD02 <sup>1</sup>	355.9	0.8862	-1.2633	0.62
SD03	-228.6	1.0877	-0.8029	0.72
SD04	-968.2	0.9138	-0.1237	0.69
SD05	-155.1	0.8368	-0.6841	0.48
SD06	-141.9	0.6136	-0.5241	0.44
SD07	-576.4	0.9615	-0.4811	0.57
SD08	231.4	1.0802	-1.2514	0.68
SD09	100.5	0.8669	-0.8952	0.56
SD10	342.0	0.8221	-1.1629	0.50
SD11	338.4	0.9823	-1.2246	0.61
SD12	-375.2	0.7775	-0.4837	0.54
SD13	-1703.4	0.8218	0.5544	0.60
SD14	162.6	0.8156	-0.9075	0.46
SD15	1211.2	0.9008	-1.8984	0.30
SD16	-594.3	0.8007	-0.3192	0.49

<sup>1</sup>: indicates older  $\text{NO}_2$ -B42F sensor type

During the calibration period, hourly ozone values (also taken from the Vondelpark station) happened to be a good proxy for the ambient  $\text{NO}_2$  concentration:

$$\text{NO}_2(t) = 44.6 - 0.40 \cdot \text{O}_3(t) \text{ in } [\mu\text{g m}^{-3}], \text{ with } R^2 \text{ of } 0.49.$$

When compared with Table 3.1, it can be seen that direct sensor readings from a fair part of the sensors cannot outperform this result. To improve the results we use additional measurements and their statistical relation to  $\text{NO}_2$ . We fit different calibration models with multiple linear regression (using OLS). The calibration models which were tested are listed in Table 3.2.

Table 3.2 Regression models for  $\text{NO}_2$  sensor calibration

Model A	$\text{NO}_2 = c_0 + c_1 \cdot S_{WE} + c_2 \cdot S_{AE}$	Linear combination of Working Electrode and Auxiliary Electrode
Model B	$\text{NO}_2 = c_0 + c_1 \cdot S_{WE} + c_2 \cdot S_{AE} + c_4 \cdot \text{RH}$	Relative humidity correction
Model C	$\text{NO}_2 = c_0 + c_1 \cdot S_{WE} + c_2 \cdot S_{AE} + c_3 \cdot T$	Temperature correction
Model D	$\text{NO}_2 = c_0 + c_1 \cdot S_{WE} + c_2 \cdot S_{AE} + c_3 \cdot T + c_4 \cdot \text{RH}$	Temperature and RH correction
Model E	$\text{NO}_2 = c_0 + c_1 \cdot S_{WE} + c_2 \cdot S_{AE} + c_3 \cdot T + c_4 \cdot \text{RH} + c_5 \cdot \text{O}_3$	Correction for temperature, RH, and ozone cross-sensitivity

Temperature and RH are taken from the DHT22 sensor. Note that there is no need to calibrate the individual T and RH sensor signals beforehand; the calibration coefficients for  $\text{NO}_2$  are determined for the specific set of all sensors in the box. However, this means that if an individual sensor is replaced, new calibration parameters for the sensor box have to be derived.

### 3.4.3 Calibration results

A complete overview of the regression coefficients and their error estimates for all models can be found in the Appendix A. The sign of the calibration parameters can be easily understood. As the electrochemical  $\text{NO}_2$  sensor loses sensitivity at higher temperatures (see the negative slope in Figure 3.7(b) for temperatures below  $30^\circ\text{C}$ ), coefficients  $c_3$  are positive to compensate for this effect. The additional sensor response due to cross-sensitivity with ozone is compensated by negative values for  $c_5$ .

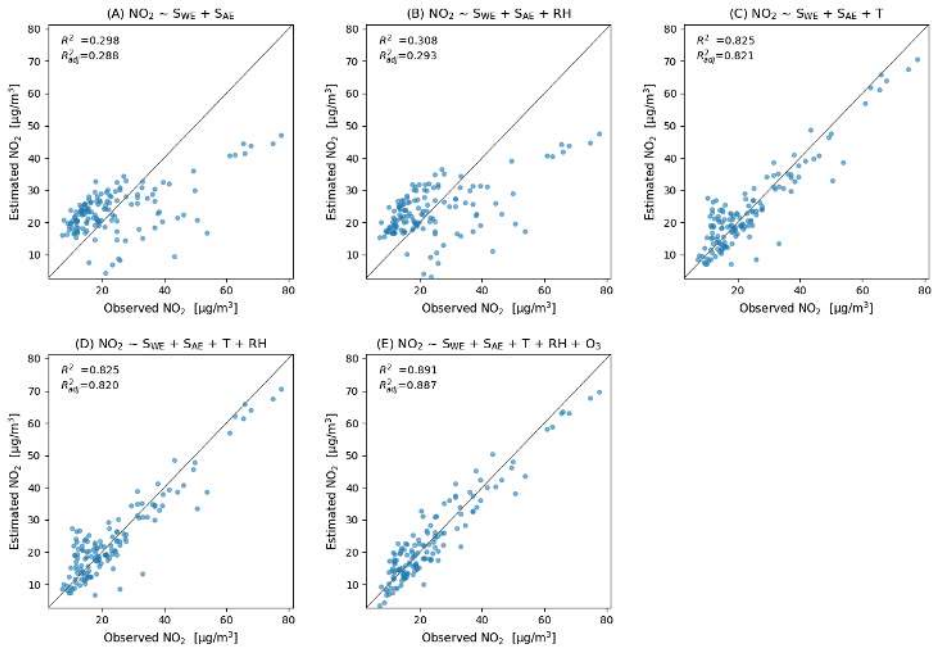


Figure 3.6(a) Calibration model results for an average performing sensor (SD15). Bottom row shows the recommended calibration by Model D (left), and the results when ozone would be included (right).

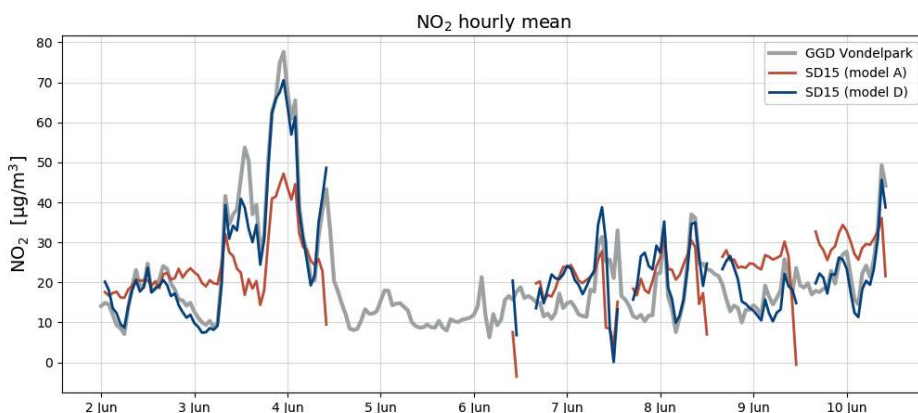


Figure 3.6(b) Time series compared to ground truth with calibration parameters of Model A and D.

From the fit results we see that Model B (including RH) performs better than Model A, but Model C (including T) outperforms Model B. When both RH and T are included (Model D) the results of Model C are marginally improved. This can be understood in terms of a strong sensor dependence on temperature, a weak dependence on RH, and the collinearity between temperature and RH. Note that measuring RH is essential for guarding the data quality of electrochemical sensors, as these sensors are very sensitive to *sudden changes* in RH, see e.g. AAN 110 (2013) and Pang et al. (2017).

The best calibration results (i.e.  $R^2$  values closer to 1) are obtained by including ozone (Model E). The ozone values were obtained from the GGD Vondelpark station, as the sensor devices do not measure ozone themselves.

As local ozone measurements were only available during the calibration periods, we used Model D for the Urban AirQ campaign, i.e. generating an  $\text{NO}_2$  value based on a linear combination of  $S_{WE}$ ,  $S_{AE}$ , T, and RH. The regression analysis of Model D and correlation with the  $\text{NO}_2$  ground truth can be found in Table 3.3.

The two worst performing sensor devices (SD02 and SD01) contain the older  $\text{NO}_2$ -B42F sensor. The newer  $\text{NO}_2$ -B43F model is designed to have higher sensitivity to  $\text{NO}_2$  and less interference of ozone. The old sensor model has indeed smaller coefficients for  $S_{WE}$  and larger correction terms for ozone (see the  $c_1$  and  $c_5$  coefficients of model E in the Supplement). This, however, can also be related to their longer operating time, as both sensors have been used in previous experiments for more than a year. Again, it can be seen that even within the same batch of sensors there is a significant spread in performance, around a median value for  $R^2$  of 0.83. Figure 3.6 shows the results for the different calibration models for the average performing sensor SD15. The time series in Figure 3.6(b) shows clearly how the performance of a typical sensor device improves when temperature and humidity are included in the calibration analysis. The adjusted  $R^2$ , which corrects  $R^2$  for the number of explanatory variables, increases from 0.29 to 0.82. Note that

$R_{\text{adj}}^2$  is only slightly smaller than  $R^2$ , as the number of observations ( $n \approx 150$ ) is relatively high compared to the number of regression variables ( $k = 2 \dots 5$ ).

Table 3.3 Fit results for regression model D.

Sensor ID	$c_0$	$c_1$ (SWE)	$c_2$ (SAE)	$c_3$ (T)	$c_4$ (RH)	$R^2$
SD01 <sup>1</sup>	790.9	0.8707	-1.5645	-0.5051	0.4513	0.62
SD02 <sup>1</sup>	589.2	0.8618	-1.4742	0.2142	0.4204	0.67
SD03	-1272.1	1.2045	-0.1492	1.2690	-0.2944	0.87
SD04	-1613.3	1.1499	0.1818	0.3200	-0.4442	0.85
SD05	-1623.1	1.1235	0.2088	1.7161	-0.4430	0.75
SD06	-824.8	1.1850	-0.5839	1.6737	-0.3069	0.81
SD07	-1217.6	1.1305	-0.1642	1.9435	0.0000	0.79
SD08	-1129.7	1.1835	-0.2705	2.2559	-0.2704	0.86
SD09	-586.3	1.1794	-0.6738	2.0415	-0.2192	0.90
SD10	-1152.7	1.1668	-0.3120	2.9112	-0.2147	0.72
SD11	-1109.8	1.1055	-0.2339	3.3191	-0.1693	0.81
SD12	-1074.9	1.0961	-0.2346	1.4954	-0.2799	0.84
SD13	-1074.6	1.1294	-0.3058	1.8671	-0.1561	0.83
SD14	8.1	1.1860	-1.1889	2.5401	0.0268	0.84
SD15	-104.5	1.8111	-1.7939	4.8373	0.0596	0.83
SD16	-1215.5	1.2551	-0.3038	2.1742	-0.1333	0.84

<sup>1</sup>: indicates older NO2-B42F sensor type

### 3.4.4 Dependency on temperature

Calibrated data without temperature filter show occasionally strong negative values, see Figure 3.7. These negative peaks coincide with internal sensor temperatures exceeding 30 °C. This behavior can be explained from the dependency of the electrochemical sensor on temperature becoming non-linear, see Figure 3.7(b): the sensitivity of the NO2 sensor decreases linearly with temperature up to around 30 degrees, while above 40 degrees the sensor gains sensitivity with rising temperatures. In these regimes, the response of the sensor cannot be described well with our multilinear regression approach. As temperatures during the measurement period only rose occasionally above 30 °C, we decided to filter these measurements out.

### 3.4.5 Startup time

When a sensor device is switched on for service, the electrochemical cell must be stabilized by the potentiostatic circuit which can take a few hours due to the high capacitance of the working electrode (AAN 105-03, 2009). Furthermore, when the sensor is transported to another environment the sudden change in RH causes an equilibrium distortion with a relaxation time of about 2h (Mueller et al., 2017). The startup-effect is translated by the calibration model as a strong positive NO<sub>2</sub> peak,

which should be filtered out. From our sensor data we estimate a stabilization time of 4 hours. Note that this startup effect should not be confused with the response time, which is determined to be less than 2 minutes in (Mead et al., 2013) and (Spinelle et al., 2015a).

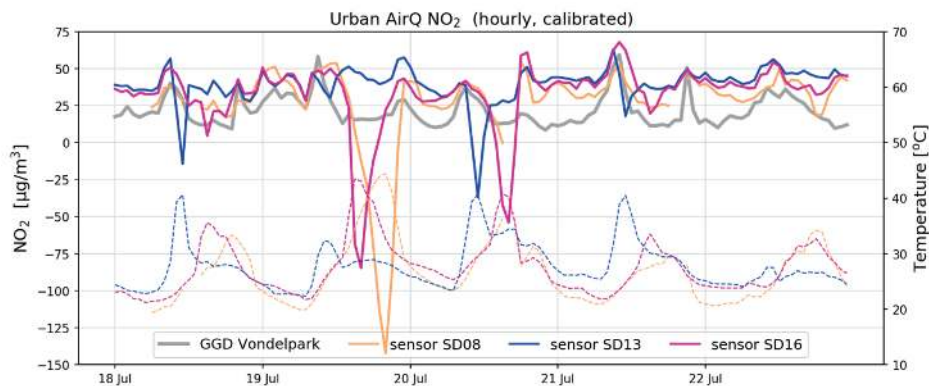


Figure 3.7(a) Examples of negative spikes in the calibrated  $\text{NO}_2$  measurements (solid line) due to internal sensor temperatures (dotted line) exceeding  $30\text{ }^\circ\text{C}$ .

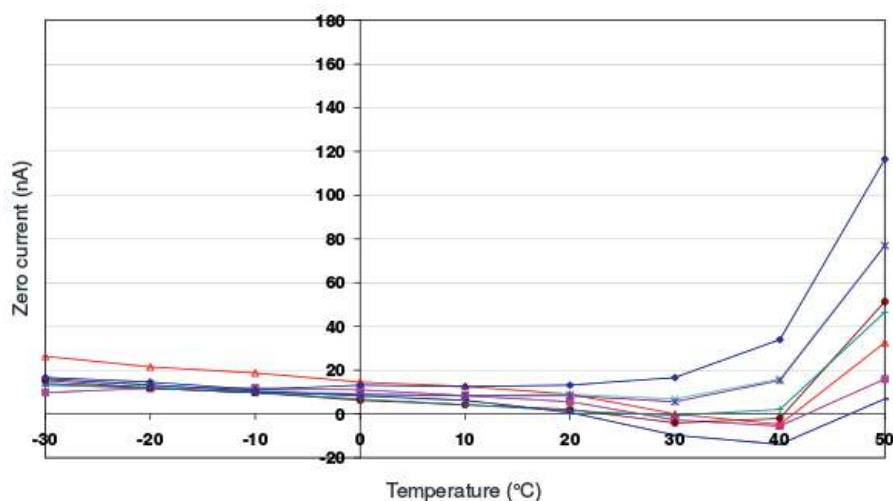


Figure 3.7(b) Variation of zero output of the working electrode caused by changes in temperature for a typical batch of electrochemical sensors. Image taken from Alphasense Data Sheet for  $\text{NO}_2\text{-B43F}$  (ADS, 2017).

### 3.4.6 Predictivity, sensor drift, and uncertainty estimation

Almost all electrochemical sensors have some degree of drift because of aging and poisoning (Di Carlo et al., 2011; Hierlemann and Gutierrez-Osuna, 2008). This becomes a serious complication when the drift is in the order of the strength of the signal of interest. The idea of keeping sensor SD03 next to the reference station

during the whole campaign was to study sensor degradation in more detail. Unfortunately, the sensor was removed temporarily from 10 to 14 July for service, when it was decided to add a PM module to the device. The increased energy dissipation after the modification (the Shinyei PPD42NS module uses a heater resistor to force a convective flow of sampling air) caused an increase of the internal device temperature by 2.5°C on average. This sudden jump in temperature disrupted the reference time series.

Instead, to assess the short-term stability of the calibration model, we use the first 60% of the measurements from the calibration period (2-7 June) to derive the regression coefficients, and predict the NO<sub>2</sub> values for the remaining 40% (8-10 June), see Table 3.4. The average RMSE increases from 6.5 to 7.0 µg m<sup>-3</sup> when the regression is used for prediction.

Table 3.4 Descriptive and short-term predictive error of model D in µg m<sup>-3</sup>

Sensor ID	2-7 June (descriptive)		8-10 June(predictive)	
	Uptime	RMSE	Uptime	RMSE
SD01 <sup>1</sup>	92h	9.25	54h	9.31
SD02 <sup>1</sup>	89h	7.95	53h	13.74
SD03	88h	5.58	53h	4.37
SD04	90h	6.00	54h	4.94
SD05	90h	7.62	53h	8.75
SD06	97h	6.36	57h	5.57
SD07	85h	7.09	52h	6.26
SD08	88h	5.95	52h	6.59
SD09	88h	4.94	52h	3.69
SD10	99h	7.44	59h	8.09
SD11	91h	6.78	53h	5.42
SD12	93h	6.08	52h	5.07
SD13	89h	6.25	54h	5.31
SD14	83h	3.96	48h	14.61
SD15	89h	6.75	52h	4.52
SD16	93h	6.06	55h	5.61

<sup>1</sup>: indicates older NO<sub>2</sub>-B42F sensor type

We assess the long-term stability of the sensors with a second calibration period after measurement campaign, again at the Vondelpark calibration site. As can be seen from the distribution of the residuals in Figure 3.8, most sensors drift significantly in the intermediate two-month period. We describe this degradation effect as a bias  $b$  between the mean of the hourly estimated NO<sub>2</sub> values  $\hat{x}_i$  and the mean of the hourly true NO<sub>2</sub>  $x_i$  during the calibration period:

$$b = \frac{1}{N} \sum_{i=1}^N \hat{x}_i - \frac{1}{N} \sum_{i=1}^N x_i \quad (4)$$

and the root-mean-square error (RMSE) of the difference between the bias corrected calibrated measurement and the ground truth. The latter is the same as the standard deviation of the residuals (SDR)  $\hat{x}_i - x_i$ :



$$\text{SDR} = \sqrt{\frac{1}{N} \sum_i ((\hat{x}_i - b) - x_i)^2} \quad (5)$$

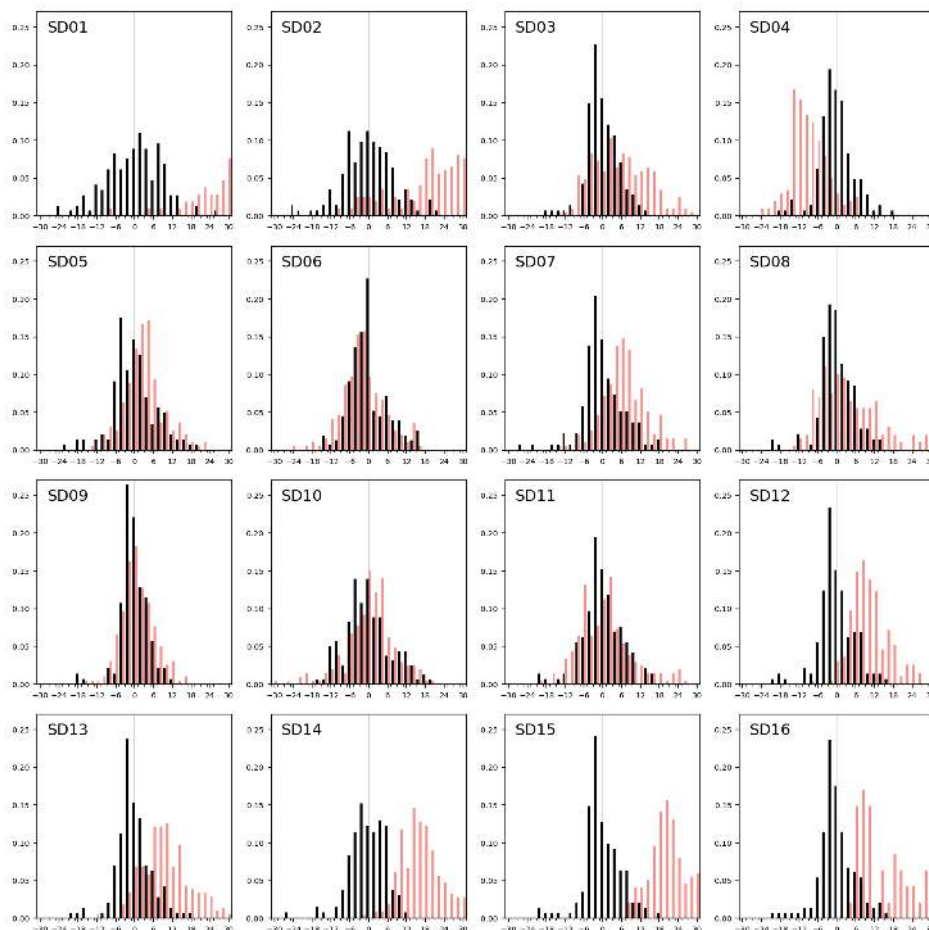


Figure 3.8 Sensor drift during two months of operation, shown as the distribution of residuals (in  $2 \mu\text{g m}^{-3}$  bins) with the reference measurements during the first calibration period (black bars) and during the second period (red bars).

As can be seen in Table 3.5, the bias is mostly positive. Note that sensor SD16 and SD01 had a limited uptime in the second period, which makes their bias and RMS calculation not very representative.

The strongest bias after two months is found for SD02 and SD01. Both are of model NO2-B42F and have been used in others experiments for more than one year. These sensors have also the largest RMSE in the first calibration period (see also Table 3.3), which is another indication of their poor performance. The range in RMSE of the remaining sensors is  $4.5 - 7.2 \mu\text{g m}^{-3}$  for the first period. The bias corrected

RMSE increases to 5.3 – 9.3  $\mu\text{g m}^{-3}$  for the second period. The latter is a more conservative yet more realistic estimation of the precision of the  $\text{NO}_2$  estimates, as they are based on measurements which were not used for calibration. Based on our results listed in the last columns of Table 3.4 and 3.5, we take 7  $\mu\text{g m}^{-3}$  as a typical uncertainty for the estimated  $\text{NO}_2$  values.

Table 3.5 Bias and random error in  $\mu\text{g m}^{-3}$  when calibrated in the first period with model D.

Sensor ID	1 <sup>st</sup> calibration period			2 <sup>nd</sup> calibration period		
	Uptime	Bias	SDR	Uptime	Bias	SDR
SD01 <sup>1</sup>	146h	-0.1	8.8	106h	40.1	18.2
SD02 <sup>1</sup>	142h	0.0	8.2	199h	21.4	12.8
SD03	141h	0.0	5.1	205h	5.6	9.3
SD04	144h	0.0	5.5	202h	-9.2	5.8
SD05	143h	0.0	7.0	192h	3.0	6.3
SD06	154h	0.0	6.0	197h	-2.1	6.8
SD07	137h	0.0	6.6	196h	6.6	6.8
SD08	140h	0.0	5.4	199h	3.1	9.1
SD09	140h	0.0	4.5	196h	0.7	5.3
SD10	158h	0.0	7.2	206h	0.2	7.9
SD11	144h	0.0	6.3	205h	0.5	8.5
SD12	145h	0.0	5.7	194h	10.1	6.0
SD13	143h	0.0	5.8	206h	9.8	7.7
SD14	131h	0.0	5.9	211h	16.6	6.9
SD15	141h	0.0	6.0	198h	21.3	6.8
SD16	148h	0.0	5.7	47h	15.6	8.7

<sup>1</sup>: indicates older  $\text{NO}_2$ -B42F sensor type

The increase of SDR is also due to a loss of sensitivity over time. The aging of the sensors can be further investigated by recalibrating the devices, i.e. determining the coefficients of regression model D, using the data of the second calibration period (see the Appendix A). All calibration coefficients of  $S_{WE}$  (the only component which has direct sensitivity to  $\text{NO}_2$ ) decrease in value, showing that all sensors suffer from sensitivity loss to  $\text{NO}_2$ . This results in lower  $R^2$  values, although the performance loss is partly compensated by the other components in the regression. The older Alphasense models  $\text{NO}_2$ -B42F suffer the largest sensitivity loss, which (although the regression tries to compensate with an increased temperature dependence) result in the worst performance loss in terms of  $R^2$ .

### 3.4.7 Weighted calibration

Taking 18  $\mu\text{g m}^{-3}$  as a typical  $\text{NO}_2$  concentration in an urban environment (Figure 3.4), the sensor drift as listed in Table 3.5 is a significant error component, even after a two-month period. It is impossible to predict the progressing bias for an individual sensor. However, using the second calibration period we can compensate for signal drift in hindsight. If  $\hat{x}_1(t)$  represents the estimated  $\text{NO}_2$  value at time  $t$  based on the first calibration period (starting at  $t_1$ ), and  $\hat{x}_2(t)$  the

estimated NO<sub>2</sub> value based on the second calibration period (ending at  $t_2$ ), the we take for intermediate times  $t_1 \leq t \leq t_2$  a weighted average of both calibrations:

$$\hat{x}(t) = (1 - f(t))\hat{x}_1(t) + f(t)\hat{x}_2(t) \quad (6)$$

Assuming that the sensor degradation is linear in time we select

$$f(t) = (t - t_1)/(t_2 - t_1) \quad (7)$$

such that  $f(t_1)=0$  and  $f(t_2)=1$ .

### 3.4.8 Validation against an independent reference station

Citizen science can be unpredictable, and we were fortunate that sensor SD04 was handed over to an Urban AirQ participant living at Korte Koningsstraat (ground floor), which happens to be 120m from another GGD station at Oude Schans (see Figure 3.1). The Korte Koningsstraat is a side street away from traffic arteries, whereas Oude Schans also classifies as an urban background location. The proximity to a reference station enabled us to perform an independent validation of the sensor measurements, as the calibration of the sensor is based on side-by-side measurements with Vondelpark station, at 3 km distance. As can be seen from Figure 3.9, the sensor readings agree very well with the official measurements. Using the weighted calibration explained in the previous section, the measurement bias largely disappears (Table 3.6). The RMSE ( $5.3 \mu\text{g m}^{-3}$ ) is comparable to the RMSE found during the calibration period. The results give confidence that our calibration method remains valid for similar urban locations, and that our assumption of sensor degradation being linear in time is acceptable.

Table 3.6 Comparison of sensor SD04 with Oude Schans station during the campaign period, according to different calibrations.

	1 <sup>st</sup> calibration	2 <sup>nd</sup> calibration	Weighted calibration
Mean NO <sub>2</sub> , GGD Oude Schans	19.96 $\mu\text{g m}^{-3}$	19.96 $\mu\text{g m}^{-3}$	19.96 $\mu\text{g m}^{-3}$
Mean NO <sub>2</sub> , sensor SD04	17.02 $\mu\text{g m}^{-3}$	22.21 $\mu\text{g m}^{-3}$	19.87 $\mu\text{g m}^{-3}$
Bias	-2.94 $\mu\text{g m}^{-3}$	2.25 $\mu\text{g m}^{-3}$	-0.09 $\mu\text{g m}^{-3}$
RMSE	6.10 $\mu\text{g m}^{-3}$	5.25 $\mu\text{g m}^{-3}$	5.20 $\mu\text{g m}^{-3}$
Correlation	0.89	0.89	0.88

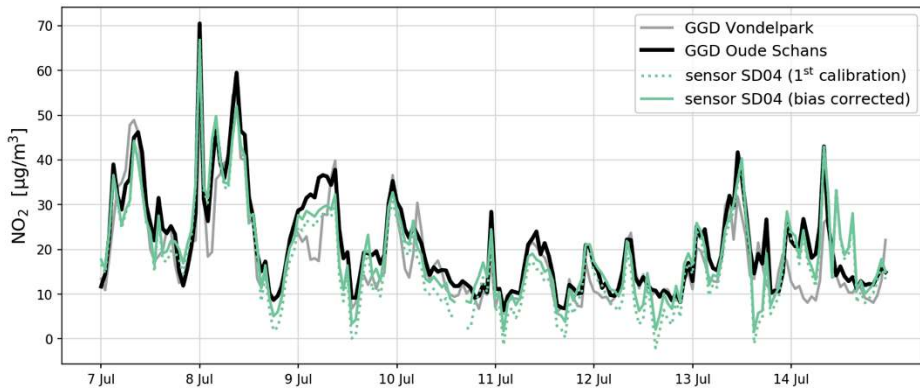


Figure 3.9(a) Comparison of sensor SD04 NO<sub>2</sub> time series with the nearby Oude Schans station (8-day snap shot), and the effect of bias correction. For comparison, measurements of Vondelpark station are also shown.

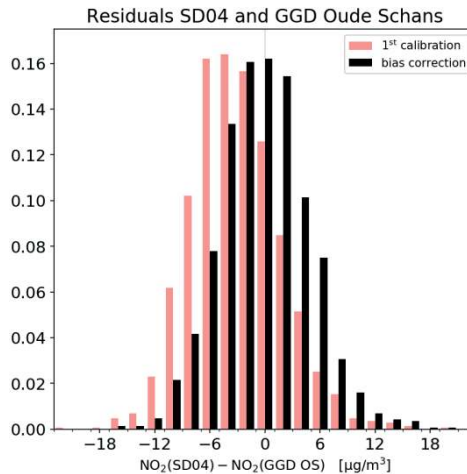


Figure 3.9(b) Distribution of residuals of NO<sub>2</sub> measurements between sensor SD04 and Oude Schans station during the campaign period, with and without bias correction.

### 3.5 Discussion

The Alphasense NO<sub>2</sub>-B4 sensor is used in many low-cost air quality applications for measuring ambient NO<sub>2</sub>. As all electrochemical NO<sub>2</sub> sensors, it is not very selective to the target gas. The sensor response can be explained well by a linear combination of NO<sub>2</sub>, O<sub>3</sub>, temperature and relative humidity signals ( $R^2 \approx 0.9$ ).

As a consequence, a linear combination of the Working Electrode and the Auxiliary Electrode alone give poor indication of ambient NO<sub>2</sub> concentrations. The accuracy varies greatly between different sensors ( $R^2$  between 0.3 and 0.7). For the Urban

AirQ campaign, temperature and relative humidity were included in a multilinear regression approach. The results improve significantly with  $R^2$  values typically around 0.8. This corresponds well with the findings of Jiao et al. (2016), who find an adjusted  $R^2=0.82$  for the best performing electrochemical  $\text{NO}_2$  sensor in their evaluation, when including T and RH.

Best results are obtained by also including ozone measurements in the calibration model:  $R^2$  increases to 0.9. Spinelle et al. (2015b) used a similar regression and found  $R^2$  ranging from 0.35 to 0.77 for 4 electrochemical  $\text{NO}_2$  sensors during a two-week calibration period, but dropping to 0.03—0.08 when applied to a successive 5-month validation period. Low  $\text{NO}_2$  values at their semi-rural site partly explains this poor performance, but most likely also unaccounted effects such as changing sensor sensitivity and signal drift.

The sensor devices were tested in an Amsterdam urban background in summertime, with  $\text{NO}_2$  values ranging from  $3 \mu\text{g m}^{-3}$  to  $78 \mu\text{g m}^{-3}$ , and median values around  $15 \mu\text{g m}^{-3}$ . During the 3-month period most sensors show loss of sensitivity and significant drift, ranging from  $-9$  to  $21 \mu\text{g m}^{-3}$ . After bias correction we found a typical value for the accuracy of the  $\text{NO}_2$  measurements of  $7 \mu\text{g m}^{-3}$ .

This error consists of several components. The reference measurements by the  $\text{NO}/\text{NO}_x$  analysers have an estimated hourly error of 3.65% (certified validation at a  $200 \mu\text{g m}^{-3}$   $\text{NO}_2$  concentration), which would contribute to  $0.5 \mu\text{g m}^{-3}$  under typical conditions. The low-cost DHT22 sensor has a reported error of  $0.5 \text{ }^\circ\text{C}$  for temperature and 2–5% for RH. For a single measurement, this would contribute to a propagated regression error of approximately  $1 \mu\text{g m}^{-3}$  and  $0.5 \mu\text{g m}^{-3}$ , respectively. It should be noted, however, that binning minute-based measurements to hourly averages removes large part of the variability, while determining the best fitting regression model for each sensor device removes large part of the remaining systematical biases. The largest part of the error term is therefore introduced by the linear regression model itself, which does not include all interfering species or meteorological quantities and is not able to describe non-linear dependencies of its variables. One should therefore be careful to extrapolate the calibration model for conditions different than the calibration period.

The validation results from Section 4.8 show that the calibration holds well for urban locations with similar  $\text{NO}_2/\text{O}_3$  ratios. Neglecting  $\text{O}_3$  as regression parameter, however, will introduce a bias at locations with different  $\text{NO}_2/\text{O}_3$  ratios found e.g. closer to emission sources. To get a better understanding of the possible impact, we compared hourly ozone measurements from the GGD authorities at Van Diemenstraat (VDS, classified as street station) against Nieuwendammerdijk (NDD, classified as urban background station) during June-August 2016. The relation can best be described by  $[\text{O}_3]_{\text{VDS}} = 0.87 [\text{O}_3]_{\text{NDD}} + 0.85$  (with 0.93 correlation), which means that ozone levels at the street station are typically 13% lower, due to titration of  $\text{O}_3$  with  $\text{NO}$ . Due to the sensor's cross-sensitivity for ozone, larger values must be subtracted from its signal when the ozone concentration increases. This explains the negative sign of the ozone coefficient  $c_5$  of model E (see Supplement). Calibration with model D will overcorrect (i.e. subtract too much) for locations which have lower ozone concentrations than at the calibration site, resulting in an

underestimation of NO<sub>2</sub> concentrations. Using typical values  $c_5 = -0.3$  and  $[O_3] = 60$   $\mu\text{g}/\text{m}^3$  (75<sup>th</sup> percentile of the distribution during the measurement camping, according to Figure 3.4) we estimate the underestimation of NO<sub>2</sub> at street side as  $0.3 \times 13\% \times 60 = 2.3$   $\mu\text{g}/\text{m}^3$ .

The found sensor accuracy after weighted calibration is good enough to provide some complementary spatial information on local air quality between reference stations. When looking at the difference between Vondelpark station and Oude Schans station (both classified as city background stations) in the period June-August 2016, 22% of the hourly measurements differ more than 7  $\mu\text{g m}^{-3}$ , and 6% of the hourly measurements differ more than 14  $\mu\text{g m}^{-3}$ . These differences increase further when considering road side stations. From this perspective, even sensor devices with an accuracy around 7  $\mu\text{g m}^{-3}$  can contribute to an improved understanding of spatial patterns. However, it must be further investigated if the calibration method used here would provide realistic estimates for peak values (such as the EU hourly limit value, 200  $\mu\text{g m}^{-3}$ ).

### 3.6 Conclusions and outlook

In this study, we examined low-cost electrochemical air quality sensors for citizen urban air quality monitoring. In other words, we evaluated an imperfect air quality sensor in an imperfect scientific experiment. In general, we found that low-cost electrochemical sensors have the potential to complement official environmental monitoring data to help answer questions from the public, which usually cannot be fully answered from official data alone. To reach the potential, however, proper measurement set-up, calibration and recalibration, and data analysis should be guaranteed.

The current generation of low-cost NO<sub>2</sub> sensors has some serious issues which trouble straightforward application. To make electrochemical NO<sub>2</sub> sensor measurements accurate, careful filtering of the raw data is necessary. There is a strong spread in sensor performance, even if the sensors come from the same batch, which make individual calibration essential. A practical calibration method is measuring side-by-side to an air monitoring station. The accuracy of the measurements can be improved by including temperature and humidity measurements from other low-cost sensors in a multilinear regression approach. It is worth noting that more advanced calibration algorithms such as by Cross et al. (2017) and Mueller et al. (2017) could give better results, but this is not the focus of this paper. It is hard to quantify an optimal length of a calibration period without having a proper understanding of the sensor degradation rate beforehand. The measurement period should be at least a few days to capture the sensors behaviour under a wide range of pollution levels and meteorological conditions. Very long calibration periods (in the order of months) will cause sensor degradation issues to interfere with the calibration results.

Startup time of sensors is estimated 4 hours. To avoid nonlinear response of the electrochemical sensor at elevated temperatures, we filter out measurements

above 30 °C. This is not a serious restriction for applicability in moderate climates such as in the Netherlands, provided that the sensor is protected from direct sunlight. However, for warmer regions or during heat waves this may reduce the data stream considerably, unless the temperature dependencies are better captured by more advanced regression models.

The calibration seems to be location independent, as long as the NO<sub>2</sub>/O<sub>3</sub> ratio is comparable. Application at a street side is likely to introduce a small positive bias. Calibration coefficients are not constant in time. During the 3-month period most sensors suffer from significant sensitivity loss and drift. The strongest drift and largest uncertainty are found for the older NO<sub>2</sub>-B42F sensors. It remains unclear if the worse performance is related to the sensor model or the longer usage in field experiments.

The sensor degradation troubles practical applications in operational urban networks. Smart re-calibration programs are essential: bringing back sensors to a calibration facility on a regular basis, or recalibrating on the spot by a travelling reference instrument. New data driven techniques, such as Bayesian networks (e.g. Xiang et al. (2016)), might offer a solution for this problem.

On the hardware side we recommend to include active ventilation to guarantee a constant air flow over the gas sensor and suppresses unwanted internal temperature changes due to heating of electronic components. To improve the NO<sub>2</sub> measurements further we recommend to include an additional low-cost ozone sensor, e.g. Ox-B431 by Alphasense. It is likely that the linear regression approach is able to resolve a significant part of the cross-sensitivity to ozone and NO<sub>2</sub>. The RH sensor signal should be used more cleverly to detect and filter sudden changes in relative humidity. Adding a local data logger is also recommended, to be able to recover data for periods when the WiFi connection to the central database is lost.

## Data availability

A complete overview of fit results for all models can be found in the supplement A. The hourly Urban AirQ sensor data, calibrated in hindsight by interpolating the calibration in time between two calibration periods, can be downloaded at <https://github.com/waagsociety/making-sensor>.

## Acknowledgements

The authors would like to thank Emma Pareschi from Waag Society who was responsible for the hardware development.





# Chapter 4

## **Formal and informal environmental sensing data and integration potential: Perceptions of citizens and experts**

This chapter is based on:

Jiang, Q., Bregt, A.K., Kooistra, L., 2018. Formal and informal environmental sensing data and integration potential: Perceptions of citizens and experts. *Science of The Total Environment* 619–620, 1133-1142.

## **Abstract**

Environmental sensing data provide crucial information for environment-related decision-making. Formal data are provided by official environmental institutes. Beyond those, however, there is a growing body of so-called informal sensing data, which are contributed by citizens using low-cost sensors. How good are these informal data, and how might they be applied, next to formal environmental sensing data? Could both types of sensing data be gainfully integrated? This paper presents the results of an online survey investigating perceptions within citizen science communities, environmental institutes and their networks of formal and informal environmental sensing data. The results show that citizens and experts had different views of formal and informal environmental sensing data, particularly on measurement frequency and the data information provision power. However, there was agreement, too, for example, on the accuracy of formal environmental sensing data. Furthermore, both agreed that the integration of formal and informal environmental sensing data offered potential for improvements on several aspects, particularly spatial coverage, data quantity and measurement frequency. Interestingly, the accuracy of informal environmental sensing data was largely unknown to both experts and citizens. This suggests the need for further investigation of informal environmental sensing data and the potential for its effective integration with formal environmental sensing data, if hurdles like standardisation can be overcome.

---

## 4.1 Introduction

Environmental issues cannot be tackled without environmental data. These data are often produced by official institutions, which provide formal sensing data from which valuable information can be derived about the state of the environment (Artiola et al., 2004; Schnebele et al., 2014b). Nowadays, however, more and more environmental data are being produced by so-called “citizen science”, using low-cost sensors as monitoring instruments. These data are known as informal environmental sensing data (Kamel Boulos et al., 2011; Kooistra et al., 2009).

Though citizens have long been involved in science, the term citizen science is still new and evolving. There is as yet no wide consensus about the definition, as underlined in a paper by members of the European Citizen Science Association (ECSA), the Citizen Science Association (CSA) and the Australian Citizen Science Association (ACSA) (Eitzel et al., 2017). The current paper defines citizen science as including not only communities but also individuals performing scientific activities ranging from posing research questions to finding answers with or without the involvement of professional scientists (Bonney et al., 2016; Haklay, 2013; Lewenstein, 2016; See et al., 2016).

Traditionally, most data collection has been done by professional scientists within projects, based on the questions that these projects posed. To implement projects, scientists usually rely on funding and cooperation, which however, typically stop after funding ends (Jalbert and Kinchy, 2016). Citizen environmental sensing is generally more loosely organised. It involves citizens’ monitoring of the environment using sensors enabled by advancements in information and communication technologies (ICTs) (Kamel Boulos et al., 2011). Examples are smart phones and the “internet of things”. Open source movements are another key aspect in citizen environmental sensing campaigns. These allow citizens to establish networks, or communities, from the local to the global level, and collect data as never done before. Using open hardware and software, citizens can even make their own tools (Carton and Ache, 2017; Hemmi and Graham, 2014; Jiang et al., 2016). These new technologies and movements are being observed by policymakers and scientists as well, spurring them to think differently about how citizen data can be utilised for improved policymaking – for more effective and efficient social impact.

Hemmi and Graham (2014) compared a bottom-up open citizen science with a closed expert-oriented approach in tackling the radiation monitoring problem after the Fukushima Daiichi accident. Their findings emphasise the merit of open citizen science, which was found to be more successful than closed expert-oriented approaches (Hemmi and Graham, 2014). The main reason for this success was the use of open data, as opposed to closed data. Data openness enabled faster evolution of the data creation and use cycle. In this case, open source and open community platforms attracted a diversity of experts and numerous citizens to contribute trustworthy open data at low cost (Hemmi and Graham, 2014). D’Hondt et al. (2013) did a citizen science experiment about noise monitoring and claimed that citizen noise monitoring (informal) can achieve the same accuracy as standard

noise monitoring (formal) if implemented properly. A one-year experiment conducted in Paris by Aumond et al. (2017) concluded also the usefulness of (informal) urban noise measuring using mobile phone. The air pollution monitoring case study conducted in Antwerp, however, considered (informal) mobile monitoring as useful in respect of spatial trend identification but also indicate the challenges of collecting sufficient data and proper data analysis (Van den Bossche et al., 2016).

Bell et al. (2015) conducted research on data quality from citizen weather stations. They observed, however, that significant instrument biases may appear in the data. Analogous research on informal citizen environmental data has been conducted in other fields as well: climate and atmospheric sciences (Muller et al., 2015), air quality (Borrego et al., 2016; Mead et al., 2013; Weissert et al., 2017), water (Little et al., 2016) and noise pollution (Maisonneuve et al., 2010). These studies have discussed or compared various aspects of formal and informal sensing data, for instance, accuracy and accessibility.

Research on experts' and citizens' perceptions of formal and informal sensing data, and possibilities for integrating the two, has been limited up to now. Perceptions, however, here defined as the way "something is regarded, understood, or interpreted" (Oxford Dictionaries, 2017), influence not only the development of environmental sensing technology but also the applications of the data produced. To understand the perceptions from experts and citizens is crucial for citizen environmental science. The experts may not have the same perceptions as described by Minkman et al. (2017) about experts' perceptions on citizen science in water resource management. Furthermore, there might be contrasting perceptions between experts and citizens which lead to conflicts, distrust and tensions rather than collaboration as Weng (2015) concluded in an ecological restoration case study.

For implementation of citizen science projects, especially co-created citizen science, it is important to understand the different perceptions of citizens and experts regarding formal and informal data sources and the potential for integrating the two, as this can raise awareness of obstacles, influence how produced data are used, help to find solutions for problems and mark how changes and shifts happens. The current research investigated perceptions of citizen scientists and experts on formal and informal environmental sensing data and the potential for their integration. To our knowledge, no such study has been done before in terms of the citizen environmental sensing topic and the method.

This chapter is organised in five sections. After this introduction, section 4.2 elaborates on the survey method used. Section 4.3 describes and analyses survey results. Section 4.4 discusses these results, together with other associated research findings and limitations. Finally, section 4.5 presents the conclusions.

---

## 4.2 Method

### 4.2.1 Online survey questionnaire

A survey questionnaire was designed for the purpose of collecting experts' and citizen scientists' perceptions of formal and informal sensing data. The survey was administered online and targeted experts and citizens from selected environmental organisations and citizen communities. Figure 4.1 presents the conceptual design underlying the survey. Thus, citizen and expert respondents answered questions about both formal and informal environmental sensing data. Each also gave their opinions about the potential for integrating the two data types to address environmental concerns.

The perception aspects investigated were derived in part from (Hemmi and Graham, 2014; Lewis and Edwards, 2016; Mead et al., 2013; Muller et al., 2015; Veregin, 1999) (Figure 4.2), and formulated in part by the authors. For instance, according to Veregin (1999), data quality components include accuracy, precision or resolution, consistency and completeness. Accuracy, consistency and completeness were directly selected. However, instead of using precision or resolution, calibration and coverage were used. Due to the particularities of citizen sensing, other aspects selected were calibration, maintenance and training and support (Mead et al., 2013). Social aspects, like trust, privacy and public awareness, were also considered important. For instance, after the Fukushima Daiichi accident, distrust was a main reason why citizens began a citizen environmental sensing community (Hemmi and Graham, 2014).

Of course, there are other important aspects as well, especially in relation to specific disciplines. We did not include more however, to prevent the questionnaire from becoming too complex, particularly for citizen respondents. For instance, the geo-information science and remote sensing community divide data accuracy further into spatial, temporal and thematic categories. Perception aspects range from quality of the data to continuity of data collection.

The perception aspects selected for use in our questionnaire were subsequently developed into questions to create the online survey (Figure 4.2 and Appendix B).

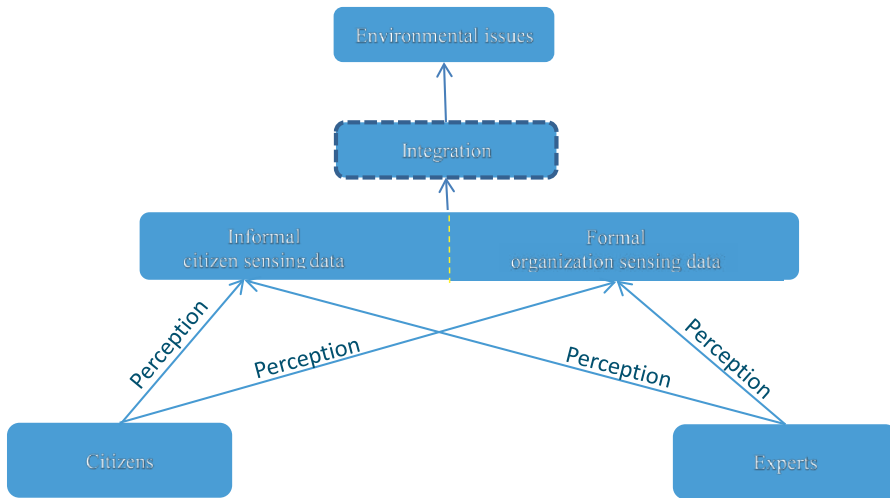


Figure 4.1. Overview of the conceptual design for the perception questionnaire.

## 4.2.2 Questionnaire design

The perception questionnaire consisted of 21 groups of questions including multiple-choice questions, Likert scale questions and semi-open questions. These were grouped into five sections: (1) introduction, (2) formal environmental sensing data, (3) citizen environmental sensing data (informal); (4) integration of formal and informal environmental sensing data; and (5) additional comments, contact information and acknowledgment. An extra item “I don’t know” was added to Likert scale questions. This accounted for the fact that some respondents might not have answers to all questions, as a “neutral” option, representing neither agreement (full agreement) nor disagreement (full disagreement), was not always sufficient.

## 4.2.3 Population and sample

The target population in this research consisted of two groups: experts and citizens. We asked respondents to indicate the group they belonged in and their role and to answer the questions accordingly.

The experts approached were from the Dutch National Institute for Public Health and Environment (RIVM), the Royal Netherlands Meteorological Institute (KNMI) and their networks and the Public Health Service of Amsterdam (GGD).



Figure 4.2. Aspects divided by sections included in questionnaire design.

The citizens approached to participate in the survey were involved in one or more of three communities. The first was the Internet-of-Things Sensemakers. This is an open online and offline community, self-described as “dedicated to creative, advanced and human-centred Internet-of-Things, sensor (networks), electronics/hardware, open hardware/source and hardware startups” (see [www.meetup.com/sensemakersams](http://www.meetup.com/sensemakersams)). As of 18 October 2017, it had 5003 members. The second community was the Amsterdam Smart Citizens Lab. This is an open online and offline urban environmental monitoring community. It brings together Amsterdam-based scientists, hackers, designers and citizens with interest in environmental issues (Jiang et al., 2016). There were 939 members registered on Meetup as of 18 October 2017. The third community was the Smart Emission Community. This was similar to the Amsterdam Smart Citizens Lab, but located in the city of Nijmegen in the Netherlands (Grothe et al., 2016). This, more dedicated community had 37 members.

It is worth noting that there were citizens from the institutes' networks and on the other hand there were experts from citizen science communities as well.

#### **4.2.4 Data collection**

As noted, the survey was conducted through an online questionnaire. Invitation e-mails were sent to experts at the selected organisations and to community members in the three specified citizen science communities. Respondents were given one month to complete the questionnaires, and reminders were sent during the period. Offline communication was used to stimulate responses, that is, researcher joined in community event.

#### **4.2.5 Data analysis**

Data visualisation tools and descriptive statistics were used for preliminary analysis. The data were stored in Microsoft Excel for further analysis and visualisation to derive more detail. For ease of comparison and interpretation, some questions were regrouped or divided up for the analysis. There were no completely open questions except those at the end asking for further comments and respondent contact details. But some of the questions did leave blanks to be filled in. These responses were examined, and important information noted in the results.

### **4.3 Survey results**

#### **4.3.1 Participants**

In total, we received 107 responses. Of these, 26 were incomplete, and thus omitted from the analysis. Of the 81 completed questionnaires, 37 were from experts and 44 were from citizens.

All of the expert respondents had experience in environmental sensing applications. Respondents indicated experience in sensing applications including monitoring of air, noise, soil, vegetation, waste, water and weather (Figure 4.3). Others noted were light monitoring and gasses monitoring. Some 41% of the citizen respondents had no experience with any environmental sensing application.

Respondents' involvement in environmental sensing varied from environmental policymaking to a passing interest (Figure 4.4). Comparing experts' and citizens' involvement, we observe that most experts were policymakers, researchers or scientists. Logically, most citizen respondents categorised themselves as citizen scientists, citizens or just interested. The hardware and software developer and data user categories were almost equally divided between experts and citizens. Journalist and social researcher were among the other attributions mentioned.



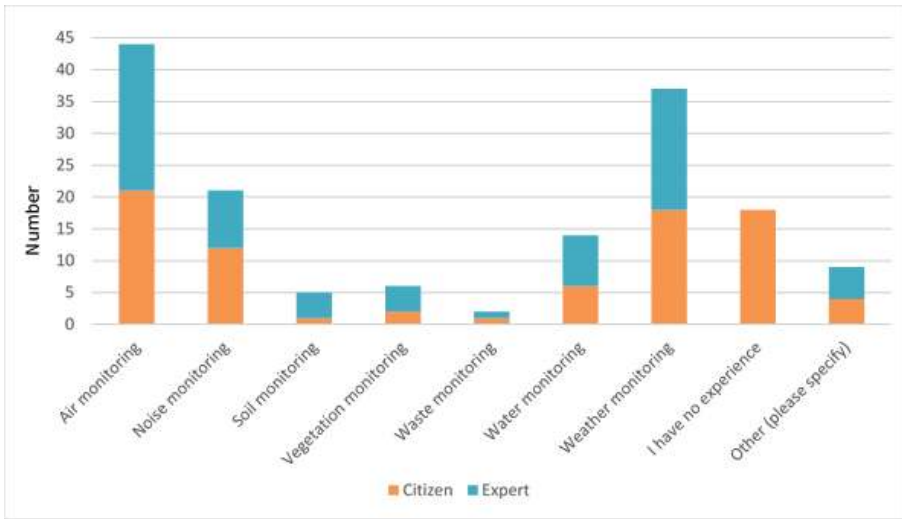


Figure 4.3. Respondents' experience with environmental sensing applications.

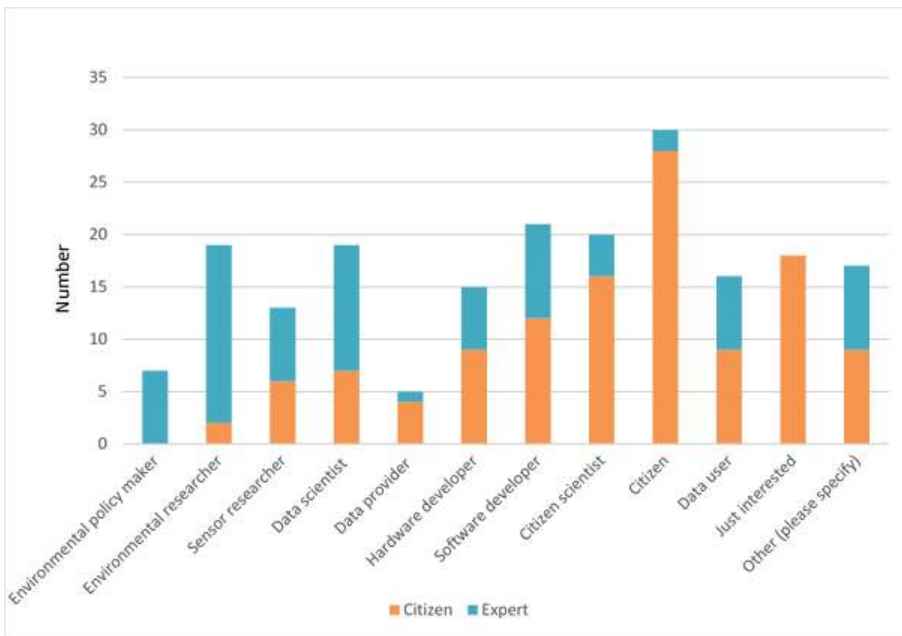


Figure 4.4. Respondents' involvement in environmental sensing.

## 4.3.2 Questionnaire analysis

### 4.3.2.1 Perceptions of formal and informal data

Figure 4.5 presents the complete overview of citizens' and experts' perceptions of formal and informal environmental sensing data. Below the different items are examined in detail.

**Accuracy.** In terms of accuracy, most experts indicated that formal data had sufficient accuracy. Herein, the experts indicated greater confidence in formal data than the citizens. However, negative perceptions of formal sensing data were also more common among experts than citizens. Some 51% of the experts considered informal data as having insufficient accuracy, compared to 25% for citizens. It is worth noting that around one third of both experts and citizens chose a neutral response regarding their perception of the accuracy of informal data. Thus, both experts and citizens seem to consider the accuracy for informal environmental data as largely an unknown.

**Accessibility.** Regarding the item accessibility, only 35% of experts and 18% of citizens indicated that formal data were easily accessible. Both experts and citizens said the accessibility of formal data needed to be improved. In contrast, both were slightly more positive regarding the accessibility of informal data, with experts being relatively more positive. Some 41% of the experts considered informal data easily accessible (Figure 4.5).

**Understanding.** For the formal data, a larger percentage of the experts (41%) compared to the citizens (25%) perceived formal data as easy to understand. Similarly, as indicated in Figure 4.5, 38% of experts considered informal data easy to understand (though 30% of the experts did not agree). Citizens had a relatively negative view on informal data. Some 25% considered it easy to understand, while 25% also did not agree. Some 27% of citizens chose a neutral answer on this question, and 23% chose "I don't know".

**Coverage.** Opinions were similar regarding formal data coverage. As shown in Figure 4.5, 51% experts and 50% citizens indicated that they thought the coverage of formal environmental data to be insufficient. Perceptions of informal data coverage were slightly more positive, especially among experts. Thus, a majority of the experts and citizens agreed that formal environmental sensing data had insufficient coverage, but they were not satisfied with the coverage of informal environmental sensing data either.

**Real time.** Interestingly, as indicated in Figure 4.5, a large proportion of citizens (34%) and experts (27%) indicated that formal data were not timely, though a substantial group of experts disagreed. Informal data, on the other hand, was considered timelier by both experts (59%) and citizens (43%). Thus, experts and citizens seemed to consider informal environmental sensing data more real time than formal environmental sensing data.

---

**Information provision.** A large proportion of both experts and citizens said that neither formal sensing data nor informal sensing data were sufficient for environmental information provision (Figure 4.5). However, it is worth noting that the experts were more positive about the information provision power of formal sensing data. Citizens, on the other hand, were more positive about informal sensing data. Many of the experts surveyed chose a neutral response on the information provision aspect of formal (43%) sensing data. This suggests that almost half of the experts were unsure whether the formal environmental sensing data provided sufficient information for users.

**Consistency.** In terms of consistency, perceptions were generally more positive among both experts and citizens about formal sensing data. Specifically, experts were quite confident about the consistency of formal environmental sensing data, as 70% agreed that the data were consistent. In contrast, they were critical of the consistency of informal environmental sensing data, with 81% disagreeing that the data were consistent.

**Completeness.** In general, formal environmental sensing data were considered to be more complete than informal environmental sensing data. Experts were most positive about the completeness of formal sensing data. Although some 43% of the experts considered formal data to be complete, another 32% of the experts and 39% of the citizens considered it incomplete. Regarding informal data, disagreement on completeness was substantial, with 78% of the experts and 50% of citizens considering it incomplete.

**Frequency.** Both experts and citizens perceived measurement frequency as being sufficient for both formal as well as informal environmental sensing data (Figure 4.5). However, experts and citizens did differ slightly in their overall views. The experts held more positive opinions of the frequency of the formal sensing data, while citizens were more positive about informal sensing data.

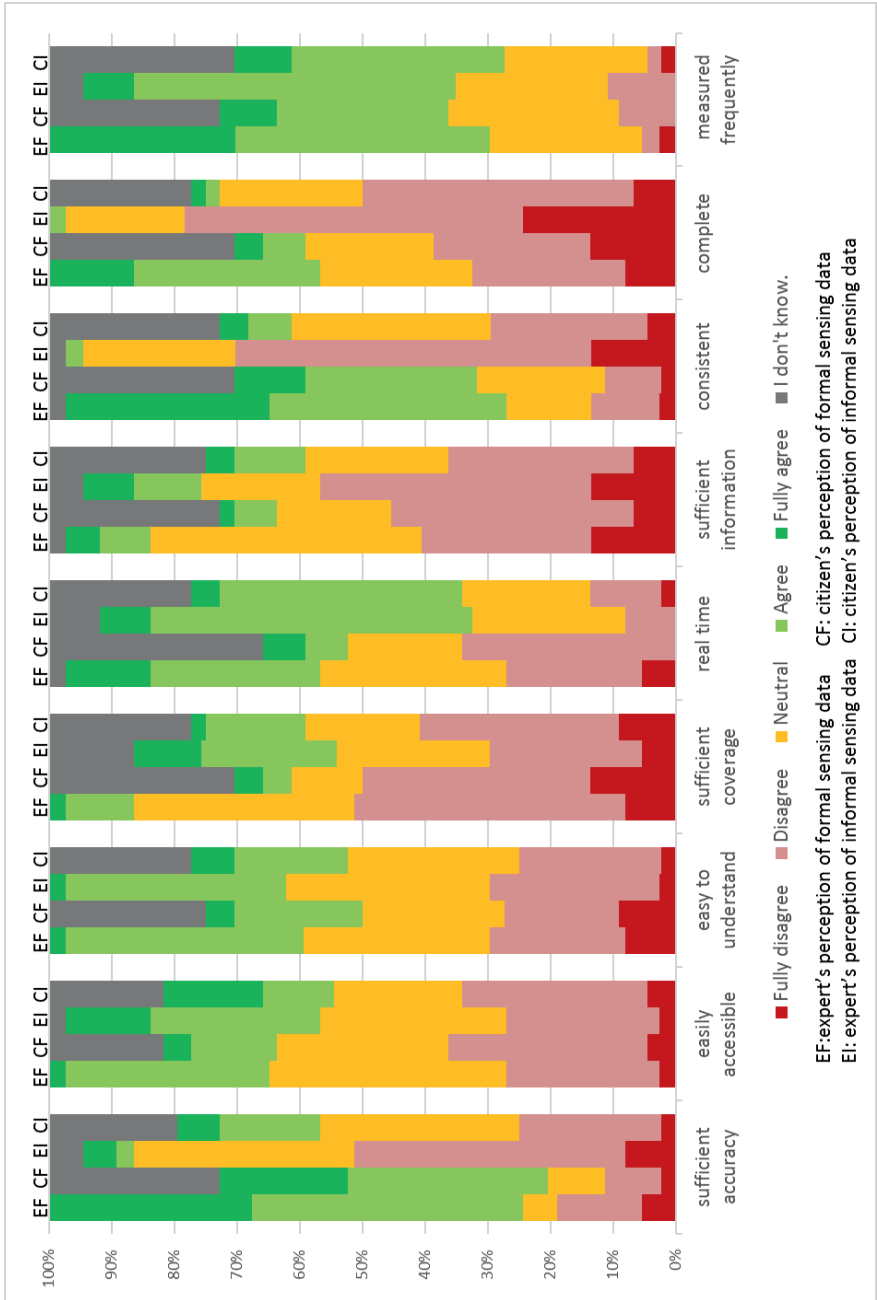


Figure 4.5. Experts' and citizens' perceptions of formal and informal sensing data.

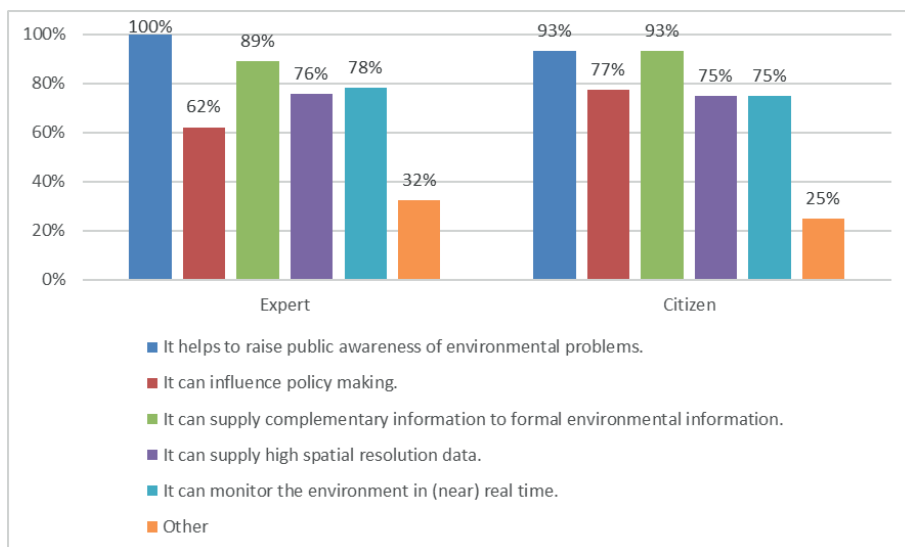


Figure 4.6. Experts' and citizens' perceptions of the value of informal environmental sensing data.

**Value of informal environmental sensing data.** A majority of the citizens and experts surveyed considered it important for citizens to be able to collect environmental data. Only a few experts disagreed with this proposition or chose the neutral response (doi:10.7910/DVN/TP0DYS). The most-acknowledged value of informal environmental sensing data was to raise public awareness of environmental issues (Figure 4.6). In addition, experts indicated its value in complementing formal data, in providing (near) real-time measurements and in improving the spatial resolution of environmental data. Citizens, however, expected these data to have relatively more impact on policymaking. Other values mentioned were citizen empowerment, lower cost, greater coverage and enhanced social bonding. Nevertheless, when it comes to trust in the data, both experts (81%) and citizens (77%) indicated having greater trust in formal environmental sensing data (Figure 4.7). However, compared to experts, citizens trusted informal environmental sensing data relatively more and formal environmental sensing data less. The majority of both experts and citizens indicated that citizen sensing data were underestimated, because it was deemed to be less trustworthy than official data (doi:10.7910/DVN/TP0DYS). Regarding the coordination of citizen sensor networks, despite 35% of experts and 27% of citizens opting for a neutral response on this question, both experts and citizens supported more coordination by official organisations (doi:10.7910/DVN/TP0DYS).

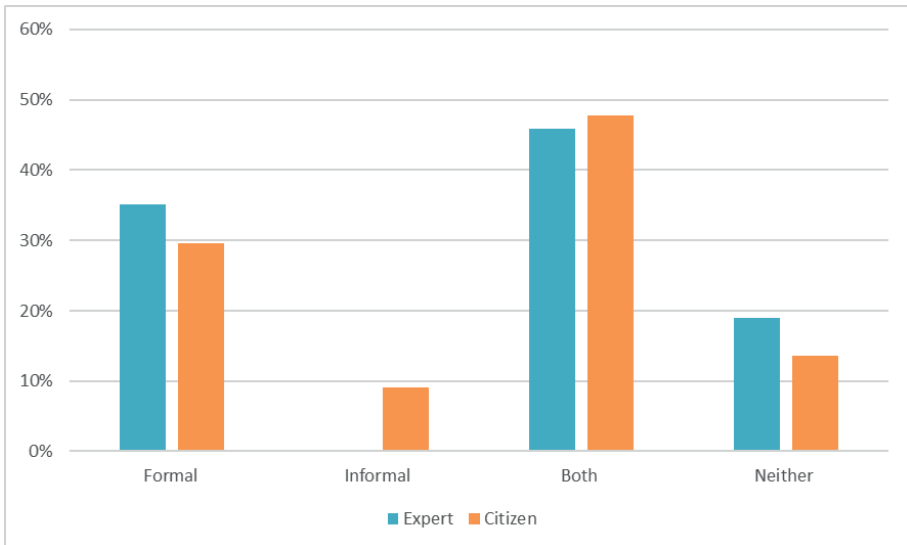


Figure 4.7. Responses to the question: “Which data do you trust?”.

**Challenges and support.** Both experts and citizens indicated that it was difficult for citizens to calibrate (doi:10.7910/DVN/TP0DYS) low-cost sensors. Support from experts was indicated as needed for sensor calibration, proper sensor use and data analysis (doi:10.7910/DVN/TP0DYS).

#### 4.3.2.2 Formal and informal data integration

As shown in Figure 4.8, spatial coverage and data quantity were the top two aspects in which experts and citizens agreed that informal sensing data could complement formal data. About half of the surveyed experts and citizens also indicated measurement frequency and accessibility as aspects in which the two types of data could complement one another. Completeness, data accuracy and consistency were considered the least promising aspects for informal sensing data to complement formal data. In addition, except for the data amount, citizens were overall more positive than experts regarding complementarity in other areas. Some respondents wrote in other aspects in which informal sensing data could complement formal data, for instance, citizen involvement itself and micro-measurements.

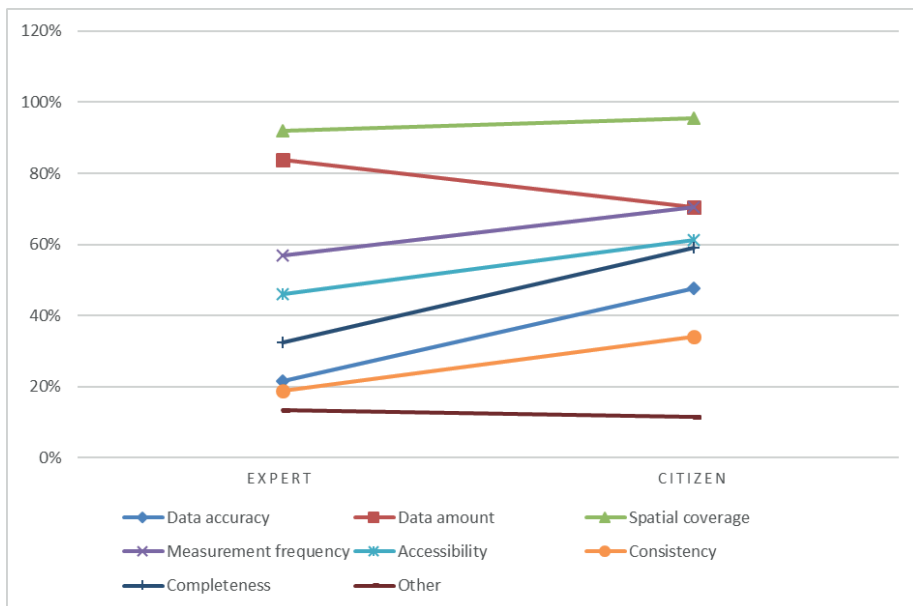


Figure 4.8. Aspects where informal sensing data can complement formal sensing data, according to surveyed experts and citizens.

In terms of integration, large percentages of both experts and citizens indicated that there was room for improvement on all of the items included: spatial coverage, data quantity, measurement frequency and data accuracy (Figure 4.9). Other aspects mentioned as having scope for improvement were completeness, verification and validation of formal data. Standardisation, was the top aspect considered as hindering the integration of informal and formal sensing data. Except that, comparatively, experts were more concerned about accuracy and metadata, while citizens were more concerned with privacy (Figure 4.10). Other aspects mentioned were sensor selection, calibration and validation.

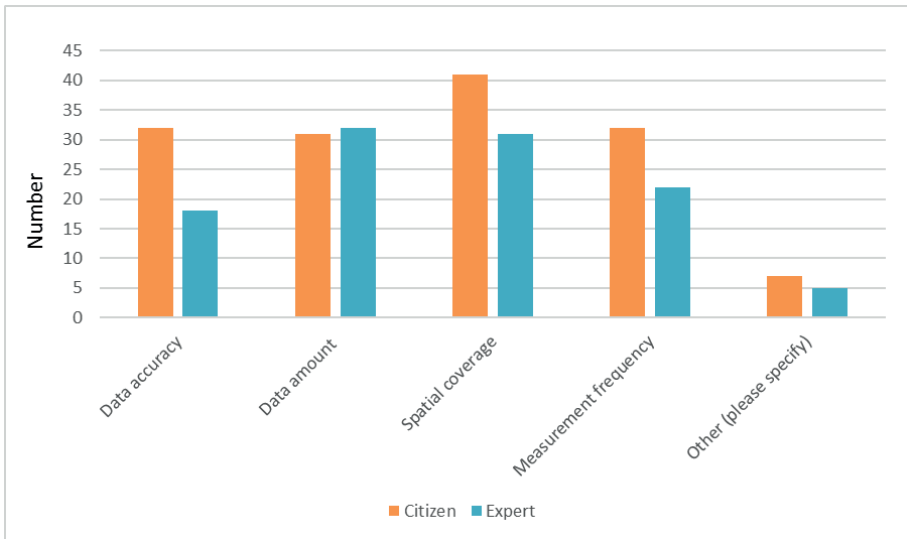


Figure 4.9. Aspects which could be improved by integration of informal and formal sensing data, according to surveyed experts and citizens.

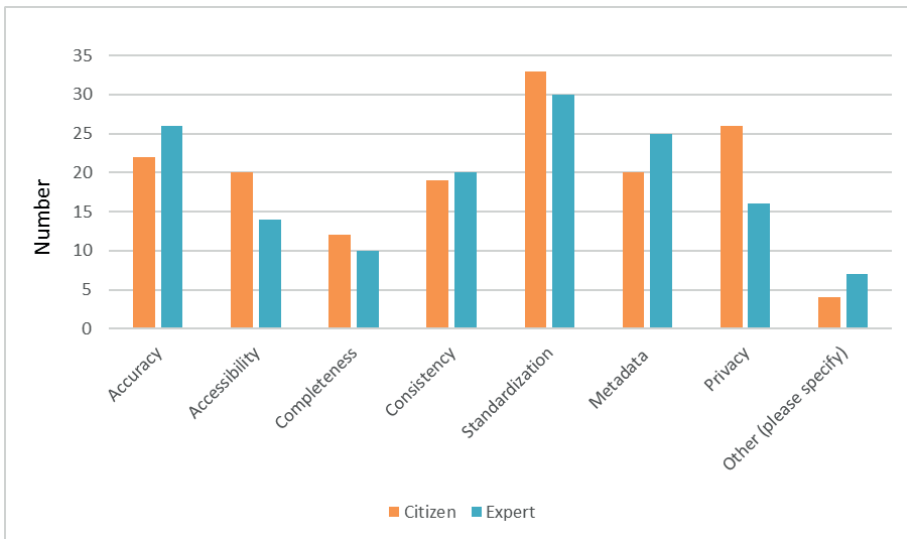


Figure 4.10. Aspects hindering the integration of informal and formal environmental sensing data, according to surveyed experts and citizens.



---

## 4.4 Discussion

Informal environmental sensing is emerging. However, attitudes towards this new socio-technical movement have been shown to be complex. Among researchers, some agree that citizen involvement is important for social aspects like engagement, education and to raise public awareness of environmental issues (Bonney et al., 2009; Jalbert and Kinchy, 2016; Jollymore et al., 2017). Others are concerned more with the technology aspect, that is, the usefulness of the data produced (Bell et al., 2013; Lewis and Edwards, 2016).

Citizens and environmental institutes are both important players in the collection of environmental sensing data. Official institutes produce formal environmental sensing data, while citizens use low-cost sensors to collect informal data. It is therefore crucial to understand the perceptions of experts and citizens regarding both types of data and the potential for integrating the two. We collected these perceptions using an online questionnaire and then analysed the results. The communities we selected were quite diverse but all related to one or more aspects of general sensor technology, citizens environmental sensing or environmental monitoring. For instance, Internet-of-Things Sensemakers community is mainly a broad sensor technology community including professionals but also interested citizens. The Amsterdam Smart Citizens Lab is mainly a community with members who want to measure the air quality, noise, water quality and so on by themselves. This is also indicated in their background as shown in Figure 4.6: expert respondents were mainly environmental policy makers, environmental or sensor researchers and data scientists; citizen respondents were mainly citizen scientists, citizens or just interested; hardware and software developers have large portion in both citizens and experts. The findings and the limitations of the current research are discussed below.

### 4.4.1 Formal and informal environmental sensing data

Formal environmental sensing data have generally been considered expensive and inadequate for capturing the spatial and temporal dynamics required in environmental monitoring. This is due to static and sparse coverage of stations (Kumar et al., 2015; Little et al., 2016). This perception was confirmed in our research. Most experts and citizens considered formal environmental sensing data do not have sufficient coverage.

In terms of citizen science in general, some authors have suggested that citizens are capable of producing data of similar quality to that of experts (Cohn, 2008; Haklay, 2013; Holt et al., 2013). Looking specifically at citizen environmental sensing, emerging technologies give particular power to citizens to conduct scientific measurements that otherwise would not have been possible. However, new technologies bring new challenges as well, like sensor calibration. According to our survey, citizens have high expectations of the training and support they need for sensor calibration, proper sensor use and data analysis. Interestingly, a substantial proportion of respondents in both groups were neutral on the question of informal sensing data accuracy. This indicates that experts shared citizens' uncertainties

regarding the quality of informal environmental sensing data. It suggests that experts and citizens could learn together by helping each other. This is in line with previous findings on public perceptions of citizen science (Lewandowski et al. (2017)).

Regarding other aspects, the majority of the surveyed experts and a big portion of citizens considered informal sensing data to be incomplete and inconsistent. More interestingly, both experts and citizens considered formal sensing data insufficient on coverage and information provision, but they were not satisfied on informal sensing data in these aspects neither. They also considered informal sensing data to have insufficient information provision power. Thus, even though both experts and citizens acknowledged the potential of informal environmental sensing data to complement formal environmental sensing data, for instance, to improve on the currently sparse coverage, the informal environmental sensing data themselves are still viewed as unable to live up to that potential.

Above all, both experts and citizens clearly perceived a shortage of formal environmental sensing data. The gap between the data available, and that which is needed, cannot, yet be filled by informal sensing data, according to both groups surveyed. Experts and citizens furthermore considered the quality of informal sensing data to be largely an unknown. This hinders applications of informal sensing data. According to Little et al. (2016), informal sensing data have the greatest potential to complement formal sensing data in regard to coverage. Our survey, however, indicated that experts and citizens do not consider such better coverage by low-cost sensors to have been achieved as yet. This is understandable, because the utilisation of low-cost sensors for citizen environmental monitoring is still quite new. How informal environmental sensing data provision should be organised, and how it should be maintained to produce complete and consistent data are still open questions. Nevertheless, the history of equipment development suggests that the equipment will become more powerful and easier to operate by citizens over time (Haklay, 2013). We therefore believe that informal environmental sensing will increasingly contribute to environmental monitoring.

The importance and value of informal environmental sensing data was recognised by our respondents, though citizens and experts were not in full agreement regarding issues of trust and how sensor networks should be coordinated. One way to resolve the issue of distrust might be by full documentation of the entire data collection process, including the hardware, software, data, analysis and so on (Hemmi and Graham, 2014). Our survey, furthermore, pointed out the difficulties experienced by citizens in data collection, suggesting that training and support from experts is crucial.

#### **4.4.2 Integration of formal and informal environmental sensing data**

Despite their differing perceptions of formal and informal environmental sensing data, both experts and citizens agreed that informal and formal sensing data could complement each other regarding spatial coverage, data quantity, measurement frequency, accessibility, completeness, accuracy and consistency (Figure 4.8 and

4.9). This is in line with Kumar et al. (2015) who thought informal sensing could identify air pollution hotspots in real time. However, according to our survey, difficulties do remain. Standardisation, accuracy, metadata, privacy, consistency and accessibility are among the primary obstacles hindering the integration of formal and informal environmental sensing data (Figure 4.10). Furthermore, there are more technical challenges during practical combining of diverse data sources including formal and informal ones because of problems such as imperfection, correlation, inconsistency and disparateness of data as described by Khaleghi et al. (2013).

To capitalise on the potential of informal environmental sensing data, the need for standardisation is urgent. Broadly, this would help resolve metadata issues, guarantee accuracy and promote accessibility. For instance, the Open Geospatial Consortium (OGC) standards could be used (Kamel Boulos et al., 2011).

#### **4.4.3 Limitations**

This study administered an online survey to examine citizens' and experts' perceptions of formal and informal environmental sensing data and the potential for integrating the two. Members of citizen science communities, experts at official institutes for environmental monitoring and networks of official institutes including experts and citizens were selected as survey respondents. However, unlike the experts, all of whom were familiar with either citizen science or environmental monitoring, the citizen communities were more diverse. Among those surveyed were members of the Amsterdam Smart Citizens Lab and the Internet-of-Things Sensemakers communities because they are open communities. The survey focused on the Netherlands. Most responses were from Dutch citizens, although some respondents had an international background. It would be relevant to scale up this survey to the European level or among an even broader international group of countries, covering a greater diversity of communities and experts from different disciplines.

#### **4.5 Conclusions**

This chapter examined the perceptions of experts and citizens regarding formal and informal environmental sensing data. These perceptions were collected via a survey questionnaire administered within citizen science communities, official environmental institutes and their networks. The aim was to gain insights on both citizens' and experts' perceptions of informal and formal environmental sensing data and integration potential. The survey indicated areas of agreement but also disagreement between citizens and experts on formal and informal environmental sensing data. Both experts and citizens surveyed were generally in agreement regarding the accuracy and consistency of formal environmental sensing data. Importantly, the accuracy of informal environmental sensing data was largely considered unknown by both experts and citizens. Despite both experts and citizens considered formal and informal sensing data were not sufficient for information provision, citizens, however, were inclined to think that informal

sensing data were more sufficient, while experts considered formal sensing data to be more sufficient. This could explain why citizens initiate various informal environmental sensing data-gathering campaigns and why experts may be concerned or critical of the data produced. More important, however, the survey showed that citizens expected to need training and support in order to produce valuable data.

The experts recognised the value of citizen sensing campaigns and the data produced. They also acknowledged the potential complementarity of formal and informal environmental sensing data. Respondents agreed that by integrating these two data sources, gains could be made mainly in data quantity, spatial coverage, measurement frequency and data accuracy. They also commented on the obstacles for using and integrating informal environmental sensing data. Standardisation was at the top of the list of obstacles, among others.

The respondents' perceptions underscored disagreements, challenges and also expectations, especially on the status of formal and informal environmental sensing data and complementary integration. The insights produced could be valuable for guiding further formal and informal environmental sensing practice and research. Furthermore, our online survey was administered in the Netherlands to selected citizen communities and official environmental institutes including their networks. It would be relevant to adapt the questionnaire to the European or a broader international level and to include more diverse communities and institutes to gain a more complete picture. For instance, the European Citizen Science Association (ECSA), the Citizen Science Association (CSA) and Australian Citizen Science Association (ACSA) could connect citizen science communities globally for a more comprehensive survey.

### **Acknowledgements**

The valuable support of Internet-of-Things Sensemakers, Amsterdam Smart Citizens Lab, Smart Emission Community, the Dutch National Institute for Public Health and Environment (RIVM), the Royal Netherlands Meteorological Institute (KNMI) and the Public Health Service of Amsterdam (GGD) is highly appreciated by the authors.

# Chapter 5

## **Approaches for Formal and Informal Environmental Sensing Data Integration**

This chapter is based on:

Jiang, Q., Bregt, A.K., Kooistra, L., 2018. Approaches for Formal and Informal Environmental Sensing Data Integration. *Sensors* (to be submitted).

## **Abstract**

Citizen environmental sensing data are increasingly produced. These so-called informal data have the potential to provide valuable information. However, how should these data interact with formal data sources? What if we integrate these data with formal sensing data? How can they be integrated? In this paper, we present a conceptual framework for formal and informal environmental sensing data integration. Five approaches are proposed for the integration. To evaluate these approaches, five success indicators were selected: information complement, increased spatial coverage, increased frequency, more timely and increased accuracy. The identified approaches are evaluated by literature study and expert validation. According to the evaluation, formal and informal environmental sensing data integration using the proposed approaches can create additional value for environmental decision making. The proposed framework and integration approaches are applicable. Calibration is currently often used to improve the data accuracy of informal environmental sensing data. Few studies explore the data merge approach but it is considered overall valuable by all experts. To reach the potential of informal environmental sensing data, further study on quality standards by purposes and data merging methods are needed.

## 5.1 Introduction

Environmental problems are global issues. These problems affect human health and could even cause mortality. It was estimated by the World Health Organization that 12.6 million deaths globally were related to environmental issues in 2012 (Prüss-Ustün et al., 2016). To know the status of the environment, finding solutions and making decisions for environmental problems, environmental monitoring is crucial. Due to legal and quality requirements, environmental monitoring networks are usually operated by official organizations with a legislative mandate. The individual monitoring stations are often expensive and labour intensive to operate. As a result, the official monitoring networks, producing formal environmental sensing data, are usually sparse in the number of individual stations. Taking air quality monitoring as an example, there are 12 stations in the city of Amsterdam, the Netherlands, measuring particulate matter (PM10, PM2.5), NO<sub>2</sub>, CO, SO<sub>2</sub>, O<sub>3</sub> and some of them only measure a few air pollutants (<https://www.luchtmeetnet.nl>).

Increasingly citizens are highly interested in monitoring their living environment. Recently, due to the development of sensor technology, more and more low cost, miniaturised and portable sensors are becoming available. These give citizens the opportunities to measure themselves the environment which they are concerned about (Huck et al., 2017; Jovašević-Stojanović et al., 2015; McKercher et al., 2017; Meier et al., 2017), resulting in informal environmental sensing data.

Official monitoring stations are expensive, huge and often sparsely distributed. Low cost, miniaturized sensors operated by citizens are emerging, but their usefulness needs to be investigated. For instance, low-cost air quality sensors have the potential to complement formal data sources (Borrego et al., 2016; Duvall et al., 2016; Kumar et al., 2015). So, can we use those data for decision making? What if we integrate formal and informal sensing data? Can we get more or better information about our environment by integrating those data? A survey about perspectives from experts and citizens on formal and informal environmental sensor data and their integration potential, showed that people have different opinions on formal and informal sensor data, but all think that integration of both can improve environmental information provision in different aspects (Jiang et al., 2018).

Formal and informal sensor data are used with a certain purpose in mind. For instance, if the data are used for a legislation purpose, the data must be collected according to certain standards. For air quality, the Air Quality Directive 2008/50/EC (EU, 2008) is the standard. Informal sensor data may not be suitable, because usually they are not collected according to the official standard (Castell et al., 2017a). On the other hand, in the case that the dynamics of air pollution in cities for personal use is the objective, data accuracy is still important but can be compromised with low budget, mobile and small size devices (Huck et al., 2017; Miskell et al., 2017). For instance, in an indicative use case, cheap, small and portable sensors can be used for personal air pollution exposure monitoring. Even though the data accuracy may be relatively poor, the purpose to capture the dynamics of personal exposure to air pollution might be achieved.

Although several studies present cases, currently to our knowledge, a framework on how to use formal and informal environmental sensing data especially the possibility to integrate the two for environmental information provision is not available.

Therefore the overall objective of this paper is to propose a conceptual framework for formal and informal sensing data integration to support decision making on environmental issues. The application of the integration framework is demonstrated by a case study on air quality monitoring in the city of Amsterdam, the Netherlands.

The proposed framework is presented in section 5.2. A literature review and an expert validation based on air quality sensor data integration are used to evaluate this framework in section 5.3. Section 5.4 discusses the results, while in section 5.5 conclusions are presented.

## **5.2 Proposed integration framework**

The proposed integration framework for formal and informal data is presented in Figure 5.1. The integration framework guides from defining the purpose to the evaluation of the results. The purpose decides which data source is needed: formal, informal or the integration of both. When integration is selected, certain approaches can be applied. The selection of using formal, informal or the integration of both should then be evaluated accordingly. To have a comprehensive view, the framework includes using informal data source or formal data source only. However, this paper will only focus on the integration of formal and informal environmental sensing data.

### **5.2.1 Define purposes**

Before considering integration of formal and informal data sources, the purpose for the environmental issue under consideration should be defined (Figure 5.1). For different purposes, there are different requirements. For example, for regulation and legal issues, often accurate, reliable and complete data are needed. In such a case, the aspect of data quality is most important. On the other hand, for citizens who are interested in their surrounding environment or their personal daily air pollution exposure which they otherwise cannot be obtained from formal data, the aspect of sufficient information is most important. Even though, it is still worth noting that the data should still be at certain data quality level since it influences personal decision making. If integration is needed for a certain purpose, the proper integration approaches can be selected as presented in Figure 5.2. A detailed description per approach is provided in section 5.2.2. The integration result can be validated by success indicators as elaborated in section 5.2.3.



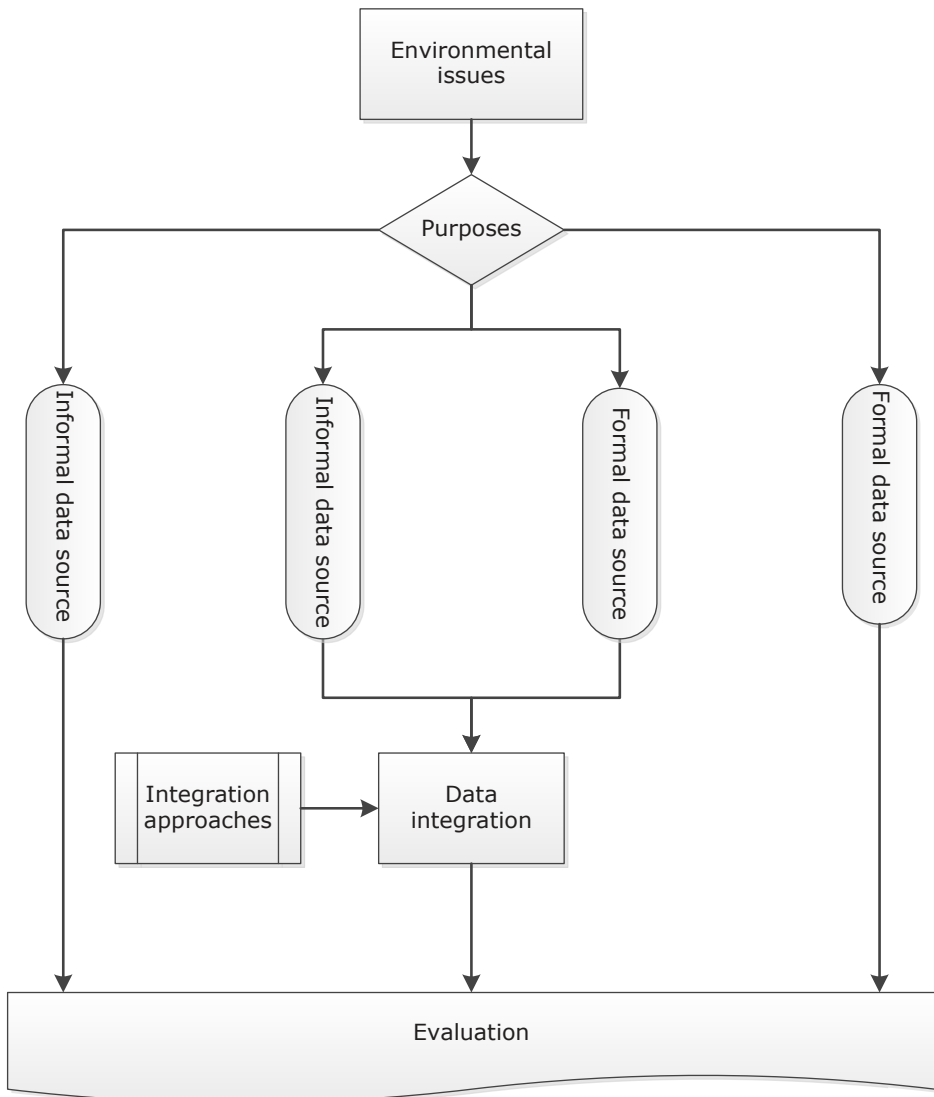


Figure 5.1 General integration framework for formal and informal data sources to support the monitoring of environmental issues

### 5.2.2 Integration approaches

The identified integration approaches (Figure 5.2) are based on the relationship between formal and informal environmental sensing data and the different integration levels (Abdelgawad and Bayoumi, 2012). The complement of formal and informal sources (approach 2 and 3) can be considered as integration at the information level. This means that the data are not merged directly. Only information derived from formal and informal environmental sensing data are of interest. For instance, formal data may give accurate measurements at certain locations of a city but does not provide the street level dynamics due to limited measurement locations. Informal data on the other hand may not give data as accurate as formal measurements, but if there is a large number and wide coverage of informal measurements informal data may give insight into street level environmental dynamics. By integrating the information derived from the two, more comprehensive information can be gained. Merging of formal and informal data (approach 4) is integration at the data level. In addition, since field calibration of informal sensing data needs formal data (approach 1), this is considered as integration too. Using informal data as formal data (approach 5) is practically how informal data are used when formal data are limited or do not even exist. Below, the five identified approaches (Figure 5.2) are described in more detail.

#### Approach 1 : Use formal data to calibrate informal data

In this approach, formal data are used for informal data calibration. Co-locating informal sensors to official environmental monitoring stations for field calibration is crucial to ensure informal sensor data quality (Duvall et al., 2016; Spinelle et al., 2015b).

#### Approach 2 : Use formal data sources to complement informal data

In this approach, formal data will be used to complement informal data at the information level. Formal data normally have higher quality, therefore formal data are considered to be used first to avoid unnecessary repeated measurements when planning informal data collection.

#### Approach 3 : Use informal data to complement formal data

In this approach, informal data will be used to complement formal data at information level. This is mainly the purpose of informal data when formal data is not sufficient to cover the demand for environmental information due to sparse measurements location or low measurement frequency. Informal data can then be used to derive local environmental information timely when this is not possible by formal data source (de Albuquerque et al., 2015; Schnebele et al., 2015).

#### Approach 4 : Merge formal and informal data

In this approach, formal and informal data will be merged to a unified format as a new data source. This is different compared to the approaches 2 and 3. In these approaches, the information from informal and formal data are integrated to derive

more comprehensive understanding of the environment which neither can achieve. But the data are not merged, informal and formal data still have different data accuracy, measurement frequency, format, standards and so on. Integrating by data merge as meant in this approach creates one unified data set. This integration may have lower absolute accuracy but a better prediction of interpolated locations compared to using only a limited number of formal data observations (Mazzoleni et al., 2017).

#### Approach 5 : Use informal data as formal data

In this approach, informal data is the only (or main) data source when there is no (or few) formal data existing locally. The informal data are used as formal data to provide environmental information (Weissert et al., 2017). On a larger scale, these informal data can be integrated with other informal and formal data for national environmental information provision.

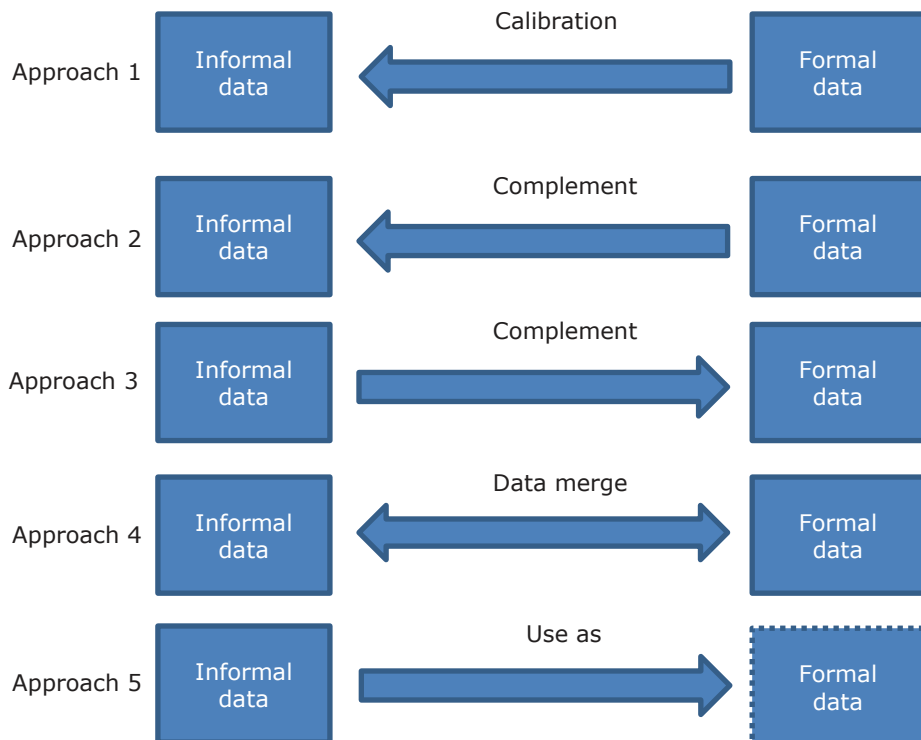


Figure 5.2 Overview of integration approaches for formal and informal sensing data evaluated in this research

### 5.2.3 Validation approaches

In this study, two validation approaches are used to evaluate the proposed integration approaches. The first one is a literature study. In this approach, a selection of formal and informal data integration studies from different environmental domains were reviewed.

Specific keywords were evaluated to search relevant references. However, due to the very diverse terminology for citizen science, as discussed by Eitzel et al. (2017), data integration, formal and informal, it resulted a long and non-exhaustive list of keywords combinations. Therefore, a selected keywords combination was first used to select relevant papers. Then the citation tracking (backward and forward) of the selected relevant papers was used to further search relevant papers. Scopus® was used as the literature database to search references. The keywords "sensor" AND "calibration" AND "citizen", "sensor" AND "citizen" AND "complement", and "sensor" AND "citizen" AND ("data assimilation" OR "data integration" OR "data fusion" OR "data merge"), in article title, abstract, keywords were used to search relevant papers. This resulted in a selection of 75 papers in total. Within these lists of papers, 8 papers were considered relevant and selected for further citation tracking. In total, 7 papers were identified and evaluated by success indicators.

For the second approach, an expert validation is used (Ligtenberg, 2006) which can be considered as a subjective validation. Six experts were selected by the following criteria: 1) the expert works in an official environmental monitoring institute; and 2) the expert is familiar or involved in citizen environmental sensing projects. They were invited to validate each of the proposed approaches for the case study by rating the defined success indicators in five levels from "Strongly disagree" to "Strongly agree". Comments were asked for each indicator of every integration approach. In this validation, an air quality case study (section 5.2.3.1) is used as a concrete example to validate the identified integration approaches. The integration success indicators are presented in more detail in section 5.2.3.2.

#### 5.2.3.1 Case study for expert validation

In the city of Amsterdam in the Netherlands, there are 12 formal air quality monitoring stations operated by the Public Health Service of Amsterdam (GGD Amsterdam) or Dutch National Institute for Public Health and Environment (RIVM). Among them, there are 10 stations measuring NO<sub>2</sub>. The project Urban AirQ was conducted in 2016 with a group of citizens, experts and researchers using 16 low-cost sensor devices. The data produced from these sensors are called informal data source as they were obtained at locations operated by citizens. The sensors were operational from June to August 2016 around two selected streets with relatively high pollution levels. The informal sensors were calibrated by collocating sensor devices to one formal station at Vondelpark before and after the measurement campaign and validated by comparing data from one informal sensor device to a nearby formal station at Oude Schans. More details on the followed calibration and validation procedures can be found in (Mijling et al., 2017). The case study is used as a specific example for expert validation.

### 5.2.3.2 Integration success indicators

In a previous study, the perceptions of experts and citizens on integrating formal and informal sensor sources was evaluated (Jiang et al., 2018). The selected indicators cover different aspects which were indicated as being relevant for evaluating the added value of integration. The following indicators were taken into account in this study:

- **Information complement:** by integration, complementary information can be derived by increased data volume.
- **Increased spatial coverage:** by integration, derived information covers more locations.
- **Increased frequency:** by integration, increased measurements derives information which otherwise not be possible. For instance, only hourly information is available without integration, after integration, information by minute become possible.
- **More timely:** by integration, information is available in a more timely way. For instance, information may only be available days after measuring for one data source, when integrating this data source with more timely data source, information may become available in seconds.
- **Increased accuracy:** by integration, derived information has better accuracy than without integration.

## 5.3 Evaluation of integration approaches

### 5.3.1 Literature study on formal and informal data integration

A total of 7 papers were selected as basis for the literature assessment. The environmental domains are air quality, flood, noise and radiation monitoring. These domains are of high concern to citizens and formal data information is often limited. We studied these applications and evaluated if the five success indicators were applicable and what result could be derived from the cases. The results are presented in Table 1 by grouping the studies according to the five approaches.

As indicated in Table 5.1, we do see formal and informal environmental integration applications fitting the proposed approaches. However, these studies are mainly focusing on validation of usefulness and potential of informal data sources. Less research is conducted on how to reach that potential and how the increasing amount of informal data sources can be used properly for policy making and personal decision making. These studies will be further discussed together with expert validation for each approach in section 5.3.3

Table 5.1 Overview of studies on formal and informal sensing data integration classified according to integration approach.

Integration approach	Domain	Application	Informal data source	Formal data source	Indicators	Reference
Calibration	Air quality	Calibration	Air quality sensors	Official air quality monitoring station	Increased accuracy	(Castell et al., 2017a)
Calibration	Air quality	Calibration	Air quality sensors	Official air quality monitoring station	Increased accuracy	(Hagan et al., 2018)
Calibration	Air quality	Calibration	Air quality sensors	Official air quality monitoring station	Increased accuracy	(Lin et al., 2017)
Informal complements formal	Radiation	Radiation Monitoring	Citizen sensing	Official radiation measurements	Information complement	(Hemmi and Graham, 2014)
Informal complements formal	Flood	Disaster management	Photos	Authoritative data	Information complement Increased coverage Increased frequency	(Schnebele et al., 2014c)
Merge	Flood	Flood information	Crowdsourced data	Official data sources	Increased accuracy	(Mazzoleni et al., 2017)
Informal used as formal	Noise	Official alternative	Participatory sensing	If not available	Increased accuracy	(D'Hondt et al., 2013)

### 5.3.2 Expert validation of integration approaches

In total, five experts assessed the five approaches. The rating from the expert validation is shown in Table 5.2 divided by approach. The mean and standard deviation (SD) for every score is calculated for the score for a range of 1 to 5 to 'strongly disagree' to 'strongly agree'. Except the rating, experts also made comments on approaches or indicators of approaches. These comments are presented in Appendix C2.

### 5.3.3 Evaluation per integration approach

#### 5.3.3.1 Approach 1: Use formal data to calibrate informal data

In this approach, formal environmental sensing data are used to calibrate informal environmental sensing data. As indicated in Table 5.2, all experts strongly agree on the increased accuracy when using formal environmental sensing data to calibrate informal environmental sensing data as this will increase the accuracy of informal sensing data. This is in line with reported studies in literature (Table 5.1). For instance, Castell et al. (2017a) compared laboratory evaluation with field evaluation using commercial low-cost sensor platforms. In the laboratory evaluation, all tested gas measurements (CO, NO, NO<sub>2</sub> and O<sub>3</sub>) had correlations higher than 0.9 with the reference values but in the field evaluation, most

calibration parameters obtained in laboratory changed significantly, and sensors became sensitive to temperature and relative humidity (Castell et al., 2017a). Informal environmental sensing data is used without calibration in some cases due to reasons such as technical complexity, lack of support and facilities. In these cases, informal environmental sensing data may only be used in a limited way, for instance, used as trend indicative measurements. Some informal environmental sensors are calibrated in laboratories in controlled environment which could improve data quality. However, the improvement is still limited since lab environment is often quite different from the real environment. Using formal environmental data as calibration source would improve accuracy of informal environmental data substantially (Lin et al., 2015; Mead et al., 2013; Spinelle et al., 2015b). By integrating formal environmental sensing data under real monitoring conditions, informal environmental sensing data can be calibrated to avoid for instance informal sensing data drift.

Table 5.2 Expert validation of proposed approaches for the case study.

Approach	Indicator	Percentage (%)					Sum	Mean	SD
		Strongly disagree	Disagree	Neutral	Agree	Strongly agree			
Calibration	Increased accuracy	0	0	0	0	100	5	5.00	0.00
Formal complements informal	Information complement	0	20	20	40	20	5	3.60	1.14
	Increased spatial coverage	0	20	40	0	40	5	3.60	1.34
	Increased frequency	0	40	40	20	0	5	2.80	0.84
	More timely	0	40	40	20	0	5	2.80	0.84
	Increased accuracy	0	20	20	20	40	5	3.80	1.30
Informal complements formal	Information complement	0	0	0	60	40	5	4.40	0.55
	Increased spatial coverage	0	0	0	40	60	5	4.60	0.55
	Increased frequency	0	0	60	40	0	5	3.40	0.55
	More timely	0	20	60	0	20	5	3.20	1.10
	Increased accuracy	0	50	50	0	0	4	2.50	0.58
Data merge	Information complement	0	0	0	60	40	5	4.40	0.55
	Increased spatial coverage	0	0	0	20	80	5	4.80	0.45
	Increased frequency	0	0	40	40	20	5	3.80	0.84
	More timely	0	0	80	20	0	5	3.20	0.45
	Increased accuracy	0	0	0	60	40	5	4.40	0.55
Informal used as formal	Information complement	0	25	0	50	25	4	3.75	1.26
	Increased spatial coverage	0	25	0	25	50	4	4.00	1.41
	Increased frequency	0	25	25	50	0	4	3.25	0.96
	More timely	0	25	75	0	0	4	2.75	0.50
	Increased accuracy	33	0	33	33	0	3	2.67	1.53

### **5.3.3.2 Approach 2: Use formal data to complement informal data**

In this approach, formal data are used to complement informal data. According to the experts' validation (Table 5.2), information complement was supported by most experts. The other experts considered formal data has low data volume and therefore disagree on the information complement. As indicated in Table 5.2, the opinion on increased spatial coverage is quite disperse. The increased spatial coverage is considered neutral if official measurements are used together with low-cost sensors (Appendix C2). Most experts chose either 'disagree' or 'neutral' on increased frequency. One reason, according to the comment, is that formal data are available in hourly value. However, the other experts agreed on this. In terms of more timely, there was no agreement. One comment was that formal data are available hourly, sensors will not be much faster (Appendix C2). More than half of the experts agreed on the increased accuracy.

There was no literature found for this approach. However, logically this should be an option. Even though it seems not a hot topic for research, it is expected to be used in practice. In this sense, the integration of formal environmental sensing data could save budget for informal sensor network building. If existing formal environment sensing stations are taking into account to complement informal environmental sensing network, the informal network can be designed in a better way. For instance, in the case study (Mijling et al., 2017), there were two official monitoring stations in the project area. They could be used to complement the informal sensing network.

### **5.3.3.3 Approach 3: Use informal data to complement formal data**

In this approach, informal environmental sensing data is used to fill the formal measurement gaps. Overall, all experts agreed on information complement and increased spatial coverage. They were less positive on increased frequency and more timely. No expert agreed on indicator on increased accuracy (Table 5.2). However, one comment was that informal data can indicate representativeness of formal data and therefore improve overall modelling accuracy (Appendix C2).

Often, formal measurement locations are planned for certain reasons from the legal monitoring purpose perspectives. However, it does not always meet the citizens' concern. In the case study, due to heavy traffic in the selected areas, citizens living around these streets are very much concerned about the air pollution at these locations. They are eager to know more details about the status of air pollution in their living areas. However, two formal environmental monitoring stations in that area supply only limited information. By installing a number of informal sensors in these areas, more detailed information can be gained to monitor the air quality status and dynamics which cannot be supplied by available formal environmental sensing data alone. As studied by Droste et al. (2017), the citizen urban air temperatures monitoring by over 10 million battery temperature readings in São Paulo, Brazil were indicated potentially valuable to complement scarce official weather observations.



#### **5.3.3.4 Approach 4: Merge formal and informal data**

In this approach, formal and informal environmental data are merged to one data set. The overall evaluation from the expert validation was positive (Table 5.2). One comment from an expert was that: “this approach is what we (might mean official environmental organizations from the expert perspective) should do, but integrating not only informal sensing data but also other relevant data sources” (Appendix C2).

When informal environmental sensing data is considered qualified to merge with formal environmental sensing data, the merged data can supply information neither formal nor informal environmental sensing data can do. For instance, the merged data can result in increased measurement locations (spatial coverage) and measurement times.

For this approach, the absolute accuracy might decrease due to the accuracy of informal data. The accuracy depends very much on the accuracy of informal environmental sensing data if we assume the formal measurements always have high accuracy even though it is not always true in all circumstances. However, even if the accuracy of informal data is lower than formal data but has sufficient accuracy, the merged data would still have better overall accuracy or modelling resolution (Mazzoleni et al., 2017) comparing to either sole formal or informal sensing data.

#### **5.3.3.5 Approach 5: Use informal data as formal data**

In this approach, informal environmental data are used as formal data. As can be observed from Table 5.1, there are some studies using this approach. However, experts are quite hesitating. Some experts chose not to rate this approach. One comment was that the informal data are not good enough so far to be used as formal data but it might be the way in some areas and countries if there is no formal alternative available (Appendix C2). A comment from an expert indicated the significance of data accuracy for this approach, it might be used somehow if data accuracy is at a relevant level for the indicated purpose (Appendix C2).

The use of informal data as formal data happens often when no formal data exist or the formal data source does not meet the demands. For instance, after the Fukushima Daiichi nuclear disaster in Fukushima Prefecture, Japan, in 2011, citizens demanded sufficient radiation information which was not enough from the formal measurements. In that situation, citizens initiated informal sensing network which was growing fast and was used as formal information provision (Hemmi and Graham, 2014).

For this approach, even though no formal data is existing, still, as commented by experts, data accuracy which determines the usefulness of the data is crucial to meet the purposes. What’s more, data accuracy is the base for other improvements. Only when the data are confirmed to be useful, the information complement, increased spatial coverage, increased frequency and more timely can be achieved (D’Hondt et al., 2013).

## 5.4 Discussion

In this part, we first discuss the proposed framework and associated integration approaches. Next, data quality and data quantity related to informal sensing are then discussed. Finally, the methods for data merge are discussed.

### 5.4.1 Evaluation of the framework

Informal environmental sensing data are increasingly produced by citizens using affordable and portable sensors which are capable to measure diverse environmental parameters. However, how these data can be used is an important question even though there are increasing studies indicating the potential of these data (Kumar et al., 2015; Moltchanov et al., 2015; Schneider et al., 2017). In this paper, a conceptual framework is proposed to study how these informal sensing data could be used (Figure 5.1). The framework starts from the purposes of using formal data source, informal data source or their integration. As indicated, the purpose decides what would be the requirement that determines which data source should be used. For instance, if education is the purpose, informal sensing data might be sufficient for learning. But for legislation purposes, formal sensing data might be required. Five integration approaches are proposed based on the relationship of the two data sources (Abdelgawad and Bayoumi, 2012) and evaluated using literature study and expert validation. The first approach is using formal data to calibrate informal data. As studied by many researchers, Borrego et al. (2016) and Spinelle et al. (2015b) for instance, and evaluated by experts in this study, this approach can increase the accuracy of informal sensing data. Questions remain on: What if these data have similar quality as formal data? Or if the data quality of informal data is slightly lower than formal data? Can they be used for policy making or personal decision making? In other words, what would be the requirement or standard for informal environmental sensing data to reach the potential mentioned in an increasing number of studies or have the impact it may have. This is an import question since many scientists and politicians are talking about the motivation of citizens to participate in citizen sensing projects. Citizens will not be motivated when told the potential on the one hand but warned it should not be used on the other hand. Or without knowing how useful these data are and how these data can be used. No literature was found about Approach 2, but logically it is valid as an approach and it is likely to be used in practice. Approach 3 gives another way on how informal sensing data can be integrated with formal data to enhance the spatial coverage and complement information, but the data accuracy will not be improved according to expert validation. Approach 4 introduces the data merge of formal and informal environmental sensing data. According to the expert validation, it will improve all aspects including accuracy. This is interesting, based on the literature study (Table 5.1) and comments from experts (Appendix C2), the accuracy increased because informal environmental sensing data can supply extra valuable information for modelling which otherwise would result in lower accuracy (de Albuquerque et al., 2015; Mazzoleni et al., 2017; Restrepo-Estrada et al., 2018). This indicates that informal environmental sensing data can be used by merging with formal data to derive information with higher quality for better decision

making. Approach 5 is even more interesting, from the literature, this approach is often applied for emergency or natural disasters when formal data are limited or not available. Some experts recognized these applications but considered informal data are far from been accepted to be used as formal data. Nevertheless, the value of informal sensing data should be recognized in those circumstances. Therefore, it is essential to make standards for quality assurance to make sure when these data are used, their potential can be achieved, and their drawback can be reduced as much as possible for decision making. For instance, the Air Sensor Guidebook (Williams et al., 2014) classifies the purposes and suggests accuracy requirement according which guide citizens how to use informal data for certain purpose.

An expert validation is used to evaluate the formal and informal environmental sensing data integration approaches. Experts do have expertise to evaluate the approaches, however, as indicated in (Jiang et al., 2018), experts may still have biased opinion. What is more, the validation was based on an air quality case study, it may not be representative for other domains.

#### **5.4.2 Data quality and big data**

According to the validation of formal and informal environmental sensing data integration approaches, it was clear that data quality of informal environmental sensing data should be known. This determines if the data are valuable for data integration. Therefore, recognized quality standards are needed. It is not yet possible for informal data to have the same quality as official data. Even though, it is likely that the quality of informal from sensors will increase. However, we cannot wait till the informal data have the same quality as formal data. . More studies should be conducted to explore the data quality requirement for different purposes. The government or policy makers should support and guide the informal data collection and applications. United States Environmental Protection Agency (EPA) gives a good example on this by giving Air Sensor Toolbox for citizen scientists, researchers and developers (<https://www.epa.gov/air-sensor-toolbox>).

Some experts are concerned about the data quality which on the one hand stimulates the development of sensor technology, and on the other hand neglect the value of the increasing amount and diversity of informal sensor data. Some experts or official organizations may see the potential of informal sensor data but are still hesitating to use or integrate those data with formal data sources. For example, the Weather Observations Website (WOW) (<http://wow.metoffice.gov.uk>) is coordinated by the UK Met Office and by Dutch national weather service the Royal Netherlands Meteorological Institute (KNMI) for a Dutch version (<https://wow.knmi.nl>). The data quality and potential for integration were studied and explored (Bell et al., 2013; Bell et al., 2015; Merkus, 2016). Such research should be encouraged, also for other environmental aspects such as air quality, water quality, and soil quality. Data quality is indeed crucial and should never be ignored, but in the coming big sensor data era, the informal sensor data can complement the sparse formal data substantially in spatial and temporal aspects. By integrating formal data and informal sensor data using the approaches

described in this paper, the integrated data can complement each other to provide more comprehensive information.

### **5.4.3 Methods for data merge**

In this chapter, we presented a conceptual framework for utilizing informal environmental sensing data. Apart from validation of the data quality as discussed in 4.1, another important aspect is the use of data merge methods. Geostatistical methods like kriging might give good interpolation for spatial data prediction (Li and Heap, 2014). Furthermore, since air quality is highly related to for example weather conditions, transportation system and land use, data sources for these parameters can help to build more complex models for better prediction (Briggs et al., 1997; Donnelly et al., 2016). To build these kind of models, machine learning algorithms might be used for calibration and to find the relationships (Esposito et al., 2016; Zheng, 2015; Zheng et al., 2013). For instance Schneider et al. (2017) and Castell et al. (2017b) integrated the informal air quality sensing data into an emission model. It proved the added value for a more timely air quality prediction. If formal data with high quality also could be integrated in this model, it may further improve the model performance. We need to think how to efficiently and effectively use formal and informal environmental sensing data with different characteristics to provide the environmental information to meet diverse demands and purposes.

## **5.5 Conclusions**

In this chapter, a framework is presented for the use of formal and informal sensing data for environmental decision making. At the most generic level the framework consists of 1) using only formal data, 2) using only informal data and 3) using an integration of formal and informal data. The last level was identified as a research gap.

Five conceptual approaches for formal and informal environmental sensing data integration have been proposed and were evaluated by experts and a literature study. It is concluded that the approach which is using formal data to calibrate informal data, was considered an important approach to increase the accuracy of informal data. The approach which is using formal data to complement informal data, may not be studied often but can be used in practice. The approach using informal data to complement formal data, was considered positively to information complement, increased spatial coverage, increased frequency and more timely, but less likely the increased accuracy. The approach which is merging formal and informal data, was considered as a goal to use information environmental sensing data and was positively rated for all five indicators. Finally, the approach which is using informal data as formal data, was partly controversial, the data quality was a concern but also considered to be useful when there is no alternative. Overall, the proposed framework is applicable to guide formal and informal environmental sensing data applications. The proposed integration approaches are valuable to improve data quality. Apart from calibration, more attention should be given to how to use these informal data. Establishing quality standards according to different

purposes is important. Best practices can also be used as guidelines. More research should be done to study and evaluate the methods on merging formal and informal data to reach the potential of informal environmental sensing data.

### **Acknowledgements**

The valuable support of the experts who made the expert validation is highly appreciated by the authors.



# **Chapter 6**

## **Synthesis**

## **6.1 Main results**

The overall objective of this thesis is to assess opportunities for informal sensor data and their integration with formal sensor data to improve environmental information provision by investigating challenges and obstacles in not only technical but also societal aspects. According to this objective, four research questions were identified and covered in chapter 2 to 5 respectively. In this section, findings for these questions are summarized and discussed. In chapter 1, I introduced the context of this research by a short story indicating the problem citizens experience when they require environmental information to decide on a healthy route for daily commuting. According to the findings in this thesis, citizens can use low-cost sensors to help them gain information about environmental issues when sensors are calibrated and used properly even though this is still challenging. Citizens and experts do have different perceptions on both formal and informal data. It indicates that application of informal environmental sensing does not only encounter technical but also societal issues. Positively, even though the accuracy of informal environmental sensing data is largely unknown for both citizens and experts, both citizens and experts consider that informal environmental sensing data can complement formal ones and integrating them can improve environmental data provision in aspects such as spatial-temporal resolution.

### **6.1.1 How does a bottom-up citizen science project develop, and can it contribute to environmental monitoring?**

Environmental information is demanded for policy-making, public service and personal decision-making. However, the information from formal environmental monitoring institutes is limited due to the cost of monitoring facilities and labour. On the other hand, citizens have been contributing to data collection in various domains such as ecology (Dickinson et al., 2010), astronomy (Schwamb et al., 2013) and agriculture (Beza et al., 2017). In terms of environmental monitoring, thanks to information and communication technology and especially the sensor development, citizens are able to measure aspects of the environment by themselves. The types of sensors to measure and monitor various environmental phenomenon are increasing, while at the same time the price and the size of sensors are decreasing. Sensors allow citizens to make measurements themselves on the environment that they are concerned about. With the open hardware and open software movements, described in chapter 2, citizens can even make or create their own sensing tools. The data produced from this citizen science movement can potentially not only answer the questions of the citizens regarding their direct surrounding but can also complement the formal environmental data for policy making and public services.

To find out how citizen environmental sensing should be organized, we started the Amsterdam Smart Citizens Lab. A framework for a bottom-up sensor development approach was proposed which involved citizens in each step from raising questions or concerns to developing strategies, making tools, measuring, data analysis and interpretation, and finally decision making or action (Figure 2.1).



---

As part of this approach, several groups were exploring different environmental issues. One group of persons including citizens, air quality experts, makers and researcher focused on the air quality issue and used this as a case study. According to experiments by Mead et al. (2013), electrochemical sensors seem promising and these were therefore selected. Subsequently, we made the sensing tools and conducted sensing experiments. According to the experiments, the selected electrochemical sensors potentially can be used for urban air quality monitoring. However, proper hardware and software design, sensor calibration and data analysis are essential for bottom-up citizen science project, which develop sensing tools by the community. As also argued by Hubbell et al. (2018), I therefore recommend broad open cooperation involving wide backgrounds, interests and expertise especially for bottom-up citizen environmental science approaches and using digital tools like Meetup (<https://www.meetup.com>) to organize and stimulate sustainable and growing citizen environmental science communities. For a sustainable community, motivation is also an important aspect, environmental concerns are usually a key driver for motivation as is shown by for example the Groningen earthquake monitor and the Airplane Monitor Schiphol communities in the Netherlands where citizens are eager to collect their own data to protect their local environment (Carton and Ache, 2017).

Based on the approach described in Chapter 2, it is demonstrated that citizens are able to design and implement local sensor networks for air quality monitoring, but how good are their sensed data? Calibration of sensors is essential for knowing the quality of citizen sensing data.

### **6.1.2 How to calibrate low cost air quality sensor data, and how valuable are these data?**

According to the experience of previous experiments (Mead et al., 2013), the sensor calibration appears to be of major importance for reliable data. Therefore, in Chapter 3 we focused on field calibration of informal sensor devices. Continuing with the air quality experiments, a follow-up project called Urban AirQ was started. The overall goal of this project was to evaluate low-cost air quality sensors for urban environment monitoring by citizens. In this project, citizens of the two most polluted streets in Amsterdam, the Netherlands, were invited to raise their concerns and questions about the local air quality. They collaborated with experts and scientists who took care of sensor calibration and evaluation. The sensor devices were installed at citizens' home for a two months period of data collection.

According to literature (Mead et al., 2013), the selected electrochemical sensors can be affected by temperature and/or relative humidity. Therefore, a sensor was added to each device to measure these parameters. The calibration was facilitated by collocating sensor devices on the roof of an official monitoring station.

In the study, low-cost electrochemical air quality sensors were examined for citizen urban air quality monitoring. In general, it was confirmed that low-cost electrochemical sensors compared to official measurements with  $R^2$  values ranging

from 0.6 to 0.9 due to individual difference. This result indicates that these informal sensors are valuable for air quality data provision. This is in line with some of the results found for the same type of electrochemical sensors compared with reference measurements at an urban traffic location in Aveiro, Portugal, in 2014 with diverse devices (Borrego et al., 2016). It is worth noting that they got quite different results with devices designed differently but using the same type of sensors. The individual difference may partly explain this, but the hardware design may also contribute to the different performance. These results show that informal sensors when well calibrated have the potential to complement official environmental monitoring data to help answer citizens' questions, which may not be answered from official data. For example, when air quality data are used for an application like finding the healthiest route for daily commuting a higher spatial temporal detail is required to prepare this assessment. However, the sensor data quality should be known and proved valuable. We demonstrated the application of a multiple linear regression model to calibrate the sensor. It proves the value of low-cost sensors for citizen urban air quality monitoring if sensors are properly calibrated and recalibrated. It is possible to gain better performance if more advanced calibration methods are used. For instance, Cross et al. (2017) applied a high-dimensional model representation method for calibration which significantly improved the sensor performance in their experiment. As indicated by Mueller et al. (2017) who compared different models for ozone and nitrogen dioxide sensor calibration, more advanced models should be studied to improve the sensor performance in different environmental conditions. The calibration of air quality sensors still has open scientific questions and is not fully studied yet (Lewis and Edwards, 2016). Standard calibration procedures are not widely established. Therefore, citizen scientists are recommended to collaborate with experts on this.

The research in Chapter 3 showed that citizens could produce valuable environmental data, but that calibration is essential. Can these citizens produced data, also called informal sensing data, be integrated with formal ones? What are the perceptions of both the experts and the citizens on formal and informal sensing data and their integration potential?

### **6.1.3 What are the perceptions of citizens and experts on formal and informal environmental sensing data and their integration potential?**

Since the potential value of citizen environmental sensing is proven, how to reach that potential? Therefore, it is important to understand how citizens and experts think about the data produced by citizens as informal sensing data and the data produced by official organizations as formal environmental sensing data (Figure 6.1).

In chapter 4, I describe the design of an online questionnaire survey to find out the perceptions of citizens and experts on formal and informal environmental sensing data and their integration potential. The experts were selected from mainly Dutch

official environmental monitoring institutes including Dutch National Institute for Public Health and Environment (RIVM), the Royal Netherlands Meteorological Institute (KNMI) and their networks and the Public Health Service of Amsterdam (GGD). The citizens approached for the survey were involved in one or more of the three sensing communities: Internet-of-Things Sensemakers in Amsterdam, Amsterdam Smart Citizens Lab and the Smart Emission Community in Nijmegen.

According to the results, both citizens and experts were quite positive about the accuracy and consistency of formal environmental sensing data (Figure 4.4). Interestingly, when asked in the questionnaire about sufficient accuracy of informal environmental sensing data, a large part of both citizens and experts chose 'Neutral'. They both considered that either formal or informal environmental sensing data were sufficient for environmental information provision. Relatively, citizens were more positive on informal sensing data while experts were more positive on formal sensing data. Positively, both experts and citizens recognized the value of informal environmental sensing and the potential to complement formal environmental sensing data in various aspects especially regarding the potential to increase the spatial coverage. Standardisation was considered the main obstacle to integrate informal sensing data with formal data. Training and support from experts was expected by citizens due to issues related to calibration and data analysis.

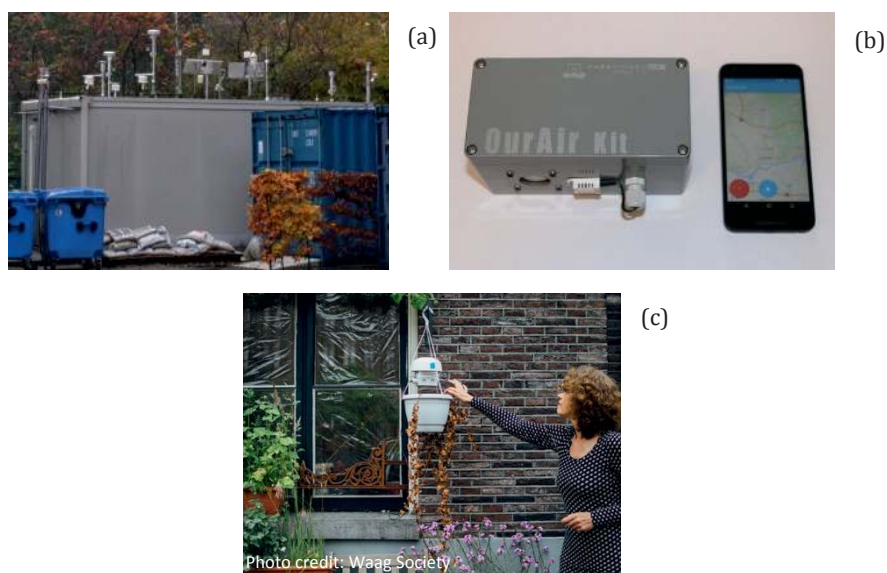


Figure 6.1 Examples of formal and informal sensing devices: (a) official air quality monitoring stations; (b) miniaturized sensor device; (c) informal sensing device measuring air quality at citizen's home in Urban AirQ project.

From the research in Chapter 4, it became clear that both citizen and experts see potential for integrating their sensor data. How can this be achieved?

#### **6.1.4 How can informal and formal sensing data be integrated to provide enhanced environmental information?**

In chapter 5, a framework for using formal and informal environmental sensing data based on purposes towards environmental issues was proposed. According to a certain purpose, either formal or informal environmental sensing data can be used. What is more, the integration of the two can also be used. The approaches to integrate the two was identified as a research gap. Therefore, five integration approaches, based on the relationship (Abdelgawad and Bayoumi, 2012) between formal and informal environmental sensing data, were proposed including: 1) use formal data to calibrate informal data, 2) use formal data source to complement informal data, 3) use informal data to complement formal data, 4) merge formal and informal data and 5) use informal data as formal data. Five success indicators were selected for evaluation of the proposed approaches. A literature study, and an expert validation using Urban AirQ project as a case study were used to validate the proposed integration approaches. According to the literature study and expert validation applying the success indicators, the proposed framework and integration approach are applicable. According to the literature study, the calibration approach (1) is currently most applied. The other approaches are gained increasing attention and have also been evaluated in other studies. However, the data quality of informal environmental sensing data is an important issue for those integration approaches. To establish quality standards for different purposes was therefore recommended.

In this chapter, the focus was on the integration of formal and informal environmental sensing data measuring the same environmental parameter. Broadly, the other data sources measuring different environmental parameters can also be integrated to solve certain environmental issue. For instance, apart from measuring NO<sub>2</sub> directly, also data on traffic intensity are useful to predict NO<sub>2</sub> distribution. Abdelgawad and Bayoumi (2012) defined this as cooperative integration. Cooperative integration can be used with formal or informal environmental sensing data respectively. It can also be applied to the proposed integration approaches. For instance, due to cross sensitivity the integration of informal NO<sub>2</sub> and O<sub>3</sub> data may improve the prediction of NO<sub>2</sub> or O<sub>3</sub> measurements (Mijling et al., 2017). For formal and informal data merge, other data sources can also be integrated to enhance modelling performance (Schneider et al., 2017).

## **6.2 Reflection and Outlook**

In the previous section, the main results of this thesis research were presented. How do they fit into the larger picture? What are expected developments? In this section I will discuss these issues by focussing on the 1) the role of proximate

---

sensing for near real-time environmental data collection, 2) development of sensor networks for smart city applications and 3) the changing role of citizens and open data.

### **6.2.1 Proximate sensing for various high spatial temporal environmental data provision**

On a global level, environmental sensing used to and still relies largely on satellite imagery. Those imagery with diverse electromagnetic bands provide large-scale environmental information, which give valuable information for environmental monitoring and decision-making. The information from satellites is continuously improving as the spatial, spectral and temporal resolutions are increasing. However, information derived from these imagery is still not sufficient for increasing environmental information demands for policy or personal decision-making. For instance, NO<sub>2</sub> detection by satellites can produce concentration information not with resolutions lower than km<sup>2</sup> (Mijling et al., 2009), it is therefore not possible to supply street level information. To complement the satellites, aerial sensing from manned aircrafts and unmanned aerial vehicles (UAVs) for instance can fill the spatial and temporal gaps. However, there are still gaps for more detailed, timely and consistent environmental information provision. That is the part where the increasing availability of sensor networks can contribute. First, with increasing local to global sensor networks, the data collected are not only proximate imagery but also more diverse data which were not available from imagery data. For instance, in situ sensors are created with increasing capabilities to measure directly weather conditions including temperature, relative humidity, wind direction, precipitation and solar radiation (Bell et al., 2013; de Vos et al., 2017; Mei et al., 2017). They can also measure gases related to climate change including CO<sub>2</sub> and methane and related to air quality including for instance NO<sub>2</sub>, O<sub>3</sub>, and SO<sub>2</sub>. Furthermore, sensor networks are also capable of characterizing parameters related to water quality (Adu-Manu et al., 2017), soil parameters (Viscarra Rossel and Bouma, 2016) and other environmental parameters like noise (D'Hondt et al., 2013). Complementary, some of those parameters can be measured indirectly or be modelled using imagery data. Miniaturised diverse sensors give new ways to measure directly increasing amounts of environmental constituents. These measurements could give valuable information for decision making not only about environment policies related to such as climate change and environmental pollution, but also connected domains like precision agriculture. Those sensors can be connected as static sensor networks with sensors distributed in houses, rivers, roadsides and so on. They can also be connected as mobile networks installed on vehicles and be embedded in smart gadgets carried by citizens. Moreover, of course, static and mobile sensor networks can be combined. Those sensor data can first complement, or be integrated with, each other to provide enhanced information.

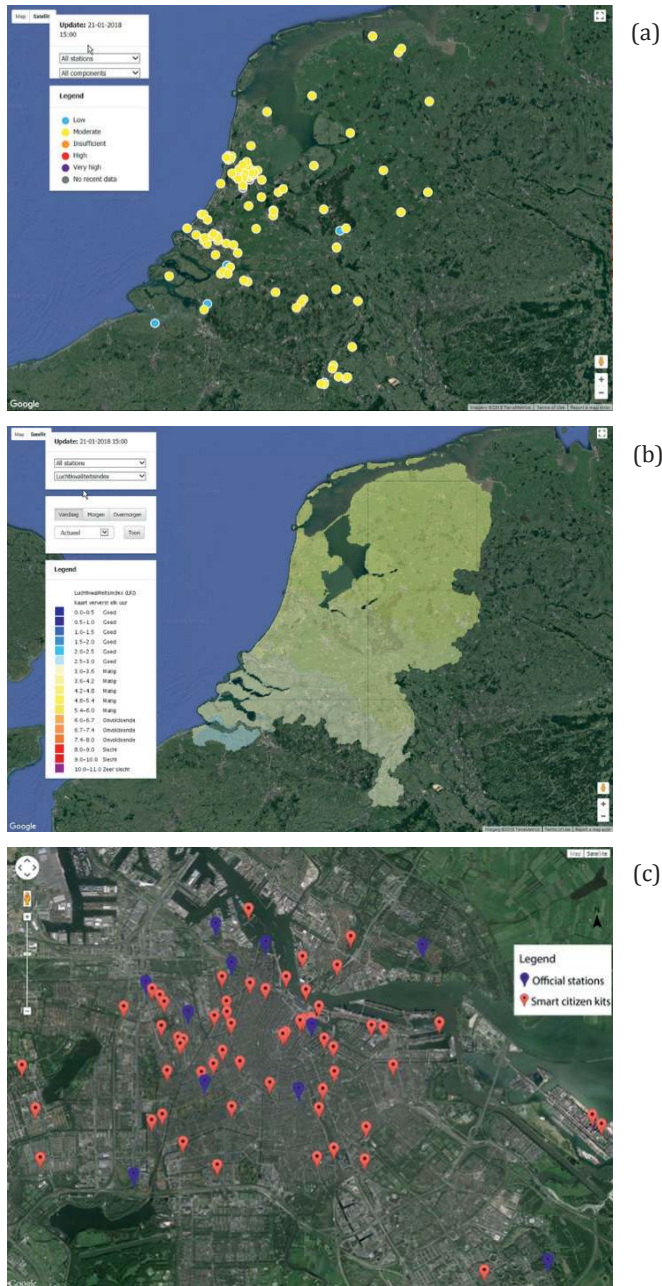


Figure 6.2 Locations of formal and informal sensing stations and modelling: (a) official NO<sub>2</sub> monitoring stations in the Netherlands; (b) The Netherlands national scale NO<sub>2</sub> modelling; and (c) formal and informal air quality measurement location during Smart Citizen Kit Amsterdam project in 2014. (Data source: <https://www.luchtmeetnet.nl> and <https://smarcitizen.me>)

Taking the Netherlands as an example, as shown in Figure 6.2, there are a limited number of official air quality monitoring stations. In this case, NO<sub>2</sub> monitoring stations are selected as an example. Those stations are also concentrated in a limited number of large cities. Policy makers or citizens who want to know the pollution level at certain locations not measured by those sparse stations are normally referring to data from modelling approaches (Beelen et al., 2009; Jato-Espino et al., 2018) (Figure 6.2). Even though, for low-cost sensing, data quality is considered lower than formal sensing data, these measured data could derive better data than modelled figures according to studies evaluating the correlation between low-cost sensing data and reference data (Borrego et al., 2016) and the correlation between modelled data and reference data (Manders et al., 2017). On the other hand, as shown in Figure 6.2 (c), informal sensing data are increasingly becoming available, integration of these low-cost sensing data into these models could improve the performance of modelling (Schneider et al., 2017). In addition, low-cost environmental sensors are often minimized and portable. Their data can be used or integrated with data from other emerging wearable devices for personal spatial temporal air pollution exposure and health monitoring (Hu et al., 2014; Jerrett et al., 2017; Nieuwenhuijsen et al., 2015).

### **6.2.2 Sensor networks, big data and smart city**

Globally, 54 percent of the population lived in cities in 2014 and it is predicted to be increased to 66 percent in 2050 (United Nations, 2015). This also creates problems. For instance, traffic jams, air pollution, waste management, water pollution, noise and so on are typical issues in cities. The smart city concept is proposed as a potential solution for these problems (Caragliu et al., 2011). The smart city, which has different definitions, either focuses on technological or human aspects (Albino et al., 2015). For both the technological and the human aspect, data are important to realise smart cities. These data including both formal and informal ones are increasingly produced by sensor networks and citizens. As in the Amsterdam Smart Citizens Lab, citizens can get together, discuss their concerns, and collaborate with each other with different societal background and expertise in order to make sensing tools to measure air quality, water quality, noise and so on in the city. Citizens are helping to make a greener, sustainable and healthier city.

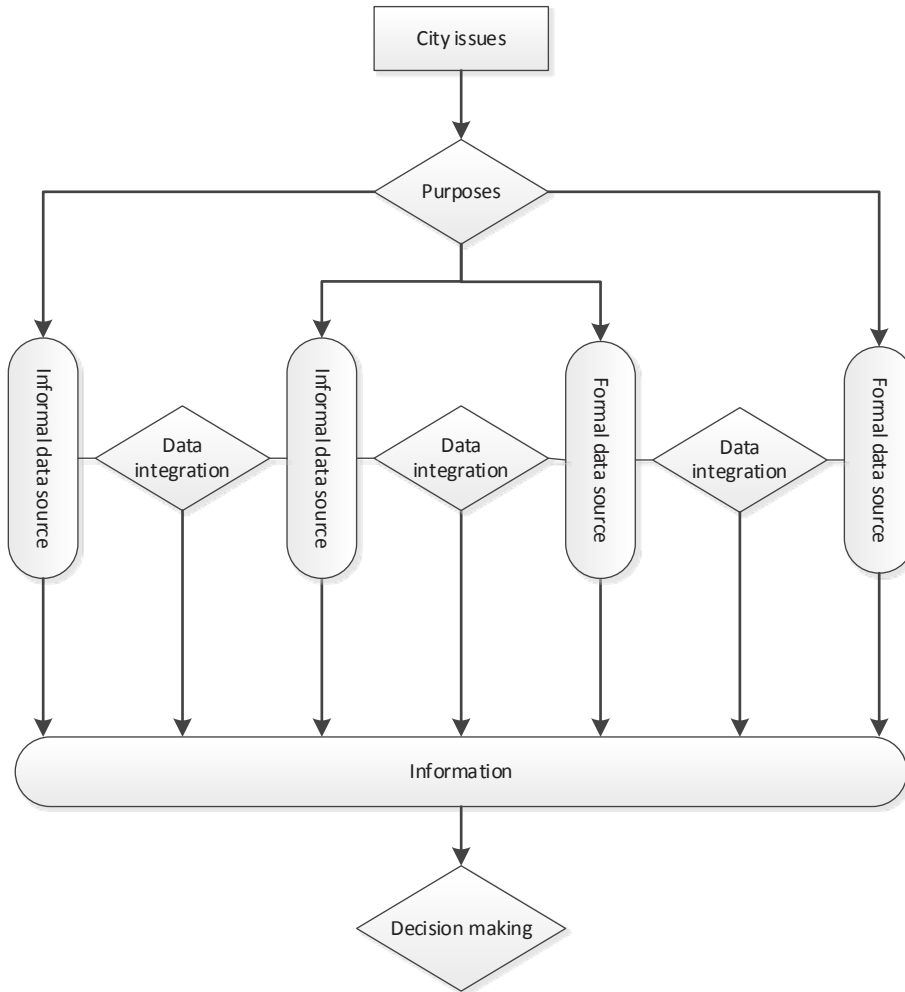


Figure 6.3 How formal and informal data sources produced in city can be integrated with each other to supply more valuable information for smart city decision making.

In this thesis, I focused on evaluation and integration approaches of informal sensor data. Thinking more broadly, sensor networks will contribute substantially to the big data era and the development of smart cities. Sensors are increasingly producing data for big data analysis. According to Gandomi and Haider (2015), volume, variety and velocity are often used to define big data regardless other Vs and evolving definitions. In this thesis, I focused mainly on the volume aspect of big data where informal environmental sensing is expected to contribute significantly. That is, for instance, informal NO<sub>2</sub> measurements data complement directly to formal NO<sub>2</sub> measurements. However, with other sensors measuring diverse environmental phenomena, which constitute the variety of big data. For instance, this relates to the use of transportation data, which measure the frequency of



vehicles including cars, trucks, ships and so on. These data are measured by other sensors than air quality sensors. But these data can be fed to models which can optimize air quality prediction (Zheng et al., 2013). Therefore, data integration becomes even more crucial. Next to the increasing data availability through active sensors in networks, also increasingly relevant data is produced passively by social media such as Twitter, Facebook and Instagram together with increasing data produced by sensors embedded in smartphones or sensors installed on light posts for smart cities. However, those data are often heterogeneous which makes integration more challenging. For instance, these data could have different types: points, texts, image (vector and raster), audio and video. They could have a different spatial temporal resolution and they could have different formats: Excel, text, API, data stream. Furthermore, there might be different processing requirements. For instance, driverless cars need to interact with other cars and environment in near real time. Consequently, the current city maps are not capable to assist this requirement, even though, Waze (<https://www.waze.com>) a community based traffic APP is contributing near real time traffic information by mainly drivers. Semantic web and linked data (Le-Phuoc and Hauswirth, 2009; Llaves et al., 2016) may help to tackle data heterogeneity issues. And the edge computing which process data locally might help to solve the velocity issues (Bibri, 2018).

As indicated in Figure 6.3, both formal and informal data produced in cities are the sources to analysis and solve city issues. Those data can be used directly to create information needed for decision-making. But they can also be integrated: formal with formal, informal with informal and also formal with informal to create enhanced information for better understanding of issues and more wise decisions to tackle issues or create solutions for smart cities. Standards are therefore critical for sharing and utilize data.

### 6.2.3 Citizen science and open science

Open science is defined by Michael Nielsen as “the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process” as cited in (European Commission, 2016). In the Oxford Dictionaries (2018), citizen science is defined as “The collection and analysis of data relating to the natural world by members of the general public, typically as part of a collaborative project with professional scientists”. However, there is no consensus on one citizen science definition and its definitions are evolving (Eitzel et al., 2017). In this thesis, the integration of informal data produced by citizens and formal data produced by experts are studied. I argue that not only the data should be integrated, the citizen science and science domain should also be integrated. That is how open science should interact with citizen science.

Citizen science and open science are two connected movements getting more and more attention but also encounter challenges. For instance, open access of publications received quite some attention and there are increasing open access options too. However, those open access publications often require author(s) to pay publication process fee rather than asking readers to pay subscription fee. The other aspect becoming more popular is open data due to increasing amount of data

been produced. Privacy is a big issue. However, open data without open algorithms may not accelerate innovation dramatically. In terms of citizen science, data quality is an issue of major concern. Sometimes, citizen science is considered as part of open science. From chapter 2 it was concluded that a successful citizen science project should be open and involve wide collaboration, even though we do recognize that an individual can also do citizen science (Bonney et al., 2016). Citizen science can either be a bottom-up citizens' initiative or top-down scientists' initiative. The openness will stimulate participation and wide collaboration among scientists, hardware and software developers, designers, domain experts, policy makers and others. Openness and wide collaboration together helps creating tools, improving data quality and establishing trust as shown in citizen radiation monitoring after Fukushima Daiichi nuclear disaster in Japan, 2011 (Hemmi and Graham, 2014). Scientists joining citizen initiatives could not only help citizens for project development but also help themselves to produce scientific results. On the other hand, scientists opening up science by open publications, methods, data and so on can also stimulate citizens participating in science. Ideally, I think science should be open to everyone and in this view citizens play a role in science development as well (Figure 6.4).

As observed in the Amsterdam Smart Citizens Lab project in chapter 2, citizens are anyway more capable to do science based on their concerns and interests by using open source tools and organize online and offline communities locally, nationally and globally. Some of them indeed are motivated by dissatisfaction about formal data sources. At the same time, some scientists and experts do question the relevance of citizen science especially the data produced by them. However, as indicated by Cohn (2008), citizens are capable to produce data with the same quality as experts. By conducting a case study, Pratihast et al. (2014) also showed that citizens can collect accurate data to supply valuable complementary information for forest change monitoring. Moreover, we observe that more scientists and experts are embracing citizen science and open their science to citizens. Scientists and experts opening their science do not only stimulate peer collaboration but also feedback, application and further development by citizens. By co-creating in science projects with citizens, scientists can not only get data but also other input like research questions, tools and method development. Citizens, on the other hand, can also benefit from the integration of citizen science and open science. For instance, the calibration and recalibration challenge of air quality sensors mentioned in chapter 3 was supported by experts and scientist and therefore produced data with reasonably quality to provide environmental information. When standard calibration procedures would become available, citizens can follow these procedures and with potential additional help from scientists produce valuable data for further big data analysis and smart city applications. After the Fukushima Daiichi nuclear disaster in Japan, 2011, the citizen radiation monitoring initiative were successful partly also because they keep the community open for diverse scientists and experts which solved many technical challenges and established trust (Hemmi and Graham, 2014).

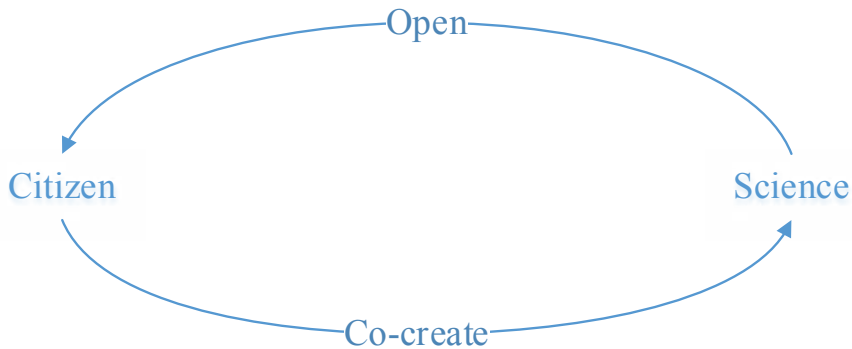


Figure 6.4 An ideal relationship between citizen (science) and (open) science.

### 6.3 Recommendations for further research

Based on the research results from my PhD research, I recommend the following research topics for further investigation:

1. Research on advanced air quality sensor calibration and data analysis for improved data quality and prediction is recommended. This includes consideration the complexity of sensor measurements involving challenges such as cross sensitivity, temperature and/or relative humidity interference (Mead et al., 2013). It would be valuable to study the behaviour of a group of sensors measuring different species of gaseous and to apply advanced calibration methods and data analysis algorithms to improve data quality and performance of prediction.
2. Research on integration methods for heterogeneous multiple data sources is recommended. Formal and informal sensing data do have differences in terms of data quality, standards, formats, etc. The integration methods studied in this thesis focused on point-sensed data from sensors only. How about measurement by different Unmanned Vehicle Systems platforms with electric noses? How about other informal data and other data type like raster, audio and video? It would be relevant to study and develop methods to integrate more heterogeneous data sources including both formal and informal ones. For instance, the informal data source can also be social media data (Sakaki et al., 2010; Salas-Olmedo et al., 2018) and the formal data source can be satellite images.
3. Research on the developments of citizen science and open science and their impact on society is recommended. Citizen science and open science have different development routes and levels in different countries due to diverse cultures and social systems. Different domains and different communities may also have different development levels, for instance citizen science in agriculture (Beza et al., 2017). It would be relevant to study these different developments and their impact on society.



## References

- AAN 105-03, 2009. Designing a potentiostatic circuit.[http://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN\\_105-03.pdf](http://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN_105-03.pdf).
- AAN 110, 2013. Environmental Changes: Temperature, Pressure, Humidity.[http://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN\\_110.pdf](http://www.alphasense.com/WEB1213/wp-content/uploads/2013/07/AAN_110.pdf).
- Abdelgawad, A., Bayoumi, M., 2012. Data Fusion in WSN, Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks. Springer US, Boston, MA, pp. 17-35.
- ADS, 2017. Alphasense Data Sheet for NO2-B43F.<http://www.alphasense.com/WEB1213/wp-content/uploads/2017/07/NO2B43F.pdf>.
- Adu-Manu, K.S., Tapparello, C., Heinzelman, W., Katsriku, F.A., Abdulai, J.D., 2017. Water quality monitoring using wireless sensor networks: Current trends and future research directions. *ACM Transactions on Sensor Networks* 13(1).
- Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E., 2002. Wireless sensor networks: A survey. *Computer Networks* 38(4), 393-422.
- Albino, V., Berardi, U., Dangelico, R.M., 2015. Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology* 22(1), 3-21.
- Artiola, J.F., Pepper, I.L., Brusseau, M.L., Brusseau, J.F.A.L.P.L., 2004. 1 - Monitoring and Characterization of the Environment, *Environmental Monitoring and Characterization*. Academic Press, Burlington, pp. 1-9.
- Aumond, P., Lavandier, C., Ribeiro, C., Boix, E.G., Kambona, K., D'Hondt, E., Delaitre, P., 2017. A study of the accuracy of mobile technology for measuring urban noise pollution in large scale participatory sensing campaigns. *Applied Acoustics* 117(Part B), 219-226.
- Balestrini, M., Diez, T., Marshall, P., Gluhak, A., Rogers, Y., 2015. IoT community technologies: leaving users to their own devices or orchestration of engagement? *EAI Endorsed Transactions on Internet of Things* 1(1).
- Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. *Science of The Total Environment* 407(6), 1852-1867.
- Bell, S., Cornford, D., Bastin, L., 2013. The state of automated amateur weather observations. *Weather* 68(2), 36-41.
- Bell, S., Cornford, D., Bastin, L., 2015. How good are citizen weather stations? Addressing a biased opinion. *Weather* 70(3), 75-84.
- Beza, E., Steinke, J., van Etten, J., Reidsma, P., Fadda, C., Mitra, S., Mathur, P., Kooistra, L., 2017. What are the prospects for citizen science in

- agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers. *PLOS ONE* 12(5), e0175700.
- Bibri, S.E., 2018. The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. *Sustainable Cities and Society* 38, 230-253.
- Bonney, R., Cooper, C., Ballard, H., 2016. The Theory and Practice of Citizen Science: Launching a New Journal. *Citizen Science: Theory and Practice*. 1(1).
- Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., Shirk, J., 2009. Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience* 59(11), 977-984.
- Borrego, C., Costa, A.M., Ginja, J., Amorim, M., Coutinho, M., Karatzas, K., Sioumis, T., Katsifarakis, N., Konstantinidis, K., De Vito, S., Esposito, E., Smith, P., André, N., Gérard, P., Francis, L.A., Castell, N., Schneider, P., Viana, M., Minguillón, M.C., Reimringer, W., Otjes, R.P., von Sicard, O., Pohle, R., Elen, B., Suriano, D., Pfister, V., Prato, M., Dipinto, S., Penza, M., 2016. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmospheric Environment* 147, 246-263.
- Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebet, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science* 11(7), 699-718.
- Cape, J.N., 2009. The Use of Passive Diffusion Tubes for Measuring Concentrations of Nitrogen Dioxide in Air. *Critical Reviews in Analytical Chemistry* 39(4), 289-310.
- Caragliu, A., Del Bo, C., Nijkamp, P., 2011. Smart Cities in Europe. *Journal of Urban Technology* 18(2), 65-82.
- Carton, L., Ache, P., 2017. Citizen-sensor-networks to confront government decision-makers: Two lessons from the Netherlands. *Journal of Environmental Management* 196, 234-251.
- Castell, N., Dauge, F.R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., Bartonova, A., 2017a. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International* 99, 293-302.
- Castell, N., Schneider, P., Grossberndt, S., Fredriksen, M.F., Sousa-Santos, G., Vogt, M., Bartonova, A., 2017b. Localized real-time information on outdoor air quality at kindergartens in Oslo, Norway using low-cost sensor nodes. *Environmental Research*.
- Chen, K.Y., Janz, K.F., Zhu, W., Brychta, R.J., 2012. Redefining the roles of sensors in objective physical activity monitoring. *Medicine and Science in Sports and Exercise* 44(SUPPL. 1), S13-S23.
- Cohn, J.P., 2008. Citizen Science: Can Volunteers Do Real Research? *BioScience* 58(3), 192-197.
- Conrad, C., Hilchey, K., 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental Monitoring and Assessment* 176(1-4), 273-291.

- Cooper, C.B., Dickinson, J., Phillips, T., Bonney, R., 2007. Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society* 12(2).
- Corke, P., Wark, T., Jurdak, R., Hu, W., Valencia, P., Moore, D., 2010. Environmental wireless sensor networks. *Proceedings of the IEEE* 98(11), 1903-1917.
- Cross, E.S., Williams, L.R., Lewis, D.K., Magoon, G.R., Onasch, T.B., Kaminsky, M.L., Worsnop, D.R., Jayne, J.T., 2017. Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements. *Atmos. Meas. Tech.* 10(9), 3575-3588.
- D'Hondt, E., Stevens, M., Jacobs, A., 2013. Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. *Pervasive and Mobile Computing* 9(5), 681-694.
- de Albuquerque, J.P., Herfort, B., Brenning, A., Zipf, A., 2015. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science* 29(4), 667-689.
- de Vos, L., Leijnse, H., Overeem, A., Uijlenhoet, R., 2017. The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrol. Earth Syst. Sci.* 21(2), 765-777.
- Delaney, D.G., Sperling, C.D., Adams, C.S., Leung, B., 2007. Marine invasive species: validation of citizen science and implications for national monitoring networks. *Biological Invasions* 10(1), 117-128.
- Di Carlo, S., Falasconi, M., Sanchez, E., Scionti, A., Squillero, G., Tonda, A., 2011. Increasing pattern recognition accuracy for chemical sensing by evolutionary based drift compensation. *Pattern Recognition Letters* 32(13), 1594-1603.
- Dickinson, J.L., Zuckerman, B., Bonter, D.N., 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics* 41, 149-172.
- Donnelly, A., Naughton, O., Misstear, B., Broderick, B., 2016. Maximizing the spatial representativeness of NO<sub>2</sub> monitoring data using a combination of local wind-based sectoral division and seasonal and diurnal correction factors. *Journal of Environmental Science and Health, Part A* 51(12), 1003-1011.
- Droste, A.M., Pape, J.J., Overeem, A., Leijnse, H., Steeneveld, G.J., Delden, A.J.V., Uijlenhoet, R., 2017. Crowdsourcing Urban Air Temperatures through Smartphone Battery Temperatures in São Paulo, Brazil. *Journal of Atmospheric and Oceanic Technology* 34(9), 1853-1866.
- Duvall, R., Long, R., Beaver, M., Kronmiller, K., Wheeler, M., Szykman, J., 2016. Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide. *Sensors* 16(10), 1698.
- Eitzel, M.V., Cappadonna, J.L., Santos-Lang, C., Duerr, R.E., Virapongse, A., West, S.E., Kyba, C.C.M., Bowser, A., Cooper, C.B., Sforzi, A., Metcalfe, A.N., Harris, E.S., Thiel, M., Haklay, M., Ponciano, L., Roche, J., Ceccaroni, L., Shilling, F.M., Dörler, D., Heigl, F., Kiessling, T., Davis, B.Y., Jiang, Q., 2017. Citizen Science Terminology Matters: Exploring Key Terms. *Citizen Science: Theory and Practice* 2(1).

- Esposito, E., De Vito, S., Salvato, M., Bright, V., Jones, R.L., Popoola, O., 2016. Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems. *Sensors and Actuators B: Chemical* 231, 701-713.
- EU, 2008. Directive 2008/50/EC of the European Parliament and the Council of 21 May 2008 on Ambient Air Quality and Cleaner Air for Europe.
- European Commission, 2016. Open innovation, open science, open to the world: a vision for Europe
- Gandomi, A., Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35(2), 137-144.
- Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal* 69(4), 211-221.
- Grothe, M., Carton, L., Van Den Broecke, J., Volten, H., Kieboom, R., 2016. Smart emission - Building a spatial data infrastructure for an environmental citizen sensor network, *CEUR Workshop Proceedings*.
- Hagan, D.H., Isaacman-VanWertz, G., Franklin, J.P., Wallace, L.M.M., Kocar, B.D., Heald, C.L., Kroll, J.H., 2018. Calibration and assessment of electrochemical air quality sensors by co-location with regulatory-grade instruments. *Atmos. Meas. Tech.* 11(1), 315-328.
- Haklay, M., 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In: D. Sui, S. Elwood, M. Goodchild (Eds.), *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*. Springer Netherlands, Dordrecht, pp. 105-122.
- Hasenfratz, D., Saukh, O., Thiele, L., 2012. On-the-fly calibration of low-cost gas sensors, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 228-244.
- 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, Part B, 268-285.
- Hemmi, A., Graham, I., 2014. Hacker science versus closed science: building environmental monitoring infrastructure. *Information, Communication & Society* 17(7), 830-842.
- Hierlemann, A., Gutierrez-Osuna, R., 2008. Higher-Order Chemical Sensing. *Chemical Reviews* 108(2), 563-613.
- Holt, B.G., Rioja-Nieto, R., Aaron MacNeil, M., Lupton, J., Rahbek, C., 2013. Comparing diversity data collected using a protocol designed for volunteers with results from a professional alternative. *Methods Ecol Evol* 4(4), 383-392.
- Horn, C.v.d., Boonstra, R., 2014. Eindrapportage Smart Citizen Kit Amsterdam Meten is weten?, *Waag Society*
- Hu, K., Davison, T., Rahman, A., Sivaraman, V., 2014. Air pollution exposure estimation and finding association with human activity using wearable



- sensor network, ACM International Conference Proceeding Series, pp. 48-55.
- Hubbell, B.J., Kaufman, A., Rivers, L., Schulte, K., Hagler, G., Clougherty, J., Cascio, W., Costa, D., 2018. Understanding social and behavioral drivers and impacts of air quality sensor use. *Science of The Total Environment* 621, 886-894.
- Huck, J.J., Whyatt, J.D., Coulton, P., Davison, B., Gradinar, A., 2017. Combining physiological, environmental and locational sensors for citizen-oriented health applications. *Environmental Monitoring and Assessment* 189(3), 114.
- Jalbert, K., Kinchy, A.J., 2016. Sense and Influence: Environmental Monitoring Tools and the Power of Citizen Science. *Journal of Environmental Policy & Planning* 18(3), 379-397.
- Jato-Espino, D., Castillo-Lopez, E., Rodriguez-Hernandez, J., Ballester-Muñoz, F., 2018. Air quality modelling in Catalonia from a combination of solar radiation, surface reflectance and elevation. *Science of The Total Environment* 624, 189-200.
- Jerrett, M., Donaire-Gonzalez, D., Popoola, O., Jones, R., Cohen, R.C., Almanza, E., de Nazelle, A., Mead, I., Carrasco-Turigas, G., Cole-Hunter, T., Triguero-Mas, M., Seto, E., Nieuwenhuijsen, M., 2017. Validating novel air pollution sensors to improve exposure estimates for epidemiological analyses and citizen science. *Environmental Research* 158(Supplement C), 286-294.
- Jiang, Q., Bregt, A.K., Kooistra, L., 2018. Formal and informal environmental sensing data and integration potential: Perceptions of citizens and experts. *Science of The Total Environment* 619-620, 1133-1142.
- Jiang, Q., Kresin, F., Bregt, A.K., Kooistra, L., Pareschi, E., van Putten, E., Volten, H., Wesseling, J., 2016. Citizen Sensing for Improved Urban Environmental Monitoring. *Journal of Sensors* 2016, 9.
- Jiao, W., Hagler, G., Williams, R., Sharpe, R., Brown, R., Garver, D., Judge, R., Caudill, M., Rickard, J., Davis, M., Weinstock, L., Zimmer-Dauphinee, S., Buckley, K., 2016. Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos. Meas. Tech.* 9(11), 5281-5292.
- Jollymore, A., Haines, M.J., Satterfield, T., Johnson, M.S., 2017. Citizen science for water quality monitoring: Data implications of citizen perspectives. *Journal of Environmental Management* 200, 456-467.
- Jovašević-Stojanović, M., Bartonova, A., Topalović, D., Lazović, I., Pokrić, B., Ristovski, Z., 2015. On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter. *Environmental Pollution* 206, 696-704.
- Kamel Boulos, M.N., Resch, B., Crowley, D.N., Breslin, J.G., Sohn, G., Burtner, R., Pike, W.A., Jezierski, E., Chuang, K.-Y.S., 2011. Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples. *International Journal of Health Geographics* 10(1), 67.

- Kelman, I., 2006. Warning for the 26 December 2004 tsunamis. *Disaster Prevention and Management* 15(1), 178-189.
- Khaleghi, B., Khamis, A., Karray, F.O., Razavi, S.N., 2013. Multisensor data fusion: A review of the state-of-the-art. *Information Fusion* 14(1), 28-44.
- Kolb, C.E., Herndon, S.C., McManus, J.B., Shorter, J.H., Zahniser, M.S., Nelson, D.D., Jayne, J.T., Canagaratna, M.R., Worsnop, D.R., 2004. Mobile Laboratory with Rapid Response Instruments for Real-Time Measurements of Urban and Regional Trace Gas and Particulate Distributions and Emission Source Characteristics. *Environmental Science & Technology* 38(21), 5694-5703.
- Kooistra, L., Thessler, S., Bregt, A.K., 2009. User Requirements and Future Expectations for Geosensor Networks – An Assessment. In: N. Trigoni, A. Markham, S. Nawaz (Eds.), *GeoSensor Networks: Third International Conference, GSN 2009, Oxford, UK, July 13-14, 2009. Proceedings.* Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 149-157.
- Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., Britter, R., 2015. The rise of low-cost sensing for managing air pollution in cities. *Environment International* 75, 199-205.
- Le-Phuoc, D., Hauswirth, M., 2009. Linked open data in sensor data mashups, *CEUR Workshop Proceedings*, pp. 1-16.
- Lewandowski, E., Caldwell, W., Elmquist, D., Oberhauser, K., 2017. Public Perceptions of Citizen Science. *Citizen Science: Theory and Practice* 2(1).
- Lewenstein, B.V., 2016. Can we understand citizen science? *Journal of Science Communication* 15(1).
- Lewis, A., Edwards, P., 2016. Validate personal air-pollution sensors. *Nature* 535(7610), 29-31.
- Lewis, A.C., Lee, J.D., Edwards, P.M., Shaw, M.D., Evans, M.J., Moller, S.J., Smith, K.R., Buckley, J.W., Ellis, M., Gillot, S.R., White, A., 2016. Evaluating the performance of low cost chemical sensors for air pollution research. *Faraday Discussions* 189(0), 85-103.
- Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software* 53, 173-189.
- Ligtenberg, A., 2006. Exploring the use of multi-agent systems for interactive multi-actor spatial planning, s.n.], [S.l.
- Lin, C., Gillespie, J., Schuder, M.D., Duberstein, W., Beverland, I.J., Heal, M.R., 2015. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. *Atmospheric Environment* 100(Supplement C), 111-116.
- Lin, C., Masey, N., Wu, H., Jackson, M., Carruthers, D., Reis, S., Doherty, R., Beverland, I., Heal, M., 2017. Practical Field Calibration of Portable Monitors for Mobile Measurements of Multiple Air Pollutants. *Atmosphere* 8(12), 231.
- Little, K.E., Hayashi, M., Liang, S., 2016. Community-Based Groundwater Monitoring Network Using a Citizen-Science Approach. *Groundwater* 54(3), 317-324.

- 
- Llaves, A., Corcho, O., Taylor, P., Taylor, K., 2016. Enabling RDF stream processing for sensor data management in the environmental domain. *International Journal on Semantic Web and Information Systems* 12(4), 1-21.
- Maisonneuve, N., Stevens, M., Ochab, B., 2010. Participatory noise pollution monitoring using mobile phones. *Information Polity* 15(1-2), 51-71.
- Manders, A.M.M., Builtjes, P.J.H., Curier, L., Denier van der Gon, H.A.C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J.J.P., Segers, A.J., Timmermans, R.M.A., Visschedijk, A.J.H., Wichink Kruit, R.J., van Pul, W.A.J., Sauter, F.J., van der Swaluw, E., Swart, D.P.J., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., van Velthoven, P., Banzhaf, S., Mues, A.C., Stern, R., Fu, G., Lu, S., Heemink, A., van Velzen, N., Schaap, M., 2017. Curriculum vitae of the LOTOS-EUROS (v2.0) chemistry transport model. *Geosci. Model Dev.* 10(11), 4145-4173.
- Masson, N., Piedrahita, R., Hannigan, M., 2015. Quantification Method for Electrolytic Sensors in Long-Term Monitoring of Ambient Air Quality. *Sensors* 15(10), 27283.
- Mazzoleni, M., Verlaan, M., Alfonso, L., Monego, M., Norbiato, D., Ferri, M., Solomatine, D.P., 2017. Can assimilation of crowdsourced data in hydrological modelling improve flood prediction? *Hydrol. Earth Syst. Sci.* 21(2), 839-861.
- McKercher, G.R., Salmond, J.A., Vanos, J.K., 2017. Characteristics and applications of small, portable gaseous air pollution monitors. *Environmental Pollution* 223, 102-110.
- Mead, M.I., Popoola, O.A.M., Stewart, G.B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J.J., McLeod, M.W., Hodgson, T.F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J.R., Jones, R.L., 2013. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment* 70, 186-203.
- Mei, B., Li, R., Cheng, W., Yu, J., Cheng, X., 2017. Ultraviolet Radiation Measurement via Smart Devices. *IEEE Internet of Things Journal* 4(4), 934-944.
- Meier, F., Fenner, D., Grassmann, T., Otto, M., Scherer, D., 2017. Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate* 19, 170-191.
- Merkus, T., 2016. Integrating Formal data and Volunteered Geographic Information-A case with amateur weather data and formal KNMI data.
- Mijling, B., Jiang, Q., de Jonge, D., Bocconi, S., 2017. Practical field calibration of electrochemical NO<sub>2</sub> sensors for urban air quality applications. *Atmos. Meas. Tech. Discuss.* 2017, 1-25.
- Mijling, B., van der A, R.J., Boersma, K.F., Van Roozendaal, M., De Smedt, I., Kelder, H.M., 2009. Reductions of NO<sub>2</sub> detected from space during the 2008 Beijing Olympic Games. *Geophysical Research Letters* 36(13), n/a-n/a.
- Minkman, E., van der Sanden, M., Rutten, M., 2017. Practitioners' viewpoints on citizen science in water management: a case study in Dutch regional water resource management. *Hydrol. Earth Syst. Sci.* 21(1), 153-167.

- Miskell, G., Salmond, J., Williams, D.E., 2017. Low-cost sensors and crowd-sourced data: Observations of siting impacts on a network of air-quality instruments. *Science of The Total Environment* 575, 1119-1129.
- 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.
- Morzy, M., 2015. ICT Services for open and citizen science. *World Wide Web* 18(4), 1147-1161.
- Mueller, M., Meyer, J., Hueglin, C., 2017. Design of an ozone and nitrogen dioxide sensor unit and its long-term operation within a sensor network in the city of Zurich. *Atmos. Meas. Tech.* 10(10), 3783-3799.
- Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., Leigh, R.R., 2015. Crowdsourcing for climate and atmospheric sciences: current status and future potential. *Int. J. Climatol.* 35(11), 3185-3203.
- Nieuwenhuijsen, M.J., Donaire-Gonzalez, D., Rivas, I., de Castro, M., Cirach, M., Hoek, G., Seto, E., Jerrett, M., Sunyer, J., 2015. Variability in and Agreement between Modeled and Personal Continuously Measured Black Carbon Levels Using Novel Smartphone and Sensor Technologies. *Environmental Science & Technology* 49(5), 2977-2982.
- Oxford Dictionaries, 2017. Oxford University Press.
- Oxford Dictionaries, 2018. Oxford University Press.
- Pang, X., Shaw, M.D., Lewis, A.C., Carpenter, L.J., Batchellier, T., 2017. Electrochemical ozone sensors: A miniaturised alternative for ozone measurements in laboratory experiments and air-quality monitoring. *Sensors and Actuators B: Chemical* 240, 829-837.
- Piedrahita, R., Xiang, Y., Masson, N., Ortega, J., Collier, A., Jiang, Y., Li, K., Dick, R.P., Lv, Q., Hannigan, M., Shang, L., 2014. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmos. Meas. Tech.* 7(10), 3325-3336.
- Pratihast, A., DeVries, B., Avitabile, V., de Bruin, S., Kooistra, L., Tekle, M., Herold, M., 2014. Combining Satellite Data and Community-Based Observations for Forest Monitoring. *Forests* 5(10), 2464.
- Prüss-Ustün, A., Wolf, J., Corvalán, C., Bos, R., Neira, M., 2016. Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks. World Health Organization.
- Restrepo-Estrada, C., de Andrade, S.C., Abe, N., Fava, M.C., Mendiando, E.M., de Albuquerque, J.P., 2018. Geo-social media as a proxy for hydrometeorological data for streamflow estimation and to improve flood monitoring. *Computers & Geosciences* 111(Supplement C), 148-158.
- Ruestow, E.G., 1996. *The Microscope in the Dutch Republic: The Shaping of Discovery*. Cambridge University Press.
- Sakaki, T., Okazaki, M., Matsuo, Y., 2010. Earthquake shakes Twitter users: Real-time event detection by social sensors, *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, pp. 851-860.

- Salas-Olmedo, M.H., Moya-Gómez, B., García-Palomares, J.C., Gutiérrez, J., 2018. Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management* 66(Supplement C), 13-25.
- Samarajiva, R., 2005. Mobilizing information and communications technologies for effective disaster warning: Lessons from the 2004 tsunami. *New Media and Society* 7(6), 731-747.
- Schnebele, E., Cervone, G., Kumar, S., Waters, N., 2014a. Real time estimation of the calgary floods using limited remote sensing data. 6(2), 381-398.
- Schnebele, E., Cervone, G., Waters, N., 2014b. Road assessment after flood events using non-authoritative data. 14(4), 1007-1015.
- Schnebele, E., Cervone, G., Waters, N., 2014c. Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Sciences* 14(4), 1007-1015.
- Schnebele, E., Oxendine, C., Cervone, G., Ferreira, C.M., Waters, N., 2015. Using non-authoritative sources during emergencies in urban areas, *Computational Approaches for Urban Environments*, pp. 337-361.
- Schneider, P., Castell, N., Vogt, M., Dauge, F.R., Lahoz, W.A., Bartonova, A., 2017. Mapping urban air quality in near real-time using observations from low-cost sensors and model information. *Environment International* 106(Supplement C), 234-247.
- Schwamb, M.E., Orosz, J.A., Carter, J.A., Welsh, W.F., Fischer, D.A., Torres, G., Howard, A.W., Crepp, J.R., Keel, W.C., Lintott, C.J., Kaib, N.A., Terrell, D., Gagliano, R., Jek, K.J., Parrish, M., Smith, A.M., Lynn, S., Simpson, R.J., Giguere, M.J., Schawinski, K., 2013. Planet hunters: A transiting circumbinary planet in a quadruple star system. *Astrophysical Journal* 768(2).
- See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M., Liu, H.-Y., Milčinski, G., Nikšič, M., Painho, M., Pödör, A., Olteanu-Raimond, A.-M., Rutzinger, M., 2016. Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. *ISPRS International Journal of Geo-Information* 5(5), 55.
- Silvertown, J., 2009. A new dawn for citizen science. *Trends in Ecology & Evolution* 24(9), 467-471.
- Snyder, E.G., Watkins, T.H., Solomon, P.A., Thoma, E.D., Williams, R.W., Hagler, G.S.W., Shelow, D., Hindin, D.A., Kilaru, V.J., Preuss, P.W., 2013. The Changing Paradigm of Air Pollution Monitoring. *Environmental Science & Technology* 47(20), 11369-11377.
- Spinelle, L., Gerboles, M., Alexandre, M., 2015a. Performance Evaluation of Amperometric Sensors for the Monitoring of O3 and NO2 in Ambient Air at ppb Level. *Procedia Engineering* 120, 480-483.
- Spinelle, L., Gerboles, M., Villani, M.G., Alexandre, M., Bonavitacola, F., 2015b. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. *Sensors and Actuators B: Chemical* 215, 249-257.

- Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., Kelling, S., 2009. eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* 142(10), 2282-2292.
- Thompson, J.E., 2016. Crowd-sourced air quality studies: A review of the literature & portable sensors. *Trends in Environmental Analytical Chemistry* 11, 23-34.
- United Nations, 2015. *World Urbanization Prospects: The 2014 Revision*, Department of Economic and Social Affairs, Population Division
- Van den Bossche, J., Theunis, J., Elen, B., Peters, J., Botteldooren, D., De Baets, B., 2016. Opportunistic mobile air pollution monitoring: A case study with city wardens in Antwerp. *Atmospheric Environment* 141(Supplement C), 408-421.
- Veregin, H., 1999. Data quality parameters. *Geographical information systems* 1, 177-189.
- Viscarra Rossel, R.A., Bouma, J., 2016. Soil sensing: A new paradigm for agriculture. *Agricultural Systems* 148, 71-74.
- Weissert, L.F., Salmond, J.A., Miskell, G., Alavi-Shoshtari, M., Grange, S.K., Henshaw, G.S., Williams, D.E., 2017. Use of a dense monitoring network of low-cost instruments to observe local changes in the diurnal ozone cycles as marine air passes over a geographically isolated urban centre. *Science of The Total Environment* 575, 67-78.
- Weng, Y.-C., 2015. Contrasting visions of science in ecological restoration: Expert-lay dynamics between professional practitioners and volunteers. *Geoforum* 65, 134-145.
- Werner-Allen, G., Johnson, J., Ruiz, M., Lees, J., Welsh, M., 2005. Monitoring volcanic eruptions with a wireless sensor network, *Proceedings of the Second European Workshop on Wireless Sensor Networks, EWSN 2005*, pp. 108-120.
- Wilderman, C.C., 2007. Models of community science: design lessons from the field, *Citizen Science Toolkit Conference*, C. McEver, R. Bonney, J. Dickinson, S. Kelling, K. Rosenberg, and JL Shirk, Eds., Cornell Laboratory of Ornithology, Ithaca, NY.
- Williams, R., Vasu Kilaru, E. Snyder, A.K., T. Dye, A. Rutter, A. Russell, Hafner, A.H., 2014. *Air Sensor Guidebook*, U.S. Environmental Protection Agency, Washington, DC
- Xiang, Y., Tang, Y., Zhu, W., 2016. Mobile sensor network noise reduction and recalibration using a Bayesian network. *Atmos. Meas. Tech.* 9(2), 347-357.
- Zerger, A., Viscarra Rossel, R.A., Swain, D.L., Wark, T., Handcock, R.N., Doerr, V.A.J., Bishop-Hurley, G.J., Doerr, E.D., Gibbons, P.G., Lobsey, C., 2010. Environmental sensor networks for vegetation, animal and soil sciences. *International Journal of Applied Earth Observation and Geoinformation* 12(5), 303-316.
- Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., Van Cooten, S., Kelleher, K., Kitzmiller, D., Ding, F., Seo, D.-J., Wells, E., Dempsey, C., 2011. *National Mosaic and Multi-Sensor QPE (NMQ) System: Description*,

Results, and Future Plans. *Bulletin of the American Meteorological Society* 92(10), 1321-1338.

Zheng, Y., 2015. Methodologies for Cross-Domain Data Fusion: An Overview. *IEEE Transactions on Big Data* 1(1), 16-34.

Zheng, Y., Liu, F., Hsieh, H.-P., 2013. U-Air: when urban air quality inference meets big data, *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, Chicago, Illinois, USA.





# Appendices

---

**Appendix A: Supplementary material for chapter 3****NO<sub>2</sub> regression model coefficients**Units  $c_0$  (Intercept):  $\mu\text{g m}^{-3}$ Units  $c_1$  ( $S_{WE}$ ):  $\mu\text{g m}^{-3}/\text{count}$ Units  $c_2$  ( $S_{AE}$ ):  $\mu\text{g m}^{-3}/\text{count}$ Units  $c_3$  (T):  $\mu\text{g m}^{-3}/^{\circ}\text{C}$ Units  $c_4$  (RH):  $\mu\text{g m}^{-3}/\%$ Units  $c_5$  ( $O_3$ ):  $\mu\text{g m}^{-3}/\mu\text{g}\cdot\text{m}^{-3}$ **Table S1 Relation sensor ID and its network ID, which is used as reference in raw data**

<b>Sensor device ID</b>	<b>WiFi chip ID</b>
SD01	1184206
SD02	14560051
SD03	55303
SD04	54200
SD05	1184527
SD06	1184739
SD07	1183931
SD08	53788
SD09	26296
SD10	1185325
SD11	1184453
SD12	717780
SD13	55300
SD14	13905017
SD15	1184838
SD16	54911

Table S2 Regression results for sensor devices

SD01 <sup>a</sup>	Intercept	S <sub>WE</sub>	S <sub>RE</sub>	T	RH <sup>b</sup>	O <sub>3</sub>	R <sup>2</sup>
Model A	1st period	455.38 ± 55.18	0.6977 ± 0.0649	-1.0835 ± 0.0970			0.471
	2nd period <sup>c</sup>	-6.04 ± 36.69	0.2475 ± 0.0488	-0.2343 ± 0.0604			0.199
Model B	1st period	715.45 ± 59.71	0.8394 ± 0.0592	-1.4811 ± 0.1001	0.5326 ± 0.0743		0.612
	2nd period <sup>c</sup>	2.24 ± 43.51	0.2469 ± 0.0490	-0.2431 ± 0.0654	0.0280 ± 0.0782		0.200
Model C	1st period	827.92 ± 87.54	0.8688 ± 0.0680	-1.5498 ± 0.1262	-1.6344 ± 0.3130		0.556
	2nd period <sup>c</sup>	-173.77 ± 64.95	0.3000 ± 0.0499	-0.1698 ± 0.0618	1.5927 ± 0.5177		0.265
Model D	1st period	790.88 ± 82.04	0.8707 ± 0.0635	-1.5645 ± 0.1178	0.4513 ± 0.0958		0.617
	2nd period <sup>c</sup>	-178.93 ± 64.10	0.3133 ± 0.0497	-0.2007 ± 0.0628	2.1055 ± 0.5715	0.1650 ± 0.0827	0.292
Model E	1st period	274.85 ± 78.12	0.3186 ± 0.0703	-0.4805 ± 0.1346	-0.5447 ± 0.2820	-0.5349 ± 0.0503	0.788
	2nd period <sup>c</sup>	56.69 ± 54.19	0.2864 ± 0.0371	-0.3343 ± 0.0490	1.4917 ± 0.4309	-0.1120 ± 0.0686	0.611
<sup>a</sup> Alphasense N02-B42F sensor; used in previous experiments for more than one year							
<sup>b</sup> RH sensor overestimates and often saturated at 100%							
<sup>c</sup> Only 42% uptime in 2 <sup>nd</sup> calibration period.							
SD02 <sup>a</sup>	Intercept	S <sub>WE</sub>	S <sub>RE</sub>	T	RH	O <sub>3</sub>	R <sup>2</sup>
Model A	1st period	355.92 ± 65.74	0.8862 ± 0.0621	-1.2633 ± 0.0921			0.615
	2nd period	303.68 ± 86.54	0.2770 ± 0.0667	-0.5599 ± 0.1034			0.133
Model B	1st period	624.53 ± 85.42	0.8686 ± 0.0583	-1.5077 ± 0.1017	0.3916 ± 0.0863		0.665
	2nd period	629.53 ± 97.17	0.3356 ± 0.0624	-0.9477 ± 0.1159	0.3625 ± 0.0615		0.265
Model C	1st period	502.09 ± 109.36	0.9007 ± 0.0624	-1.4001 ± 0.1229	-0.5684 ± 0.3410		0.623
	2nd period	68.85 ± 147.75	0.2973 ± 0.0671	-0.3864 ± 0.1357	0.8454 ± 0.4327		0.150
Model D	1st period	589.20 ± 105.35	0.8618 ± 0.0596	-1.4742 ± 0.1174	0.2142 ± 0.3720	0.4204 ± 0.1000	0.666
	2nd period	34.28 ± 123.80	0.4429 ± 0.0584	-0.6025 ± 0.1161	2.8976 ± 0.4263	0.5956 ± 0.0651	0.407
Model E	1st period	-87.90 ± 101.40	0.3690 ± 0.0645	-0.2424 ± 0.1460	0.1739 ± 0.2770	-0.6170 ± 0.1234	0.816
	2nd period	-174.15 ± 107.47	0.4075 ± 0.0496	-0.3524 ± 0.1023	3.8518 ± 0.3769	0.2585 ± 0.0672	-0.3428 ± 0.0390
<sup>a</sup> Alphasense N02-B42F sensor; used in previous experiments for more than one year							

<b>SD03</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>AE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-228.65 ± 137.58	1.0877 ± 0.0578	-0.8029 ± 0.1113			0.720
	2nd period	-470.06 ± 98.31	0.8521 ± 0.0388	-0.4193 ± 0.0772			0.709
Model B	1st period	-1335.96 ± 157.68	1.2551 ± 0.0482	-0.1132 ± 0.1127	-0.6560 ± 0.0686		0.832
	2nd period	-991.61 ± 161.21	0.8898 ± 0.0386	-0.0591 ± 0.1168	-0.1618 ± 0.0404		0.730
Model C	1st period	-972.80 ± 115.40	1.1445 ± 0.0410	-0.3343 ± 0.0878	1.7279 ± 0.1455		0.862
	2nd period	-913.18 ± 132.27	0.8192 ± 0.0375	-0.0765 ± 0.1031	0.8840 ± 0.1867		0.738
Model D	1st period	-1272.13 ± 137.05	1.2045 ± 0.0425	-0.1492 ± 0.0979	1.2690 ± 0.1867	-0.2944 ± 0.0798	0.874
	2nd period	-1050.59 ± 159.66	0.8448 ± 0.0410	0.0095 ± 0.1172	0.6707 ± 0.2328	-0.0758 ± 0.0497	0.741
Model E	1st period	-818.09 ± 120.96	0.8961 ± 0.0487	-0.1706 ± 0.0782	0.5898 ± 0.1678	-0.5387 ± 0.0695	0.920
	2nd period	-728.05 ± 108.84	0.8202 ± 0.0275	-0.1908 ± 0.0795	1.0731 ± 0.1579	-0.2465 ± 0.0350	0.885
<hr/>							
<b>SD04</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>AE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-968.20 ± 145.13	0.9138 ± 0.0538	-0.1237 ± 0.1254			0.689
	2nd period	-371.22 ± 144.45	0.9786 ± 0.0500	-0.6833 ± 0.1329			0.668
Model B	1st period	-1729.95 ± 119.61	1.1641 ± 0.0430	0.2736 ± 0.0939	-0.5386 ± 0.0444		0.848
	2nd period	-1190.28 ± 141.99	1.0625 ± 0.0413	-0.0659 ± 0.1236	-0.4225 ± 0.0414		0.783
Model C	1st period	-1044.89 ± 110.06	1.0490 ± 0.0427	-0.2245 ± 0.0954	1.4562 ± 0.1412		0.823
	2nd period	-864.22 ± 116.48	0.9909 ± 0.0378	-0.3182 ± 0.1048	1.5499 ± 0.1269		0.811
Model D	1st period	-1613.28 ± 153.33	1.1499 ± 0.0445	0.1818 ± 0.1204	-0.4442 ± 0.0896		0.850
	2nd period	-1055.65 ± 131.76	1.0203 ± 0.0384	-0.1723 ± 0.1144	-0.1639 ± 0.0561		0.819
Model E	1st period	-1129.35 ± 115.34	0.8046 ± 0.0426	0.1830 ± 0.0848	-0.3285 ± 0.1936	-0.3671 ± 0.0308	0.926
	2nd period	-848.14 ± 97.58	0.8909 ± 0.0298	-0.1992 ± 0.0836	1.5326 ± 0.1378	-0.3227 ± 0.0427	0.904

<b>SD05</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-155.10 ± 197.19	0.8368 ± 0.0743	-0.6841 ± 0.1768			0.478
	2nd period	475.82 ± 194.53	0.9137 ± 0.0542	-1.2719 ± 0.1730			0.602
Model B	1st period	-1953.53 ± 246.66	1.1485 ± 0.0672	0.5047 ± 0.1881	-0.9840 ± 0.1050		0.680
	2nd period	-805.01 ± 261.61	1.0611 ± 0.0538	-0.3549 ± 0.2090	-0.6526 ± 0.0988		0.677
Model C	1st period	-1056.05 ± 162.02	1.0371 ± 0.0562	-0.1946 ± 0.1340	2.3488 ± 0.2045		0.732
	2nd period	-983.97 ± 191.54	0.9821 ± 0.0414	-0.2015 ± 0.1588	2.3771 ± 0.1997		0.774
Model D	1st period	-1623.07 ± 222.70	1.1235 ± 0.0592	0.2088 ± 0.1715	1.7161 ± 0.2649	-0.4430 ± 0.1245	0.754
	2nd period	-1162.98 ± 221.80	1.0114 ± 0.0452	-0.0756 ± 0.1771	2.1686 ± 0.2386	-0.1564 ± 0.0989	0.777
Model E	1st period	-1079.04 ± 158.48	0.7104 ± 0.0522	0.2328 ± 0.1174	0.5648 ± 0.2032	-0.8305 ± 0.0906	0.886
	2nd period	-1067.82 ± 174.06	0.8927 ± 0.0371	-0.0218 ± 0.1389	2.4442 ± 0.1887	-0.4412 ± 0.0818	0.863

<b>SD06</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-141.88 ± 158.37	0.6136 ± 0.0607	-0.5241 ± 0.1168			0.440
	2nd period	437.30 ± 151.50	0.8025 ± 0.0589	-1.2130 ± 0.1582			0.493
Model B	1st period	-931.37 ± 123.99	1.2158 ± 0.0619	-0.4780 ± 0.0800	-0.7288 ± 0.0555		0.740
	2nd period	-300.44 ± 174.06	0.9395 ± 0.0566	-0.7145 ± 0.1600	-0.4714 ± 0.0692		0.592
Model C	1st period	-639.87 ± 102.28	1.0652 ± 0.0470	-0.6367 ± 0.0721	2.3781 ± 0.1504		0.790
	2nd period	-581.47 ± 122.97	0.9636 ± 0.0413	-0.5853 ± 0.1151	2.6484 ± 0.1756		0.768
Model D	1st period	-824.79 ± 106.47	1.1850 ± 0.0529	-0.5839 ± 0.0695	1.6737 ± 0.2198	-0.3069 ± 0.0728	0.812
	2nd period	-666.44 ± 134.13	0.9811 ± 0.0427	-0.5242 ± 0.1212	2.4866 ± 0.2035	-0.0941 ± 0.0604	0.771
Model E	1st period	-463.82 ± 73.02	0.8150 ± 0.0426	-0.4419 ± 0.0459	0.8318 ± 0.1531	-0.5519 ± 0.0499	0.923
	2nd period	-592.51 ± 107.94	0.8732 ± 0.0358	-0.4531 ± 0.0976	2.6967 ± 0.1647	-0.2927 ± 0.0522	0.853

<b>SD07</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH<sup>a</sup></b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-576.41 ± 188.25	0.9615 ± 0.0716	-0.4811 ± 0.1520			0.573
	2nd period	-239.15 ± 155.74	0.8866 ± 0.0486	-0.6834 ± 0.1418			0.645
Model B	1st period	-576.41 ± 188.25	0.9615 ± 0.0716	-0.4811 ± 0.1520			0.573
	2nd period	-239.15 ± 155.74	0.8866 ± 0.0486	-0.6834 ± 0.1418			0.645
Model C	1st period	-1217.57 ± 144.34	1.1305 ± 0.0528	-0.1642 ± 0.1110	1.9435 ± 0.1678		0.788
	2nd period	-977.93 ± 145.57	0.8717 ± 0.0393	-0.0987 ± 0.1284	1.6673 ± 0.1647		0.769
Model D	1st period	-1217.57 ± 144.34	1.1305 ± 0.0528	-0.1642 ± 0.1110	1.9435 ± 0.1678		0.788
	2nd period	-977.93 ± 145.57	0.8717 ± 0.0393	-0.0987 ± 0.1284	1.6673 ± 0.1647		0.769
Model E	1st period	-578.07 ± 142.70	0.7891 ± 0.0606	-0.3243 ± 0.0934	1.7254 ± 0.1405	-0.2656 ± 0.0330	0.858
	2nd period	-495.36 ± 120.13	0.7724 ± 0.0316	-0.3963 ± 0.1025	2.3365 ± 0.1401	-0.2254 ± 0.0198	0.863
<sup>a</sup> RH sensor not working							
<b>SD08</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	231.44 ± 103.68	1.0802 ± 0.0639	-1.2514 ± 0.1086			0.677
	2nd period	428.20 ± 110.91	1.0221 ± 0.0609	-1.3582 ± 0.1103			0.600
Model B	1st period	-521.55 ± 174.37	1.1806 ± 0.0618	-0.7141 ± 0.1443	-0.4831 ± 0.0937		0.730
	2nd period	141.16 ± 175.82	1.0604 ± 0.0631	-1.1578 ± 0.1454	-0.0651 ± 0.0311		0.609
Model C	1st period	-798.25 ± 114.09	1.1319 ± 0.0454	-0.5061 ± 0.0995	2.4721 ± 0.2100		0.840
	2nd period	-941.92 ± 168.22	0.9603 ± 0.0505	-0.2244 ± 0.1480	2.5145 ± 0.2593		0.731
Model D	1st period	-1129.69 ± 139.87	1.1835 ± 0.0454	-0.2705 ± 0.1136	2.2559 ± 0.2085	-0.2704 ± 0.0716	0.855
	2nd period	-983.10 ± 189.26	0.9685 ± 0.0534	-0.1975 ± 0.1586	2.4876 ± 0.2659	-0.0127 ± 0.0265	0.731
Model E	1st period	-725.55 ± 113.06	0.8481 ± 0.0478	-0.2249 ± 0.0860	1.2801 ± 0.1849	-0.4709 ± 0.0577	0.918
	2nd period	-685.96 ± 131.35	0.8376 ± 0.0377	-0.2914 ± 0.1089	2.5194 ± 0.1824	-0.1211 ± 0.0196	0.874

<b>SD09</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	100.52 ± 221.21	0.8669 ± 0.0671	-0.8952 ± 0.1979			0.556
	2nd period	407.81 ± 127.41	0.9154 ± 0.0458	-1.1897 ± 0.1159			0.675
Model B	1st period	-1138.92 ± 172.00	1.1781 ± 0.0498	-0.1707 ± 0.1407	-0.8205 ± 0.0609		0.810
	2nd period	-132.85 ± 146.09	1.0685 ± 0.0488	-0.8851 ± 0.1171	-0.2933 ± 0.0477		0.729
Model C	1st period	-332.23 ± 109.76	1.1460 ± 0.0353	-0.8613 ± 0.0965	2.4841 ± 0.1183		0.895
	2nd period	-504.18 ± 113.38	1.0011 ± 0.0334	-0.5837 ± 0.0943	2.0206 ± 0.1492		0.834
Model D	1st period	-586.25 ± 132.75	1.1794 ± 0.0358	-0.6738 ± 0.1103	2.0415 ± 0.1799	-0.2192 ± 0.0687	0.903
	2nd period	-688.42 ± 119.56	1.0694 ± 0.0368	-0.4885 ± 0.0944	1.8326 ± 0.1522	-0.1460 ± 0.0380	0.846
Model E	1st period	-383.42 ± 107.85	0.8973 ± 0.0424	-0.5253 ± 0.0892	1.1754 ± 0.1726	-0.4695 ± 0.0613	0.939
	2nd period	-498.89 ± 100.31	0.9728 ± 0.0319	-0.5403 ± 0.0778	2.1983 ± 0.1309	-0.2250 ± 0.0323	0.897

<b>SD10</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	342.04 ± 94.07	0.8221 ± 0.0657	-1.1629 ± 0.1206			0.503
	2nd period	417.68 ± 78.62	0.8047 ± 0.0546	-1.2119 ± 0.1009			0.520
Model B	1st period	-89.45 ± 187.91	0.9168 ± 0.0738	-0.8859 ± 0.1583	-0.2824 ± 0.1071		0.524
	2nd period	103.71 ± 118.52	0.8641 ± 0.0558	-0.9951 ± 0.1164	-0.2487 ± 0.0717		0.547
Model C	1st period	-847.45 ± 133.34	1.1001 ± 0.0566	-0.5102 ± 0.1108	2.9678 ± 0.2803		0.712
	2nd period	-784.93 ± 122.97	0.8745 ± 0.0432	-0.3272 ± 0.1113	3.2652 ± 0.2889		0.706
Model D	1st period	-1152.70 ± 175.33	1.1668 ± 0.0611	-0.3120 ± 0.1325	2.9112 ± 0.2760	-0.2147 ± 0.0820	0.725
	2nd period	-862.03 ± 131.60	0.8947 ± 0.0449	-0.2759 ± 0.1154	3.1490 ± 0.2968	-0.0950 ± 0.0593	0.710
Model E	1st period	-825.25 ± 115.40	0.7707 ± 0.0478	-0.1058 ± 0.0867	1.8251 ± 0.1930	-0.4975 ± 0.0564	0.886
	2nd period	-622.53 ± 103.17	0.8094 ± 0.0352	-0.3689 ± 0.0890	3.2492 ± 0.2283	-0.2528 ± 0.0475	0.830

<b>SD11</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>ΔE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	338.42 ± 80.88	0.9823 ± 0.0665	-1.2246 ± 0.1025			0.608
	2nd period	748.59 ± 74.96	0.9642 ± 0.0547	-1.5368 ± 0.0924			0.622
Model B	1st period	0.26 ± 133.88	1.0444 ± 0.0675	-0.9995 ± 0.1229	-0.2995 ± 0.0961		0.633
	2nd period	752.43 ± 95.23	0.9629 ± 0.0587	-1.5387 ± 0.0973	0.0038 ± 0.0575		0.622
Model C	1st period	-962.71 ± 126.96	1.0735 ± 0.0485	-0.3309 ± 0.1070	3.4356 ± 0.2980		0.799
	2nd period	30.62 ± 145.29	1.0385 ± 0.0526	-1.0668 ± 0.1198	1.8190 ± 0.3228		0.674
Model D	1st period	-1109.75 ± 139.25	1.1055 ± 0.0495	-0.2339 ± 0.1128	3.3191 ± 0.2972	-0.1693 ± 0.0709	0.807
	2nd period	33.02 ± 143.00	0.9974 ± 0.0539	-1.0453 ± 0.1182	2.2205 ± 0.3501	0.1582 ± 0.0580	0.686
Model E	1st period	-480.10 ± 118.32	0.7539 ± 0.0490	-0.3363 ± 0.0839	1.6813 ± 0.2670	-0.3806 ± 0.0560	0.895
	2nd period	99.69 ± 109.82	0.9454 ± 0.0416	-1.0242 ± 0.0907	2.4400 ± 0.2692	-0.0973 ± 0.0494	-0.2625 ± 0.0222

<b>SD12</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>ΔE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-375.21 ± 197.57	0.7775 ± 0.0611	-0.4837 ± 0.1851			0.545
	2nd period	-406.98 ± 191.77	0.8879 ± 0.0500	-0.5767 ± 0.1841			0.652
Model B	1st period	-1332.74 ± 156.87	1.1032 ± 0.0497	0.0257 ± 0.1345	-0.6993 ± 0.0561		0.783
	2nd period	-1248.39 ± 178.05	0.9608 ± 0.0414	0.0870 ± 0.1644	-0.4312 ± 0.0437		0.770
Model C	1st period	-819.17 ± 126.64	1.0416 ± 0.0420	-0.4203 ± 0.1154	2.0988 ± 0.1400		0.824
	2nd period	-800.71 ± 148.10	0.9405 ± 0.0379	-0.3286 ± 0.1402	1.6465 ± 0.1364		0.804
Model D	1st period	-1074.88 ± 140.40	1.0961 ± 0.0430	-0.2346 ± 0.1219	1.4954 ± 0.2136	-0.2799 ± 0.0770	0.840
	2nd period	-1012.78 ± 166.26	0.9545 ± 0.0377	-0.1466 ± 0.1541	1.2583 ± 0.1985	-0.1562 ± 0.0589	0.811
Model E	1st period	-595.45 ± 113.66	0.7813 ± 0.0435	-0.2757 ± 0.0908	0.8578 ± 0.1697	-0.4865 ± 0.0605	0.912
	2nd period	-701.86 ± 121.46	0.8586 ± 0.0280	-0.3051 ± 0.1111	1.6906 ± 0.1460	-0.2922 ± 0.0434	-0.2300 ± 0.0172



<b>SD13</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-1703.40 ± 201.83	0.8218 ± 0.0583	0.5544 ± 0.1554			0.605
	2nd period	-1088.31 ± 189.21	0.8631 ± 0.0504	-0.0632 ± 0.1732			0.642
Model B	1st period	-1826.17 ± 148.81	1.1334 ± 0.0515	0.3588 ± 0.1156		-0.5732 ± 0.0523	0.788
	2nd period	-1161.34 ± 190.56	0.8856 ± 0.0497	0.0550 ± 0.1729		-0.1936 ± 0.0589	0.660
Model C	1st period	-872.76 ± 146.63	1.1012 ± 0.0437	-0.4577 ± 0.1269	2.3418 ± 0.1732		0.829
	2nd period	-968.33 ± 167.16	0.8761 ± 0.0445	-0.1315 ± 0.1532	1.1078 ± 0.1454		0.722
Model D	1st period	-1074.57 ± 179.99	1.1294 ± 0.0458	-0.3058 ± 0.1490	1.8671 ± 0.3032	-0.1561 ± 0.0822	0.834
	2nd period	-999.93 ± 174.21	0.8800 ± 0.0450	-0.1057 ± 0.1584	1.0664 ± 0.1587	-0.0381 ± 0.0582	0.723
Model E	1st period	-594.35 ± 134.76	0.7795 ± 0.0444	-0.2874 ± 0.1062	1.0126 ± 0.2282	-0.4704 ± 0.0645	0.916
	2nd period	-505.72 ± 107.36	0.8246 ± 0.0271	-0.4485 ± 0.0964	2.1700 ± 0.1113	-0.2329 ± 0.0363	0.901

<b>SD14</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>VE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	162.64 ± 165.94	0.8156 ± 0.0903	-0.9075 ± 0.1248			0.459
	2nd period	-3.20 ± 202.78	0.8580 ± 0.0540	-0.8237 ± 0.1811			0.556
Model B	1st period	369.33 ± 139.19	1.0602 ± 0.0807	-1.2825 ± 0.1134		-0.6434 ± 0.0819	0.636
	2nd period	-1011.65 ± 198.00	1.0253 ± 0.0480	-0.1369 ± 0.1663		-0.4382 ± 0.0452	0.695
Model C	1st period	19.56 ± 91.93	1.1888 ± 0.0544	-1.1987 ± 0.0709	2.4905 ± 0.1454		0.837
	2nd period	-1147.64 ± 153.38	0.9569 ± 0.0366	-0.0244 ± 0.1311	2.1478 ± 0.1342		0.802
Model D	1st period	8.09 ± 97.95	1.1860 ± 0.0552	-1.1889 ± 0.0766	2.5401 ± 0.2039	0.0268 ± 0.0770	0.837
	2nd period	-1278.51 ± 159.07	0.9905 ± 0.0383	0.0621 ± 0.1333	1.8680 ± 0.1693	-0.1217 ± 0.0460	0.809
Model E	1st period	114.64 ± 71.51	0.8144 ± 0.0527	-0.8532 ± 0.0635	1.2001 ± 0.1929	-0.4387 ± 0.0705	0.915
	2nd period	-844.54 ± 120.58	0.9049 ± 0.0287	-0.1972 ± 0.0992	2.2316 ± 0.1266	-0.2564 ± 0.0350	0.898

<b>SD15</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>AE</sub></b>	<b>T</b>	<b>RH<sup>a</sup></b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	1211.20 ± 242.16	0.9008 ± 0.1180	-1.8984 ± 0.2883			0.298
	2nd period	1455.17 ± 155.20	1.2443 ± 0.0810	-2.4648 ± 0.1843			0.554
Model B	1st period	911.69 ± 319.97	0.9893 ± 0.1330	-1.7240 ± 0.3122	-0.2561 ± 0.1797		0.308
	2nd period	1455.17 ± 155.20	1.2443 ± 0.0810	-2.4648 ± 0.1843			0.554
Model C	1st period	-166.53 ± 139.22	1.8265 ± 0.0748	-1.7541 ± 0.1448	4.8106 ± 0.2373		0.825
	2nd period	-438.20 ± 143.92	1.4576 ± 0.0516	-1.1488 ± 0.1363	3.6043 ± 0.2039		0.830
Model D	1st period	-104.50 ± 169.26	1.8111 ± 0.0786	-1.7939 ± 0.1576	4.8373 ± 0.2413,	0.0596 ± 0.0921	0.825
	2nd period	-438.20 ± 143.92	1.4576 ± 0.0516	-1.1488 ± 0.1363	3.6043 ± 0.2039		0.830
Model E	1st period	-56.70 ± 134.13	1.2676 ± 0.0865	-1.2255 ± 0.1397	3.1038 ± 0.2705	-0.3226 ± 0.0356	0.891
	2nd period	-217.54 ± 133.72	1.2729 ± 0.0539	-1.1467 ± 0.1228	3.7105 ± 0.1844	-0.1401 ± 0.0207	0.863

<sup>a</sup> RH sensor breaks down after July 25

<b>SD16</b>	<b>Intercept</b>	<b>S<sub>WE</sub></b>	<b>S<sub>AE</sub></b>	<b>T</b>	<b>RH</b>	<b>O<sub>3</sub></b>	<b>R<sup>2</sup></b>
Model A	1st period	-594.31 ± 220.12	0.8007 ± 0.0704	-0.3192 ± 0.1976			0.485
	2nd period <sup>a</sup>	-254.68 ± 307.78	0.3469 ± 0.0885	-0.1361 ± 0.2747			0.275
Model B	1st period	-1537.42 ± 194.12	1.1674 ± 0.0655	0.1164 ± 0.1584	-0.5503 ± 0.0550		0.696
	2nd period <sup>a</sup>	-1053.52 ± 346.39	0.5320 ± 0.0926	0.3510 ± 0.2752	-0.2220 ± 0.0601		0.452
Model C	1st period	-1045.41 ± 129.96	1.2206 ± 0.0476	-0.4227 ± 0.1144	2.4971 ± 0.1466		0.829
	2nd period <sup>a</sup>	-1118.84 ± 294.51	0.5547 ± 0.0805	0.3426 ± 0.2357	1.3564 ± 0.2612		0.558
Model D	1st period	-1215.51 ± 146.15	1.2551 ± 0.0490	-0.3038 ± 0.1229	2.1742 ± 0.1972	-0.1333 ± 0.0555	0.836
	2nd period <sup>a</sup>	-1156.53 ± 316.09	0.5629 ± 0.0846	0.3693 ± 0.2498	1.2518 ± 0.3962	-0.0290 ± 0.0819	0.560
Model E	1st period	-623.06 ± 135.29	0.8844 ± 0.0575	-0.3786 ± 0.0993	1.5146 ± 0.1753	-0.2937 ± 0.0482	0.894
	2nd period <sup>a</sup>	-553.67 ± 329.07	0.7349 ± 0.0897	-0.2996 ± 0.2928	1.7739 ± 0.3817	-0.2115 ± 0.0894	0.663

<sup>a</sup> Only 18% uptime in 2<sup>nd</sup> calibration period

## Appendix B: Supplementary material for chapter 4



### Questionnaire on Formal and Informal Environmental Sensing Data

#### Introduction

Small and low cost sensors provide alternative approaches for citizens to collect environmental data (**informal**). These data potentially can complement the data collected by official organizations (**formal**). This survey is designed to analyse the opinions from citizens and experts on: 1. environmental sensing data collected by citizens and official organizations respectively, 2. the integration of these data.

The survey results are planned to be part of a scientific paper.

Your privacy is guaranteed.

Best regards,  
Qijun Jiang, Arnold Bregt and Lammert Kooistra  
Wageningen University & Research

**\* Please choose one role (expert or citizen) and fill in the questionnaire according to this role.**

- I am an expert on environment sensing.
- I am a citizen interested in environmental sensing.

**\* Do you have experience on following environmental sensing applications. (Tick one or more boxes)**

- Air monitoring
- Noise monitoring
- Soil monitoring
- Vegetation monitoring
- Waste monitoring
- Water monitoring
- Weather monitoring
- I have no experience
- Other (please specify)

**\* What is your role related to environmental sensing? (Tick one or more boxes)**

- Environmental policy maker
- Environmental researcher
- Sensor researcher
- Data scientist
- Data provider
- Hardware developer
- Software developer
- Citizen scientist
- Citizen
- Data user
- Just interested
- Other (please specify)



## Questionnaire on Formal and Informal Environmental Sensing Data

### II. Citizen environmental sensing data (informal)

**\*What do you think about citizen environmental sensing data (informal)?**

	Fully disagree	Disagree	Neutral	Agree	Fully agree	I don't know.
It has sufficient accuracy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is easily accessible.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It has sufficient coverage.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is real time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is sufficient to get environmental information.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is produced by small sensors.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is produced by mobile sensors.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is consistent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is complete.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is measured frequently.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**\* It is important that citizens collect environmental data by themselves.**

Fully disagree	Disagree	Neutral	Agree	Fully agree	I don't know.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**\* What is the value of environmental sensor data provided by citizens (informal)? (Tick one or more boxes)**

- It helps to raise public awareness of environmental problems.
- It can influence policy making.
- It can supply complementary information to formal environmental information.
- It can supply high spatial resolution data.
- It can monitor the environment in (near) real time.
- Other (please specify)

**\* Citizen sensing offers an alternative way to collect environmental data.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

**\* Citizen sensing data are underestimated because of lower trustworthiness compared to official data.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

**\*It is difficult for citizens to calibrate low cost sensors.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

**\*It is difficult for citizens to maintain low cost sensors.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

**\*Citizens need training or support from experts in following aspects.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

Sensor calibration                       

Proper use of sensors                       

Data analysis                       

**\*Do you think the citizen sensing data can contribute to better understanding of our environment.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

**\*Citizen sensing data need to be analysed by experts to derive meaningful environmental information.**

Fully disagree    Disagree    Neutral    Agree    Fully agree    I don't know.

**\*For which aspect(s) do you think citizen sensing data (informal) could complement formal data?  
(Tick one or more boxes)**

- Data accuracy
- Data amount
- Spatial coverage
- Measurement frequency
- Accessibility
- Consistency
- Completeness
- Other (please specify)

## Questionnaire on Formal and Informal Environmental Sensing Data

### III. Integration of formal and informal environmental sensing data

**\*Integration of citizen sensing data (informal) with formal data can improve data in the following aspects. (Tick one or more boxes)**

- Data accuracy
- Data amount
- Spatial coverage
- Measurement frequency
- Other (please specify)

**\*The following aspects hinder the integration of citizen sensing data (informal) and formal environmental sensing data. (Tick one or more boxes)**

- Accuracy
- Accessibility
- Completeness
- Consistency
- Standardization
- Metadata
- Privacy
- Other (please specify)





If you have additional comments, please let us know:

If you want to be informed about the results of this survey, please enter your name and email address.

Full name

Email address

**Thank you for your contribution!**

We are deeply grateful for your participation. For further questions, please contact [qijun.jiang@wur.nl](mailto:qijun.jiang@wur.nl)

## Appendix C: Supplementary material for chapter 5

### Appendix C1: Expert validation formal and informal environmental sensing data integration

Name:

Date:

#### Introduction

The objective of this expert validation is to validate the conceptual approaches for formal and informal environmental sensing data integration. The informal environmental sensing data are measured by citizen(s) using sensors. The formal environmental sensing data are measured by official environmental institutes. The proposed five integration approaches illustrated below are individually elaborated later on (Figure 1).

Five indicators to be evaluated are:

**Information complement:** by integration, complementary information can be derived by increased data volume.

**Increased spatial coverage:** by integration, derived information covers more locations.

**Increased frequency:** by integration, increased measurements derives information which otherwise not be possible. For instance, only hourly information is available without integration, after integration, information by minute become possible.

**More timely:** by integration, information is available in a more timely way. For instance, information may only be available days after measuring for one data source, when integrating this data source with more timely data source, information may become available in seconds.

**Increased accuracy:** by integration, derived information has better accuracy than without integration.

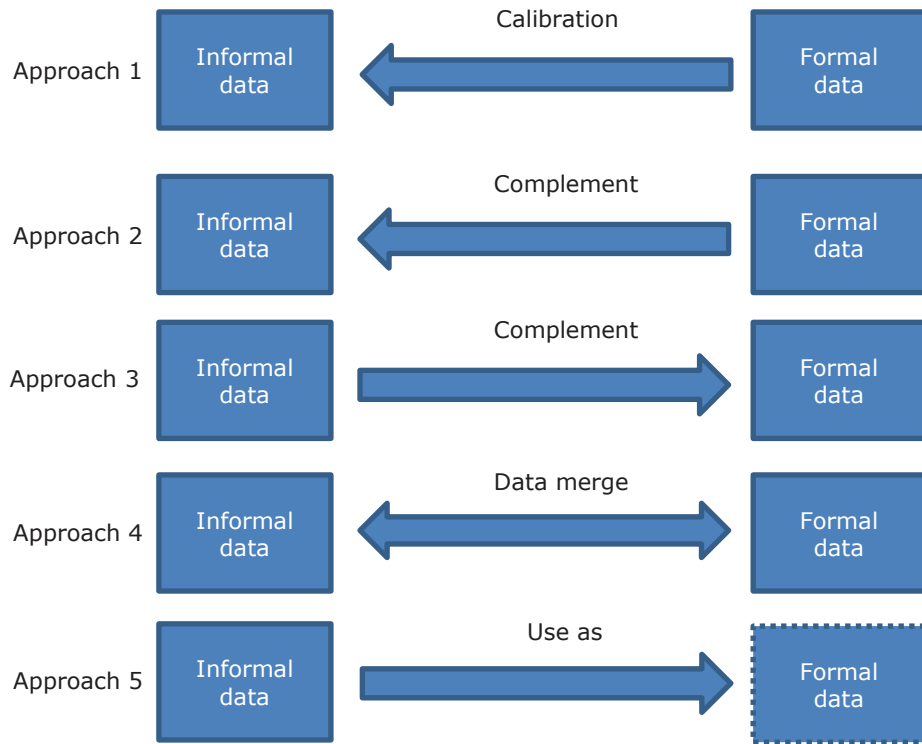


Figure 1: Overview of five conceptual approaches for formal and informal environmental sensing data integration

### Case study: Formal and informal air quality sensing data integration

In city of Amsterdam, there are 12 formal air quality monitoring stations operated by GGD Amsterdam or RIVM. Among them, there are 10 stations measuring  $\text{NO}_2$ . A project called Urban AirQ was conducted in 2016 with a group of citizens, experts and researchers using 16 low-cost sensor devices. The data produced are called informal data source. They were measured from June to August 2016 around two selected most polluted streets. The data were calibrated by collocating sensor devices to one formal station at Vondelpark before and after measurement campaign and validated by comparing data from one sensor device to a nearby formal station at Oude Schans.

We would like you to validate the value of the five formal and informal environmental sensing integration approaches for the air quality case study presented above by answering the questions for each approach.

#### Approach 1: use formal data to calibrate informal data

In this integration approach, formal data are used for informal data calibration often by co-locating sensors to official environmental monitoring station(s) in real environment as applied in the case study.

How do you rate the indicator in the table for this integration approach 1?

Indicator	Rating	Comment
Increased accuracy	1 2 3 4 5	

Rating:

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

### **Approach 2: Use formal data to complement informal data**

In this approach, formal data will be used to complement informal data. Formal data normally have higher quality, therefore formal data are considered to be used first to avoid unnecessary repeated measurements when planning informal data collection. In the case study, data coming from official station located in the research area can be integrated with informal data coming from distributed sensor devices.

How do you rate the indicators in the table for this integration approach 2?

Indicators	Rating	Comment
Information complement	1 2 3 4 5	
Increased spatial coverage	1 2 3 4 5	
Increased frequency	1 2 3 4 5	
More timely	1 2 3 4 5	
Increased accuracy	1 2 3 4 5	

Rating:

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

### **Approach 3: Use informal data to complement formal data**

In this integration approach, informal data will be used to complement formal data. This is mainly the function of informal data when formal data is not sufficient to provide needed environmental information due to sparse measurement locations and/or low measurement frequency. Therefore, these informal data are used to derive local environmental information timely when this is not possible by formal data sources. This is also the main purpose of the case study.

How do you rate the indicators in the table for this integration approach 3?

Indicators	Rating	Comment
Information complement	1 2 3 4 5	
Increased spatial coverage	1 2 3 4 5	
Increased frequency	1 2 3 4 5	
More timely	1 2 3 4 5	
Increased accuracy	1 2 3 4 5	

Rating:

1. Strongly disagree
2. Disagree
3. Neural
4. Agree
5. Strongly agree

#### **Approach 4: Merge formal and informal data**

In this approach, formal and informal data will be merged to a unified format as a new data source. This is different compared to the above approaches. In the earlier mentioned approaches, the information from informal and formal data are integrated to derive more comprehensive understanding of the environment which neither can achieve. But the data are not merged, informal and formal data still have different data accuracy, measurement frequency, format, standards and so on. Integrating by data merge create an unified data set. In the case study, it means merging the formal and informal data sources as one data set and use it for applications.

How do you rate the indicators in the table for this integration approach 4?

Indicators	Rating	Comment
Information complement	1 2 3 4 5	
Increased spatial coverage	1 2 3 4 5	
Increased frequency	1 2 3 4 5	
More timely	1 2 3 4 5	
Increased accuracy	1 2 3 4 5	

Rating:

1. Strongly disagree
2. Disagree
3. Neural
4. Agree
5. Strongly agree

#### **Approach 5: Use informal data as formal data**

In this approach, informal data are the only (or main) data source when there are no (or a few) formal data existing locally. The informal data are used as formal data

to provide environmental information. On a larger scale, these informal data can be integrated with other informal and formal data for national environmental information provision. In the case study, there are official stations measuring at a few locations in the city. To street level in the study area, the informal data are considered to be used as formal data to fill the gap if this approach is applied.

How do you rate the indicators in the table for this integration approach 5?

Indicators	Rating	Comment
Information complement	1 2 3 4 5	
Increased spatial coverage	1 2 3 4 5	
Increased frequency	1 2 3 4 5	
More timely	1 2 3 4 5	
Increased accuracy	1 2 3 4 5	

Rating:

1. Strongly disagree
2. Disagree
3. Neural
4. Agree
5. Strongly agree

**Appendix C2: Comments made by the experts in the expert validation**

Approaches	Indicators	Comments
Approach 1	Increased accuracy	<p>Without the formal measurements the sensor data would be meaningless.</p> <p>Low-cost sensors still suffer from severe issues such as drift, cross-sensitivity and sensitivity loss. In-field calibration next to a formal station is essential before application.</p>
Approach 2	Information complement	<p>Data volume formal data is low.</p> <p>More instruments mean more information. With one remark, the sensor data has to be reliable and accurate. In Urban AirQ the sensors were still not so good. In the near future they will likely improve strongly.</p>
	Increased spatial coverage	<p>Data volume formal data is low.</p> <p>In principle, more instruments mean more spatial coverage. But it makes more sense to use the official measurements for calibration and then you have to put a sensor at each official location and the official measurements do not add spatial coverage to the sensor measurements</p>
	Increased frequency	<p>Formal measurements are formally available in hourly values, but if you want you can get minute values, sensors are not going to add much. Unless you compare with formal NO<sub>2</sub> measurements with Palmes tubes which give month values. If these were complemented by sensors with minute values it would increase frequency considerably.</p> <p>This strongly depends on the response time of the low-cost sensor, and the required integration time to suppress measurement noise sufficiently.</p>
	More timely	<p>Formal data is available every hour, sensors are not going to be much faster.</p> <p>Both raw formal and informal data can be delivered near-real time. However, for both data streams holds that an analysis based on calibrated data will take more time.</p>

	Increased accuracy	Using formal data for on the fly calibration. Without the formal measurements the sensor data would be meaningless
Approach 3	Information complement	More instruments mean more information. With one remark, the sensor data has to be reliable and accurate. In Urban AirQ the sensors were still not so good. In the near future they will likely improve strongly.
	Increased spatial coverage	More instruments mean more spatial coverage.
	Increased frequency	Formal measurements are formally available in hourly values, but if you want you can get minute values, sensors are not going to add much. Unless you compare with formal NO <sub>2</sub> measurements with Palmes tubes which give month values. If these were complemented by sensors with minute values it would increase frequency considerably.  This strongly depends on the response time of the low-cost sensor, and the required integration time to suppress measurement noise sufficiently. For e.g. Palmes tubes the integration time is by far longer than the formal data frequency (weeks versus hours).
	More timely	formal data is available within one hour at <a href="http://luchtmeetnet.nl">luchtmeetnet.nl</a> .  Formal data is available every hour, sensors are not going to be much faster.  Both raw formal and informal data can be delivered near-real time. However, for both data streams holds that an analysis based on calibrated data will take more time.
	Increased accuracy	Without the formal measurements the sensor data would be meaningless. The other way around, the sensors may give you information about how representative a formal measurement point is. This increases the overall accuracy, in the



		<p>sense that model calculations may also become more accurate.</p> <p>I don't think adding informal data with formal data for a given location will improve the accuracy of the formal measurement.</p>
Approach 4	Information complement	When using appropriate data assimilation techniques
	Increased spatial coverage	When using appropriate data assimilation techniques
	Increased frequency	
	More timely	
	Increased accuracy	When using appropriate data assimilation techniques
Approach 4	<p>This is what we should do, but the integration should also include other information, like emission numbers. There is already an air quality model that combines measured data, emission numbers, chemical processes, weather information etc. Sensor data needs to be an integral part of this system and then it will enhance the spatial coverage and increase accuracy etc.</p>	
Approach 5	Information complement	Uncertain (i.e. informal) measurements are better than no measurements at all. However, given their large error bar and calibration issues, they should be used with great care.
	Increased spatial coverage	Uncertain (i.e. informal) measurements are better than no measurements at all. However, given their large error bar and calibration issues, they should be used with great care.
	Increased frequency	
	More timely	
	Increased accuracy	Depends strongly on the way informal data are calibrated on the fly (e.g. with the help of air quality models, or network analyses). If that is successful than accuracy issues might be tackled to an extent that informal data can be treated as formal data in one way or another.

		Uncertain (i.e. informal) measurements are better than no measurements at all. However, given their large error bar and calibration issues, they should be used with great care.
Approach 5	<p>I have a problem with seeing it like that for the Dutch situation where we have so much information from other sources combined into good models that we can use to fill in the gaps, see my comment above. Sensor data may be an important part of the system, but as for now they are just not good enough to come anywhere near the status of formal data.</p> <p>But I can imagine situations where there is very little information available (in Africa or South America or some parts of Asia) that sensor data may take up the role of formal data source, because there is simply no alternative.</p>	
Overall	<p>...to judge the different approaches it is really important to know the quality of both the formal and informal sensors. As informal sensors are, at the moment, in general less accurate, they first need to be corrected: the more informal sensors are available and the more ancillary data (for example about the environment) is available the better this correction and filtering of data will work. To achieve this, our strategy is data merging (approach 4). For example, we combine formal data, informal data and site characteristics (like solar radiation from satellites, AHN, Topographic information, actual land cover from satellites, etc.) to correct WOW stations. And within a pilot study regarding air quality we combine both formal and informal AQ-measurements with other data sources (traffic, AQ-models, environment) to assess the biases in informal networks and correlate the informal measurements with formal and ancillary data. For approach 4 we also use approach 1 first. Approach 5 is still far future for meteorological applications.</p>	

## Summary

Citizens contributing to science started even before scientist became a profession. Traditionally, citizens mainly contribute to data collection through their observations. With the development of sensor technology, and information and communication technology (ICT), citizens are increasingly able to use low-cost sensors to collect environmental data they could not do before. However, how good is the quality of the data they collect with low-cost sensors? How should the collected data be used? What is the role of traditional environmental monitoring facilities operated by authorized environmental organizations? These are examples of questions, which need to be answered in the changing landscape of environmental data collection in which citizens play an increasing role.

The environmental data collected by citizens using low-cost sensors are defined as informal environmental sensing data and the environmental data collected by official environmental organizations are defined as formal environmental sensing data in this thesis. The overall objective of this thesis is to assess opportunities for informal sensing data and their integration with formal sensing data to improve environmental information provision by investigating challenges and obstacles regarding technical and societal aspects. The thesis was guided by four main research questions: 1) How does a bottom-up citizen science project develop, and can it contribute to environmental monitoring? 2) How to calibrate low cost air quality sensor data, and how valuable are these data? 3) What are the perceptions of citizens and experts on formal and informal environmental sensing data and their integration potential? 4) How can informal and formal sensing data be integrated to provide enhanced environmental information?

In chapter 2, a bottom-up approach to develop sensing tools for urban environmental data collection by citizens is presented. By conducting an air quality monitoring case study, different steps of the bottom-up approach from selecting sensors, developing sensor devices to doing preliminary experiments were conducted. Low-cost electrochemical NO<sub>2</sub> sensors were proven to have the potential to provide valuable environmental information for urban environmental monitoring. However, it was clear that except proper hardware and software design, sensor calibration and data analysis were also important to reach the potential to provide valuable environmental information. All these aspects were actually quite challenging for citizens, especially the calibration, which is still technically complicated, and needs facilities for support. Therefore, broad cooperation including participants with a diverse background and expertise is recommended for such bottom-up citizen sensing communities. At the same time, fully utilizing offline and online facilities was recommended for a sustainable citizen sensing community.

As indicated by the preliminary experiments with electrochemical sensors in chapter 2, a more serious sensor calibration was needed to evaluate how useful the low-cost air quality sensing data are. Therefore, sensor calibration is the focus of chapter 3. An updated version of an informal sensor device was developed by mainly adding a temperature and relative humidity sensor to examine the influence of the changing environmental conditions. In total, 16 of these sensor devices were developed and installed on top of an official air quality monitoring station for calibration. This was conducted before the measurements campaign at citizens' homes started. A multiple linear regression model was developed for sensor calibration. Then, a measurement campaign lasting two months was conducted at locations selected based on citizens' concerns. After that, a recalibration was conducted to study the data drift and improve the calibration model. According to the results, the sensors have individually a varying performance. Therefore, sensors should be individually calibrated. A field calibration with an official measurement station under real environment conditions is essential and needs more study for different environmental situations. As part of the data analysis, data outliers need to be filtered properly. Sensor aging and data drift are problems, which need to be considered seriously for long measurement period. Recalibration is therefore considered crucial as demonstrated.

In chapter 3, it was shown that low-cost informal sensors were useful to provide valuable environmental information. However, is the value really recognized by citizens and experts? What are their opinions? Can formal and informal environmental sensing data be integrated? In chapter 4, a perception study was conducted to answer these questions. An online questionnaire was prepared to investigate the perceptions of citizens and experts on formal and informal environmental sensing data and their integration potential. Experts were mainly from official environmental monitoring organizations. Citizens were mainly from three citizen sensing related communities. According to the results, citizens and experts do have agreement. For instance, both citizens and experts were positive on formal environmental sensing data in terms of data accuracy and consistency. However, there were disagreements too. For instance, citizens and experts have different opinions on measurement frequency. Interestingly, a large part of both citizens and experts chose 'Neutral' about the data accuracy of informal environmental sensing data. Even though both experts and citizens recognized the value of informal environmental sensing data, it is still far to that potential due to challenges in aspects like standardization, accuracy, metadata and privacy.

In chapter 5, a conceptual framework for using formal and informal environmental sensing data for different purposes towards environmental issues is proposed. The integration of formal and informal environmental sensing data was found to be a research gap. Five approaches are proposed for integration: use formal data to calibrate informal data, use formal data source to complement informal data at information level, use informal data to complement formal data at information level, merge formal and informal data and use informal data as formal data. To evaluate these approaches, five success indicators were selected. A literature study and expert validation was used to evaluate the five approaches based on the success indicators. According to the evaluation, the proposed integration framework is

applicable. The calibration approach currently is used most often to evaluate the informal data quality. Very few studies have started to explore the data merging approach. In this study, the data quality was found important for data integration. A standard for data quality assurance according to purposes (legislation, education or personal decision making for instance) is recommended.

Chapter 6 discusses the main results of the thesis and reflects on the broader context. The research questions are discussed comprehensively. The reflection focuses on proximate sensing for various high spatial and temporal resolution environmental data provision, sensor networks, big data and smart cities, and citizen science and open science. Finally, recommendations for further research are presented.



## Acknowledgements

First, I would like to thank my promotor Arnold and co-promotor Lammert. Thank you, Arnold, for giving me the opportunity to do the research at Wageningen University & Research. Thank you for your inspiration and guidance during my PhD. Thank you for your patience and great support at the moments when I was lost. Thank you, Lammert, for being my co-promotor. Thank you for pushing me and motivating me running fast in my PhD journey. Thank you for your advices and great support.

I would like to express my gratitude to my collaborators: Bas, Christine, Dave, Edith, Emma, Frank, Gijs, Hester, Ivonne, Joost, Matthijs, Rob and Stefano. It was lucky to work with you in various projects which contributed to the PhD thesis.

Thanks to all my colleagues at Laboratory of Geo-information Science and Remote Sensing: Agnieszka, Aldo, Alvaro, Anne, Antoinette, Arend, Arun, Astrid, Benjamin B, Benjamin D, Benjamin K, Beyetmante Demas, Brice, Corne, Dainius, Danaë, Daniela, Devis, Diego, Eliakim, Erika R, Erika S, Eskender, Frans, Giulia, Gustavo, Harm, Jaap-Willem, Jalal, Jan C, Jan V, Joao, Johannes B, Johannes E, Johannes R, John S, John V, Jose, Juha, Kalkidan, Kim, Konstantin, Loïc, Lukasz, Marcello, Maria, Marian, Marston, Martin, Mathieu, Michael, Na, Nandika, Niki, Patric, Peter, Philip, Ricardo, Richard, Roberto, Roland, Ron, Rosa, Sabina, Samantha, Sarah, Shivangi, Simon, Srirama, Sylvain, Sytze, Titia, Truus, Valerio, Willy and Yang. Thank you for your help, the coffee/tea breaks, chats, drinks we had and so on. Special thanks to my paranymphs: Konstantin and Na.

I would like to thank my MSc students Thomas and Imke as well. It was my pleasure to work with you. I also learnt a lot from you.

Thank you, Margherita and Patrycja, for our science and art communication and collaboration. Thank you for the artwork design reflecting my research topics.

During my PhD research, I did some projects collaborating with quite some people from the Amsterdam Smart Citizens Lab, the European Citizen Science Association (ECSA) and other communities. There are too many names to be mentioned. Thank you for your cooperation.

最后，我要感谢我的家人。感谢父母给予的磨砺、爱与支持。感谢姐姐、姐夫和弟弟及其女友对父母的照顾和对我的理解，也要感谢侄女。感谢岳父岳母对我的鼓励与支持。感谢妻子的爱与陪伴，感谢儿子让我在家基本无法工作及带给我的快乐。

## List of publications

### Peer reviewed publications

**Jiang, Q.**, Kresin, F., Bregt, A.K., Kooistra, L., Pareschi, E., van Putten, E., Volten, H., Wesseling, J., 2016. Citizen Sensing for Improved Urban Environmental Monitoring. *Journal of Sensors* 2016, 9.

Mijling, B., **Jiang, Q.**, de Jonge, D., and Bocconi, S., 2018. Field calibration of electrochemical NO<sub>2</sub> sensors in a citizen science context. *Atmospheric Measurement Techniques*.

**Jiang, Q.**, Bregt, A.K., Kooistra, L., 2018. Formal and informal environmental sensing data and integration potential: Perceptions of citizens and experts. *Science of The Total Environment* 619–620, 1133-1142.

Eitzel, M.V., Cappadonna, J.L., Santos-Lang, C., Duerr, R.E., Virapongse, A., West, S.E., Kyba, C.C.M., Bowser, A., Cooper, C.B., Sforzi, A., Metcalfe, A.N., Harris, E.S., Thiel, M., Haklay, M., Ponciano, L., Roche, J., Ceccaroni, L., Shilling, F.M., Dörler, D., Heigl, F., Kiessling, T., Davis, B.Y., **Jiang, Q.**, 2017. Citizen Science Terminology Matters: Exploring Key Terms. *Citizen Science: Theory and Practice* 2(1).

### Other scientific contributions

**Jiang, Q.** ; Kresin, F. ; Kooistra, L. ; Bregt, A.K. (2015). Citizen sensing for improved urban environmental monitoring. In: *3<sup>rd</sup> Science for the Environment Conference*, Aarhus, Denmark, 1-2 October 2015.

**Jiang, Q.** ; Kooistra, L. ; Bregt, A.K. (2016). How to make use of citizen environmental sensing data? In: *First International ECSA Conference 2016 Citizen Science – Innovation in Open Science, Society and Policy*, Berlin, 19-21 May 2016.

**Jiang, Q.** ; Kooistra, L. ; Bregt, A.K. (2017). Is Citizen Environmental Sensing Valuable? In: *Netherlands Centre for Geodesy and Geo-Informatics (NCG) symposium*, Delft, The Netherlands, 2 November 2017.



## **Short biography**

Qijun Jiang was born and grew up in China. He studied Forest Resources Conservation and Recreation (Ecotourism) and achieved his bachelor's degree at Sichuan Agricultural University, China. He continued his study in Forest Resources Management (Forest Resources and Environmental Information Systems) and achieved his master's degree at the same university. During his master's studies, Qijun Jiang conducted a lot of forest inventory field work and data analysis using geographic information system and remote sensing. After that, he was awarded a scholarship to conduct PhD research in a new direction at Wageningen University and Research. By accomplishing his PhD, Qijun Jiang has built up expertise in sensor device development and calibration, as well as citizen science and environment monitoring. He developed his interests in smart city, open science, robotics and other domains as well.

Qijun Jiang is active in citizen science communities. He collaborated widely with Dutch and international citizen science communities in various ways.

His current research interests are related to sensing and geo-information science. This includes developing sensor hardware and software to collect, process, analyse and visualise data for domain applications.

## PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities).



### Review of literature (6 ECTS)

- Amsterdam Smart Citizens Lab meetings

### Writing of project proposal (4.5 ECTS)

- Sensor data integration for high spatial temporal resolution environmental information provision

### Post-graduate courses (4.9 ECTS)

- Spatio-temporal analysis and big-data processing using free and open source software; SENSE (2013, 2014)
- Spatial sampling for mapping; PE&RC and SENSE (2015)

### Competence strengthening / skills courses (3.5 ECTS)

- Improve your writing; Wageningen in'to languages
- Competence Assessment; WGS
- Entrepreneurship in and outside Science; WGS

### PE&RC Annual meetings, seminars and the PE&RC weekend (2.4 ECTS)

- PE&RC First years weekend (2013)
- PE&RC Day (2013)
- PE&RC Midterm weekend (2014)
- PE&RC Last years weekend (2017)

### **Discussion groups / local seminars / other scientific meetings (10.3 ECTS)**

- People first in tropical forests? Are we on the right track? (2013)
- Workshop smart inspectors: unmanned aerial systems for environmental monitoring (2013)
- Wageningen PhD symposium (2013)
- Intelligent sensor networks (2013, 2014)
- WGS PhD Workshop carousel (2015)
- National PhD day: 'Enhance your career!' (2015)
- Agro food robotics (GreenVision) (2015-2017)
- One wageningen robotics (2016)
- Making sense workshop: co-designing and generative tools (2016)
- Netherlands Centre for Geodesy and Geo-Informatics (NCG) symposium (2016)
- Landscape dynamics (2017)
- Open science: the national plan and you (2017)
- Citizen science lab: air pollution (2018)

### **International symposia, workshops and conferences (4.8 ECTS)**

- Science for the environment (2015)
- First International ECSA Conference (2016)
- Netherlands Centre for Geodesy and Geo-Informatics (NCG) symposium (2017)

### **Supervision of MSc students (3 ECTS)**

- Integrating Formal data and Volunteered Geographic Information





Qijun Jiang is supported by China Scholarship Council (CSC) for his PhD research.

The Amsterdam Smart Citizens Lab is supported by the municipality of Amsterdam and Creative Industries Fund NL. The Urban AirQ project was partly funded by a 2016 Stimulus Grant from the Amsterdam Institute for Advanced Metropolitan Solutions (AMS). The Urban AirQ project is also part of Making Sense, funded by European Union's Horizon 2020 research and innovation programme.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by Paula Baggen

Printed by ProefschriftMaken||DigiForce

