European Commission

HORIZON 2020

GUIDANCE DOCUMENTS ON MEASUREMENTS & MODELLING OF NOVEL AIR QUALITY POLLUTANTS:

URBAN MAPPING & CITIZEN SCIENCE

RI-URBANS

Research Infrastructures Services Reinforcing Air Quality Monitoring Capacities in European Urban & Industrial AreaS (GA n. 101036245)

By

vito UNIVERSITY OF BIRMINGHAM DCMR milieudienst Rijnmond

07/11/2024

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

Authors: Gerard Hoek (Utrecht University, UU), Jules Kerckhoffs (Utrecht University, UU), Martine van Poppel (Flemish Institute for Technological Research, VITO), Jelle Hofman (Flemish Institute for Technological Research, VITO), Roy Harrison (University of Birmingham, UoB), Sef van den Elshout (DCMR Environmental Protection Agency)

Reviewers: Tuukka Petäjä (UHEL), Xavier Querol (CSIC), Elli Suhonen (FMI)

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

Table of Contents

ABBREVIATIONS

ACTRIS	Aerosols, Clouds and Trace gases Research InfraStructure
AQ	Air quality
AQMN	Air quality monitoring network
AQMS	Air quality monitoring stations
BC	Black carbon
COV	Coefficient of variation
EEA	European Environmental Agency
EURO	European emission standards
GDPR	General data protection regulation
GIS	Geographic information systems
HRE	High resolution exposure
IDW	Inverse distance weighting
LCS	Low-cost sensor
LDSA	Lung deposited surface area
LUR	Land use regression
OPC	Optical particle counters
PM	Particulate matter
$PM_{2.5}$	Mass concentration of particles <2.5 µm
PM_{10}	Mass concentration of particles <10 µm
ppb	Parts per billion
RI-URBANS	Research Infrastructures Services Reinforcing Air Quality Monitoring Capacities in European Urban & Industrial AreaS EU project
UFP	Ultrafine particles

CHEMICAL SPECIES

NO_2	Nitrogen dioxide
NO_x	Nitrogen oxides ($NO+NO_2$)
O_3	Ozone
Pb	Lead

1. ABOUT THIS DOCUMENT

This document describes methods that air quality agencies, air quality researchers and other groups can use to develop fine spatial resolution maps of urban air pollution derived from monitoring. The methods described in this document are complimentary to routine monitoring with reference equipment at one or a few monitoring sites across the city. The described methods are also complimentary to deterministic dispersion models which are often applied by air quality agencies for regulatory purposes. One rationale to use more empirical methods is uncertainty on magnitude or location of emission sources. Combination of deterministic model and empirical models may further reduce uncertainty, as these methods have uncertainties in different domains. The spatially-resolved data collected by the methods presented, can be used to evaluate and improve deterministic dispersion models.

A more detailed description of methods and experiences with applying these methods can be found in RI-URBANS D13 (D2.5) and D14 (D2.6). D13 (D2.5) outlines the methods, distinguished into mobile/ fixed monitoring and with / without citizen involvement. These methods have been applied in pilot studies in RI-Urbans. The lessons learned in these pilots are included in the recommendation section of this document. These monitoring approaches were subsequently tested in three RI-URBANS pilot cities as part of WP4. Specifically, pilots were conducted in the cities Birmingham (UK), Rotterdam (The Netherlands) and Bucharest (Romania). The approaches included mobile monitoring with and without involvement of citizens and with and without fixed site measurements or low-cost sensors. The lessons learned from these pilot city initiatives are compiled in this ST derived from D14 (D2.6).

This is a RI-URBANS/ACTRIS guidance for this specific service tool that is part of the RI-URBANS deliverable D46 (D6.1, containing guidance for all service tools provided in the project) with the support for publication from AXA Research Fund to build up the final dissemination D55 (D7.6). Any dissemination of results must indicate that it reflects only the author's view and that the European Commission is not responsible for any use that may be made of the information it contains.

2. BACKGROUND ON MONITORING APPROACHES

Air quality is measured routinely through fixed air quality monitoring stations (AQMS). These stations include high-quality monitors that fulfill the data quality requirements as set in the European Air Quality Directive (2008/50/EC). Whereas a network of these fixed stations gives information on temporal trends of air quality, the density of the network is not sufficient to give information on air quality at street level. Some pollutants, especially traffic-related (e.g. UFP, BC & NO_x), can show a very high spatial and temporal variability within a city or neighborhood. Other sources than traffic such as woodburning may also result in high variability While established urban networks of fixed site monitors have spatial densities on the order of 1-10 km^2, concentrations of air pollutants can vary significantly within 10-100 m from roadways. It is difficult to extend the density of the network of AQMS due to their high installation and maintenance cost.

Improved spatiotemporal resolution of multi-component air quality data is critical for improved understanding of the connection between human exposure to air pollutants and health effects. To assess the impact of air quality on health it is important to have fine-scale data on air quality exposure. Advances in sensor technologies and the availability of portable and low-cost sensing devices give rise to new opportunities for mobile monitoring and denser fixed sensor networks.

Different approaches can be used to assess exposure to pollutant concentrations including Ultrafine aerosol particles (UFP), Black carbon (BC), Nitrogen dioxide (NO_2), particle mass concentrations below 2.5 µm ($PM_{2.5}$) at fine spatial scale. Mobile sensing platforms and fixed (low-cost) sensor networks can be used as complementary tools to data from fixed regulatory AQMNs (Morawska et al., 2018; Hofman et al. 2022; WMO, 2024), to map pollutant concentrations at a higher spatial density to provide more localized insight in the state and management of air quality.

A major application of these concentration maps is exposure assessment in health studies of long-term exposure to air pollution. High-resolution pollution maps can be of interest for local authorities in a variety of applications; i.e., hot spot detection, new AQMS localization, model validation, evaluation of policy measures. In this text we discuss mobile measurements and sensor networks to generate high resolution exposure (HRE) maps. Such measurements can be collected by citizens or can be collected by research institutes or AQMN. Involving citizens in data collections requires a simple and straightforward monitoring instrument and user-friendly methodology. When using more complex/expensive instruments and/or methods, there is a need for experienced staff to perform the measurements. We make a distinction between **mobile/fixed measurements** and experimental designs **with/without citizens** (see Figure 1).

The collected data can **be processed and analyzed** using only measured data or using interpolation /modelling techniques like Land Use Regression (LUR)-based or machine learning models. The selected techniques used for data processing have an impact on the required data collection approach.

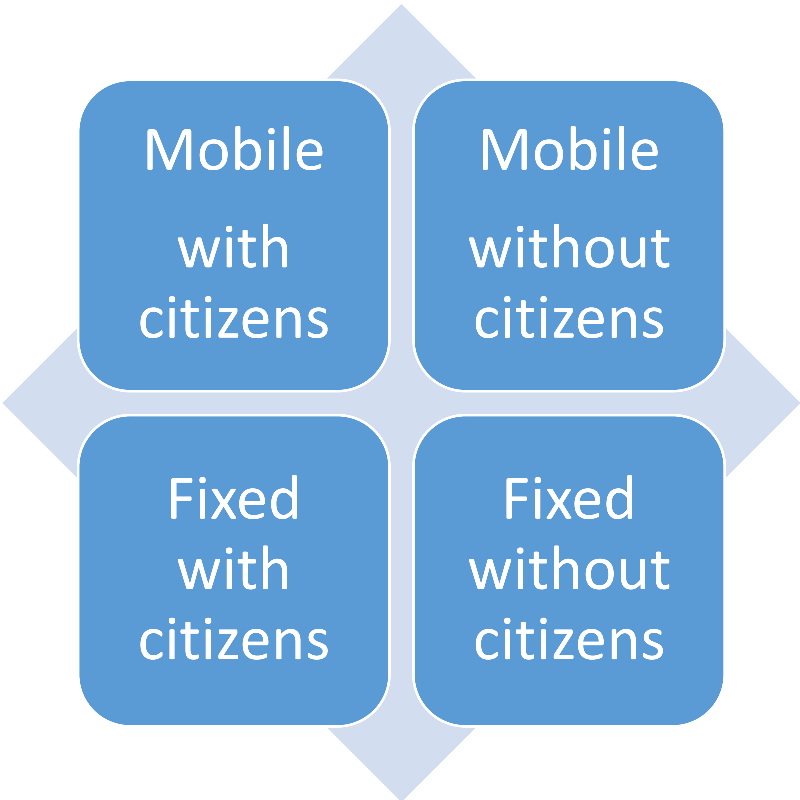

***Figure 1**: Schematic overview of different approaches for collecting data for high-resolution exposure mapping as discussed in this deliverable.*

3. USE CASES OF HIGH-RESOLUTION CONCENTRATION MAPS

The users of the high-resolution concentration maps/data are:

- Researchers.
- Epidemiologists.
- Local authorities (municipalities/cities).
- Regional authorities (e.g., exploiting AQMN).
- Stakeholders working on navigation (e.g., Garmin, TomTom, Polar): e.g. (healthy) routing applications.
- Stakeholders working on health technology/wearables (e.g., Apple, Fitbit, Garmin, Polar): such as smartwatch, trackers.
- Citizens who are curious of the air pollution in their environment.
- Real estate agencies that provide air quality labels for their real estate.
- Other end-users who are interested in the variability of air pollution in high-resolution.

Cities and AQMN can also be interested in using these high-resolution maps/data for:

- Hot spot detection (e.g., for selection of relevant new AQMS locations or targeted abatement policies.
- Evaluation of policy measures (when comparing data collected before and after introduction of the measures, e.g., traffic measures).
- In support of evidence-based policies: e.g., target areas for policy measures to reduce exposure (e.g., location of facilities for vulnerable groups like schools or hospitals) .
- Evaluation/improvement of dispersion models.
- Exposure assessments in health studies.
- Implementation of route-planner apps which can help citizens understand and reduce their exposures to air pollution.
- Gap-filling for certain pollutants not monitored by the networks (e.g., UFP, BC, LDSA), using virtual sensor proxies.

Local authorities (municipalities/cities) are important stakeholders and can also play a role in the recruitment of citizens. The involvement of citizens is first discussed. We will focus here on recommendations on how citizens can be motivated and engaged and how results/feedback can be communicated.

In the COMPAIR project (wecompair.eu) the use of sensors and citizen science approaches have been explored for policy makers, citizens and researchers.

4. CITIZEN INVOLVEMENT AND MOTIVATION: GENERAL RECOMMENDATIONS AND BEST PRACTICES

The involvement of citizens in air quality monitoring is a growing practice, recommended by the European Environmental Agency (EEA, 2019), that can bring important benefits to both science and society. From a scientific point of view, citizen science experiments in air quality monitoring can help to obtain reliable, up-to-date, cost-efficient and high-resolution air quality data in a timely manner. Depending on the number of volunteers, citizen science experiments can be used to measure air pollutants in large areas at high spatial resolution (e.g., street level), and complement data from the official urban/traffic air quality stations, which are often insufficient in number. Citizen science may also motivate the development of innovative solutions to air quality problems.

In addition to scientific benefits, experiments involving citizens can promote positive changes in society by, for instance, increasing public awareness to air quality problems, and influencing environmental activism and policymaking (Huyse et al., 2019). Citizen science experiments also motivate public learning about science and scientific methods (Perelló et al., 2021), and enhance retention of information in comparison with traditional learning methods. Through air quality citizen science projects, citizens learn about the current state of air quality in their community and how they can influence it ("seeing is believing"). Dissemination of scientific results (e.g., knowledge) through communities and participants' social media, further contributes to public environmental awareness and engagement, scientific interest and future public inclusion. An important advantage of citizen science experiments is that they may lead to long-term changes in public environmental perceptions and behavior, which may then contribute to air quality improvements. For instance, participants may opt for environmentally friendly actions such as the use of public transportation, advocate for increasing the number and size of green spaces at the community level. We also refer to other best practices on citizen science like Citizen science toolkit. https://making-sense.eu/publication_categories/toolkit/and the toolkit from the CitiesHealth project: https://citizensciencetoolkit.eu/.

4.1 Recommendations on how to motivate and engage citizens

Motivating participants is essential for the success of any citizen science project. Below, we list some examples on how to motivate public participation:

Educating participants on the air pollution levels they are exposed to, for instance, during daily commuting routes, and on the effects that this exposure may have on health may motivate citizens to participate in air quality studies and take actions to decrease air pollution.

Educating participants on the complexity of air pollution, like various pollutants, meteorological impacts (e.g., wind, temperature inversions), chemical reactions (e.g., photochemistry) and source contributions (cross-boundary, regional, local) in order to manage citizen expectation.

Public engagement increases when the values of citizens and their community concerns are taken into consideration (Vohland et al., 2021, and references therein). For this reason, whenever possible, citizens should participate in every step of the project and be considered research team members rather than users. Citizens and communities may contribute to the formulation of research questions, search for scientific methods and co-develop project rules. For example, in the xAire project in Barcelona citizens co-designed the study by deciding on which streets the NO_2 Palmes tube passive samplers should be placed (Perelló et al. 2021). In GroundTruth 2.0 citizens were involved in defining the measurement requirements and set-up, and execution of the measurements, in close consultation with the scientists and the city council (Van Poppel et al. 2024). Diverse entry-points and levels of commitment may increase citizen engagement (Vohland et al., 2021). As an example, the CitieS-Health project developed in Barcelona offered citizens three levels of participation: one in which citizens only answered health-related questionnaires and provided their residential address for air pollution exposure estimation; one in which, in addition to the above mentioned, they shared their geolocalization (to be able to use spatiotemporal air pollution predictions); and another one in which, in addition to the previous information, they carried a passive tubes to measure NO_2 concentrations (Basagaña et al., 2020).

Citizen science studies should not require from the participants advanced skills or too much preparation as prerequisite. As participants often lack a scientific background, citizen science projects should rely on easy-to-learn tasks, and ready-to-use, easy-to-understand monitoring sensors/sampling/indicators rather than complex ones. One example of this is citizen science contribution in Healthy Outdoor Premises for Everyone (HOPE, Petäjä et al. 2022) project in Helsinki that distributed simple to use multicomponent sensors to the citizens with automatic data collection and upload to cloud service and visualization. This allowed the citizens to explore air quality in their close vicinity with ease (Rebeiro-Hargrave et al. 2022).

Involving as many actors as possible in the citizen science experiments (e.g., public and private sectors, policy-makers, etc.) improves synergies and supports the development of efficient air quality regulations, which may further motivate future citizen participation.

Communication is key to obtaining realistic expectations from citizens and avoiding disappointment, e.g., on data quality, usability of data, own impact. Workshops or other community events can be useful to clarify to citizens the research motivations, to discuss what will be done with the information collected, and to clarify what they will obtain (e.g., a report) and when, and to manage citizens' expectations. In addition, data sharing, communication of project results in a timely manner, and acknowledging citizens participation, particularly in scientific publications, are suggested to increase satisfaction and motivation of participants for future collaborations (de Vries et al., 2019). Sharing preliminary results and/or personal reports (personalized results) with the participants during the project, instead of only at the end of the project, may also increase motivation. In the pilot Rotterdam of RI-URBANS, a lunch talk (for the participants) was organized to give feedback on the outcome; because not all measurements were collected on the same time, we also had to explain the differences in personal results related to differences in background concentrations.

4.2 Recruitment process

Recruitment of participants can be done for instance through announcements on social media, newspapers and outdoors e.g., in metro stations or similar public places. Recruitment can also be done through workplace (if employees are the target group) like in the pilot Rotterdam of RI-URBANS; this makes communication for recruitment, experimental phase and feedback much easier.

Researchers are recommended to investigate motivations and obstacles for participation, and offer tasks that require different levels of engagement based on participants interest, availability, experiences and motivation (Vohland et al., 2021). While the level of participants' experience should be taken into consideration, inclusive communication strategies should be used. For instance, the use of words like "unskilled" may demotivate engaged participants (Vohland et al., 2021).

Simple training should be provided for all the participants explaining project objectives, methodology guidelines, data quality requirements and instrumental usage. A short general introduction to air quality can also help; the level needs to be fine-tuned for the target group.

To avoid frustration, demotivation and participation withdrawals, researchers should be aware of the participants motivations and clearly explain how the project goals align with their expectations. Especially, it should be clear that participating in a research study may not directly lead to changes in environmental policies. One way to improve the level of citizen commitment is to organize an event midway through the citizen science project with a target to discuss initial results, tackle technical challenges and provide an update on the scientific objectives and recent findings. This also underlines to the participating citizens the continuing commitment and progressive insights of the scientific team. To conclude the action, it is important to organize a closing event.

4.3 Ethics for citizen science experiments

Just like with any other research involving humans, ethics is a major issue for citizen science studies, as these studies often collect personal data, geolocation and individual exposure to air pollutants. Ethical implications vary depending on the nature of the study. E.g., it is not the same to participate in measuring air quality in a given public location than measuring personal exposure to pollution, which includes addresses and routes that can lead to identifying the participant. When health data are including ethical issues are even more complicated.

In general, before taking part on the study, all the participants should sign an **informed consent form** which contains the purpose of the study, all the potential risks and benefits the participant are exposed to during the study, the type of personal data collected and their freedom to decide how the data will be used. The form should also contain information on who will have access to the data (name and position of the responsible personnel) and contact information of the investigators.

The participants should be aware of their freedom to withdraw their participation from the study at any moment (Kocman et al., 2019) and to request elimination of their personal data from the study or limit the usage of their data for future studies (Basagaña et al., 2020).

Consent forms may be dynamical in nature, being updated whenever the researchers or participant researchers feel that new information should be added (Vohland et al., 2021). “Dynamic consent forms” allow for adaptation of consent forms to issues that may arise throughout the project.

Regarding data security, citizen science studies should comply with the **General Data Protection Regulation (GDPR)** and any other local or institutional regulations to ensure protection of personal data. For this purpose, all personal data collected by the project (including questionnaires and sensors) should be anonymized before data analysis to avoid identification of the participants. Similarly, project results should be released in an aggregated form to preserve participants identity. Personalized reports sent to individual participants are possible and often an important motivation to participate in a study. Identifiable information should be kept in a password secured environment and accessed only by authorized personnel. All the project personnel should be trained in how to handle confidential personal data, and sign confidentiality agreements (Basagaña et al., 2020). In addition to protecting personal data, project photos that allow identification of participants should not be published without previous consent (Vohland et al., 2021 and references therein).

Other ethical considerations include treating participants as research partners and not just as data providers or “free labour” by involving them in the decisions, and ensure that the study is designed in a way that minimizes the risk that potential biases of the researchers or participating citizens can influence the results (Froeling, 2022).

When participants contribute in a study to collect data for an aggregated map (and not for own personal exposure assessment), and their names are not linked to the collective dataset (concentration and GPS) there are less GDPR issues.

4.4 Representativeness and diversity in citizen science

Diversity in citizen science recruitment improves inclusiveness, and brings varied experiences and perspectives to the project which may result in technological development and innovative solutions. Whenever possible, participants should be selected from different socio-economic backgrounds, cultural status, educational level, location, ethnicity, disabilities and gender. Workshops where successful female scientists talk about their experience in science may be a good way to motivate female participation in citizen science (Vohland et al., 2021). Inclusion of including different ethnic groups may be difficult especially if language problems play a role.

The development of tools that facilitate the participation of people that would otherwise have difficulties in participating in the study are also recommended. The D-NOSES project (funded by Horizon 2020) for example, complemented the odour pollution data collection method, traditionally done via a smartphone application, with the use of “Odour diaries” aiming at increasing the participation of people that find difficult handling technology (e.g., the elderly). Alternatively, training on the use of mobile applications and participation in different languages could be offered to improve inclusiveness (Vohland et al., 2021; and references therein). In general, data collection strategies should adapt to the varying community capabilities (e.g., socio-economic levels) and concerns (Vohland et al., 2021).

4.5 Data collection

Air quality citizen science studies often require citizens to measure air pollutants or report on the participants' perception of air quality (e.g., D-NOSES (Vohland et al., 2021)). Because citizen science participants often lack scientific background and often the air quality sensors used in these experiments are less accurate than reference methods, data quality is a common concern in citizen science projects. This however, may be less of an issue in co-created citizen science projects, with involvement of researchers and research grade instruments. Citizen science is not synonymous with the use of low-cost sensors (Froeling, 2021). The pilot in Rotterdam of RI-URBANS showed the potential of using mid-range portable instruments in citizen science.

Data quality has a direct influence on the project impacts (Kocman et al., 2019; and references therein). One should not get the impression that a citizen science study equals a study with low-cost sensors. Citizen science studies can indeed use research-level equipment and, in any case, the data obtained must be as accurate, complete and as relevant as possible. Of course, this needs to be balanced with budgetary constraints, but one should ensure that the project will provide data that is useful, otherwise there is no point in conducting the monitoring campaign.

Other considerations apart from quality of the data come to play when deciding the sensors or instruments to be used, for example that they are light weight and small (if they need to be carried out by citizens for some period of time), and that its instructions of use are easy to understand (Kocman et al., 2019; and references therein). Instruction manuals were prepared for the participants of the Rotterdam pilot and instrument procedures/handling were kept as little labor intensive as possible.

Researchers and citizens are encouraged to co-create standards for data collection, deciding for instance the location at which sensors should be placed, for example to capture certain sources, to have good geographical variability, or to cover places that are frequently visited by citizens or certain vulnerable populations (e.g., children or the elderly). Other issues such as the time or frequency of measurement can also be decided in collaboration between scientists and citizens. Citizens can also help in making the study protocol or instructions more understandable. In addition, researchers and participants may co-develop strategies to avoid misconduct of research participants. Finally, the project should make all the efforts to generate data under the FAIR (findable, accessible, interoperable, reusable) principles.

Low-cost sensors for Air Pollution monitoring

Since air quality may be measured at a large number of locations in citizen science studies, common budgetary constraints will point towards the use of low-cost sensors or passive samplers. Nowadays there exists a wide variety of low-cost sensors for measuring a wide range of pollutants, including gaseous (e.g. NO_2, O_3) but also particulate matter (PM) (deSouza, 2022; Kang et al., 2022; Kumar et al., 2015; Weijers et al., 2021).

While low-cost sensors are still unavailable for ultrafine particle measurements (Morawska et al., 2018), PM concentrations can be measured by light scattering low-cost sensors. These sensors however are only capable of measuring particles starting from 0.3 µm in diameter, as smaller particles do not scatter light sufficiently and therefore cannot be detected by the photometer (Rai et al., 2017). It is important to notice that the algorithms used by some sensors to convert signals to particle size are potential sources of error for particle size classification (Rai et al., 2017; and references therein). For this reason, the use of size selective mechanisms, such as impactors and filters, to select particle size (e.g., < 2.5 µm) before entering the instrument are recommended (Rai et al., 2017).

In terms of performance, the inter-sensor variability is usually rather small while the comparability with the reference method can be significant. As environmental variables such as temperature and relative humidity may affect sensors performance, users are advised to calibrate the instruments under conditions that are similar to those found at the experimental location (Rai et al., 2017). In general, the overall performance of PM sensors is considered reliable when the instrument is properly used and calibrated (Kocman et al., 2019, Hofman et al., 2022; Rai et al., 2017). Monitoring devices are discussed further in Chapter 5 (mobile) and 6 (fixed). The reason that the PM sensors seem to work well is the fact that they mainly capture PM2.5 which has very low spatial and temporal gradients. This makes the sensor interesting for the citizen but the value added for the scientist is often limited.

For NO_2, passive samplers are still the method of choice if the interest is in long-term exposure, as they are more reliable than the current state of the art of low-cost NO_2 sensors. The same applies to ammonia monitors.

4.6 Data validation

Quality control of the sensors or passive samplers used for citizen science studies should be carried out before, during and after the measurement campaigns. The sensors and/or samplers used in the study should be collocated and compared between themselves and against a reference instrument before the start of the study to ensure their performance is acceptable for data collection (Hofman, Peters, et al., 2022). During data collection, it is also advised to perform continuous (e.g. keep 3 sensors co-located next to AQMS) or periodic (e.g. periodic co-locations of used sensors during project) intercomparisons to assess their performance after some use. Finally, after the measurements are finished all the sensors should be collocated to evaluate their performance at the end of the study and identify potential malfunctioning.

Data quality may be validated for instance using testing protocols that aim to estimate the performance of air quality sensors used in the citizen science study in comparison with reference instrumentation. Examples of parameters that can be tested are: linearity between concentrations measured by a reference instrument and those measured by the sensor; accuracy; precision (variation in concentrations from simultaneous measurements); response time to changes in concentrations; lowest detection limit (lowest reliable concentration measured) and detection range; effects of changes in temperature and relative humidity on measured concentrations; and interferences caused by the presence of other air pollutants (Morawska et al., 2018).

4.7 Dissemination of results

Dissemination is an essential step to inform the general public about the study and the importance of the results to society (Arévalo Nieto et al., 2016). Because the project results should be disseminated to different types of audience (e.g. scientific community, general public, policy-makers, etc.) it is recommended that different divulgation and communication strategies are used for each audience group (Arévalo Nieto et al., 2016).

Dissemination of project results and lessons learned is usually done through scientific publications, conferences, social media, newspapers, etc. While these forms of dissemination are efficient to inform the general public, they are “one-way messages” that do not encourage feedbacks from the participants.

Specific activities for dissemination of citizen science project results should be put into practice to allow discussion and questions from all the participants in the study and other audience that could be interested. For instance, it can be done through science cafes or lunchtalks, where participants can ask questions, ask for clarifications, or suggest actions. Other options could be online webinars/workshops or informal/festive meetings with time for questions and discussion. Interactive dissemination events are often more inclusive in terms of local participants and stakeholders; hereby facilitating behavior change and local impact in communities more easily than scientific papers or reports. In the RI-URBANS Rotterdam pilot, we had an interactive DCMR lunchtalk to disseminate results and collect feedback from participants on ease of use of instrumentation and their experiences with regard to monitoring, awareness raising and/or behavioral change.

The COMPAIR project showed that citizen science can be a powerful catalyst for behavioral change; participants who were actively involved in data collection demonstrated a greater awareness of environmental issues and were likely to alter their behaviors (such as reducing car use and supporting local environmental policies). (source: D6.2

pathways to behavioral change report COMPARE, draft version). COMPAIR concludes that empowering citizens through science not only enhances their understanding of environmental issues but also motivates them to adopt more sustainable practices. The project recommends that future initiatives continue to focus on community-driven approaches, ensuring that participants have a strong sense of ownership over the data and outcomes. Additionally, the findings suggest that integrating citizen science into broader environmental strategies can significantly amplify the impact of such efforts.

5. MOBILE MONITORING

5.1 Objectives and data needs

A mobile platform provides the possibility to sample spatially diverse environments in a limited time and with a limited number of (costly) monitoring devices. Together with advancements in air monitoring instrumentation, such as higher time resolution and greater portability, these platforms can capture the high variability of air pollutants in space and time in a complex urban terrain. At the same time, mobile measurements usually consist of only a few seconds of data per street segment, needing temporal aggregation to be representative for the long-term/average exposure. Fixed sensors/instruments provide high-resolution time series of air quality data, representative for one location, but offering low spatial coverage. So mobile platforms can serve more locations but need more repetitions in time, whereas fixed sensor networks provide temporal profiles but at a limited number of locations. Figure 2 lists the two approaches.

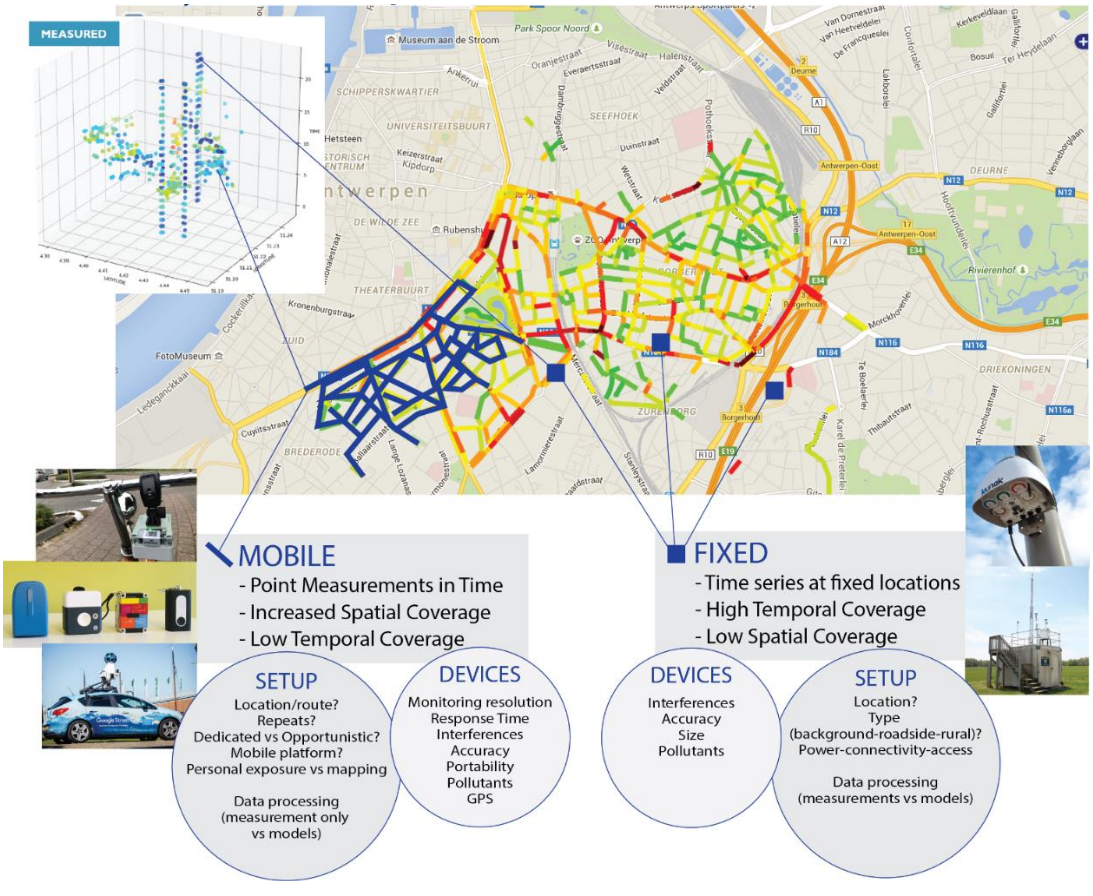

Figure 2: *Difference between mobile and fixed air quality measurements in terms of monitoring setup and device requirements.*

In this section, we describe the objectives and data needs for mobile mapping in terms of exposure assessment. The approach of mobile measurements and the requirements for the set-up and instruments might be different when one wants to make an air quality map for exposure assessment as compared to a map for hotspot detection. Data collected during mobile measurements can be used (i) directly, (ii) to validate dispersion models, (iii) as input for mapping applications with data-driven models (LUR, machine learning or others). Here, we consider mobile measurements when data is collected by instruments that are in motion; opposite to data that is collected by portable instruments that are moved from one location to another to collect stationary data over short periods.

5.2 Monitoring strategy

It is very important to think in advance about the monitoring strategy (mobile/fixed; figure 2) related to the use of data before data collection is started. Slight changes in the data collection scheme might result in considerable improvement of the final results. At the same time, the required (personnel) efforts need also be taken into account e.g. opportunistic data collection can result in a wealth of data with limited staff resources though interpretation might be more complicated.

In this section, we reflect on the data needs on different levels (measurement scheme, monitoring devices, data processing, ...) as function of how data will be used for and how they will be processed. At the same time, the data processing itself is important for the final result.

The monitoring setup for mobile measurements (**Figure**), i.e. air quality measurements performed on a mobile platform (walking person, bicycle, car, tram, …), needs careful consideration as the applied mobile platform, measurement timing (e.g. hours of the day), location and routing will determine the representativity of the data for the intended application. Thinking about high-resolution exposure maps, it will be important to collect enough monitoring data in both space and time to represent the average air pollution exposure. Moreover, while one can opt for a **dedicated monitoring approach** strictly defined in terms of considered route/number of repeats/hours of the day, an **opportunistic approach** could allow for the collection of more data (in time and space) using existing mobile platforms (eg service fleet vehicles/city servants/…) or citizens performing their normal (commuting) activities, for data collection. However, design of each approach (e.g. study population) has also an impact and the research question to be answered will determine if opportunistic data collection is a feasible option. Mobile instruments and IoT sensors are diverse and specific requirements are applicable for mobile monitoring (e.g. high monitoring temporal resolution (depending on the mobile platform used), low response time, relevant pollutants, GPS functionality, portability). In order to obtain time and space representative data, the collected point measurements will need to be processed, either (i) by aggregating /averaging the measurements (ii) by using measured data and statistical techniques to interpolate between time/space instances (IDW/Kriging) or (iii) by training models (LUR or machine learning), often in combination with contextual data (e.g. meteo, traffic, spatial) to explain the observed measurements and extrapolate to unknown time/space instances.

Compared to mobile measurements, fixed air quality monitoring is more straight forward in terms of setup and instrumentation. As the location is fixed and continuous data is being collected, the approach is by definition representative for that location. However, this representativity of primary importance depends on the considered source environment (urban background, roadside, rural), wind direction, practical requirements (power, connectivity, space), … and is sometimes also a function of practical constraints (power supply, fixing material...). Fixed measurements do not require a high (<1 min) monitoring resolution and are more stable/controlled in terms of turbulence, vibrations, temperature, … Again, data can be reported per location combined with dispersion modeling to add spatial information, or inter/extrapolated by mathematical techniques or LUR/machine learning models. Direct interpolation in urban areas often does not result in informative maps.

5.3 Targeted versus opportunistic approach and involving citizens or not

Mobile monitoring can be performed by **repeating predefined fixed routes** or using a more **opportunistic approach**, using a carrier that performs the measurements during its day-to-day activities without intervening with these activities. When using the opportunistic approach, constrains are needed in the post-processing (e.g. map matching, number of repeats) in order to obtain robust exposure data. Mobile monitoring can be performed with or without the help of citizens.

Mobile crowd sensing refers to a broad range of community sensing mainly participatory sensing and opportunistic sensing; Brahem et al. (2021) explain *participatory sensing* as data collection by citizens who measure their own exposure and observe their own environment. When data collection is entirely automatic, the data collection is referred to as *opportunistic* sensing. In the remainder of this document, we refer to opportunistic sensing not strictly to approaches where all data collection is entirely automatic but relate it to the way the data is collected in time and location (not predefined in terms of routing/data coverage).

Van den Bossche et al. (2016) defined opportunistic mobile monitoring as data collection making use of existing carriers to move measurement devices around. The movement of the carriers (the travelled route) is uncontrollable from the point of view of the researcher, as it is not designed and performed with the data collection in mind as primary goal. The data collection takes advantage of **existing mobile infrastructure** or **people's common daily routines**. This contrasts with targeted mobile monitoring, which is a coordinated, goal driven approach in which the mobile measurements are **deliberately planned and carried out with a specific purpose in mind**. The carrier can be citizens, a certain professional group (e.g. city wardens, home nurses, taxi drivers), but also a vehicle (postal van, bus, tram...). In the Rotterdam pilot, we explored the approach of employees collecting data during daily commuting from and to work.

Opportunistic mobile monitoring is a promising approach to collect large data sets that give useful additional information at a reasonable cost compared to classical data collection methods. But, depending on the set-up of the data collection, such new data can lead to new challenges in data processing and interpretation. Campbell et al. (2008) described opportunistic way of data collection already in 2008 as 'opportunistic people-centric sensing' where small devices were carried by individuals in their daily activities to collect information related to human activity and to the environment around them (Campbell et al., 2008; Kumar et al., 2015). This approach was utilized in Helsinki in HOPE project, where 100 volunteer citizens carried air quality sensors and made observations during their normal movement within the city (Rebeiro-Hargrave et al. 2022). This provided information on the local air quality but also about urban mobility.

Opportunistic data collection can take different forms (Van den Bossche et al. 2016). Firstly, they can vary according to the degree of human interaction they need. Possible human interactions are related to carrying the measurement system, the operation and maintenance of the measurement system and to the data collection and handling. Examples of campaigns that can run independently for long periods without human interaction after initial set-up are those based on sensors mounted on vehicles such as cars, buses or trams. The more human interaction the data collection needs, the more the user-friendliness of the instrument and the motivation of the people involved become important issues. In this deliverable, we make a distinction between data collection **'with' or 'without' citizens**, referring to the carrier. Secondly, the data collection can follow a repeated structure along the same routes and/or within the same time frame or can be rather unstructured. Whereas opportunistic approach is not goal-driven (see above), depending on the involved carrier, the data collection can be rather structured (e.g. commuting to work).

A priori careful consideration of the spatial and temporal variations is needed. Temporal variation tends to be higher than the spatial variation so at least you have to consider sufficient replication of the temporal component in each relevant spatial location.

The choice of targeted versus opportunistic monitoring holds some consequences for the processing and interpretation of the data. The advantage of a targeted approach (fixed route and sampling period) is that all sections along the route are measured 'quasi-simultaneously' during the same days, seasons,... etc. which makes it easier to compare different datapoints in space and also makes it easier to perform background scaling to e.g. yearly average values. A drawback is the workload: when citizens are involved, they have to drive/walk the route in addition to their normal activities and when a vehicle is used, a driver needs to be paid to do the measurements.

The opposite is true for the opportunistic approach. From the perspective of the researcher, there is no control over the specific location and time of the measurements. This could result in sampling bias where certain urban microenvironments are underrepresented or absent in the data. The same holds for the time of sampling. A bias in time can appear in the case of data collection by commuters; the measurements are mainly limited to rush hours and no data will be available during working nor non-working (night time) hours. For the same group, we can also have a sampling bias in locations when e.g. a group of people from one company is selected because all of them commute to the same location and each takes the same route every day (the location bias is than dependent on where they live). Another example is a postman or a parking warden who will not frequently enter green zones such as parks. Also, a time/space bias is possible: this means that data is sampled at the same locations for specific time slots; this can be an additional issue compared to the time or space bias where we just have a lack of data on certain locations or time slots. An example is a postal van or bus driving the same trajectory and always sampling at point A during morning peak hours and at point B during quiet moments (e.g. noon or evening). This might give a wrong idea of the spatial pattern and resulting in an underestimation of concentrations at point B and overestimation at point A. When using an opportunistic approach, the expected sampling time and route needs to be evaluated in advance and data processing techniques including scaling for varying background concentrations needs to be considered.

As a consequence, there will be different measurement conditions for different locations, hindering the comparability of the results between these locations. This is a major problem, as it complicates the data interpretation (the comparison of the measured concentrations at the different locations), making it less evident to use the results for air quality mapping. To cope with the time/space bias, data extrapolation can be performed using mathematical (Kriging, IDW, ...) or modelling (LUR, machine learning) techniques but for a reliable result, the input data needs to be sufficient. Or use a continuous background monitoring station.

Both targeted and opportunistic approaches can be used with or without involving citizens. A targeted approach (having a fixed monitoring route) without citizens may result in a larger requirement in resources. The advantage of not involving citizens is that (when using a car) more expensive and accurate monitoring equipment can be used and the monitoring is performed in a more standardized way.

The advantage of involving citizens is to raise awareness on AQ. This is also referred to as participatory monitoring.

Finally, the sampling can also be biased by the weather conditions, e.g. when the data collection stops when it rains; this is not only true for opportunistic approaches (e.g. when the commuter takes the car instead of the bicycle on rainy days) but is also true when the monitoring equipment is not fully protected from rain.

5.4 Requirements for monitoring devices

Requirements for monitoring devices need special attention when used for mobile data collection, used for unattended use over several periods, or used by citizens who do not have specialized knowledge on air quality and measurements. Also important for the interpretation of results is to know whether the observed concentration gradients are within the instrument uncertainty or not Data quality

Data quality is important since collecting data with sensors without knowing the data quality is not useful. As miniaturized and/or portable air quality instruments are often condensed/simplified/cheaper versions of regulatory-grade instruments, proper knowledge/understanding on the inherent instrument uncertainty and precision and associated sensitivities of the applied instrumentation (preferably based on local validation campaigns) is needed for an accurate interpretation of the obtained results. For some use cases (e.g. awareness raising, personal measurements, …), the associated data quality is of less importance than for other use cases (e.g. sensor networks where sensors are compared to each other, …). However, a certain minimum data quality is also needed for these applications, to avoid misleading interpretations.

In some citizen science studies, data collection of pollution is done as a way of making citizens aware of air pollution, rather than collecting data to be used in scientific studies. The problem is that (in general) these data cannot be used to construct AQ and exposure maps. In addition, there is a risk that the data quality is not good enough to give the citizens the expected feedback; in this case expectation management is very important (). In this document, we focus on collecting data with sufficient quality to construct exposure maps and try to engage citizens at the same time, rather than only awareness raising.

Although low-cost sensors have evolved in the last decade, Snyder et al. (2013) indicate that many commercially available low-cost sensors have not been challenged rigorously under ambient conditions, including both typical concentrations and environmental factors.

Small PM sensors are typically nephelometers or particle counters based on optical measurements. The resulting error of PM sensor measurement devices is dependent on the sensor technology, the calibration algorithm and the calibration aerosol used (Hagan and Kroll, 2020). The error also depends on the environmental conditions and particle size and type.

Most low-cost PM sensors measure particles via light scattering. Sampled particles intercept a beam of light and the scattered light is measured and correlated to a PM concentration. Typically, sensors using the optical measurement principle can be broken down into two main types, nephelometers and Optical Particle Counters (OPC). Nephelometers measure the particles as an ensemble, gathering light scattered by all particles across a wide range of angles, typically 7°-173° to avoid pure forward and backward scattering (Hagan and Kroll, 2020; and references therein). The total scattering amplitude is then correlated to a mass measurement made by a reference instrument (nephelometers that measure scattered light at a single angle are sometimes referred to as photometers; and can be considered as a subclass of nephelometers.) OPCs, by contrast, detect particles individually, providing information on their number and size. Light scattered by each individual particle is measured and each pulse is assigned to a size bin based on its total light intensity, resulting in a histogram which is converted to a mass loading once the entire distribution has been measured. While these technologies have been around for decades, they have recently become available at much lower cost due to the availability of small, inexpensive light sources and electronic components. OPCs can be split into low-cost OPCs and higher-cost OPCs; since higher-cost OPCs use more expensive electronics and optics, they can measure smaller particles; the typical size range is 0.38 – 17 µm for lower cost and 0.1 – 17 µm for higher-cost OPCs (Hagan and Kroll, 2020).

The most important sources of uncertainty for PM sensors are related to (Hagan and Kroll, 2020):

- High relative humidities. Hygroscopic growth of particles at high relative humidity (around >75%) results in overestimation of particle mass due to growth of particles. All types of optical particle sensor suffer from this interference and the error ranges from 100% to a few hundred %, depending on the hygroscopic properties of the aerosol. This can be solved at least partly by using an in-line dryer or applying a correction algorithm.

- Changes in aerosol optical properties, when the sensor is calibrated using an aerosol with different optical properties. The impact of aerosol optical properties is most important for low-cost OPCs and of medium importance for higher-cost OPCs and nephelometers. The effect is especially relevant (for low-cost OPCs) when the aerosol has strong absorbing properties and when small particles become undetectable with inexpensive optical detectors (due to the small amount of scattered light).
- The particle size distribution. This is very important for low-cost OPCs and nephelometers. In this respect, the ability of a sensor to measure small particles is very important. Since higher-cost OPC are able to measure smaller particles and are typically calibrated for different sizes of aerosol, they can better assess PM mass for different size distributions. In environments where small particles (<300 nm) comprise a large amount of particle mass, low-cost OPCs will be subject to significant error. In environments in which the underlying aerosol size distribution is highly variable (e.g. in urban environments), low-cost OPCs and nephelometers will struggle to measure PM mass correctly.

In addition to the sources of uncertainty summed above, some sensors do not estimate the PM_{coarse} fraction correctly. Some sensors use an algorithm to estimate the PM_{coarse} (PM_{10} – $PM_{2.5}$) based on $PM_{2.5}$ concentrations. This can be sufficient to estimate PM_{10} concentrations in urban areas but might result in high uncertainties at locations close to sources characterized by a high amount of coarse dust (Vercauteren, 2020). We will not discuss this further in detail since this is mainly relevant at locations with specific sources with more PM_{coarse} and not related to exposure of UFP (which is our primary focus).

NO_2 sensors rely on different measurement principles (electrochemical or metal oxide) compared to the chemiluminescence principle of the reference equipment (Hofman, Nikolaou, et al., 2022). Electrochemical sensors convert a chemical reaction (reduction at sensing electrode and balancing oxidation at counter electrode) of the pollutant of interest in a quantifiable electrical current, while metal oxide sensors rely on the gas reaction with semiconductor material, resulting in free electrons. Electrochemical sensors are currently most advanced in detecting ambient (parts per billion; ppb) NO_2 concentrations, but suffer from sensor specific activity (nA/ppb response) and have shown to be cross-sensitive to other oxidizing pollutants (e.g. O_3) and environmental conditions (temperature and relative humidity) (Hofman, Nikolaou, et al., 2022). Moreover, the sensor's electrolyte (responsible for ion transportation) will age naturally as a result of exhibited temperature and humidity variability; with low humidity (<60%) resulting in drying out of the electrochemical cell affecting sensor response time, and with high humidity (>60%) leading to water absorption and dilution of the electrolyte influencing the sensor characteristics and potentially leading to leakage and resulting corrosion of the sensor pins (Raninec, 2021).

Accurate portable devices are available for pollutants such as UFP and BC; however, these are so-called "mid-range instruments" with a cost of approximately 5000 – 10000 euro, complicating wide-scale/large-number deployments. These devices will be further discussed in the detailed description of selected methods. They mostly have a good and acceptable data quality. Recurrent co-location of the instruments used in one campaign and co-location with reference-type (more expensive) instruments is recommended.

5.4.1 Particle sensors

Pollutants that show a large spatial variability (UFP, BC, NO_x) lend itself most for mobile monitoring. For other health relevant pollutants (such as $PM_{2.5}$, O_3), an increased spatiotemporal monitoring coverage compared to regulatory air quality monitoring networks does not contribute as much.

A variety of particle monitors, sensing devices and sensors is available on the market to determine the particle mass concentration in air samples. The reference method is the gravimetric method where particles are actively collected on a filter medium. Size-selective heads are used to sample a predefined fraction of particles (e.g. PM_{10} or $PM_{2.5}$). Because filter-based methods have a low time resolution (usually 24-hour), continuous (automatic) equivalent methods are used in AQMN including instruments using e.g. Beta attenuation, light scattering or oscillating microbalance technology to measure particle concentrations in near-real-time (~1-minute to 1h averages).

PM sensors are widely used in studies although they measure undifferentiated PM and they miss the finest fraction (ultrafine particles).

5.4.2 Mobile use

When using monitoring devices or sensors for mobile use, this sets specific challenges. What needs to be considered is specifications in terms of:

- Fast enough response time.
- High enough time resolution (1-10 seconds).
- Data needs to be associated with geographical information (GPS).
- Portable (function of 'carrier' platform).
- Capacity to adapt to fast changing environments (interferences; vibration, turbulence, housing, …) including movement from indoors to outdoors.

A fast response time is needed when collecting mobile AQ data. When sensors/instruments measure at a time resolution of 1 minute, this means that when driving at a low speed of 15-20 km/h (e.g. by bike), a single measurement point will take 250 - 333 m. When increased to 10 second monitoring resolution, samples will be taken every 42-56 m. At a walking pace (5 km/h), the spatial resolution becomes 14 – 83 m, at respectively, a 10 - 60 second resolution. When using the monitoring equipment on a mobile platform like a tram, bus, car, ... where travelling speeds vary from 30 to 120 km/h the response time and monitoring resolution are even more important.

The requirement for portability depends on the 'carrier'. When involving citizens this is more important, compared to a mobile carrier like a car or truck.

An important issue with mobile monitoring is the fast-changing environment; when sensors used are affected by interferences, this can make more difficult the interpretation of results. Potential interferences need to be co-measured and corrections need to be applied. It is recommended that sensors/instruments are co-located/evaluated prior to use at a stationary location.

5.4.3 Use by citizens

In general requirements that need to be considered for citizen science use are: Usability, user-friendliness, fool-proof, feedback (good measurements and results).

Specific requirements are there when collecting data by citizens. Based on the workshop with local authorities/citizens within the COMPAIR project (VMM, 2022) include:

- Data communication: preferential automatic data uploading (NB-IoT, WIFI, LTE-M, …) (100%), otherwise 12.5% daily upload, 37.5% weekly upload.
- As autonomous as possible (power on/off, required communication/intervention handling).
- Portability; weight, size, easy to carry/attach, casing/backpack, …
- Noise: as silent as possible.
- Notification: option to mute.

- Optimized battery usage, which is a compromise between sampling time, data upload frequency and individual sensor power requirements.
- Easy charging: daily charging is considered ok.
- Privacy: capacity to anonymize data.

5.5 Data processing: direct and model-based mapping

Mobile monitoring data can be used for direct mapping or as input for empirical models. It is important to know in advance which data processing technique will be used to optimize the data collection. Not only whether or not a model is used, but also the type of model used can have impact on the data collection requirements.

When direct mapping is used, it is important to have a good spatial and temporal coverage (representativity). This is also true when models are developed but to a lesser extent as spatiotemporal dependencies are learned from the available dataset.

5.5.1 Completeness of the dataset

Mehanna et al. (2022) defined three parameters for **completeness of datasets**: sensor completeness, temporal completeness and spatial completeness. Sensor completeness is defined as a quality facet that captures the extent to which the measurements of a given sensor are complete over a certain sampling period (including also aspects of data transmission and quality of the data). The authors focus on personal exposure measurements but similar concepts can be applied when using mobile measurements for mapping which is the focus of this deliverable. Also Van den Bossche et al. (2015) discussed temporal and spatial coverage of mobile data. Temporal completeness and temporal coverage (or representability), and spatial completeness and spatial coverage are similar concepts. When both mobile and fixed measurements are plotted on a 3D graph with x/y representing geographical coordinates and z the time (Figure 3), the difference between fixed time series and denser mobile point measurements becomes visible. Ideally, an air quality value for each time and space instance is available (inferred in **Figure**).

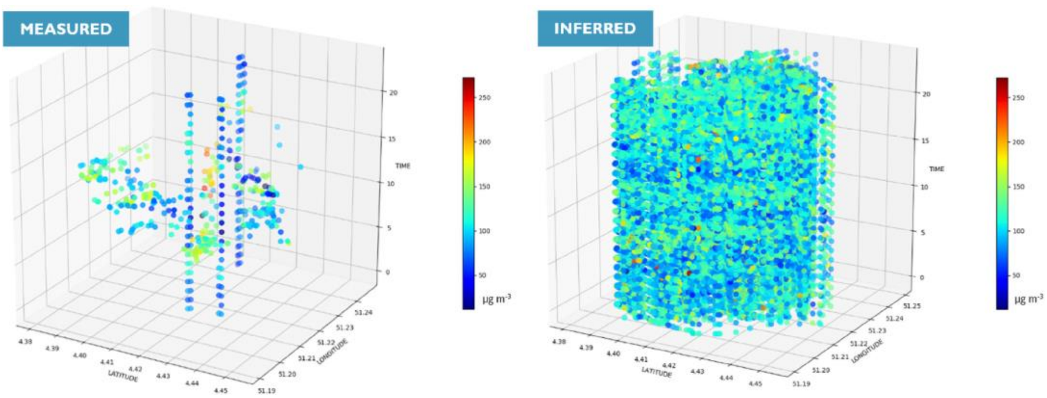

Figure 3: *Concept visualisation of fixed and mobile sensor measurements represented as a sparse data matrix in space (x-y plane) and time (z plane). Right: Inferred air quality values by a machine learning model (from Hofman et al. (2022)).*

Temporal completeness (according to Mehanna et al., 2022) characterizes the way a given period of time is covered by the collected measurements. Evaluation of the temporal completeness can go with different assumptions; e.g. assuming a uniform distribution over time, versus distributed measurements considering the variation of pollutant levels at different times of the day, month or year. **Spatial completeness** is defined by Mehanna et al. (2022) as the extent to which data sufficiently represents a specific spatial area and it characterizes the coverage of this area. In other words, spatial completeness indicates how sufficient and comprehensive the current measurements are for a particular area. Also, different assumptions can be made for the spatial coverage; e.g. assuming a uniform distribution of the measurements over the study area or considering the variation of pollutant levels in the different cells of the area of study.

In the following, we prefer the last assumption (considering variation in concentrations as function of time and location) since we want to use the resulting maps for exposure assessment and try to assess as much as possible the spatial and temporal dynamics. From now on we will use the terms temporal coverage and spatial coverage.

Temporal coverage has different components; for simplicity, we make distinction between a) time of the day, b) day of week, c) season of the year, d) years (see also Vanden Bossche et al., 2015):

The time of day that data are collected affects the final. In most cases, we want to represent a daily average of the concentrations where a typical diurnal pattern has times with higher and lower concentrations (related to peak hours and/or periods with better and worse dispersion characteristics). In some cases, the interest is in a typical hour (e.g. peak hour air pollution when looking at traffic sources) but otherwise the collected data needs to be representative for the exposure time that is considered.

The day of the week can also have an effect on pollutant concentrations with typical lower concentrations during weekend days compared to weekdays. In order to have a good coverage over an entire week, both week and weekend days need to be considered. However, when we want to maximize the observed spatial patterns it might be advisable to collect only data during working days. In some cases (e.g. when applying an opportunistic approach with employees or an opportunistic approach using cars like postal vans) only weekday measurements are achievable due to practical constraints.

The season (or month) in which data are collected also needs to be considered. Concentration differences between seasons can be explained by differences in sources and differences in meteorological conditions (dilution conditions as function of boundary layer height, affecting receptor sites as function of wind direction, washing out pollutants by rain/snow, secondary formation of pollutants as function of sunlight or temperature...). Typically, data are collected at the time when health study is performed. It might be tricky to use data from another year, especially if there is a long-time lag between data collected and health study. For example, more stringent emission limit values can result in reduced concentrations (e.g. reduction in Pb concentrations of petrol cars, sulfur content in fuel, more stringent EURO norms for cars and HD vehicles), but also local measures (like LEZ in cities) can affect the local AQ. To assess the air quality in a city, it is recommended that measurements be taken in different seasons. However, this is not always possible due to constraints, such as the budget or availability of volunteers (Van Poppel et al., 2024).

When looking at annual average exposure, it is recommended to extrapolate the measured average concentration to annual values; this can be done by using data of a fixed AQMS of the network nearby (Van den Bossche et al., 2015). In the Rotterdam pilot, data processing techniques were tested to rescale the data to representative ‘annual’ values. We refer to Deliverable 2.6 for further information.

Spatial coverage means that the study area needs to be mapped and exposure of all participants in the study area needs to be assessed. To generate high-resolution maps large quantities of data are required to include the range of possible meteorological conditions and the range of local air quality conditions (depending on local sources, e.g. traffic intensity), and to counter occasional and exceptional events. It is important to assess whether enough repetitions are made in relation to the goals of the monitoring campaign.

Sensor completeness is related to sensor quality and Mehanna et al. (2022) suggest an approach to improve data completeness by adding information about the quality of the measuring sensors. Whereas this is needed, we do not further discuss their approach here since it is function of instruments used. However, data quality control needs to be considered when collecting data (irrespective of the instrument used). Data quality, sensor uncertainty (comparability against REF) and precision (comparability against other sensor), and data quality control will be addressed when discussing the different tools further in this document.

5.5.2 Data processing for direct mapping

In general, when data is processed for direct mapping, individual data points are averaged over street segments. Details on data processing used in the airQmap approach are given below (see airQmap).

Van den Bossche et al. (2015) investigated how mobile monitoring can be used as an additional tool to acquire air quality data at a high spatial resolution. The study was based on 256 and 96 runs (repetitions) along two fixed routes (2 and 5 km long). They investigated the impact of temporal variability on the representativeness, and developed a methodology to map urban air quality using mobile monitoring. They stated that a limited number of mobile measurements may only represent a snapshot and not be representative and evaluated the number of repetitions needed with and without scaling e.g. for background concentrations. Different data processing methods were compared. They showed that using a trimmed mean and applying background normalization decreased the required number of repetitions for the same resolution and uncertainty level; it was shown that– using a trimmed mean and applying background normalization – 24-94 repeated measurement runs (depending on location type with a median of 41) are required to map the BC concentrations at a 50 m resolution with an uncertainty of 25%, compared to 33-141 runs without applying the trimming and normalisation. When relaxing the uncertainty to 50%, these numbers reduce to 5-11 (median of 8) runs compared to 8-39 runs. Reducing the length of the street segment (increasing spatial resolution) resulted in an increased number of repetitions required to obtain the same uncertainty level. When performing mobile measurements, it is important to take out extreme events (that influence the average concentrations and are not 'representative'), especially when a low number of repetitions is performed. Event detection algorithms can be used to remove extreme outliers; e.g. Hagler et al. (2012) used the running coefficient of variation (COV) method. In this method, a running 5 s standard deviation of the BC concentrations is calculated and divided by the mean concentration of the entire sampling period. The 99th percentile of the calculated COV is used as a threshold and all data points with a COV above this threshold are removed along with the data points 2 s before and after. This method has the risk that it can mask hotspots where peaks occur systematically.

Spatial aggregation is needed to smooth the data at different spatial levels (routes, streets, segments). Van den Bossche et al. (2015) showed that mapping at a spatial resolution up to 50 m is feasible for BC and a higher spatial resolution of 20 m can be obtained with a slightly increased uncertainty.

Quality control of GPS data needs some attention when processing mobile measurements (see also Van den Bossche et al., 2016). The processing and filtering of the raw measurements of the GPS device include: filtering for incorrect or unreliable GPS locations, map matching (allocating deviating GPS locations to associated street segment/map) and spatial aggregation. The reason for unreliable GPS data can be indoor periods during the day or outdoor moments with a very bad reception (e;g. travelling by subway, or in street canyons). The filtering of unreliable GPS data can be done based on minimal number of satellites (threshold). The accuracy of GPS data has consistently improved in recent years. When a dedicated route is used it is much easier to filter unreliable data. Map matching is needed when the GPS is slightly off track, which occurs often in urban environments. A way to enhance this is by assuming the measurements are always performed on the streets. This can be done by matching the locations for each individual run and street segment to the shortest distance (max 30 m) of the selected street sections.

Whereas some studies collect data during a limited period resulting in an aggregated map, collecting data using opportunistic approach or platforms (like trams, postal vans...) can result in continuous update of input data. Real-time dynamic pollution maps can provide personalized exposure information.

Van den Bossche et al. (2020) showed that spatio-temporal models (see below) can also provide a dynamic pollution map. When data are collected continuously using an opportunistic approach, a model can be constructed that is continuously updated with these new data. The R^2 and Explained Variance (EV) of the different Cross Validation (CV) schemes can then be regarded as the predictive ability of the model under different circumstances.

5.5.3 Models based on monitoring

In the review papers of Jerrett et al. (2005) and Hoek (2017), different classes of models to derive intra-urban exposure assessment were identified and also Xie et al. (2017) described different data modelling techniques ranging from geostatistical techniques, Gaussian models, linear regression, artificial intelligence to compressed sensing, to assess pollutant concentrations and personal exposure. Not all models are based on measurements but we sum them up here for completeness. The model and data-processing techniques used can be generally distinguished as:

- The simplest models are proximity-based assessments: In these models proximity to a pollution source is measured; most commonly used to assess exposure to traffic-related air pollution, where distance to road and traffic counts are the main indicators for exposure estimates; these models do not use measured data but use proximity data to explain variability in pollution levels, and are *out of scope for this study.*

- Spatial interpolation methods (sometimes referred to 'geostatistical models', Jerrett et al., 2005) estimate concentrations at unmeasured locations by measured concentrations at neighboring locations. They can be based on deterministic and stochastic geo-statistical techniques. Four interpolation methods are commonly used in air pollution estimation and assessment: spatial averaging, nearest neighbor, inverse distance weighting and kriging approach (Xie et al., 2017). In urban areas, interpolation alone generally is not a useful tool, because of the large number of local sources. For regional background assessment, interpolation can be useful.

  Spatial averaging calculates the mean of pollutant measurements from the nearby monitoring stations (located within a predefined grid, a country, or even a city). *Nearest neighbor* assigns the pollutant measurements of the closest monitoring station to the unmeasured location, regardless of the actual distance between them. The first two do not consider spatial variability of the concentrations as they only consider one monitoring station or do not consider distance to neighboring monitoring stations to calculate the unknown concentration. Therefore, they are no longer commonly used. *Inverse Distance Weighting (IDW)* is a deterministic method for spatial interpolation and calculates the value at the unknown locations as the weighted average of the measurements at the monitoring stations, using inverse distance as weighing factor. IDW approaches are applied at different spatial levels. *Kriging* is also a weighted combination of measurements at surrounding monitoring stations. Kriging is the most common geostatistical technique used in the air pollution field. Kriging assigns weights at each concentration by exploiting the spatial correlation among the observed measurements. It generates the estimate and standard deviation. Kriging models exploit spatial dependence in the data to develop continuous surfaces of pollution. IDW considers the distance and more stations but is not suitable to be used on e.g. an urban scale where very high spatial gradients may exist (e.g. street canyons). Also, Kriging assumes a homogeneous terrain where concentrations are only determined by the distance to the nearest AQMS, while this is not the case in real-life. Geostatistical modelling requires a dense network of sampling sites.

- Land-Use Regression (LUR) models use measured pollution concentrations at locations in the study areas to predict concentrations at unmonitored locations based on land use types within buffers around the locations. They are based on the principle that the pollutant concentrations at any location depend on the environmental characteristics of the surrounding area. The models are developed through construction of multiple regression equations describing the relationship between the pollutant measurements at the monitoring stations and the predictor variables usually obtained through Geographic Information Systems (GIS), such as traffic intensity, road length, distance to the major road, road type, population density, land cover, wind speed, etc. A dense monitoring network is required to cover the different land-use parameters. LUR is a *common technique* to assess spatial variation in air pollution to estimate exposure to air pollution in epidemiological studies (Jerrett et al., 2005; Hoek et al., 2008; Brauer et al., 2008; Beelen et al., 2014) or health studies (e.g. Dons et al. 2014; Hoek et al., 2011)). Therefore, the basic principles are explained here in more detail.

- *Input data*

LUR models require AQM data at multiple locations across the study area. Typically, stationary monitoring is used at 20-100 locations (Hoek et al., 2008). However, Basagana et al. (2012) proposed that LUR models for complex urban settings should be based on a much larger number of measurement sites (> 80 in their study). Mobile monitoring can be a way to improve the spatial resolution of the measurements. In addition to AQM data collected, the LUR model uses predictor variables including traffic, population and land-use variables in buffers with variable sizes.

- *Spatio-temporal resolution*

In most cases, the LUR models focus on the spatial variation in (annual) average concentration and do not include a temporal dimension. However, in many applications, temporal variability is an important factor for exposure. To incorporate the temporal dimension in LUR models, different approaches are used in literature. One approach is temporal adjustment of annual average model output (Brauer et al., 2008; Wu et al., 2011; de Nazelle et al., 2013), in which the annual average exposure at each location is adjusted to temporal variations in air pollution concentrations. Another approach is to develop separate models for each typical hour (Dons et al., 2013) or for each time period (Hasenfratz et al., 2015; Mueller et al., 2016). A third approach is to include time-dependent data that are related to the temporal variability of the air quality in the model (e.g. Maynard et al., 2007; Ragettli et al., 2014 and references in Van den Bossche et al., 2020). In Van den Bossche et al. (2020) opportunistic measurements by city wardens are used to build a real-time pollution map. Possible time-dependent variables include meteorological variables (wind speed and direction, temperature) and air pollution measurements at fixed site monitoring stations.

- Machine Learning (ML) Models: Due to the observed correlation over both space and time, data matrices of air quality data are in some cases considered 'low rank' and thus explainable by statistical/numerical techniques (Asif et al., 2016; Udell & Townsend, 2019). The underlying low rank and slowly time-varying structure of the air quality data can be leveraged to create numerical models that facilitate an effective spatiotemporal extrapolation, enabling the prediction of air quality at unmonitored locations (Paliwal et al., 2020). Machine learning (ML) approaches allow for training of underlying dependencies based on large air quality datasets and supplied context information (traffic, meteorology, street type, speed limit), hereby enabling data inference or matrix completion (**¡Error! No se encuentra el origen de la referencia.**) in both s pace and time (Hofman, Do, et al., 2022). These statistical approaches are *data driven* and include models based on copula functions and neural network models. The term 'data-driven spatial prediction methods' is to distinguish them from the dispersion models; data driven models do not rely on underlying physical processes. Examples of machine learning models to predict air quality include Do et al. (2019; 2020), Lim et al. (2019), Qin et al. (2022; 2021). Models of Do et al. and Qin et al. were recently applied and validated based on diverse mobile (bicycle, Google car, postal van) datasets resulting in comparable model performances, ultimately depending on the applied instrumentation (sensor performance) and acquired spatiotemporal data coverage (Hofman, Do, et al., 2022).

5.6 Conclusions on mobile measurements

Mobile monitoring can be applied to generate a hyperlocal exposure air pollution map in a city or region. The map can be generated based on measured data only (data-only approach) or by combining measured data with models often incorporating geographical predictor variables. Given the large temporal and spatial variability of air quality concentrations, mobile monitoring has some challenges because of the spatio-temporal nature of the collected dataset. Care should be taken during data collection and/or data processing for a good data interpretation.

Deriving the map from only measurements can be achieved by measuring every single street segment a lot of times. This is possible in a small area, like a couple of streets or a neighbourhood (for example to measure pollution before and after an intervention policy), and can also be interesting to study specific trajectories (e.g. to compare commuting traffic or routes to school). However, for a regular (European) city, this takes a large amount of time and effort. Nevertheless, examples exist in Antwerp (BE), Mechelen (BE), Oakland (US), Amsterdam (NL) and Copenhagen (DK). For large (national) cohorts, it takes too much time to measure. A limited number of repetitions to overcome this huge workload can create uncertainty in average concentration levels. Therefore, in most studies that use mobile monitoring, the measurements are supplemented with a model. The numbers of repetitions required for data-only mapping depends on the study area (variability of sources) and on the data processing (e.g. background correction has shown to reduce the number of required repetitions). It also depends on the targeted representativity of the map (e.g. daily average or peak exposure).

One more potential disadvantage of the data-only approach is the fact that the resulting map consists of only road segments and not the locations where people live. Depending on the use case, there may be a need to translate on-road measurements to residential addresses. Work in the Netherlands has documented that UFP and BC models derived from mobile monitoring represented contrast in residential concentrations well without further procedures. Furthermore, only a modest overestimation was found. On-road measurement translation could be done for example by transfer learning, a machine learning method that uses long-term/residential concentration distributions to better estimate residential exposure. The biggest advantage is the small-scale variation (<50m) and hotspots of air pollution it can detect, whereas LUR models generally have more difficulty characterising small-scale variation.

Data-only mapping seems viable from a policy standpoint, where small-scale variation can be detected and acted upon with interventions. From an epidemiological standpoint, especially regarding large multi area cohorts, it makes more sense to use models. It is evident that with just a few drives on a road segment (1-4 drives) we are not able to characterise long-term average concentrations, but a LUR model can easily achieve a good correlation with only 1 or 2 drives. However, some studies found that the ability of data-only mapping surpasses the LUR model at about 5 drives when it comes to predicting long-term average (on-road) concentrations (Messier et al., 2018). Future research should verify if this holds in other geographical areas and for other pollutants. For example, UFP is more variable than NO_2 in urban environments and mobile measurements might need more repeats to achieve a stable average. Areas where non-traffic sources are dominant, with a stronger dependence on wind direction may require a different amount of mobile data.

The balance between data-only and model maps also depends on how extensive and detailed predictor variables are available. More and better predictors are likely to increase the performance of LUR models. Mobile monitoring complements traditional air quality data with more variation in contextual information (traffic, urban topology, road types, …), when compared to stationary AQMS, which is important information to train spatiotemporal models.

For monitoring campaigns that intend to create **LUR models**, the number of repeats is less of an issue, as only a few repeats on a large number of road segments are needed for model development. It is important that there is enough spatial coverage by including all different spatial characteristics (road type variation) of the domain. Similar considerations apply to representing other relevant sources in the neighbourhood, e.g. airports, harbours. Next, one needs to consider the time of day (daytime, night-time, rush hour), day of the week and season the road segments are measured. Multiple studies found that with a limited number of repeats and limited amount of street segments robust external predictions can be made. In a mobile setup, street segments with similar characteristics serve as pseudo repeats, meaning LUR models can be developed based on street segments with mobile measurements only measured once, as long as coverage and distribution of all predictor variables is similar to the prediction sites. It is important that the situation during the measurements is representative (e.g. no road closure/deviation, or very polluting vehicles that passes).

Most LUR models based on mobile data use the road segments as spatial aggregation for their models, defined as stretch of road from one intersection to the next. Length is mostly between 100 and 300m in urban environments. A few studies evaluated the impact of spatial resolution in a mobile monitoring campaign and found very little difference between the performance of LUR models that were based on segments where concentrations were averaged over 50, 100 and 200m. As the goal of mobile monitoring is often to find fine-scale variation of air pollution it is best to keep the spatial resolution as low (< 200m) as possible. It is important to aggregate homogeneous points, possibly favouring street segments over regular grids.

Data-only mapping can also be combined with LUR models in a **mixed-model framework**. Here, a LUR model is used to create a base map, and with more and more measurements more local variation can be added to the map. Since models will never perfectly predict concentration levels there will always be a moment where measurements are more precise than models when measuring long enough. By creating a mixed model with a LUR model as the fixed effects and all road segments as random effects, both LUR model and all measurements influence the predicted concentration per road segment. The more precise the on-road measurements are (i.e., less variation in measurements), the more this influences the output. This approach requires repeated measurements on all roads of interest.

The resulting dataset can also be used to validate and improve **dispersion models**. Van Poppel et al., 2024 showed how mobile data can be used for model validation. The study showed that that dispersion models can underestimate concentrations at (some) traffic locations, related to absence of good traffic data. On the other hand, using mobile data for dispersion model validation also has some concern like synchronisation of modelled and measured data; in the study of Van Poppel et al. (2024), dedicated model runs were performed to map modelled with measured data; Another concern is the occurrence of specific events during the mobile measurements, which can be accommodated by repeated measurements.

The performance of **machine learning models** is very sensitive to the representativity of the air quality dataset. model performance still relies on the spatial representativity (spatial monitoring coverage) of the mobile measurements. Accurate and representative data in both space and time is, therefore, needed to properly train the models and provide reliable results (Hofman, Do, et al., 2022). As a result, machine learning models will only be useful when a vast amount of mobile data is collected (e.g. opportunistic/automated sampling on service fleet vehicles).

Data collection can be done in a **dedicated** (pre-defined route) or **opportunistic** (monitoring devices go around with platforms on day-to-day activities) way. The first requires less data processing since data during one run is collected (more or less) simultaneously, whereas the latter will need to consider that data at different locations is not collected simultaneously and may have different background.

It is clear that there are not many monitoring campaigns that include citizens for data collections. An advantage of including citizens is the benefit of their involvement, awareness raising. A disadvantage is the difficulty in controlling the data collection and its quality.

6. STATIONARY SENSOR NETWORKS

A complementary network of stationary (low-cost) sensors can give a better spatial resolution and also temporal resolution than the regulatory monitoring networks. Dense sensor networks can also be combined with models to extrapolate data and/or to improve the data quality of the sensors. Dense sensor networks can also be used to evaluate dispersion models. In general, similar data processing techniques can be used as describe for mobile monitoring. As already mentioned in the previous chapter, care should be taken to assure a good data quality of the sensors. **Table** shows an overview of different projects where fixed monitoring tools are used.

Table 1: *Overview of projects where sensor networks are used and of which selected methods will be further described in this deliverable.*

Name project	Short description	Partner involved	Reference/link
CAPTOR	Low-cost ozone sensors for fixed deployment	CSIC	https://www.captor-project.eu/en/ Ripoll et al., 2019. STOTEN, 651, 1166-1179 https://cordis.europa.eu/project/id/688110/es;
Breathe London		UU, ICL	https://www.breathelondon.org
PANACEA	Low-cost PM2.5 sensors for monitoring air quality	NOA	https://www.iqair.com/greece/attica/athens/panacea-national-observatory-of-athens
HOPE Healthy Outdoor premises for everyone	AQT530 mid-cost sensor network supporting the regulatory air quality monitoring and air quality modeling	UHEL	https://ilmanlaatu.eu/briefly-in-english ; Petäjä, T. et al. (2021) Added value of supporting air quality observations with the use of Vaisala AQT530 sensor as a part of a sensor network, Frontiers in Env. Sci., https://doi.org/10.3389/fenvs.2021.719567.
RI-URBANS*	Virtual sensor for BC	CSIC, UHEL	J. Rovira, J.A. Paredes-Ahumada, J.M. Barceló-Ordinas, J. García Vidal, C. Reche, Y. Sola, P.L. Fung, T. Petäjä, T. Hussein, M. Viana (under review) Non-linear models for black carbon exposure modelling using air pollution datasets. Environmental Research, *submitted*.
Curieuzeneuzen	Network with diffusive tubes and citizen involvement	VITO	De Craemer et al., 2020. https://viewer.curieuzeneuzen.be/
Samen Meten	Network with low cost sensors of especially NO_2 and $PM_{2.5}$ / PM_{10}.	RIVM/UU	Wesseling J, de Ruiter H, Blokhuis C, Drukker D, Weijers E, Volten H, Vonk J, Gast L, Voogt M, Zandveld M, van Ratingen S, Tielemans E. Development and Implementation of a Platform for Public Information on Air Quality, Sensor Measurements, and Citizen Science. Atmosphere 2019, 10, 445.

* This approach uses only stationary reference instruments and is not further discussed in this document

The most established low-cost sensors are diffusion tubes or badges. In various countries, including the UK and the Netherlands, cities have set up networks of especially NO_2 diffusion tubes. Extensive calibration procedures have been developed, involving co-location with regulatory grade monitors. These diffusion tubes have good performance but can only provide integrated measurements (weekly to monthly averages). Depending on the use case, this may however be sufficient (e.g. epidemiological studies of long-term air pollution exposure). Low-cost real-time sensors have several limitations due to their nature. In their current state while they are able to provide very frequent measurements, they lack the accuracy of the substantially more expensive regulatory grade instruments and are greatly affected by extreme meteorological conditions (mainly high relative humidity). Thus, constant calibration and data evaluation is needed. Low-cost sensors (OEMs) have typical lifetimes of 1-2 years and long-term performance evaluations are still scarcely reported. Regardless of that, they can open opportunities of measurements that were not feasible before due to their portability and low cost.

Using a spatially dense network can help in measuring and understanding the effect of sources that are usually "lost in the big picture", such as the effect of hyper-local sources of pollution (e.g. a restaurant, a fire or very local combustion source). Additionally, this very dense measuring network can also help in understanding the evolution of the emissions in short ranges within the urban topography and their importance in local conditions. This cannot be achieved using the existing network of expensive regulatory grade instruments, as the measuring points are rather limited and in most cases in a great distance between them.

Co-location of sensors with reference-grade analysers (at AQMS) can improve the data quality. Due to their low cost, sensors can also be used in citizen science projects. However, a good calibration and follow-up of their performance and a good expectation management are important.

Not all sensors are available at low cost; some sensor systems cost a few 1000 euros per unit, include a yearly data/cloud platform subscription and also the set-up, follow-up, maintenance and interpretation of the results needs to be considered.

7. COMBINED APPROACHES

While stationary monitoring offers high-resolution temporal air quality data (time series), mobile data captures the spatial heterogeneity while the point measurements are not very representative over longer time periods. Combining both approaches (fixed + mobile) yields valuable high-resolution data across both time and space that can serve as input for spatiotemporal models (e.g. Hofman et al.(2022)) or can be used to construct maps correcting for variability in background concentrations (Van den Bossche et al., 2015).

Mobile data collection can be combined with stationary data, using stationary data of different origin. These stationary data can come from:

- Reference data (from stationary AQMN) along the mobile route or in its surroundings.
- Sensor networks (using stationary sensors) with different locations in the same area as where mobile monitoring is performed.
- Measurements with stationary sensors where the mobile instruments are used to collect data (stationary) when people are e.g. at home, at work, ... A disadvantage of this approach is that there are no continuous time series available since stationary data are lacking during the mobile data collection.

Each of these approaches has its own advantages and disadvantages and requires a different way of combining the data.

The stationary data can be used to supplement info for the mobile data in different ways:

- To give additional information on the representativity of the mobile data in terms of concentrations; to give info on variation during one run (between start and stop) or over different repetitions.
- To scale the mobile collected data (for background variations); this can be done within different runs, between different runs or to compare data sets that are collected in different time periods/seasons.
- To calibrate sensors when co-located (for a short time).

- To adapt calibration algorithms over time (when sensor is co-located with reference site during the mobile campaign), correcting for sensor behavior over time.

We have to note that scaling or correcting data for varying background concentrations has some limitations and cannot correct for e.g. seasonal local sources.

A combined approach using stationary and mobile measurements can provide insight in air quality dynamics and contributing sources at neighbourhood level. While stationary measurements provide data of proven value (in the case of AQMN), mobile measurements (using low-cost sensors) can provide street level exposure information and be used by non-scientific personnel. Many shortcomings are expected to come with such an approach. Mobile measurements are very sensitive to very local emissions (such as cars passing by, activities with a very local footprint etc.), which while being a significant factor in personal exposure can still bias the long-term (community) exposure maps. This should be explained to the involved participants and considered in the monitoring strategy (enough repeated measurements), data processing, and dissemination.

Stationary measurements are often used for calibration and validation, model development, background normalisation. Many mobile use cases can, therefore, be regarded as combined approaches. Stationary data can be derived from existing data sources (e.g. AQMS), but can also be collected using a dedicated (temporary) network or fixed sensor site. An example of a combined approach with dedicated fixed sensor network is the Birmingham pilot. This study will help in better understanding not only the strengths and weaknesses in using low-cost sensors (as their performance will be evaluated using the nearby BAQS, as well as against each other), but for more demanding applications as the one presented here and the real-life challenges arising with citizen involvement as well.

8. RECOMMENDATIONS: Lessons from RI-URBANS pilot studies

One overall message from the three pilot studies is that substantial new information can be learned by campaigns of mobile monitoring or low-cost sensor (LCS) fixed site indoor and outdoor monitoring of a limited duration (weeks to months) with a proper design. This makes these methods useful as an addition to routine monitoring with reference grade instruments at a small number of fixed sites. Obviously, the longer the campaign, the more precise estimates of air pollution patterns can be obtained.

8.1 Monitoring strategy and instrumentation

Detailed frequent quality control is essential, both in studies employing LCS and in studies using reference grade instruments. Comparisons with routine monitoring stations prior to, during and after mobile monitoring studies is important, preferably in representative pollutant and meteorological conditions (validation area is similar to application area). This is not always feasible for pollutants that are sparsely measured such as ultrafine particles and BC in the Birmingham pilot. Also, intercomparison of the LCS or mobile equipment used (if relevant) against each other is recommended in order to determine between-sensor-uncertainties (important for multiple sensor/network applications).

The Birmingham pilot also delivered a large number of practical considerations to make low-cost monitoring near homes possible, including the observation that connection to the mains was a limitation and therefore (car) batteries were a viable solution.

Experiences with specific mobile instruments and citizen monitoring were obtained. The AE-51 portable BC monitor provided useful information in both Birmingham (UK) and Rotterdam (The Netherlands). The frequent required maintenance of the DiSCmini instrument was burdensome for citizens and schools, rendering the instrument less useful for longer-term monitoring. The PM sensor needed much less attention. Moreover, out-of-the-box performance of portable PM and BC sensors showed to be better than current-available mobile NO_2 sensors (Hofman et al., 2024).

Comparability of different LCS sensors is an issue when comparing data. In the Birmingham pilot, it was found that calibration of the OPC-N3 against each other, and then a calibration of one of them against the research grade instruments had the best results.

In the three pilots, mobile campaigns were performed with cars, bicycles and pedestrians. All these options resulted in informative campaigns. It depends on the research question and the size of the study area which platform is most suitable. Cars lend themselves more easily to include non-portable reference grade instruments such as the AE33 aethalometer instead of the cheaper and portable AE51, while bicycle/pedestrian applications allow for more participants and, therefore, denser data collection.

As there is well-documented seasonal variability, conducting monitoring in at least two seasons is recommended if the interest is in long-term average exposure. If the interest is in specific sources e.g. wood burning than a season-specific campaign could be sufficient. Data-processing can further 'translate' the average measurement values to more representative long-term average exposure.

In none of the pilots, it was feasible to perform mobile monitoring at night, hence the measured concentrations primarily reflect daytime averages. In the Rotterdam car pilot, a limited number of monitoring days were performed in the weekend, but as in the other pilots mostly were performed in weekdays.

For mobile measurement approaches, the data collection can be done in a dedicated or opportunistic way with their specific pros and cons (see section 5). In the Rotterdam pilot, we showed an approach which was in between, with participants (and their trajectories) selected by the scientists and at the same time data collected during everyday activities (commuting).

8.2 Involvement of citizens

The approach of involving citizens in mobile monitoring of air pollution worked well, both in Birmingham and Rotterdam. In Rotterdam, these were employees of the DCMR and municipality of Rotterdam. These were more knowledgeable in terms of air quality than the average citizen. In addition, it made it easy to have a local coordinator from DCMR organizing and supervising the campaign. In Birmingham citizens were invited from the population, sometimes with a reward for participation added. The latter increased the response rate, but also resulted in students primarily interested in the reward.

The Birmingham pilot obtained useful experiences for interacting with citizens for LCS monitoring in or near their home. Building trust between citizens and researchers is an important issue. Respecting anonymity is another requirement for both citizens and schools and other organizations. Providing relevant feedback is important as citizens often participate because they are interested in the topic.

The Birmingham indoor AQ showed that people were interested in the AQ of the spaces they spend most of their time in, and the factors that affect their quality of life. Simultaneous collection of data from as many indoor environments as possible. In Birmingham, houses as well as the three classrooms had significant differences between them, even though they were located very close to each other.

Awareness raising, identification of hotspot locations, more representative personal exposure assessments in urban environments/during specific activities (e.g., commuting) were identified as added value of involving citizens in the Rotterdam pilot.

8.3 Data processing

For calibration of the PM LCS, careful consideration of the shape of the low cost sensor – reference grade measurements relationship is needed. A simple linear relationship is not optimal. Consideration of outliers related to high humidity is another key consideration.

As mobile measurements consist of point measurements in time and are affected by fluctuating background concentrations, In the Rotterdam pilot we applied a rescaling method to remove some of the temporal variation, using data from a continuous background air quality monitoring station. The rescaling method can be applied to 1) extrapolate measured averages to long-term exposure and/or to 2) rescale the data collected at another period in time (e.g. different time of the day)

GPS needs to be of good quality and generally is <10m in commercially available portable air quality sensors (Hofman et al., 2024). The geographical resolution that is supported in the Rotterdam pilot was about 50 m, even though several of the measurement were 1-second measurements.

Mobile measurements were assigned to the nearest road segment using GIS tools. This added a limited amount of uncertainty. Careful map matching of the original GPS readings is needed in addition to scripted procedures to avoid potential large errors, as we observed on a single day in the Rotterdam car pilot, and in the bicycle data.

8.4 Modelling strategy

The monitoring in itself were useful to document variability in space, e.g. between different homes in Birmingham using longer-term monitoring with LCS or along routes frequently driven by DCMR employees. In addition, the monitoring served as data to derive empirical models. This was done in Birmingham using machine-learning, and in Rotterdam using supervised linear regression, based on a variety of land use predictors.

Machine-leaning models in Birmingham performed better when predicting $PM_{2.5}$ than PM_{10} data, which were a lot harder to model, possible related to the LCS sensor used as low-cost PM sensors are not able to quantify the coarse (2.5-10µm) particulate matter fraction (Hofman et al., 2024; Vercauteren, 2021).

In the Rotterdam pilot, we in addition to the models based upon mobile monitoring, had results from dispersion modelling from our partner DCMR. Both types of models were moderately correlated. Integration of the two approaches is optimal. Moreover, discrepancies between mobile measurements and modelling results in some cases showed weaknesses in terms of model assumptions (impact of urban ventilation in open areas which was not considered inside the model).

8.5 Results

Mobile monitoring was an effective tool for characterizing spatial variation in the three pilot cities Rotterdam, Birmingham and Bucharest. Substantial variation was found within 1x1 km grid cells in all cities (Deliverable 27, WP4.3 and figures included from the Bucharest pilot).

In the Rotterdam pilot, we identified the impact of motorized traffic as in many previous mobile monitoring studies. In addition, by linking wind direction data, source locations and the individual measurements (that is, not averages across measurement days), we also found influences of the harbor, industrial area and airport on the measured air pollution variables.

Indoor measurements were usually very consistent without significant discrepancies from the research grade instruments' measurements. A simple linear regression between the OPC-N3 sensor in Birmingham and the research grade instrument used was enough to greatly improve the data.

Opportunistic data collection in the Rotterdam citizen pilot, resulted in inadequate number of repeated runs on most of the trajectories in order to derive representative long-term average maps based on measurements alone. The minimal number of required repeats for long-term average representativity, based on the subsampling analysis, varied between 24 and 54 (depending on the road segment) to be within 25% of the mean when considering the raw BC values, or between 22 and 39 when applying additional post-processing (winsorizing and/or background normalization).

9. CONCLUSIONS

Different approaches can be used to assess exposure to pollutant concentrations including ultrafine aerosol particles (UFP), black carbon (BC), nitrogen dioxide (NO_2), particulate mass below 2.5 µm ($PM_{2.5}$) at high spatial resolution for epidemiological studies and other applications. Mobile sensing platforms and fixed (low-cost) sensor networks can be used as complementary tools to data from fixed regulatory AQMN to map pollutant concentrations at a higher spatial density. These data are needed to obtain a better estimate of exposure and related health outcomes.

We make a distinction between **mobile/fixed measurements** and experimental designs **with/without citizens**. The collected data can **be processed and analyzed** using only measured data or using interpolation/modelling techniques like Land Use Regression (LUR)-based or machine learning models. The selected techniques used for data processing may have an impact on the required data collection approach.

Each of the approaches has strengths and weaknesses. When selecting a method, the user needs to define the aim of the data collection and other considerations e.g. one may prefer to engage citizens as part of awareness training on AQ (Figure 4).

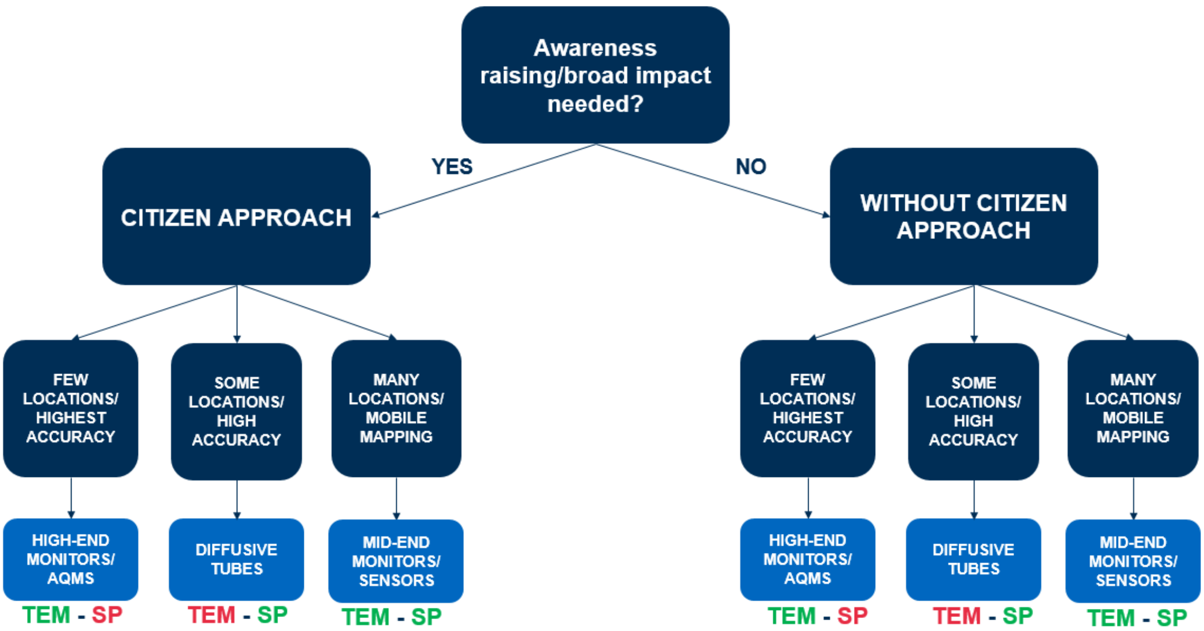

Figure 3: *Pathways for selecting methods for urban mapping. TEM refers temporal and SP to spatial. Green implies that the method performs well, red that is performs poorly for the spatial or temporal aspect*

Mobile monitoring can be used to map pollutants at a high spatial resolution with a limited number of instruments (in contrast to stationary monitoring) and can also use high-end or mid-end instruments exhibiting higher data quality than sensors. Mobile monitoring has some challenges because of the spatiotemporal nature of the collected dataset. Care should be taken during data collection and/or data processing in order to obtain representative results.

Low-cost fixed sensor networks have several limitations, especially for the real-time sensors which have shown varying performances. Good performance has been documented for low-cost diffusive samplers. Diffusive samplers only provide weekly to monthly averages, but this may be sufficient for specific use cases. If so, diffusive samplers are the method of choice. While real-time sensors are able to provide very frequent measurements, they lack the accuracy of the substantially more expensive regulatory grade instruments and are greatly affected by extreme meteorological conditions (mainly high relative humidity). Therefore, a proper calibration and validation approach is needed. We recommend co-location performance evaluation to evaluate intra- and inter-sensor uncertainty and continuous calibration/validation under representative pollutant and meteorological conditions to compensate for seasonal effects from e.g. temperature and relative humidity. Regardless of that, they provide sensing opportunities that were not feasible before due to their portability and low cost. Using a spatially dense network can help in measuring and understanding the effect of sources that are usually "lost in the big picture", such as the effect of hyper-local sources of pollution.

Overall conclusions are:

- Mobile monitoring can collect more datapoints but gives a snap-shot (no time trends). Repeated measurements, and associated sensitivity analysis, are needed to obtain representative results
- Post-processing (data cleaning, rescaling) and model approaches have shown to be viable approaches to translate mobile measurements into actionable long-term exposure data.
- Targeted mobile monitoring makes it easier to compare different locations (collected at the same time) compared to opportunistic approaches (but are often more constrained in space and time coverage)
- Opportunistic approaches mostly require less effort in data collection, result in large datasets but need proper processing to obtain representative results
- Involving citizens can significantly contribute to awareness raising and obtained impact from the collected results, but will require more time and effort for communication, logistics and dissemination.
- Sensors can be deployed at a lower cost compared to mid- and high-end instruments but suffer from lower data quality. If a high temporal resolution is no requirement, diffusion samplers may be the preferred low-cost approach.
- Sensor calibration and compensation can significantly improve data quality
- Different approaches for sensor calibration are used, varying from co-location calibrations next to AQMS or reference sensors, real-time network calibrations based on sensor-AQMS comparisons, or purely data-driven models including known covariates (e.g. temp, RH, O_3).

10. REFERENCES

Asif, M. T., Mitrovic, N., Dauwels, J., Jaillet, P., 2016. Matrix and tensor based methods for missing data estimation in large traffic networks, *IEEE Trans. on Intel. Transp. Syst.*, *17*(7), 1816-1825; https://doi.org/10.1109/TITS.2015.2507259.

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., Bartonova, A., 2017. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?, *Environ. Intern.*, *99*, 293-302; https://doi.org/10.1016/j.envint.2016.12.007.

deSouza, P. N., 2022. Key Concerns and Drivers of Low-Cost Air Quality Sensor Use [Article], *Sustain.*, *14*(1), 584; https://doi.org/10.3390/su14010584.

Do, T., Nguyen, D., Tsiligianni, E., Aguirre, A., Manna, V., Pasveer, F., Philips, W., Deligiannis, N., 2019. Matrix Completion with Variational Graph Autoencoders: Application in Hyperlocal Air Quality Inference, IEEE Intern. Conf. on Acou., Speech and Signal Proces. (ICASSP), Brighton, UK, pp. 7535-7539; https://doi.org/10.1109/ICASSP.2019.8683787.

Do, T. H., Tsiligianni, E., Qin, X., Hofman, J., La Manna, V. P., Philips, W., Deligiannis, N. (2020). Graph-Deep-Learning-Based Inference of Fine-Grained Air Quality from Mobile IoT Sensors, *IEEE IoT J.*, 1-1; https://doi.org/10.1109/JIOT.2020.2999446.

Hofman, J., Do, T. H., Qin, X., Bonet, E. R., Philips, W., Deligiannis, N., La Manna, V. P., 2022. Spatiotemporal air quality inference of low-cost sensor data: Evidence from multiple sensor testbeds, *Environ. Modell. & Softw.*, *149*, 105306; https://doi.org/https://doi.org/10.1016/j.envsoft.2022.105306.

Hofman, J., Nikolaou, M., Shantharam, S. P., Stroobants, C., Weijs, S., La Manna, V. P., 2022. Distant calibration of low-cost PM and NO_2 sensors; evidence from multiple sensor testbeds, *Atmosp. pollut. res.*, *13*(1), 101246; https://doi.org/https://doi.org/10.1016/j.apr.2021.101246.

Hofman, J., Peters, J., Stroobants, C., Elst, E., Baeyens, B., Laer, J. V., Spruyt, M., Essche, W. V., Delbare, E., Roels, B., Cochez, A., Gillijns, E., Poppel, M. V., 2022. Air Quality Sensor Networks for Evidence-Based Policy Making: Best Practices for Actionable Insights [Article], *Atmosp.*, *13*(6), 944; https://doi.org/10.3390/atmos13060944.

Hofman, J., Lazarov, B., Stroobants, C., Elst, E., Smets, I., Van Poppel, M., 2024. Portable Sensors for Dynamic Exposure Assessments in Urban Environments: State of the Science, In Preprints: Preprints.

Huyse, H., Bachus, K., Merlevede, T., Delanoeije, J., Heidi Knipprath, H., 2019. *Societal Impact of the Citizen Science Project "CurieuzeNeuzen Vlaanderen"*, H.K. Leuven; https://hiva.kuleuven.be/nl/onderzoek/thema/klimaatendo/p/projecten/CurieuzeNeuzen_Vlaanderen_EN.

Kang, Y., Aye, L., Ngo, T. D., Zhou, J., 2022. Performance evaluation of low-cost air quality sensors: A review, *Sci. of The Tot. Environ.*, *818*, 151769; https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.151769.

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, *Environ. Intern.*, *75*, 199-205; https://doi.org/10.1016/j.envint.2014.11.019.

Lim, C. C., Kim, H., Vilcassim, M. J. R., Thurston, G. D., Gordon, T., Chen, L.-C., Lee, K., Heimbinder, M., Kim, S.Y., 2019. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in Seoul, South Korea, *Environ. Intern.*, *131*, 105022; https://doi.org/10.1016/j.envint.2019.105022.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G. S. W., Jayaratne, R., Kumar, P., … Williams, R., 2018. Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?, *Environ. Intern.*, *116*, 286-299; https://doi.org/10.1016/j.envint.2018.04.018.

Paliwal, C., Biyani, P., Rajawat, K., & Sutaria, R. (2020, 2020). Scalable Spatio-Temporal Measurements and analysis of Air pollution data using Vehicle Mounted Sensors. Air Sensors International Conference - Virtual Fall Series, Qin, X., Huu Do, T., Hofman, J., Rodrigo, E., Panzica La Manna, V., Deligiannis, N., Philips, W., 2021. Street-level Air Quality Inference Based on Geographically Context-aware Random Forest Using Opportunistic Mobile Sensor Network, Intern. Conf. on Innov. in Artif. Intel. (ICIAI 2021), pp. 221-227; https://doi.org/10.1145/3461353.3461370.

Qin, X., Do, T. H., Hofman, J., Bonet, E. R., La Manna, V. P., Deligiannis, N., Philips, W., 2022. Fine-Grained Urban Air Quality Mapping from Sparse Mobile Air Pollution Measurements and Dense Traffic Density, *Rem. Sens.*, *14*(11), 2613; http://dx.doi.org/10.3390/rs14112613.

Udell, M., Townsend, A., 2019. Why are big data matrices approximately low rank?, *SIAM J. on Math. of Data Sci.*, *1*(1), 144-160; https://doi.org/10.1137/18M1183480.

Van Poppel, M., Peters, J., Vranckx, S., Van Laer, J., Hofman, J., Vandeninden, B., Vanpoucke, C., Lefebvre, W., 2024. Exploring the Spatial Variability of Air Pollution Using Mobile BC Measurements in a Citizen Science Project: A Case Study in Mechelen, *Atmosp.*, *15*(7), 757; http://dx.doi.org/10.3390/atmos15070757.

Vercauteren, J., 2021. Performance evaluation of six low-cost particulate matter sensors in the field, LIFE VAQUUMS Report; https://vaquums.eu/sensor-db/tests/life-vaquums_PMfieldtest.pdf/view.

VMM, 2022. *COMPAIR (Community Observation Measurement & Participation in Air Science), Co-creatieve workshop, Burgerwetenschap, luchtkwaliteit & data*, Flanders Environment Agency (VMM); https://en.vmm.be/projects/compair.

Weijers, E., Vercauteren, J., van Dinther, D., 2021. *Performance evaluation of low-cost air quality sensors in the laboratory and in the field*, VAQUUMS; https://vaquums.eu/sensor-db/tests/protocols/life-vaquums_testprotocol_final.pdf/view.

WMO, 2024. Integrating Low-cost Sensor Systems and Networks to Enhance Air Quality Applications. (GAW Report No. 293), World Meteorological Organization (WMO); https://wmo.int/publication-series/integrating-low-cost-sensor-systems-and-networks-enhance-air-quality-applications.