Measuring and Modeling Air Quality in Smart Cities

Authors: Jelle Hofman, Valerio Panzica La Manna, Jana Muylaert

ABSTRACT

Air pollution is currently regarded as top environmental threat for our health and top priority at the EU policy agenda (Zero Pollution EU Green Deal (1)). Moreover, more stringent EU regulations and newly emerging limit values can be expected following the WHO Guideline values. To face these challenges, cities will need to be prepared to provide their contribution in more fine-grained air quality assessments and pollution reductions. IoT and low-cost sensors, together with proper calibration and algorithmic approaches, can be used as a complementary tool for improved air quality assessments and evidence-based policy support. Engaging citizens in this transformation process is beneficial on multiple aspects: citizen science campaigns can extend the network (distributing deployment and maintenance effort), increase awareness and understanding on air quality dynamics and initiate behavioral change ultimately reducing air pollution. This report outlines the potential and limitations of air quality sensors including technical properties, differences with regulatory instruments, sensitivities and applications, data insights and analytics and the need for calibration and validation. While extending and complementing their environmental monitoring networks, cities should use sensor data as an additional tool to inform and properly engage all stakeholders in fighting air pollution together.

1. INTRODUCTION

Although air quality has improved substantially over the past decades, it is still regarded as the biggest environmental health risk in Europe, affecting people's health and the ecosystems we rely on (2). It is an invisible threat affecting every single one of us. According to the European Environmental Agency (EEA), 77% of the European urban population is exposed to fine particulate matter ($PM_{2.5}$) concentrations exceeding the World Health Organisation (WHO) guideline values, while 10% of the monitoring stations are reporting EU limit value exceedances for nitrogen dioxide (NO_2) (2). Worldwide, more than 80% of people living in urban areas are still exposed to air quality levels exceeding WHO guideline values (3). Especially in urban environments where both pollution sources and people affected by pollution are concentrated, air pollution tends to peak. As pollution levels can vary dramatically over short distances or time instances (4–9), a high monitoring resolution in both space and time should be pursued to accurately estimate population exposure. Since traditional air quality monitoring stations are rather costly and cumbersome, cities typically only deploy few at representative locations (e.g. roadside, urban background, ...). To properly assess people's exposure to air pollution, there is an urgent need for higher granularity.

Thanks to advances in low-cost sensing and Internet of Things (IoT) technologies, various affordable and small-sized air quality sensors and a wide range of monitoring applications are available today, capable of detecting representative ambient concentration levels of the most important urban pollutants; i.e. PM, NO_2 and ozone (O_3) (10–18). Applications range from indoor to outdoor, fixed to mobile, community-based to personal. However, measurement principles of ~10-600 euro sensors are typically less conditioned, accurate and selective compared to ~30 000 euro regulatory equipment, leading to lower data quality and sensitivity towards meteorological variables. Proper interpretation, validation and calibration is therefore crucial in order to make air quality sensors a smart solution.

2. What is Air Quality?

Air pollution originates from both man-made and natural sources and can be released directly into the atmosphere (primary emissions) or can form as a result of chemical reactions with so-called precursor substances. While some pollutants are directly emitted and short-lived, other pollutants are formed at a later stage and/or can travel large distances (so-called transboundary pollution). When focusing on urban environments, most important outdoor air pollutants considered by the EEA and WHO are particulate matter (PM), with aerodynamical diameter smaller than 10 µm (PM_{10}) and 2.5 µm ($PM_{2.5}$), nitrogen dioxide (NO_2) and ozone (O_3), as these compounds most significantly impact our health.

Within those environments pollution sources include road traffic, residential heating (e.g. wood burning), industrial sources, airports, constructions works, natural sources (Saharan dust), resulting in exhaust (PM, NO_2) vs non-exhaust (e.g. wearing brake pads and road dust) and directly emitted (PM) vs secondary-formed (e.g. O_3) pollution in our atmosphere.

The European Union has established health-based standards and objectives at hourly, daily and yearly level (https://ec.europa.eu/environment/air/quality/standards.htm), mainly driven by the economic feasibility. The WHO provides additional more stringent Air Quality Guideline values (https://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/), purely based on health evidence from epidemiological studies (Table 1).

Table 1: EU standards and WHO guideline values for PM2.5, PM10, NO2 and O3 according to Directive 2008/50/EC and WHO 2000 and WHO 2005.

	Legislation	Averaging period	Concentration (µg m^{-3})	Allowed exceedances
$PM_{2.5}$	EU	year	25	
	WHO	year	10	
		day	25	max 3 days/year
PM_{10}	EU	day	50	max 35 days/year
		year	40	
	WHO	day	50	max 3 days/year
		year	20	
NO_2	EU	hour	200	max 18 hours/year
		year	40	
	WHO	hour	200	-
		year	40	
O_3	EU	Highest 8h-avg of day	120	
	WHO	Highest 8h-avg of day	100	

To make these air quality levels understandable and more informative for the general public, the air quality index (AQI) is defined to provide information on the current air quality situation (multiple pollutants) in one condensed index (Table 2). Different countries have their own air quality indices, corresponding to different national air quality standards. In Belgium (BELAQI), the concentration scales are defined based on relative risks (RR; health impact of 10 µg m^3 increase), documented in the Health Risks of Air Pollution in Europe (HRAPIE) report of the WHO. The actual index (1-10) is defined by the worst pollutant level ($PM_{2.5}$, PM_{10}, O_3 and NO_2).

Table 2: BELAQI index scale, classification and associated pollutant concentration scales of 24h-averaged PM_{10}, $PM_{2.5}$ and daily max hourly O_3 and NO_2 concentration (irceline.be)

Index	Classification	PM_{10} daily mean ($\mu g/m^3$)	$PM_{2.5}$ daily mean ($\mu g/m^3$)	O_3 max 1-hourly mean per day ($\mu g/m^3$)	NO_2 max 1-hourly mean per day ($\mu g/m^3$)
1	Excellent	0 - 10	0 - 5	0 - 25	0 - 20
2	Very good	11 - 20	6 - 10	26 - 50	21 - 50
3	Good	21 - 30	11 - 15	51 - 70	51 - 70
4	Fairly good	31 - 40	16 - 25	71 - 120	71 - 120
5	Moderate	41 - 50	26 - 35	121 - 160	121 -150
6	Poor	51 - 60	36 - 40	161 - 180	151 - 180
7	Very poor	61 - 70	41 - 50	181 - 240	181 - 200
8	Bad	71 - 80	51- 60	241 - 280	201 - 250
9	Very bad	81 - 100	61 - 70	281 - 320	251 - 300
10	Horrible	>100	>70	>320	>300

3. SENSING

Currently, air quality is monitored via a network of automated regulatory monitoring stations (more than 150 in Flanders with an average cost of ~150 000 euro/station) regulated in terms of accuracy and data capture. These stations report pollutant concentrations at a near real-time (hourly) resolution. Hourly-updated measurements can be consulted via https://www.vmm.be/data/actuele-luchtkwaliteit or https://www.irceline.be/nl/luchtkwaliteit/metingen. This automated network is extended with a semi-automatic network including daily, weekly or monthly collector samples (e.g. passive NO_2 samplers) that need lab analysis. To predict pollutant concentrations in between measurement locations, environmental agencies rely on air quality models that are able to predict pollutant concentrations based on the monitored concentrations at the regulatory stations, and known line/point emissions sources, their dispersion and transformation via meteorological and urban topology data. In Flanders, the ATMO-street is used to simulate pollution dispersion over multiple spatial scales; combining a regional interpolation model (RIO; 4x4 km²), a Gaussian dispersion model of line and point emission sources (IFDM), and a street canopy model (OSPM) to simulate the reduced ventilation effect in so-called "street canyons". As these models are rather computation intensive, the lower-resolution RIO-IFDM model (100x100m) is currently applied at an hourly resolution to map pollutant concentrations in Flanders (https://www.vmm.be/data/actuele-luchtkwaliteit). The high-resolution ATMO-street model is not applied in real-time but used offline to simulate high-resolution yearly averaged maps (https://www.vmm.be/data/luchtkwaliteit-in-je-eigen-omgeving). Within the VLAIO Dencity project, a near-real-time implementation of ATMO-street was developed for the Antwerp region (https://dencity.marvin.vito.be/). The automated monitoring network is strictly regulated in terms of instrumentation, accuracy and data capture at the EU and WHO level and forms the absolute ground-truth for air quality assessments.

In addition to this existing toolset for air quality assessment, new sensor technologies, Internet of Things (IoT) applications and data analysis tools (e.g. Artificial Intelligence (AI)) are opening up new possibilities for environmental monitoring and analysis (Figure 1). These technologies and data-driven tools have enormous potential by building upon existing monitoring infrastructure, increasing its monitoring granularity in space and time, and provide more integrated insights in interrelated environmental processes by training dependencies between multiple data streams (e.g. regulatory monitoring network, satellite, traffic, citizen science sensors, ...). IMEC explores this field by (1) **benchmarking the performance** of off-the-shelf air quality sensors for PM and NO_2, (2) developing distant **calibration algorithms improving the data quality from air quality sensors**, by relying on the available reference monitoring network, and (3) developing **data-driven inference models**, interpolating air quality readings from fixed and mobile sensors in both space and time.

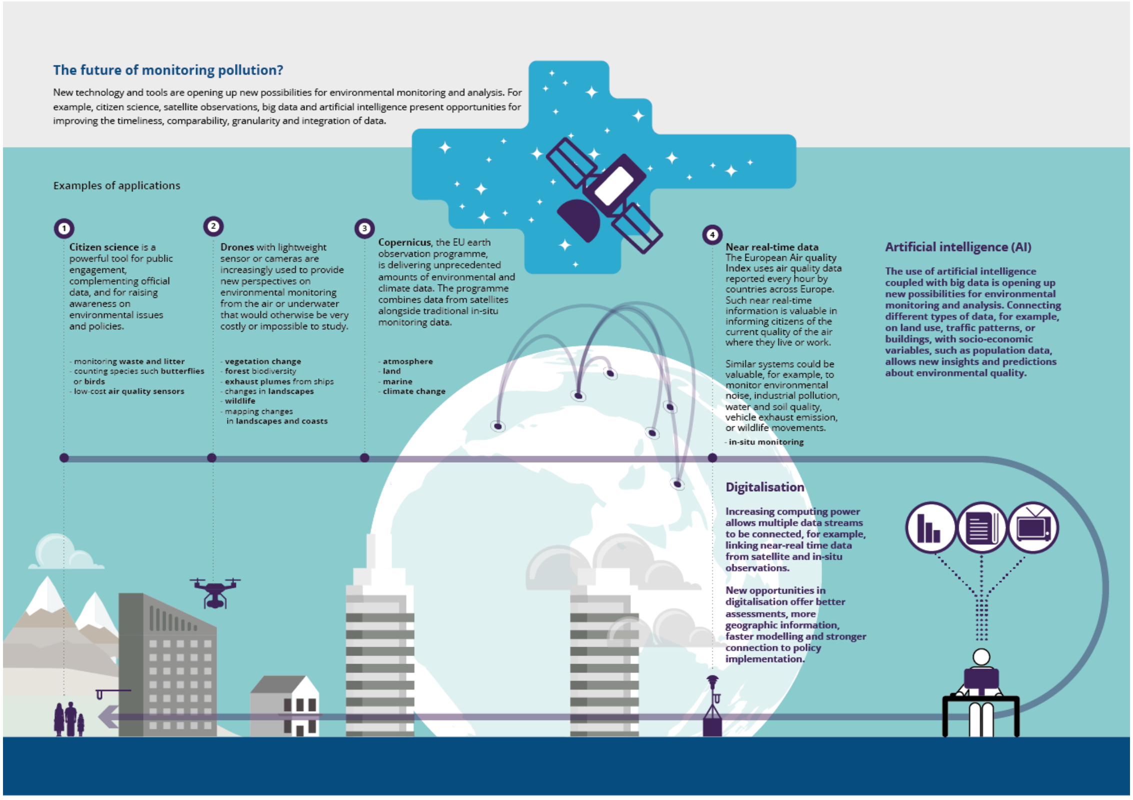

Figure 1: The future of environmental monitoring (EEA)

As a user, when thinking about an air quality sensor application in your city, it is important to consider the fitness-for-purpose. As starting point, try to evaluate **why you want to deploy air quality sensors? What are the actual questions you want to answer?** Doing so, consider that the final sensor solution should depend on its **intended purpose** and **required accuracy**. If you want to raise awareness on air quality in general and personal exposure of citizens, air quality sensors can be distributed showing the effect of indoor cooking processes, wood burning, candles, indoor ventilation, healthy routing, behavioral habits, ... As people will evaluate their own sensor in response to their behaviour, they will very rapidly get acquainted with pollutant behaviour, rush hour peaks, indoor air quality, ... without the need of a highly accurate or precise sensor. If, on the other hand, you are interested in creating a more granular air quality monitoring network, multiple sensors will need to be installed in weather-proof (IP65) housings, with dedicated power (solar, battery, adapter) and connectivity (LoRa, NB-IoT, GPRS, ...)

solutions, showing high data capture, precision (low inter-sensor variability) and high accuracy. If you would be interested in mobile sensing, power and connectivity requirements will need to be optimized towards a high monitoring resolution (~1-10 seconds), while the sensor design needs to be small and versatile.

In any case, the performance of your final sensor solution will depend on:

(i) the actual sensor
(ii) the applied hardware and casing design
(iii) the chosen power and connectivity solution
(iv) Data processing & calibration (§4)

Herewith it is important to make the distinction between "sensor" (actual sensing unit), "sensor system" (sensor(s), hardware, casing, connectivity and power) and final "sensor solution" (dashboard, quality monitoring, calibration) as shown in Figure 2:

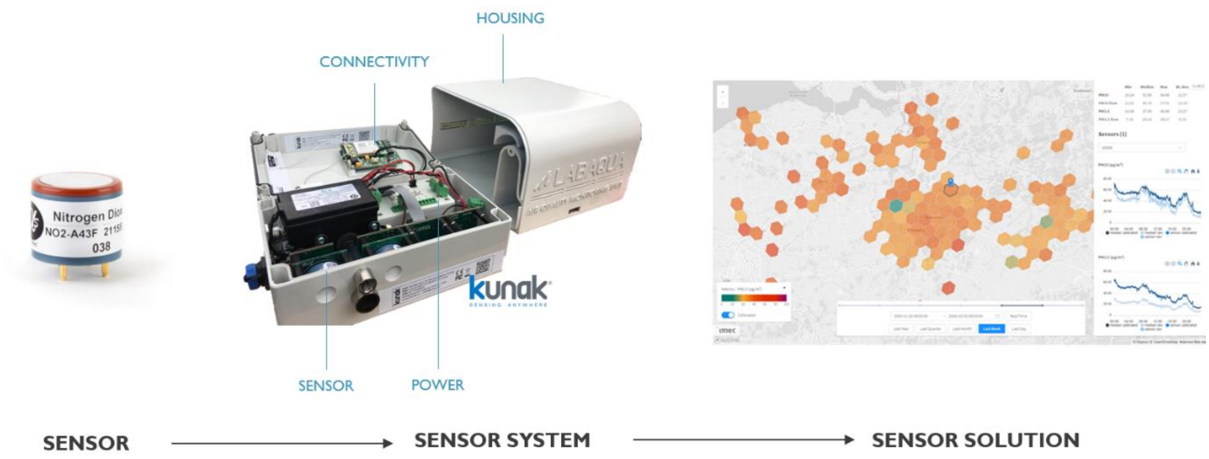

Figure 2: Difference between sensor, sensor system and sensor solution

Within the next paragraphs, we will go over the selection criteria for the sensor and sensor system one-by-one:

3.1 THE ACTUAL SENSOR

In order to select the sensor of interest, it is important to focus on the **relevant pollutants.** As mentioned earlier, pollutants considered most important in terms of outdoor health are currently particulate matter ($PM_{2.5}$ and PM_{10}), nitrogen dioxide (NO_2) and ozone (O_3). These are legislated, meaning that measured concentrations can be interpreted in terms of severity. From these, $PM_{2.5}$ and especially NO_2 are impacted by local sources (e.g. traffic, combustion, ...), therefore showing highest spatial variability. Traffic is contributing to ~40% of the total NO_2 emissions which is why NO_2 is regarded as a typical traffic tracer. When focusing on indoor environments, penetration of outdoor pollution (PM, NO_2, O_3) might be relevant, however indoors there will be no traffic emissions and photochemical O_3 formation. Indoor sources (cooking, dust resuspension, candles, incense, paint, chemicals,...) are known to impact ambient PM and Volatile Organic Compounds (VOCs) concentrations, while CO_2 (due to respiration) is an important indicator for indoor ventilation (<900 ppm considered as well-ventilated) and has shown to exert cognitive effects.

Once the monitored pollutants are decided, you can choose the actual sensors. Although no official sensor validation protocol yet exists, many 3rd party institutes have come up with their own sensor evaluation protocols and benchmarking studies which allow for comparison between different sensors; including Air Quality Index Project (http://aqicn.org/sensor/), RIVM (https://www.samenmetenaanluchtkwaliteit.nl/), AQSPEC (http://www.aqmd.gov/aq-spec), AIRLAB (http://www.airlab.solutions/en/news/results-international-challenge-2019), US EPA Toolbox (https://www.epa.gov/air-sensor-toolbox), VAQUUMS (https://vaquums.eu/), ... A wealth of literature studies evaluated various air quality sensors under controlled lab or real-life outdoor conditions (10,12,15,19–29). Frequently reported performance metrics include R^2, MAE, RMSE, Uexp, … An online advisory platform with examples/tools for citizens on how the get started with their own air quality project was created in the Interreg Zulu project (https://hoemeetiklucht.eu/).

Air quality sensors for particulate and gaseous pollutants typically consist of optical particle counters, based on laser scattering, and electrochemical gas sensors converting a chemical reaction (oxidation at sensing electrode and reduction reaction ay counter electrode) of the pollutant of interest in a quantifiable electrical current (Figure 3). Metal oxide gas sensors exist as well, relying on the gas reaction with semiconductor material, resulting in free electrons.

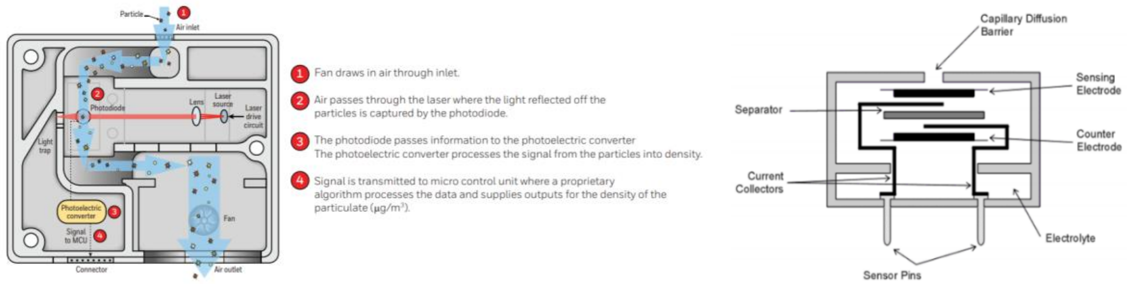

Figure 3: Working principle of particle sensors (left) and electrochemical NO_2 sensors (right)

When comparing ~30 000 euro reference equipment to ~20-600 euro low-cost air quality sensors, there are some important differences to consider:

- *Sensitivity towards temperature/relative humidity*: Particle and gas sensors typically exhibit sensitivity towards temperature and relative humidity. This pathway is however very different for particle sensors compared to electrochemical sensors. For particle sensors, relative humidity physically impacts ambient particle sizes, by growing particles via condensation. This physical effect is similar for all particle sensors and controlled by conditioning (drying) the sampled air in reference instruments while this is not the case for low-cost sensors. Relative humidity impacts electrochemical sensors very differently by changing the humidity equilibrium between the sensor electrolyte and outside air, causing the sensor to dry out, resulting in a changing sensor response. Cross-sensitivity towards other pollutants, e.g. O_3 for NO_2 sensors also exists as both pollutants have oxidative potential, triggering the electrochemical sensor.
- *Detection range/limits*: Low-cost particle sensor have no particle size cut-offs (impactor, cyclones) and lower detection range when compared to reference analyzers. While low-cost sensors are detecting particles from 0.3 µm onwards, reference equipment is able to detect particles as small as 0.18 µm. This might seem like a small difference but is rather significant when looking at particle number size distributions (Figure 4).

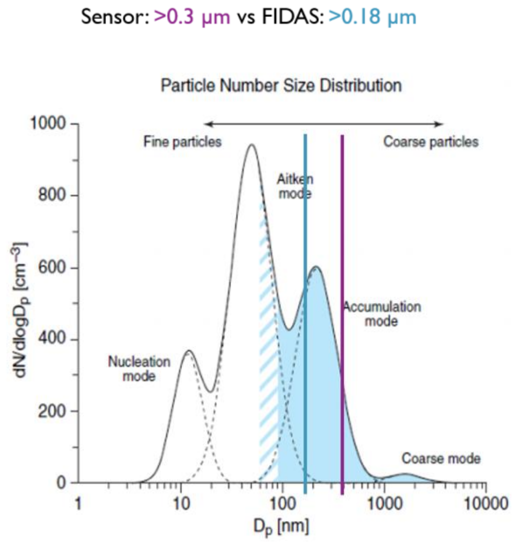

Figure 4: Lower detection limit of low-cost sensors (purple), compared to equivalent instrumentation (blue), visualized on a typical particle number size distribution (particle counts (# cm^{-3}) vs particle diameter (Dp))

Gas sensors are often sensitive in the parts per million (ppm) range, while ambient NO_2 and O_3 concentrations are typically in the 0-100 parts per billion (ppb) range. This is important as there currently are only limited NO_2 sensors available that are capable of quantifying ambient NO_2 concentrations.

- Low-cost sensors tend to drift over time because particle sensor get clogged with dust while the electrolyte inside electrochemical sensors tends to dry out. Both processes will cause changing sensor responses over time. Therefore, sensor suppliers typically provide lifetimes of ~1 year in the sensor specifications before sensors need replacing.

Besides literature studies based on existing comparative platforms, projects and studies, IMEC benchmarked several particle and gas sensors in the field next to reference (equivalent) equipment (Palas FIDAS 200S) or regulatory stations in Belgium (VMM) and the Netherlands (RIVM). Particle sensors included Sensirion SPS30, Alphansense OPC-N2 and OPC-N3, NovaFitness SDS011 and gas sensors include the Alphasense NO2A43F and showed a wide range in performance and sensitivities. VITO has a dedicated sensor validation center (https://aqsensors.vito.be/en/sensor-validation-center) and is setting up a sensor test protocol for Flanders, based on the European CEN WG42 protocol.

In order to choose a specific sensor that best fits your needs, a first selection can be made from the **sensor spec sheet**, containing information on size, power requirement, detection range, measurement resolution, ... Next, sensor selection can be based on **reported performance from independent benchmarking studies, online platforms or own experiments**. Keep in mind that the listed specifications can be very different from the performance reported in the field.

3.2 Which hardware to use?

In order to have a functional sensor solution, you need the actual sensors, printed circuit boards (PCB's), microprocessor and a solution to log or transmit the sensor data. This hardware has to be combined in a sensor housing to shield the hardware from unwanted influences of the indoor/outdoor environment. The chosen size and format of the final sensor solution will depend on its intended application. For portable applications, a sensor shouldn't necessarily be weatherproof but preferably small and versatile, while not being too power-hungry (as it will probably operate on batteries). Applications requiring continuous outdoor monitoring will need robust and weatherproof (IP65) housing. Sensors need proper exposure to outside air, either via openings in the housing or tubing to draw air in from outside. Because of the sensitivity towards temperature and relative humidity, sensors cannot be too close to the hardware (which is warming up as well) or proper ventilation of the sensor box should be foreseen. Make sure to include temperature and relative humidity sensors in the environment where the sensors are exposed (internal when sensors are inside a vented box or external when sensors are exposed outside of the sensor box).

As particle sensors rely on a sampling flow created by a small fan, ideally laminar free-flow conditions should be created. When tubing is applied, antistatic tubing is preferred while tubing length should be limited as much as possible to avoid particle scavenging. Turbulence in the sampling flow should be avoided, e.g. by putting the sensor in a dedicated housing. IMEC for example applies Kunak mobile air quality sensors on top of delivery vans (bpost) in Antwerp, relying on a dedicated housing to create laminar flow conditions over the membranes of electrochemical sensor membranes (Figure 5).

Figure 5: IMEC OCTA platform including Alphasense PM, NO_2 and O_3 sensor and LoRa antenna (left), mobile Kunak sensor hardware and dedicated housing for optimized flow conditions (middle) and picture of a deployed Kunak sensor on the roof of a Bpost van (right)

3.3 Power and connectivity

Several power and connectivity solutions are available determining the capabilities of the intended sensor solution (Table 1). Although air quality sensors have become less power consuming (<300mA at 5V), inclusion of a particle sensor typically limits the battery life as the laser and fan remain power-hungry features. Gas sensors require constant powering as electrochemical cells need a warmup period of ~4 hours in order to produce reliable readings. Current available off-the-shelf platforms consist of battery- and solar panel powered or grid powered devices.

If a local internet connection (e.g. wifi, ethernet) is no option, LoRa or SigFox are relatively cheap and reliable connectivity solutions as long as no high monitoring resolution is needed. GPRS (3G, 4G,...) and NBIoT can be used to send larger data packages or at a higher monitoring resolution (higher bandwidth). An overview of the existing IoT connectivity solutions and most important features can be found via http://iotfactory.eu/nl/iot-kenniscentrum/overzicht-van-iot-netwerken/. In practice, current available off-the-shelf platforms in most cases provide multiple connectivity solutions.

Table 3: Required specifications for dedicated sensor solutions

	Portable outdoor	Mobile continuous outdoor	Portable indoor	Fixed continuous outdoor
Low-power sensors	X	X (if battery)	X	
Weatherproof housing		X		X
Sampling resolution	High (~10 sec)	High (~10 sec)	High (~10 sec)	Low (5-15 min)
Power	Battery	Battery – 5V car connector	Battery	GPRS, wifi,
Connectivity	Bluetooth, GPRS, NBIoT	Bluetooth, GPRS, NBIoT	Bluetooth, GPRS, wifi	LoRa, GPRS, NBIoT
...				

Optimizing your air quality monitoring needs towards the selection criteria listed above should enable you to come up with an optimized sensor solution for the intended monitoring purpose. Off-the-shelf sensor platforms with integrated hardware/power/connectivity are available as well with a wide range in pricing (~100-17280 euro/3 years), depending on the considered pollutants, application, housing, included maintenance contract, ... An overview of commercially available platforms is provided in Annex 1.

4. DATA PROCESSING

When using IoT sensors for air quality monitoring purposes, some data processing steps are crucial to guarantee reliable air quality data. Today, many cities/projects are applying air quality sensors without knowledge on sensor sensitivities and data accuracy (see Sensing). When only considering raw sensor data, interpretation can lead to faulty conclusions, growing criticism and societal fatigue. This is why deployments of air quality sensors in cities should be accompanied with (1) **proper expectation management** on the possibilities and limitations of low-cost air quality sensors to all involved stakeholders and (2) a **calibration and validation protocol** to guarantee the data quality of such sensors over time. In practice, this can be achieved by locating ~3 sensor boxes next to one/multiple regulatory monitoring station(s) to compare the sensor and calibration performance over time.

4.1 CALIBRATION & VALIDATION

To cope with sensor sensitivities to meteorological conditions and sensor drift over time as explained in 3.1, correction and calibration models are needed to guarantee reliable data. Three different calibration approaches exist, namely lab-, field- and cloud calibration. During a **lab calibration**, a sensor response is tested under controlled conditions of e.g. target/interfering gas, temperature and relative humidity, but this does not mimic real-world conditions. During **field calibration**, the sensor response is tested/calibrated under changing real-world conditions. This approach has shown a better performance for outdoor applications, when compared to the lab calibration but is only valid under the evaluated calibration conditions and does not consider drift over time. **Cloud calibration** relies on data fusion of sensor data with external data sources (e.g. regulatory air quality data, environmental data, ...). This allows for network-wide corrections/calibrations and continuous calibration to cope with sensor drift over time. Different calibration approaches and associated performances have been reported in literature for PM (25,30–32) and NO_2 (14,33–37). RIVM is experimenting with online calibrations in their citizen science platform "Samen Meten" (https://www.samenmetenaanluchtkwaliteit.nl/). Within the VLAIO Dencity project (https://www.imeccityofthings.be/nl/projecten/dencity), we started experimenting with existing calibration features reported in literature and new data-driven approaches to improve the data quality from existing sensors. This by comparing the sensor performance in terms of accuracy, linearity and correlation before and after calibration (Table 2). Moreover, data quality in terms of data capture, precision and accuracy are compared to provisional data quality objectives proposed by VMM (Annex 2), as a future EU certification scheme for sensors (Class 1 –3 sensors) is being developed by CEN TC 264 WG42.

Traditional sensor calibration, by co-locating sensors temporary next to reference equipment under controlled (lab calibration) or outdoor (field calibration) conditions, corrects for the sensor sensitivities but is logistically cumbersome. Especially, since recurrent calibrations are needed to cope with sensor drift over time. IMEC developed a **cloud calibration approach**, calibrating sensors (at any location) in real-time based on reference measurements from available regulatory monitoring networks. Our algorithm provides **daily calibration factors**, considering a **35-day sliding training window**, automatic selection of **best suitable reference station** depending on the considered pollutant variability and sensor calibration based on time periods with **representative (nighttime) concentrations** (Figure 5).

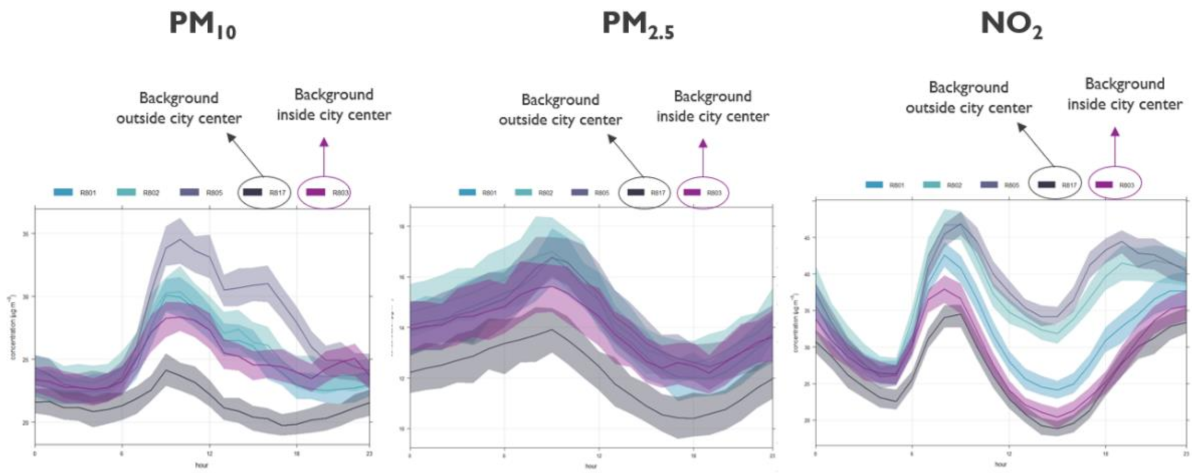

Figure 5: Experienced diurnal PM and NO_2 variability at different reference monitoring locations (VMM) in Antwerp. The shown hourly data is averaged over > year (April, 2018 - July, 2019)

	PM2.5	ACCURACY				LINEARITY		CORRELATION		
		95%CI_30	RMSE	MBE	MAE	SLOPE	INTERCEPT	R²	SPEARMAN	PEARSON
PM2.5 HOURLY	AVG RAW	128.86	13.70	-3.11	8.17	0.88	6.61	0.64	0.86	0.78
	AVG CAL v4	46.09	6.22	0.78	3.73	0.88	0.93	0.70	0.86	0.82
PM2.5 DAILY	AVG RAW	87.60	12.41	-2.65	7.92	0.95	6.00	0.68	0.87	0.81
	AVG CAL v4	26.53	4.75	0.88	2.99	0.90	0.75	0.77	0.87	0.86
PM10 HOURLY	AVG RAW	128.21	22.35	-5.31	13.62	0.50	16.61	0.39	0.70	0.62
	AVG CAL v4	63.92	11.79	-0.81	7.59	0.79	6.22	0.53	0.79	0.72
PM10 DAILY	AVG RAW	69.80	18.12	-4.98	12.48	0.58	15.16	0.44	0.70	0.65
	AVG CALv4	39.57	9.18	-0.89	6.44	0.80	5.99	0.63	0.77	0.79

Table 2: Data quality improvements of the cloud calibration algorithm, developed by IMEC-NL. Averaged results of 8 PM sensors at 8 different reference stations in 2 different cities (Antwerp, Ghent). Improvements are calculated for both $PM_{2.5}$ and PM_{10}, for hourly- and daily-averaged sensor data. Colors indicate data quality objectives for the expanded uncertainty proposed by the VMM, with red ("inadequate"; 100-150%), yellow ("sensitizing"; 50-100%) and green ("supplementary"; <50%). For more information on the applied performance metrics, see Annex 2.

Distant calibration algorithms have been developed for $PM_{2.5}$, PM_{10} and NO_2 and tested on different sensor types (NovaFitness SDS011, Sensirion SPS30, Alphasense OPC-N2/3, Alphasense A43F) in different deployments in Antwerp, Ghent, Dordrecht and Eindhoven (Figure 6). The resulting sensor performances before and after calibration can be found in Annex 2, while a scientific journal paper has been submitted (38).

Thanks to its cloud implementation, this calibration approach can be provided "as a service" on top of existing sensor networks in any city, for any sensor. The performance of the calibration will depend on the considered pollutant ($PM_{2.5}$, PM_{10}, NO_2, BC), the applied sensor system and granularity of the available reference network.

Figure 6: Sensor deployments to validate IMEC-NL's sensor solutions in Antwerp, Dordrecht and Eindhoven

To show the potential of scalable calibration algorithms, IMEC implemented its calibration algorithm for PM (v4.0) in the cloud, improving sensor data from all Belgian NovaFitness SDS011 sensors, openly available via (https://sensor.community/nl/)(Figure 7). From the time series graphs in Figure 7, it can be observed that the distant calibration corrects for the general sensor underestimation, while retaining the measured sensor variability.

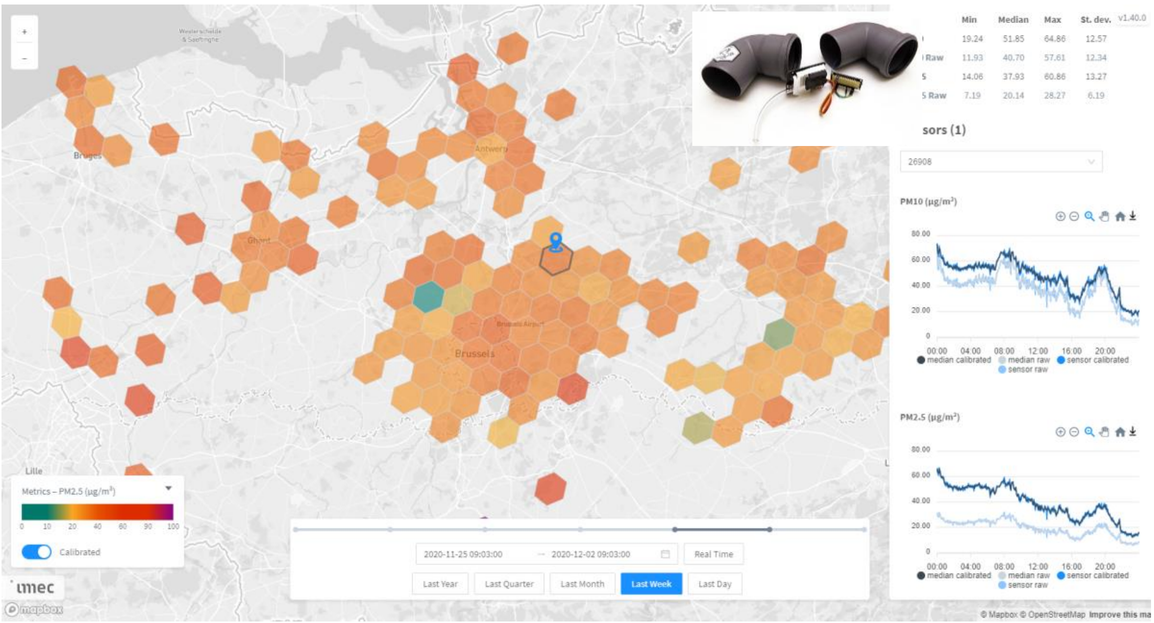

Figure 7: Visualization dashboard showing IMEC-NL's real-time calibrated sensor data from openly available PM sensors (https://sensor.community/nl/)

4.2 Visualisation

When visualizing air quality data, it is important to consider (1) the type of collected sensor data (fixed sensor network vs personal exposure with portable sensor) and (2) the general spatiotemporal variability of air quality data. **Sensor network visualizations** can include spatial variation by mapping sensor locations with associated color scales showing (near) real-time concentration levels, while the temporal variation can be visualized by plotting sensor data from the last hours/days, when selecting a certain sensor (Figure 8).

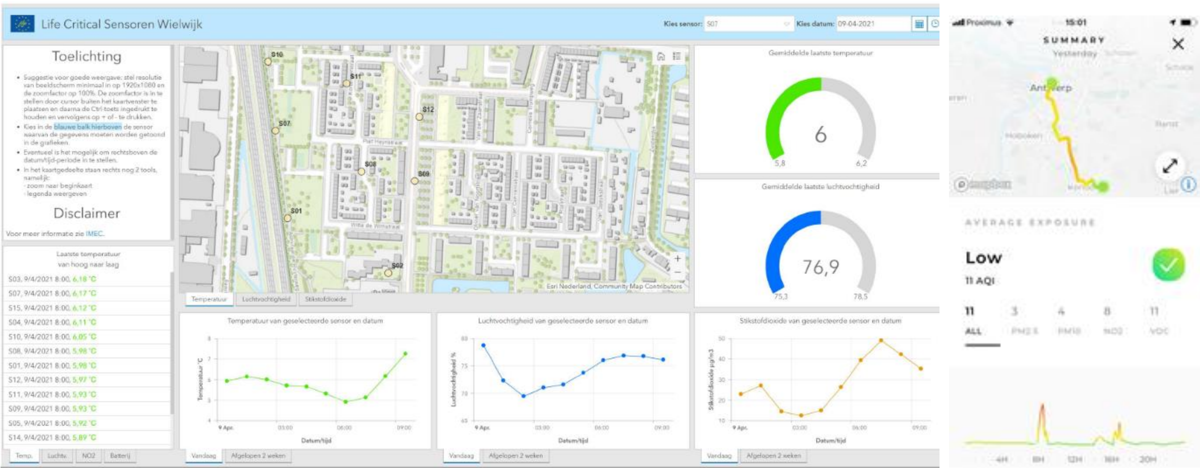

Figure 8: Left: map visualization of a fixed sensor network in Wielwijk, Dordrecht (NL) with temporal variation of the past hours (left), and right: daily exposure quantified by wearable sensor in Antwerp on a map and temporal graph

In order to generalize the measured air quality levels, the color scale can be based on the Air Quality Index, while available nearby regulatory monitoring stations can be shown as well for comparison as this data is typically openly available (e.g. in Belgium: https://github.com/irceline/open_data). Additional analytics can include time variation graphs (see §5), number of exceedance days (based on limit values), local contribution (subtracting background concentrations), spatial statistics,... To obtain air quality estimates between measured locations or time instances, spatiotemporal interpolation, physicochemical or data-driven models can be applied.

When dealing with **trajectory measurements** with portable sensors, the user is not interested in city-wide maps, but rather in the experienced pollutant levels along their trajectories (Figure 8). These pollutant levels will be impacted by a general background concentration and local sources encountered along the trajectory. In this case, a visualization of the experienced pollutant levels along the trajectory, possibly normalized for the background concentration, is more meaningful. Additional analytics might include hotspot locations along the trajectory, number of experienced peaks, inhaled dose (based on the assumed/quantified ventilation rate), ...

In any case, it is important to consider that to date, sensor readings are not regarded accurate enough by the government/EU/WHO to be legally binding in terms of limit/target values. Therefore, adding a disclaimer or additional information section on the advantages/limitations of this sensor data and expected accuracy of the collected sensor data is worthwhile when exposing sensor data to the wide public.

5 From data to insights

When properly accounting for the current limitations of low-cost air quality sensors in terms of data processing and maintenance (see §4), collected air quality data from IoT sensors complements regulatory data in creating more granular monitoring networks, detecting new hotspot locations, measuring the impact from policy interventions, improving current air quality models (e.g. traffic emission factors), developing new air quality features (i.e. healthy routing applications), ... Moreover, besides the obvious air quality-related advantages, application of air quality sensors has shown to result in more participation between citizens and policy makers, co-creation initiatives creating a wider support base for policies (e.g. https://www.samenmetenaanluchtkwaliteit.nl/), awareness raising and creates opportunities for more evidence-based policy making.

In the following paragraphs, we will touch upon potential insights and actions that can be built upon collected IoT sensor data.

5.1 Insights

5.1.1 Sensor readings

Mapping data from sensor networks and measured trajectories (§4.2) provides insights in drivers of air pollution (e.g. pollution gradients along busy roads, peak exposure during daily commutes or at wood stoves) and a better understanding on the impact from local microenvironments on the resulting air quality (e.g. urban green areas, street canyons, traffic intersections, ...). To understand the typical temporal variability of a measured pollutant, so-called time variation graphs (R openair package; https://davidcarslaw.github.io/openair/) can be generated showing the pollutant variability during the day, week, and year (Figure 9). From such graphs, the typical pollutant behaviour (e.g. rush hour peaks, wood burning events, seasonal trends, annual events, ...) can be derived at that location and compared to available background, roadside and/or hotspot regulatory monitoring locations for source attribution purposes. By comparing sensor data with wind field data, source directionality can be retrieved (Figure 10). Moreover, based on historical time series, air quality predictions can be made for upcoming hours (which might be of relevance for sensitive groups).

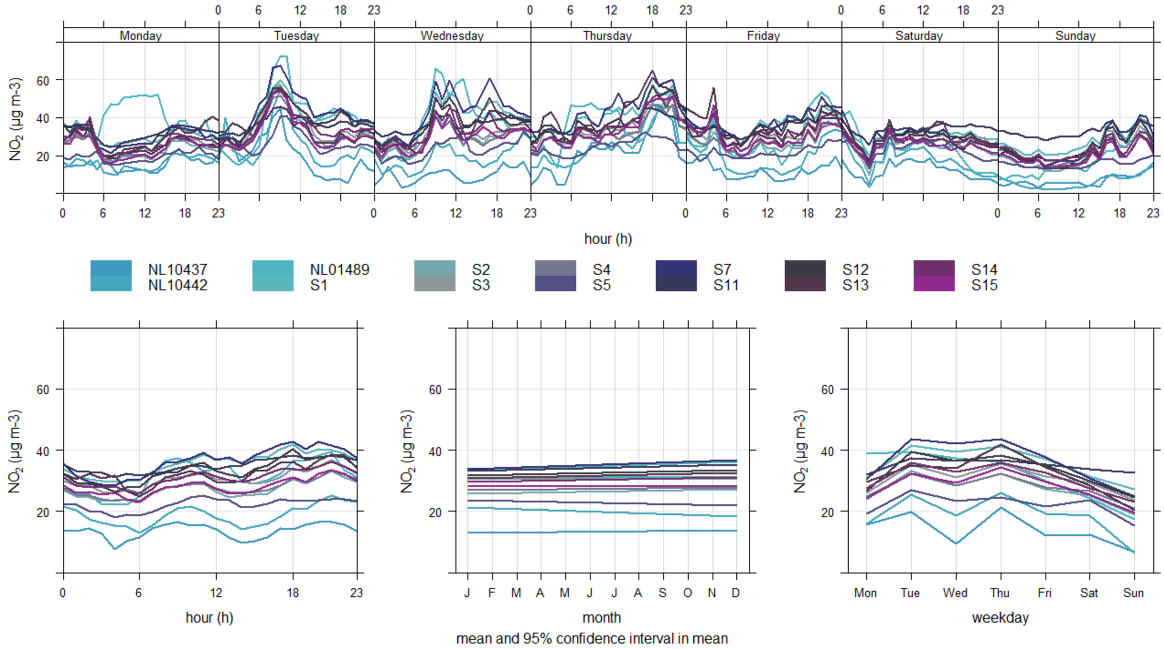

Figure 9: Time variation graphs (R openair package) of NO_2 data collected from a supplementary sensor network in Dordrecht (S1-S15) and 3 nearest reference monitoring locations (NL10437, NL01489, NL10442)

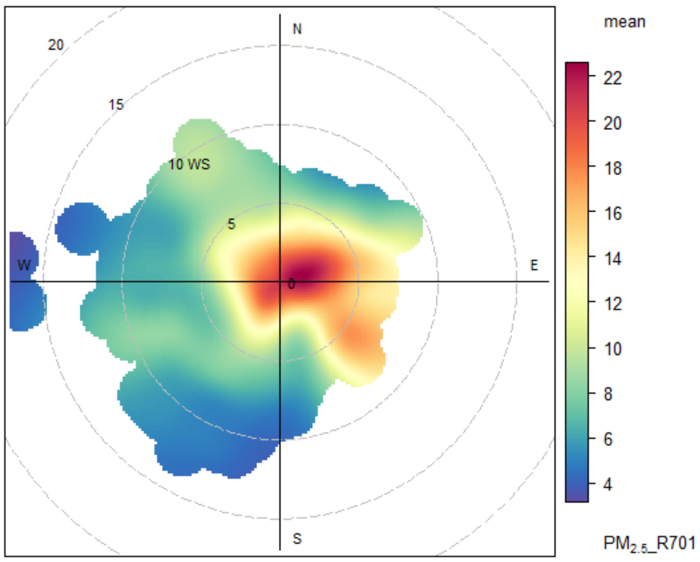

Figure 10: polar plot (R openair package) showing influence of nearby sources on exhibited pollutant concentrations

By fusing data from multiple domains (traffic, meteorology, air quality, …), underlying dependencies can be disentangled to better understand air quality dynamics and improve air quality by steering underlying processes. For example, when deploying sensors on traffic intersections, monitoring air quality, traffic flow/congestion and traffic lights, the operation of the traffic lights can be optimized towards better traffic flow and improved local air quality.

5.1.2 Sensor Inference

In addition to insights from air quality readings at measured locations, a variety of tools exists to interpolate air quality readings between monitoring locations, including traditional spatial interpolation techniques, physicochemical models, Computational Fluid Dynamics (CFD) or data-driven machine learning techniques. An overview of existing air quality models is provided in Annex 3. Traditional interpolation techniques (e.g. Kriging) fall short as they consider every measurement to be representative for the same geographical area, while this is not always the case in real-life. Rural background stations are typically representative for a wider area than for example urban roadside stations. On the other hand, physical models simulating pollutant dispersion at multiple spatial resolutions, from regional background to local emissions, are typically computation intensive which makes them unsuitable for (near) real-time simulations. This is why IMEC is focusing on data-driven techniques which do not require a lot of inventory data (emission factors, emission sources, traffic intensity, ...), since they rely on openly available data (satellite data, POI, route type, meteorology, ...) and are less computation demanding. When considering Antwerp, where a fixed and mobile (bpost vans) air quality sensor testbed is complementing the existing regulatory monitoring network (VMM), a sparse matrix of data points is collected in space and time (Figure 11). Due to the observed associations in both space and time (Figure 5), data matrices of air quality data can be considered low rank and thus explainable by statistical/numerical techniques (39,40). 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 (41). Machine learning (ML) approaches allow for training of these 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 in both space and time (Figure 11).

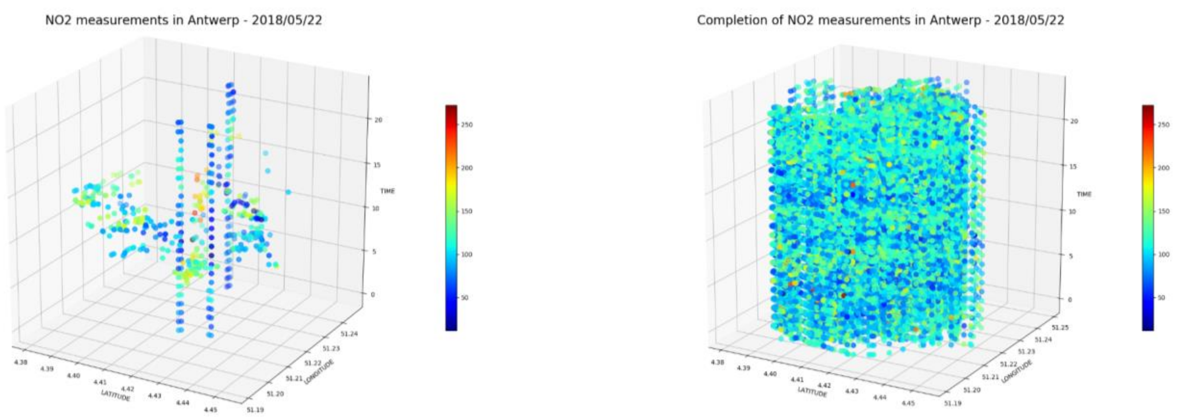

Figure 11: Matrix of collected fixed and mobile NO_2 measurements in Antwerp (22/5/2018) and completed matrix based on the AVGAE inference model. The X and Y axis denote latitude and longitude (location), while the Z axis denotes time

Together with research groups of UGhent (IPI) and the VUB (ETRO), IMEC-NL explored this potential by developing two machine learning models, trained and tested on mobile air quality datasets from Antwerp, BE (bpost vans; https://www.imeccityofthings.be/nl/projecten/bel-air), Utrecht, NL (snuffelfiets; https://snuffelfiets.nl/) and in Oakland, US (Google car; https://blog.google/products/maps/lets-clear-air-mapping-our-environment-our-health/).

The **AVGAE model** is a deep learning model based on variational graph autoencoders (VGAE), incorporating geographical dependencies by considering the road network (graph), while training time-variant dependencies based on additional context information (42,43). The **Geographical**

Random Forest (GRF) model allows for spatial non-stationarity in the relationship between a dependent and a set of independent variables, by considering both a global and local model. This technique adopts the idea of Geographically Weighted Regression (GWR), moving a spatially weighted window over the observations. While a global model is trained on random sensor and context data subsets, a local model was trained on subsets from nearest neighbors (geographical + feature neighbours). Both models rely on large air quality datasets and openly available sources of context information (POI, road type, meteorology, traffic, street canyon index, ...).

When conducting a temporal validation exercise following the JRC FAIRMODE protocol (44), predicting pollutant concentrations at reference locations without using this reference data during model training, both models perform well in terms of accuracy and correlation (Table 3). Comparing our observations to reported performance metrics (r= 0.73-0.9, RMSE 8.29-18.93 and MBE: -5.61-0.94 (NO_2)) of the current physical model applied for policy making in Flanders, Belgium (ATMOStreet (45,46)), the considered data-driven techniques seem to approach the state-of-the-art in terms of performance; with slightly lower correlations (r = 0.63 - 0.77) and slightly higher accuracies (RMSE = 3.49 - 3.99 and MBE = -1.05 - 1.99) between the predicted concentrations and the reference station readings (47).

Table 3: Temporal validation performance of hourly AVGAE and GRF predicted $PM_{2.5}$ concentrations for June 2020, at two reference locations in Utrecht (NL10636, NL10643)

Model	Station	MAE	MBE	RMSE	IA	Acc	Corr	NMB	NMSD	MQI
AVGAE	NL10636	3.14	1.79	3.99	0.82	0.64	0.72	-0.21	-0.07	0.20
AVGAE	NL10643	2.83	0.15	3.73	0.75	0.58	0.63	-0.02	-0.35	0.19
GRF	NL10636	2.88	1.99	3.8	0.79	0.67	0.77	-0.23	-0.34	0.43
GRF	NL10643	2.74	-1.05	3.49	0.79	0.6	0.73	0.15	-0.37	0.44
AVGAE	**Avg**	**2.99**	**0.97**	**3.86**	**0.79**	**0.61**	**0.68**	**-0.12**	**-0.21**	**0.20**
GRF	**Avg**	**2.81**	**0.47**	**3.65**	**0.79**	**0.64**	**0.75**	**-0.04**	**-0.36**	**0.44**

Time series graphs of the AVGAE and GRF predicted $PM_{2.5}$ concentrations are provided in Figure 12 and illustrate that predicted $PM_{2.5}$ concentrations (M_i) agree reasonably well with the reported reference data (O_i) for each of the considered models at each location, mostly falling within the uncertainty bounds of the reference equipment (RMSU; provided by FAIRMODE) and the MQI limits defined by FAIRMODE as twice (k=2) the reference measurement uncertainty (44). Similar model validation exercises have been conducted on different mobile datasets, Snuffelfiets (Utrecht, NL), postal vans (Antwerp, BE) and Google car (Oakland, US) (48).

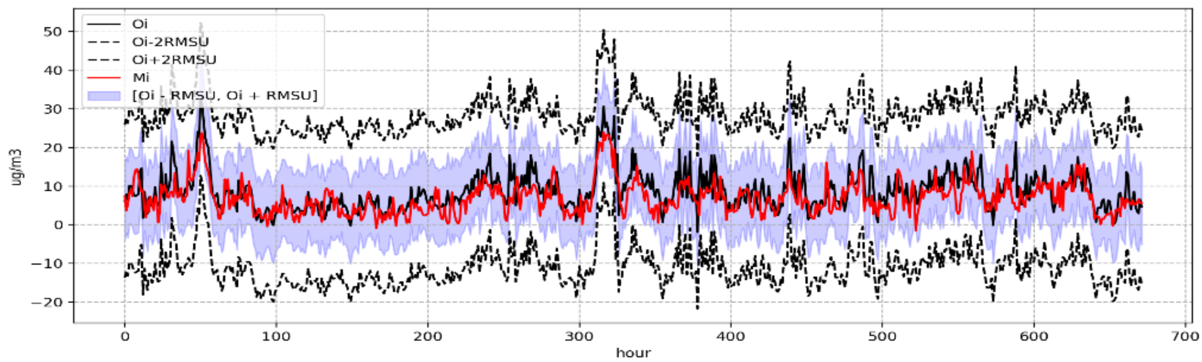

Figure 12: Time series graph of AVGAE predicted (M_i) and RIVM reported (O_i) $PM_{2.5}$ concentrations at Kardinaal De Jongweg (NL10636) with associated uncertainty bounds of the reference equipment (RMSU; purple) and model quality limits (2 x RMSU; dashed lines) defined by FAIRMODE

IMEC applied the AVGAE and GRF model to predict street-level concentrations for the city of Antwerp based on the collected data from mobile sensors (bpost vans; Figure 13), fixed sensors and the available regulatory monitoring network (VMM). This data-driven approach integrates all available air quality data of a city to map pollutant variability both in space and time.

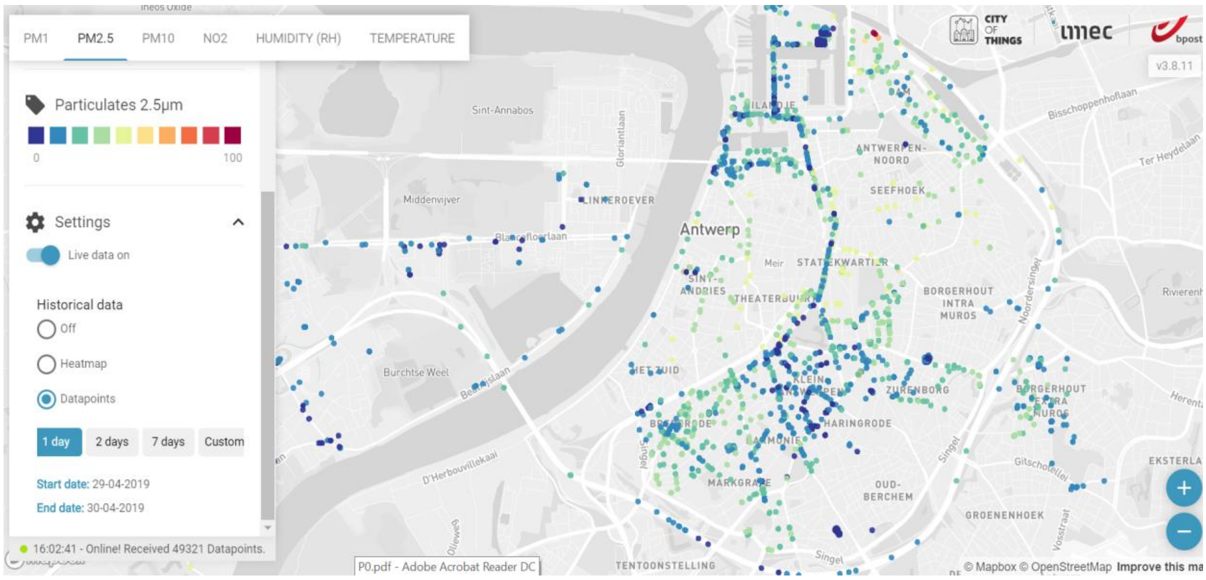

Figure 13: Mapping Air Quality (MAQY) dashboard visualizing mobile air quality measurements (PM_1, $PM_{2.5}$, PM_{10} and NO_2) collected during one day (29/4/2019) by postal vans in Antwerp (BE)

5. Learnings & Actions

Air pollution is currently regarded as top environmental threat for our health and top priority at the EU policy agenda (Zero Pollution EU Green Deal (1)). Moreover, more stringent EU regulations and newly emerging limit values can be expected following the WHO Guideline values. To face these challenges, cities will need to be prepared to provide their contribution in more fine-grained air quality assessments and pollution reductions. IoT and low-cost sensors, together with proper calibration and algorithmic approaches, can be used as a complementary tool for improved air quality assessments and evidence-based policy support. Engaging citizens in this transformation process is beneficial on multiple aspects: citizen science campaigns can **extend the network** (distributing deployment and maintenance effort), **increase awareness and understanding** on air quality dynamics and **initiate behavioral change** ultimately reducing air pollution.

Based on the outlined advantages and limitations of air quality sensors, following take-aways are crucial to reach a fit-for-purpose sensor solution:

- A sensor solution includes **sensor, connectivity and power!**
- Optimize your sensor solution based on the intended use case: **fit for purpose**
- Sensors are sensitive and **need calibration**
- Make sure to consult available **3rd party sensor performance studies** when selecting your sensor/solution
- No uniform **certification** framework (yet!)
- Make sure to include a **validation step** in your use case to have an idea about the associated data quality

When implementing **a proper calibration approach and maintenance program**, air quality sensors complement the existing monitoring networks towards more fine-grained air quality monitoring. This to create an **improved air quality map** of your city, increase the understanding of the air quality variability in your city, detect new hotspot locations with air quality exceedances or clean air environments to reside in, organize recreational events, sports/leisure activities, ...

Based on a **more granular air quality map**, priority areas can be defined for dedicated policy measures or urban transformation, while healthy routing applications can be developed for active commuters, and traffic can be redirected from high-polluted areas. More fine-grained air quality maps will contribute to better **exposure assessments** when merged with other data sources, e.g. with mobility flow data (e.g. https://www.imeccityofthings.be/nl/projecten/cityflows) to see where people are impacted most by bad air quality and which city neighborhoods need transformation priority.

Moreover, policy measures (low emission zone, one way street, bicycle streets, redirected traffic, ...) can be **optimized with sensor data** (e.g. traffic light optimization) while building more evidence to create a larger citizen support base (**evidence-based policy making**).

Finally, IoT sensors are key in **awareness raising**, as they have enabled **personalization of environmental data**. Via co-creation initiatives, citizens are informed about main drivers of air quality and will **change their behaviour** more easily when seeing the impact on their personal devices. This can impact car use, wood burning, commuting routes, indoor ventilation, ... Several community-based initiatives to measure PM with sensors have emerged in various Belgian cities, e.g. Brussels (https://influencair.be/), Leuven (https://www.vmm.be/evenementen/presentaties/studiedag-lucht-samen-meten-en-weten/maarten_reyniers_leuvenairvmmma.pdf), Roeselare (https://www.klimaatswitch.be/airsl), https://www.luchtpijp.be/ ,... In the Netherlands, farmers and citizens are currently adopting sensors from RIVM to evaluate their efforts e.g. in reducing stable emissions (https://www.rivm.nl/boeren-en-buren) or personal exposure while cycling (https://snuffelfiets.nl/). Similar campaigns can be launched to raise awareness about personal air quality exposure during commuting, wood burning, cooking, ...

While extending and complementing their environmental monitoring networks, cities should use sensor data as an additional tool to inform and properly engage all stakeholders in fighting air pollution together!

AKNOWLEDGEMENTS

This work was supported by VLAIO and has only been possible thanks the fruitful research collaboration between Flanders Environmental Agency (VMM), VITO, IMEC vzw, IPI (UGhent), ETRO (VUB) and IMEC-NL and contributions from various people including Valerio Panzica La Manna, Wilfried Philips, Nikos Deligiannis, Jana Muylaert, Martha E. Nikolaou, Tien Huu Do, Xuening Qin, Niek Bosch, Sharada Prasad Shantharam, Sander Weijs, Esther Rodrigo, Jan Adriaenssens and Gert Degreef.

Bibliography

1. EU Green Deal. Zero Pollution Action Plan [Internet]. 2021 [cited 2021 May 12]. Available from: https://ec.europa.eu/environment/strategy/zero-pollution-action-plan_it

2. EEA. Air quality in Europe 2019. Luxembourg: European Environment Agency; 2019.

3. WHO. Air pollution [Internet]. [cited 2020 Jul 16]. Available from: https://www.who.int/health-topics/air-pollution#tab=tab_1

4. Pirjola L, Lähde T, Niemi JV, Kousa A, Rönkkö T, Karjalainen P, et al. Spatial and temporal characterization of traffic emissions in urban microenvironments with a mobile laboratory. Atmos Environ. 2012 Dec;63:156–167.

5. Pattinson W, Longley I, Kingham S. Using mobile monitoring to visualise diurnal variation of traffic pollutants across two near-highway neighbourhoods. Atmos Environ. 2014 Sep;94:782–792.

6. Peters J, Van den Bossche J, Reggente M, Van Poppel M, De Baets B, Theunis J. Cyclist exposure to UFP and BC on urban routes in Antwerp, Belgium. Atmos Environ. 2014 Aug;92:31–43.

7. Hofman J, Samson R, Joosen S, Blust R, Lenaerts S. Cyclist exposure to black carbon, ultrafine particles and heavy metals: An experimental study along two commuting routes near Antwerp, Belgium. Environ Res. 2018 Apr 4;164:530–538.

8. Kumar P, Patton AP, Durant JL, Frey HC. A review of factors impacting exposure to PM 2.5 , ultrafine particles and black carbon in Asian transport microenvironments. Atmos Environ. 2018 May;187:301–316.

9. Int Panis L, de Geus B, Vandenbulcke G, Willems H, Degraeuwe B, Bleux N, et al. Exposure to particulate matter in traffic: A comparison of cyclists and car passengers. Atmos Environ. 2010 Jun;44(19):2263–2270.

10. Feenstra B, Papapostolou V, Hasheminassab S, Zhang H, Boghossian BD, Cocker D, et al. Performance evaluation of twelve low-cost PM2.5 sensors at an ambient air monitoring site. Atmos Environ. 2019 Nov;216:116946.

11. Rai AC, Kumar P, Pilla F, Skouloudis AN, Di Sabatino S, Ratti C, et al. End-user perspective of low-cost sensors for outdoor air pollution monitoring. Sci Total Environ. 2017 Dec 31;607-608:691–705.

12. Feinberg S, Williams R, Hagler GSW, Rickard J, Brown R, Garver D, et al. Long-term evaluation of air sensor technology under ambient conditions in Denver, Colorado. Atmos Meas Tech Discuss. 2018 Feb 8;1–18.

13. Castell N, Dauge FR, Schneider P, Vogt M, Lerner U, Fishbain B, et al. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? Environ Int. 2017 Feb;99:293–302.

14. Munir S, Mayfield M, Coca D, Jubb SA, Osammor O. Analysing the

performance of low-cost air quality sensors, their drivers, relative benefits and calibration in cities-a case study in Sheffield. Environ Monit Assess. 2019 Jan 22;191(2):94.

15. Karagulian F, Barbiere M, Kotsev A, Spinelle L, Gerboles M, Lagler F, et al. Review of the Performance of Low-Cost Sensors for Air Quality Monitoring. Atmosphere. 2019 Aug 29;10(9):506.

16. Snyder EG, Watkins TH, Solomon PA, Thoma ED, Williams RW, Hagler GSW, et al. The changing paradigm of air pollution monitoring. Environ Sci Technol. 2013 Oct 15;47(20):11369–11377.

17. Morawska L, Thai PK, Liu X, Asumadu-Sakyi A, Ayoko G, Bartonova A, et al. Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? Environ Int. 2018 Apr 26;116:286–299.

18. Kumar P, Morawska L, Martani C, Biskos G, Neophytou M, Di Sabatino S, et al. The rise of low-cost sensing for managing air pollution in cities. Environ Int. 2015 Feb;75:199–205.

19. Hapidin DA, Saputra C, Maulana DS, Munir MM, Khairurrijal K. Aerosol chamber characterization for commercial particulate matter (PM) sensor evaluation. Aerosol Air Qual Res. 2019;19(1):181–194.

20. Austin E, Novosselov I, Seto E, Yost MG. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor. PLoS One. 2015 Sep 14;10(9):e0137789.

21. Hofman J, Depestel G, Lenaerts S, Samson R. Dust chamber evaluation of different low-costPM2.5 sensors in view of novel air quality monitoring strategies. EMAS.

22. Jiao W, Hagler G, Williams R, Sharpe R, Brown R, Garver D, et al. Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost

sensor performance in a suburban environment in the southeastern United

States. Atmos Meas Tech. 2016 Nov 1;9(11):5281–5292.

23. Badura M, Batog P, Drzeniecka-Osiadacz A, Modzel P. Evaluation of Low-Cost Sensors for Ambient $PM_{2.5}$ Monitoring. Journal of Sensors. 2018 Oct 31;2018:1–16.

24. Lewis AC, Lee JD, Edwards PM, Shaw MD, Evans MJ, Moller SJ, et al. Evaluating the performance of low cost chemical sensors for air pollution research. Faraday Discuss. 2016 Jul 18;189:85–103.

25. Wang Y, Li J, Jing H, Zhang Q, Jiang J, Biswas P. Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement. Aerosol Science and Technology. 2015 Nov 2;49(11):1063–1077.

26. Williams R, Kaufman A, Hanley T, Garvey S. Evaluation of Field-deployed Low Cost PM Sensors [Internet]. Research Triangle Park, NC, USA 27709: EPA; 2014 Dec [cited 2019 Oct 22] p. 76. Report No.: EPA/600/R-14/464. Available from: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NERL&DirEntryId=297517

27. Liu H-Y, Schneider P, Haugen R, Vogt M. Performance Assessment of a Low-Cost PM2.5 Sensor for a near Four-Month Period in Oslo, Norway. Atmosphere. 2019 Jan 22;10(2):41.

28. Tagle M, Rojas F, Reyes F, Vásquez Y, Hallgren F, Lindén J, et al. Field performance of a low-cost sensor in the monitoring of particulate matter in

Santiago, Chile. Environ Monit Assess. 2020 Feb 10;192(3):171.

29. Borrego C, Costa AM, Ginja J, Amorim M, Coutinho M, Karatzas K, et al. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. Atmos Environ. 2016 Dec;147:246–263.

30. Hojaiji H, Kalantarian H, Bui AAT, King CE, Sarrafzadeh M. Temperature and Humidity Calibration of a Low-Cost Wireless Dust Sensor for Real-Time Monitoring. 2017 IEEE Sensors Applications Symposium (SAS). IEEE; 2017.

31. Drajic DD, Gligoric NR. Reliable Low-Cost Air Quality Monitoring Using Off-The-Shelf Sensors and Statistical Calibration. ELEKTRON ELEKTROTECH. 2020 Apr 25;26(2):32–41.

32. Badura M, Batog P, Drzeniecka-Osiadacz A, Modzel P. Regression methods in the calibration of low-cost sensors for ambient particulate matter measurements. SN Appl Sci. 2019 Jun;1(6):622.

33. Mijling B, Jiang Q, de Jonge D, Bocconi S. Field calibration of electrochemical NO_2 sensors in a citizen science context. Atmos Meas Tech. 2018 Mar 5;11(3):1297–1312.

34. Mijling B, Jiang Q, de Jonge D, Bocconi S. Practical field calibration of electrochemical NO_2 sensors for urban air quality applications. Atmos Meas Tech Discuss. 2017 Apr 4;1–25.

35. Kim J, Shusterman AA, Lieschke KJ, Newman C, Cohen RC. The BErkeley Atmospheric CO_2 Observation Network: field calibration and evaluation of low-cost air quality sensors. Atmos Meas Tech. 2018 Apr 6;11(4):1937–1946.

36. Sun L, Westerdahl D, Ning Z. Development and Evaluation of A Novel and Cost-Effective Approach for Low-Cost NO_2 Sensor Drift Correction. Sensors (Basel). 2017 Aug 19;17(8).

37. Malings C, Tanzer R, Hauryliuk A, Kumar SPN, Zimmerman N, Kara LB, et al. Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. Atmos Meas Tech. 2019 Feb 11;12(2):903–920.

38. Hofman J, Nikolaou M, Stroobants C, Weijs S, Shantharam SP, La Manna VP. Distant calibration of low-cost PM and NO2 sensors; evidence from multiple sensor testbeds . Env Sc &Tech. 2021;

39. Udell M, Townsend A. Why are big data matrices approximately low rank? SIAM Journal on Mathematics of Data Science. 2019 Jan;1(1):144–160.

40. Asif MT, Mitrovic N, Dauwels J, Jaillet P. Matrix and tensor based methods for missing data estimation in large traffic networks. IEEE Trans Intell Transport Syst. 2016 Jul;17(7):1816–1825.

41. Paliwal C, Biyani P, Rajawat K, Sutaria R. Scalable Spatio-Temporal Measurements and analysis of Air pollution data using Vehicle Mounted Sensors. UC Davis; 2020.

42. Do T, Nguyen D, Tsiligianni E, Aguirre A, Manna V, Pasveer F, et al. Matrix Completion with Variational Graph Autoencoders: Application in Hyperlocal Air Quality Inference. IEEE; 2019.

43. Do TH, Tsiligianni E, Qin X, Hofman J, La Manna VP, Philips W, et al. Graph-Deep-Learning-Based Inference of Fine-Grained Air Quality from Mobile IoT Sensors. IEEE Internet Things J. 2020;1–1.

44. Janssen S, Guerreiro C, Viaene P, Georgieva E, Thunis P. Guidance

Document on Modelling Quality Objectives and Benchmarking.

45. Irceline. Validatie luchtkwaliteitsmodel' ' ATMO - Street (Vlaanderen) voor NO_2 in 2017 . Irceline; 2017.

46. Lefebvre W, Van Poppel M, Maiheu B, Janssen S, Dons E. Evaluation of the RIO-IFDM-street canyon model chain. Atmos Environ. 2013 Oct;77:325–337.

47. Hofman J, Do TH, Qin X, Rodrigo E, Nikolaou ME, Philips W, et al. Spatiotemporal Air Quality Inference of Low-Cost Sensor Data; Application on a Cycling Monitoring Network. In: Del Bimbo A, Cucchiara R, Sclaroff S, Farinella GM, Mei T, Bertini M, et al., editors. Pattern recognition ICPR international workshops and challenges: virtual event, january 10–15, 2021, proceedings, part VI. Cham: Springer International Publishing; 2021. p. 139–147.

48. Hofman J, Do TH, Qin X, Bonet ER, Philips W, Deligiannis N, et al. Spatiotemporal Air Quality Inference of Low-Cost Sensor Data: Evidence from Multiple Sensor Testbeds . Env Mod Soft. 2021;