APPENDIX 10

BREATHE LONDON LEARNINGS, RECOMMENDATIONS AND LESSONS LEARNED

APPENDIX 10

Breathe London Learnings, Recommendations and Lessons Learned

The Breathe London project demonstrated the ability to successfully deploy and manage a network of lower-cost sensors and conduct a large-scale mobile monitoring campaign across a major global city. The purpose of this document is to share the learnings, recommendations and lessons learned from the initial planning and deployment of the monitoring platforms and the two subsequent years of operating the stationary sensor network. These not only highlight the success of the project but equally, the challenges faced and largely overcome by the project consortium (Appendix 1) in implementing a monitoring campaign of this scale. These learnings and recommendations are coupled with the lessons learned throughout the project and together are intended to support the efforts of other global cities who are looking to undertake similar hyperlocal monitoring campaigns.

1. Overall Project

Capabilities of the sensor technology

Breathe London demonstrated that with rigorous QA/QC procedures, lower-cost sensor systems can yield reasonably accurate, precise and stable data when used in real-world conditions and can help cities learn about their specific pollution challenges and opportunities. For example, analysis of Breathe London data produced insights that were complementary to findings from London's reference network, demonstrating that mobile monitoring and lower-cost sensor networks can identify and characterise pollution levels in order to assess and evaluate policy interventions (ULEZ Assessment Report).

In addition, by collecting data at previously unmeasured locations and deploying high time resolution mobile monitoring, Breathe London produced new hyperlocal insights at unprecedented spatial resolution, including leveraging the on-road data to characterise not only total concentrations but also concentration ratios to provide insights directly into transport emissions. Overall, it was found that hourly measurements at individual fixed monitoring sites and 1-10 second mobile measurements on individual streets are less accurate than reference network data, but when aggregated in space and/or time it is possible to find ways to generate the statistical power to estimate trends. When looking at insights across the network, data capture may not be an issue, but specific monitoring sites can have valid data capture lower than 75% which can prohibit site-specific analysis over longer periods (see **Figure 1** for AQMesh network monitor coverage). Findings from Breathe London could help in planning how many "hot spares" (i.e. spare, calibrated monitors) may be needed to enable data continuity.

Overall, the Breathe London network performance maintained a high operational rate between March 2019 and June 2020 with more than 80 pods in the network reporting at least 75% valid hourly data (**Figure 1**). The increase in number of active pods with valid data in early 2019 reflects correcting power supply and other performance issues in the early months of deployment. This demonstrated how vital it is to get as much information as possible on sensor performance ahead of procurement, as well as incorporating time to test and validate instrumentation. Sensor replacement and pod maintenance costs are also key considerations for understanding the overall project budget.

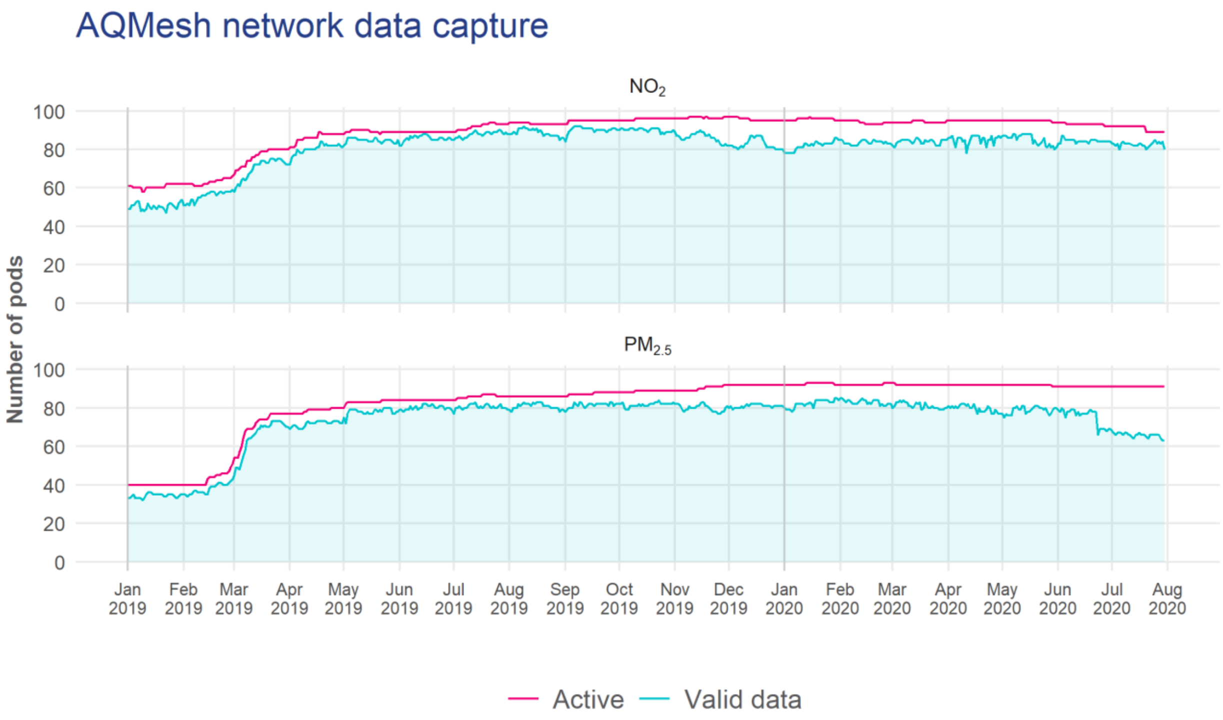

FIGURE 1. Number of AQMesh pods with > 75% valid hours of NO_2 or $PM_{2.5}$ data each day (blue line). Also shown is the maximum number of operating or active pods in the network (red line). Note that the coverage criterion is met for generally over 80% of the instruments deployed although some degradation is seen later in the project in June 2020 for $PM_{2.5}$

Developing and improving a network calibration method

The development, application and verification of the novel network calibration method demonstrated the powerful value of leveraging cloud-based technology to remotely scale a dense network of lower-cost sensors. Breathe London provided an opportune testbed for the method due to the density of the existing reference network, which enabled the extensive use of physical co-locations for validation purposes. The method was used to scale NO_2 and $PM_{2.5}$ measurements against an AQMesh pod that had been co-located at a reference monitoring site using a scale separation technique that distinguished between local, nearby, regional and far field pollution sources. The method can significantly reduce operating costs as it reduces the need for physical co-locations across the 100 AQMesh pods and improves the data quality assurance and control procedures. Co-locations were instead carried out to verify the network calibration method and calibrate sensors in the field following a sensor failure. This method is vital to managing dense, lower-cost networks and future monitoring campaigns should employ similar techniques. Details about this method are available in the Network Calibration Methodology report and in a forthcoming publication.

Measurement of CO_2

The measurement of carbon dioxide (CO_2) by both the stationary sensor network and mobile monitors enhanced the insights produced by the project.[1] The reference network in London does not measure CO_2 and this unique dataset presented the project team with the opportunity to study the relationship of CO_2 to nitrogen oxides (NO_x) and other health-impacting pollutants to improve both the understanding of London's air quality and the modelling tools used to assess it. Knowledge of the ratio of pollutant to CO_2 concentration enhancements enables separation of emissions per unit of fuel burned locally from other sources of emissions. The spatial and temporal pattern of observed emission ratios provides critical information to inform policy intervention decisions. Additionally, when compared to bottom-up accounting of emission sources such as those input into air dispersion models, observed ratios enable a key additional test of emission inventories.

Improving air quality model accuracy

Breathe London demonstrated that the assimilation of empirical data from the stationary sensor network and mobile monitors into an existing air quality model can be used to verify and improve emissions estimates and hence model accuracy. CERC's ADMS-Urban modelling software was used throughout the Breathe London project to simulate pollution measured by the Breathe London stationary sites, the Breathe London mobile instruments and the reference networks. Comparisons between modelled and measured data provided new insights into areas, or pollution hotspots, where London emissions data need to be improved and areas where model refinements were needed (see Appendix 9). Comparisons also played an important part in the QA/QC of the Breathe London measurements themselves. This work produced recommendations for improvements to London's air quality model, including updated baseline maps of air quality and source attribution data by activity sector. See Appendix 6 for more details.

Developing an interactive, publicly available data platform

Breathe London developed a creative platform to share near-real time data directly from the project, overlay reference network data and mobile monitoring data, and even explore modelled source apportionment data - all in the form of an interactive map. The display of such data was found to be helpful to external stakeholders, including residents and community

---

[1] Unintended rebasing of CO2 sensors in the stationary network caused periodic baseline shifts which affect the absolute CO2 measurements. Early investigation indicates that the local contributions are only affected in a minimal way, so that the emission index calculations which rely on the local concentrations are not influenced significantly. Some additional work is needed to minimise this effect if absolute CO2 levels are to be derived reliably (Appendix 7 *Use of emission ratios for the stationary Breathe London network*).

groups. It improved the richness and relevance of information available to Londoners as the only local resource with this type of air quality information available in one place. An iterative design process presented many learnings including:

- the value of adding historical data, including diurnal trends, to allow quick analysis of pollution patterns
- interactive features such as search functions and location finder
- keep the data simple - focus on two or three priority pollutants to display - and leave the rest to access via downloadable datasets or API's
- include policy-relevant layers such as Clean Air Zone boundaries
- host the platform in the Google Cloud to ingest large volumes of data and enable user-friendly performance

To enable replication of our data platform in other monitoring campaigns, the Breathe London interactive map was powered by design with an open-source data platform that is freely available to other cities around the world that wish to build their own.

Building in sufficient time to test and prepare equipment prior to deployment

The project was originally designed to obtain a sufficient baseline of data before the implementation of the Ultra Low Emission Zone (ULEZ) in April 2019. As a result, the timeline between securing contractual agreements with technical partners and sensor deployment was extremely tight. More time to calibrate and perform additional quality control on the evolving sensor technology would have been beneficial. The team also encountered delays in obtaining permissions to install equipment at suitable sites and had to make compromises in the choice of monitoring locations.

Future projects require careful planning and enough lead-times for optimising the sampling strategy and preparing and testing equipment. Accelerated timelines could compromise the quality and utility of data gathered, especially with newer technology. The simultaneous deployment of the stationary and mobile monitoring components created additional pressures, and one way to lessen strain on a project's capacity would be to phase these deployments.

Insights to longer-term performance of stationary sensor-based networks

Running a lower-cost stationary network of this size for this long, enabled insights to the longer-term performance of the sensors that can not only help (1) improve the accuracy of lower-cost sensors measurement in the future but (2) aid in managing the operation of a similar network.

1) *Improve accuracy:* A gradual upward drift of NO_2 measurements was identified (likely caused by ozone cross-interference), biased high PM2.5 measurements were observed during periods of high relative humidity, and biased high NO_2 readings occurred during periods of very high temperatures (> 30-35 °C). An algorithm was developed to correct for the ozone cross-interference and applied to the project's full ratified dataset.

Although all of these performance issues were not fully resolved and corrected for within the span of the Breathe London project, the identification of these environmental effects emphasises the importance of region-specific evaluation of sensing technology, and rigorous QA/QC processes to identify periods of anomalous instrument performance.

2) *Operation of the network:* Breathe London experienced more than 100 sensor failures over 2-years and from this learned that it is important to consider the manufacturer's recommendation for the lifespan of each field-deployed sensor and further research empirical evidence demonstrating how long they perform accurately in practice so that performance issues can be anticipated and planned for i.e. adequate budget and spare parts.

Building in adequate resources from the outset

Breathe London required considerable resources to build, test, deploy and maintain both the stationary sensor network and the mobile data collection platforms – substantially more resources than the project consortium initially anticipated. However, it is important to recognise that future projects could benefit from many of the advances made within this project, leading to significant efficiency and cost savings in the future.

All project partners provided additional resources and support to ensure that both the stationary and mobile data collection remained on schedule. Ideally, a project of similar scale to Breathe London should have multiple team members dedicated full-time to maintenance and operation of the data collection platforms, with regular assessments of whether more resources are needed to maintain high data quality. **Table 1** below provides a description of the roles of those involved in Breathe London and the number of full-time staff needed for each the stationary and mobile monitoring campaigns as well as a summary of their duties. These estimates provide a guide for other global cities to scale their resources according to the spatial coverage, duration and type of monitoring campaign.

TABLE 1. Estimated staff needs for stationary and mobile monitoring campaigns

Roles	Number of staff		Duties
	Stationary	Mobile	
Senior Manager	1		Overseeing all aspects of the project including relevant reporting.
Project Manager	1	1	Project management of each monitoring scheme.
Operations Manager	1	1	Logistics for all aspects of the procurement, installation, maintenance and co-location of instruments.
Communications Manager	1		Addressing and producing all communications (i.e. print, social media) for monitoring campaigns.
Communications and Project Coordinator	1		Addressing and coordinating responses to all correspondence. Coordinating meetings and assisting managers.
Air Quality Scientists	2 - 4		Designing monitoring scheme, developing QA/QC to data, assessing pollution and performance of monitoring network, and ratifying data. This includes at least one Senior Scientist.
Data Analysts	3		Applying QA/QC to data, assessing pollution data, and producing datasets. This includes at least one Data Scientist.
Maintenance Technician	1	0.25	Conducting installations, maintenance and repairs.
Drivers	-	2 - 4	For two cars on mainly weekday, daytime shifts only (approximately 8-hour shifts).
Website Designer	0.5 - 1		Designing and building website including visualisation of data. Ensuring platform and relevant data-feeds are functioning and up to date.

TABLE NOTES:

- This list does not include staff management and resources (Finances, Human Resources, Payroll, Line Management, etc) nor does it account for other overheads.
- This is a list for running networks of similar scale to Breathe London, communicating results, and undertaking policy and pollution assessments, not to develop methods such as the cloud-based network calibration or undertake modelling activities.
- Costs will depend on the rates for these specialists in the cities that the project is deployed.

<u>Understanding acceptable levels of uncertainties of lower-cost sensors</u>

All sensors, including reference monitors, have a level of measurement uncertainty. The Breathe London project pushed the boundaries of lower-cost sensors to meet a higher data quality objective than what these sensors are conventionally intended for. As such, the project consortium employed an intensive data verification process, involving a substantial number of pods using transfer standards ("gold" pods). This level of verification was pursued due to the likely scrutiny of provisional data released on the Breathe London website given the public interest in London's air quality. There remains a learning curve to understand the full potential, and limitations, of these types of new networks. For instance, the verification process employed is costly and erodes some of the cost advantages of using this type of sensor. Conversely, Breathe London partners developed novel network calibration methods that reduce the need for the physical co-location process and thus reduce these costs.

Aligning expectations around release of provisional data

The desire to provide the public with data that was both near real-time and robustly validated was a key tension of the Breathe London project. Although it is common practice for regulatory networks to release provisional data, there are additional uncertainties associated with data from lower-cost monitors. Prior to the initial data release, Breathe London partners had not fully discussed the risks of providing provisional data that had undergone only minimal validation and were subject to change after further evaluation. The project consortium elected, in this case, to err on the side of caution by pursuing a more rigorous verification process prior to the initial data release, which led to a delay from the project's expected timeline.

Project partners should clearly define levels of data quality at the outset and agree on what is acceptable for public release, and with what caveats, at different stages of the project.

Recognising the challenge of successfully assessing policy interventions

Assessing policy interventions with air pollution data takes time, expertise and suitable data quality, so it is essential to factor these into the timeline and resources. When designing a monitoring network, consider policy aims from the outset to ensure the most appropriate method of stationary or mobile data collection. However, using lower-cost technology brings greater instrument uncertainty, compared to traditional reference-grade air quality monitors, which needs to be taken into consideration when determining the type of monitoring necessary to achieve the objectives and when analysing the data. The counterpoint is that innovations such as the inclusion of CO_2 for emission ratio determination and assessment of the impacts of interventions can reduce uncertainties caused by meteorological variability.

Finally, the setup of the air quality model needs to be aligned to ensure that the output is optimal in determining the impacts on air quality of specific policy interventions which may be difficult to distinguish from other trends in pollutant concentrations. Model limitations should be understood, both in input data and in the model descriptions of the complex physics and chemistry in urban areas, as these may impact the ability to answer policy questions or test scenarios. Projects should carefully consider the resolution/number of receptors for modelling, source categories that can be modelled based on existing emissions inventories, the time it takes to run simulations (i.e. days or weeks), the time needed to analyse, interpret results, and present to the public or other relevant stakeholders.

2. Stationary Network

Sensors and equipment

Installation

Time is required for asset owners to review the necessary records on the features, location and current usage of the street furniture before installation can be agreed i.e. is a commando socket already installed, is the post already being used for signage, are there planned works in the area that could remove the post, age of the infrastructure, length of time since a structural survey has been undertaken, etc. Weight restrictions, which vary by type of street furniture, may also apply. Weight can become an issue when using solar powered pods as the solar panel adds substantial weight. In those cases, it was found that the borough may require an engineer to run a weight test and give approval, with an associated fee.

The technical specification for installation should also be clear from the outset of the project to include specific power requirements i.e. an external commando socket or wiring directly to the mains power supply inside the lamppost. Costs may be associated with installation activities from contracting engineers and electricians to upgrading the power supply and even paying for electricity to power the pods, therefore, payment for works should be arranged well in advance to avoid delays to scheduling. The AQMesh pods are powered by a trickle feed, therefore their energy use is minimal, approximately 43.8 kilowatt hours (kWhrs) annually.

Installation with solar power

There are three main power options for the AQMesh sensor pods used in Breathe London: solar, battery and mains power supply. By including sensors for a larger number of pollutants the pods required mains or solar power supply as battery power was insufficient. Due to the height of the surrounding buildings, many locations in central London do not receive enough sunlight to power the pod using a solar panel. Additionally, pods with a solar panel were only allowed to be installed at sites with a recent structural survey, owing to the extra weight of the equipment. This survey incurs additional costs, which when added to the cost of the solar panel, made upgrading the mains power supply in the lamppost (see ‘Installation’ above) more cost effective. As a result of limited sunlight, power availability and costs, only 13 pods were installed with a solar panel, compared to the 40 that were originally planned.

Power supply

The project advanced the understanding of how electrical supplies and equipment emitting electromagnetic fields/radiofrequency (EMF/RF) can detrimentally affect lower-cost sensors and sensor systems and how to compensate and eliminate these. Within a few weeks of the initial deployment it was discovered that measurements at a portion of pods displayed “pulsing” artifacts that took some time to diagnose. Eventually the problem was attributed to EMF/RF interference from the power supply, necessitating retrofits in those pods to have filtered power supply. To be safe, most of the pods were retrofitted in the network. This learning demonstrates the need for a long field-testing period for future campaigns and when deploying sensors in new ways - AQMesh are often battery or solar powered but the unique needs and constraints of

Breathe London meant that mains power supply was preferred.

Sensor failures

Over the course of the Breathe London monitoring campaign, there were more than a hundred instances of sensor failure, after which the sensor or particle counter needed to be replaced. These failures occurred most during the winter months. Additionally, newly replaced sensors require new calibrations as the old calibration for the failed sensor is no longer valid. This experience shows the need to allocate resources for ongoing maintenance of pods, installation of new sensors, and timely re-calibrations in order to maximize data capture.

Sensor turnover

Lower-cost sensor technology is rapidly evolving and sensor manufacturers frequently update their products. This can create complications for a large-scale sensor network whereby the sensors deployed at the start of monitoring may be discontinued and replaced with a newer model during the lifetime of the project. The same sensor may not be available for repair/replacement, which could result in the network employing different models of sensors with differing data quality or calibration settings. This can introduce unknown errors or uncertainty to the results. Furthermore, manufacturer firmware algorithms can be updated which can directly impact data quality. Proactively ask manufacturers about their planned product update upfront during the procurement process, and insist any firmware updates be pre-approved by the project team before they are applied. As a contingency, consider keeping a stock of spares - of the same type of sensors you start with - keeping in mind the finite shelf life of some lower-cost sensors.

Ozone interference

The duration of the Breathe London pilot project (2 years) enabled new insights to be gained that will help improve the accuracy of lower-cost sensors in the future, more than what could be garnered in a shorter study period. The project consortium identified a gradual upward drift of NO_2 measurements, which was made clear over the long time series, and identified how ozone is likely interfering with the accuracy of the readings. At some sites the upward drift might obscure longer-term trends in concentrations. In response, the science team developed an algorithm to improve data accuracy by correcting for ozone cross-interference, and applied it to the project's full ratified dataset. This technique will support the effective application of lower-cost sensors that may encounter similar issues in other geographies going forward.

Environmental effects

Breathe London QA/QC processes revealed two issues that affected $PM_{2.5}$ and NO_2 data quality during certain environmental conditions. First, PM measurements appeared biased high during periods of high relative humidity and fog. Second, NO_2 readings appeared to be biased high during periods characterised by very high temperatures (> 30-35 °C), observed in London during the summer periods. Although these effects were not able to be fully resolved and corrected for within the span of the Breathe London pilot phase, the identification of these environmental effects emphasises the importance of region-specific evaluation of sensing technology, and rigorous QA/QC processes to identify periods of anomalous instrument performance prior to their implementation.

Monitoring Plan

Local knowledge and political buy-in

When deploying a sensor network, it is important to understand the local political structure of the city and obtain buy-in with important decision makers at all levels. In the case of London there are 33 local authorities that make up the Greater London study area. The process of agreeing installations took several months and involved many discussions determining the final placement of the pods. Factors such as safe accessibility, installation costs, power supply, and management of street furniture needed to be considered in identifying a suitable location. Often the locations suggested by the boroughs were the most suitable, highlighting the importance of bottom-up, local knowledge when deploying a stationary monitoring network.

Site selection for pods

The location and position of a pod can affect the measurements. In general, schools tend to be set back from roads, so a monitoring campaign that targets mainly schools will deliver lower readings compared to some reference monitors which tend to be at a mixture of locations including on kerbs and road-sides. This is true for a range of host sites, including residences, office buildings, and others where pods are located a few meters or more from the road. As the Breathe London network serves multiple purposes, the chosen locations comprise different types of sites. The choice of locations should fit the monitoring goals and objectives.

Site Permissions

Identifying the correct, responsible party for each site early in the planning stages allows liaison to commence as soon as possible and locations to be finalised. The overall assessment and feasibility of a location can take weeks and sometimes even months. In the case of Breathe London, agreeing the location of new pods required several steps, from agreeing a location with the air quality team in the local authority, to determining if the owner of the identified asset (e.g. lamppost) was the street lighting team at the borough, Transport for London (TfL) or a third party provider to agreeing the installation with the correct asset owner.

Microscale siting

Positioning of the pods can potentially impact the representativeness of the measurements. Due to the limited options for mounting pods on buildings or street furniture, large networks may have to compromise on locations. To assess possible sampling issues at sites that did not follow siting guidelines set out in the European Union (EU) directives for reference instruments, a microscale siting study was conducted at three sites to better understand the potential effect of pod siting (see Appendix 2). This showed the effects to be minimal, at least for the sites tested, so that an important point of lower-cost, small sensors is that they can in fact be sited in places that are impossible for traditional reference instruments, generally outweighing potential disadvantages associated with microscale siting.

Latent demand for pods

Throughout Breathe London many residents and schools concerned about air pollution in their neighbourhoods enquired about the installation of additional pods for integration into the network There are several challenges in meeting these requests that go beyond the costs of procuring and installing the additional pods including the costs and resources associated with on-going pod maintenance, data QA/QC and data analyses that are required to ensure consistent data quality with those pods already a part of the Breathe London network. These challenges can certainly be overcome if anticipated and planned for. In future campaigns, the interest and commitment from the local community could be factored into the monitoring strategy in order to accommodate requests to expand the network and further leverage local knowledge about where air pollution is of greatest concern.

Data quality assurance and quality control

Pre-deployment co-location at reference monitoring sites

Ideally, adequate time should be built-in prior to deployment to ensure that all sensors can be co-located at reference sites representative of their ultimate placement location and that as many pollutants are measured by the system as possible. However, a key output from the Breathe London project was the cloud-based network calibration methodology which yielded results comparable to physical co-location. It is highly recommended that future projects test the performance of at least a subset of the sensors to be deployed in order to address any unexpected compatibility or suitability issues.

Developing a robust data QA/QC protocol

The Breathe London science team developed ongoing transfer standard protocols to ensure that the fixed network remained traceable to reference grade instruments and correctly calibrated. The quality assurance and quality control procedures identified outliers and direct remedial actions to maintain high levels of data capture. The procedures formed a living document throughout the duration of the project to ensure that as new insights were gained, for example the creation of new data flags to identify suspect data, best-practices could be tested and updated. The QA/QC protocols are documented in Appendix 2.

Scaling pod data using unratified reference network data

Data from reference grade monitors in London are ratified relatively infrequently (typically every 6 - 12 months). Since empirical scaling factors for the pods may be determined prior to ratification of the reference network data, there may be errors in the reference data that subsequently necessitate correction of sensor scaling factors. Future studies that use reference site networks to calibrate individual pods or transfer standards (‘gold pods’) should consider the lags in obtaining availability of ratified data, and develop a plan to adjust scaling factors if unratified data is used. Alternatively, future projects of lower-cost networks could be accompanied by a “super reference site” that has its data ratified more frequently to eliminate this concern.

Designated gold pods and spare pods

A valuable capability is the ability to move a subset of pods within the network based on project needs. Designating a subset of pods for gold pod calibrations or as transfer standards will help maintain network performance and for testing and validating any cloud-based calibration methodology. It is recommended that future projects consider maintaining several calibrated, spare pods to allow anomalous sites to be investigated or to replace the pods in the network that have been otherwise rendered inoperable to ensure continuity. The ability to move spare pods within the network can determine whether atypical results are a data/measurement issue or local air pollution issue (i.e. potential hotspot).

3. Mobile Monitoring

Vehicles, sensors and equipment

Space and power

Breathe London demonstrated that it is possible to operate a comprehensive suite of research-grade monitors in a small vehicle albeit with some compromises in instrument configuration and power supply. Google's available fleet was limited to small cars and as a result, space and power were limited. To address these issues the standard 100 ampere (A) alternator was replaced with a 130 A alternator, instrument choices and operating modes were modified to reduce power draw, and custom installation was arranged to fit all the equipment safely and securely. Additionally, a larger auxiliary battery would be preferred, but this was not possible due to safety concerns. The amount of power available on the mobile platform can limit the data collection shifts and coverage. It is important to consider equipment with lower power requirements, smaller size, and less weight - something that instrument companies are continually developing to make equipment more suitable for mobile deployment.

Parking

Finding a suitable, secure site for the cars to park is a key consideration. Due to the daily equipment checks and calibration needs, it was important that the parking location be easily accessible by the staff conducting those tasks. Additionally, there is value in parking the cars in proximity to reference monitoring sites to enable long-term evaluation of instrument performance, which can eliminate the need for intentional co-locations to be scheduled.

Spare parts

Instruments utilise equipment, such as pumps, more intensively on mobile platforms compared to laboratory settings and therefore more prone to wear and tear. It is helpful to have spare items of specialist parts on hand (e.g. certain air pumps, wires, batteries, laptops, etc.) as these may take longer for delivery and may not always be in stock. In addition, many anticipated and unanticipated issues with the vehicles themselves surfaced during Breathe London (e.g. flat tyres, repairs) that caused unexpected delays.

Thermal cut off

Mobile monitoring instruments in Breathe London were powered continually, even when the cars are parked overnight, which presents challenges of temperature management. This strategy extends instrument life and reduces daily start-up procedures. In the case of Breathe London it also provided valuable co-location data for QA/QC as the overnight location of the cars was within meters of a reference monitor site. However, the instruments create their own heat, and need to be cooled during warm months in order to stay in safe operation. Putting in place adequate cooling systems, both on the mobile platform and at the NPL offices where the vehicles parked overnight, was critical in protecting the instruments and ensuring that they operate effectively. It is recommended that a thermal cutoff be employed to power down instruments in the event of cooling system failure, as occurred during deployment.

Driver training and manuals

While drivers need to conduct only basic operational procedures (e.g. turning on instruments), these can impact data collection and other parts of operations. Therefore, it is important to provide effective training and a user-friendly instruction manual for drivers, including basic troubleshooting procedures.

Monitoring plan

A mobile sampling strategy needs to be adaptive and flexible enough to accommodate either an over or underprediction of spatial coverage, especially in cities with dense traffic congestion. Sampling strategies should also be tailored to local air pollution concerns and incorporate local knowledge and city-specific data sources when possible.

It is also important for future project teams to understand the trade-off between spatial and temporal coverage with mobile platforms and that a single visit to a location provides limited utility for most analyses. As a minimum, Breathe London found that hyperlocal mobile analyses required at least five drive passes over monitoring sites of interest. An understanding of sampling uncertainty as a function of number of visits and acceptable levels of sampling uncertainty for desired analyses can help guide the monitoring plan (see Appendix 3 for more detail).

Selection of areas to sample

Where to sample

At the outset of the project, drive polygons were selected to cover the full ULEZ area in central London. For areas outside of the ULEZ, a stratified random sampling strategy was developed to capture data in representative locations across the large GLA study area. Determining representative locations required local knowledge about geospatially resolved socioeconomic data that could be integrated with locally relevant geographic boundaries. This was accomplished by using the UK Index of Multiple Deprivation and postcodes for boundaries (London's air quality model was also used to sample in both high and low pollution areas). However, it is important to note that the selection of *where* to drive is highly dependent on the aims of the project. Initial planned areas were modified before data collection to include other areas of interest for the city. As the monitoring progressed, road coverage was reduced within polygons and/or reduced polygon areas in order to reach sufficient drive coverage in priority areas.

When to sample

Future mobile monitoring projects should consider when to drive/sample based on project goals and objectives, and design drive shifts to cover certain hours of the day, days of the week, and months of the year as needed. Options to consider are staggering shift start and end times to get more even coverage of all hours of the day and deciding whether overnight or weekend monitoring is needed.

Transects

In addition to drive polygons, the initial sampling plan included transects across the city or between polygons. This was intended to highlight spatial variability in the pollution profile of the city as the car drives from east to west or north to south. The shorter transects between polygons also allow data collection at areas of interest as the drivers are commuting between polygons. Because actual drive coverage was less than predicted, transect driving was limited in practice other than transiting to/from NPL offices to central London.

Consider local conditions in planning spatial coverage

Because of the need to collect data across all boroughs of the Greater London Authority (GLA), the required transiting to and from the mobile monitoring base at the National Physical Laboratory (NPL) offices, and traffic congestion within central London, the spatial coverage of mobile data collection was less than initial predictions based on previous work in other cities. As a result, adjustments were made to the sampling plan. Daily drive plans were optimised to have drivers cover clusters of closely spaced polygons, drivers' shifts, and placements were re-configured to minimise lost time, journeys between polygons were strategically used as transects, and polygon boundaries were adjusted. Future mobile monitoring campaigns should consider these strategies, and factor in transit time and local congestion when developing a mobile data collection plan.

Data quality assurance and quality control

Data Logging

Logging data correctly is vital to the overall data quality and the ability of the dataset to be used for intended analyses. The project developed fast collection data logging capabilities and improved the team's use of GPS positioning technology and as a result of these learnings, strongly recommend the following are considered:

- Test and ensure all instruments, laptop, and GPS are logging the same timestamp and the source of that timestamp is a GPS satellite atomic clock to minimise drift.
- Test and ensure the correct time response data for each instrument on the mobile platform (don't rely solely on stated time response in the manual as instruments can behave differently in the field, and can be affected by the sampling inlet system)
- Create a system for automatic alerts for data flags to alert the operations team of instrument issues as soon as possible (Breathe London used AirSense to address this issue)
- Ensure that instruments are set to appropriate time averaging (generally the minimum possible) and other settings (e.g. calibration periods on instruments and reporting interval) that could affect data capture and analyses. Look at timeseries of the measurements for signs of unwanted time-

average/smoothing because some manufacturers do not explicitly specify that averaging is used. Instrument reporting and averaging are not necessarily the same. Avoid changing instrument settings midway through the project as this could block the ability to evaluate temporal trends.

- Keep data on the onboard data logger long enough so that data is not lost if there are any issues with automated uploading of data to the cloud. Or alternatively, institute alerts to ensure data has been manually uploaded before it is overwritten.
- Avoid relying on internal instrument storage as they all have individual clocks and are difficult to keep in sync. A central data collection system with real time synchronisation to a cloud time server is the best solution. The system should not only collect measurement data, but corresponding instrument status and metadata, where available. Where internal storage is unavoidable, have a plan for synchronizing locally stored data with the GPS timestamp and test it prior to initial deployment.
- Create a fully automated data pipeline from instrument to final destination in the cloud. Avoid manual steps like saving data files from instruments locally and uploading large files at a later date. Properly size databases, bandwidth, and computers so that all data can be streamed. Avoid logging systems that require use of proprietary software to extract encrypted data.
- Any manual data entry that cannot be avoided should be checked for transcription errors by an independent reviewer.
- Standardise file formats and contents with machine-friendly structures like clearly labeled rows and columns and key-value pairs. Avoid structures that rely on data being located in an absolute cell position (such as in an Excel spreadsheet). Users can easily put values in the wrong cell making it a challenge for a machine to read the data.

The result of such learnings was that information on the challenges encountered was shared with instrument manufacturers with the aim of creating improved technologies for mobile use in the future. An open dialogue with manufacturers is important in ensuring that a monitoring campaign further contributes to the development and advancement of the technology.

Sampling systems

Much was learned about how the configuration and design of sampling systems affect the results obtained. The team made recommendations as to how to overcome some of these effects in future designs, including lag and response times and particle losses in inlets. In general, for avoiding particle losses, it is recommended that metallic tubing (e.g., stainless steel, copper) be used instead of plastic (e.g., Teflon). Losses in tubing are also dependent on particle size: large particles (> 1 micron) are lost when going through sharp bends in tubing ("inertial loss"), while small particles (< 100 nm) are lost simply due to diffusional mechanisms. To avoid inertial losses, the amount of sharp bends in the tubing should be avoided. Copper tubing is easy to bend without using any tools. To avoid diffusional losses, tubing of wider diameter (e.g., ½" or ¾" outer diameter) may be used to reduce the proximity of air flow to tubing walls. However, this can also smear instantaneous peaks in signal and increase lag times. Lastly, shorter the tubing length, smaller the overall particle losses.

Considerations for Emission Ratio Analyses

An instrument firmware update that occurred in the $PM_{2.5}$ monitors in April 2019 resulted in a change in the time response of measurements causing the concentrations monitored to be spatiotemporally broader than the CO_2 measurements. This created a discontinuity in the time series of $PM_{2.5}$:CO_2 emission rations, which negatively impacted the ability to assess trends. The lesson learned here is that for accurate emission ratio calculations, the instrument's rate of data capture and time response for both pollutant and CO_2 must be aligned. Changes in instrument configurations should be avoided if the goal is an assessment of temporal trends

Calibrations

Mobile instruments should be calibrated weekly to minimise drift over time and maximise the ability to conduct temporal trend analysis. Calibrations should include collecting data for at least five-minute periods every week at zero and span reference concentrations. Conduct multipoint calibrations and longer-term calibrations (~1 hr) at reference points of zero, span, and ideally a few points in between at the beginning, middle, and end of the project to characterize the bias, precision, and accuracy of each sensor - essential statistics for the interpretation of any insights from the data. Record the start and end time of calibration periods so that the data is easily retrieved later for critical uncertainty calculations. Automatically log calibration data rather than rely on manual documentation. Set up auto alerts for calibration data that is missing or out of range and address issues immediately before continuing with monitoring.