aclima

Mobile Ambient Air Pollution Measurement Quality Assurance System

March 10, 2023

Version 3.1

Table of Contents

1.0 Introduction

Mobile air pollution mapping is a flexible method to measure concentrations of a broad range of air pollutants and greenhouse gases over large geographic areas at high spatial resolution. Aclima has developed a platform for sampling ambient air pollutants at the hyperlocal level. Aclima-operated fleet vehicles are equipped with the custom-designed onboard mobile sensing system, the Aclima Mobile Node (ANM), which measures CO, CO_2, NO, NO_2, O_3, CH_4, C_2H_6, TVOC, BC, and $PM_{2.5}$.

The data collected on-road enable a variety of high-resolution spatial analyses that support different use cases. A fundamental output is the creation of maps of typical air pollution concentrations at block-by-block resolution that show areas of persistently high or low levels of individual pollutants. The other case study uses time-resolved data from several individual drives of the same location to identify areas where pollution concentrations vary substantially from local levels, indicating a probable local emissions source. Outlined below are Data Quality Objectives (DQOs) for data products based on these two analysis approaches as well as the specific pollutants to which the data product applies.

1. **Show areas that have relatively higher and lower pollution in a geographic region at high spatial-resolution**:
   Data Quality Objectives:
   - Produce yearly or seasonal estimates of ambient pollution concentrations from measurements balanced over the contract time period and at diverse times of day and night, weekdays and weekends, to adequately address seasonal and diurnal variations in the data.
   - Data are spatially distributed throughout the entire user-defined area.
   - Annual or seasonal estimates delivered with credible intervals at the contracted spatial resolution sufficient to enable assessment of the significance of differences in pollution levels.
   - Support pollution estimates at ~100 m road segments (sometimes referred to as "address-level").
   Pollutants:
   - O_3, NO[1], NO_2, CO, CO_2, CH_4, $PM_{2.5}$, and BC

2. **Identify probable locations of air toxic emissions sources**
   Data Quality Objective:

---

[1]See section 5.6.3 for details on the limitations of the NO sensor, which has a relatively high detection limit relative to typical ambient concentrations.

- Produce geo-located clusters of TVOC source indications to support follow-up investigation by identifying enhancements above background in the form of peaks in the 1-Hz sensor output that have a signal to noise ratio of at least 3.

Pollutants:

- TVOC, BTEX

3. **Show areas that indicate biogenic methane emissions sources**
   Data Quality Objective:
   - Produce geo-located areas where enhancements in methane indicate biogenic sources by using to ratio between methane to ethane

   Pollutants:
   - CH_4, C_2H_6

These DQOs are largely qualitative goals that provide the foundation for the types of insights that the Aclima data products are designed to support. A critical aspect of quality assurance underlying these DQOs is characterizing and maximizing the measurement quality of the sensors, however, our confidence in these data products will depend on a number of additional factors, including but not limited to our sampling strategy, the number of samples collected for features of interest (i.e. road segment or other spatial length scale), and magnitude and variability in pollution concentration and meteorology over the contract period. Therefore, the overall approach to quality assurance must include processes and metrics that address all potential sources of uncertainty, and success in meeting the data quality objectives cannot be measured by a single set of measurement quality objectives (MQOs). Further, many of the factors impacting data quality are subject to real world operational challenges (i.e. driver absences) or outside the control of even a perfectly executed QA plan (i.e. strong wildfire events impacting air quality over short time periods). Therefore, we cannot guarantee any specific maximum confidence interval (or precision) around individual atmospheric concentration estimates, but we can provide data that are high quality with well-characterized performance parameters to support aggregation and analysis of our data products, which can be used to support decision making in various use cases.

With an understanding that the quality of Aclima's data products involves many complex factors, the measurement quality of the sensors is critical. This document outlines the Quality Assurance processes for the devices, measurement, and sampling that provide the foundation upon which Aclima's data products are produced. The document covers Aclima mobile monitoring methods, equipment, and sampling methodology from device assembly to 1-Hz time-resolved data verification for each device and vehicle. We also include information on sensor calibration and data review processes.

2.0 Mobile Monitoring Methods and Equipment

This section outlines how vehicles are outfitted with Aclima's custom-designed mobile monitoring platform, as well as the reference equipment and laboratories used to calibrate and assess sensor performance.

2.1 Aclima Mobile Measurements

Aclima mobile fleet vehicles sample ambient air from a custom-designed, window-mounted inlet anchored in the rear passenger side window of the vehicle. The mobile platform has two primary sampling lines that transfer air from the ambient atmosphere to the internal sensors, one for gases and one for $PM_{2.5}$. Additional sampling inlets are used for additional measurement devices.

To minimize changes in concentrations of CO, CO_2, NO, NO_2, O_3, and TVOC due to chemical reaction with or loss on tubing surfaces, these gases are sampled through non-reactive PTFE (polytetrafluroethylene) tubing with PFA (perfluoroalkoxy alkane) fittings. These gases are measured within the Aclima Mobile Node (AMN). Sample lines for CH_4 and C_2H_6, and air toxics consist of PTFE tubing with stainless steel fittings. Methane and C_2H_6 are measured by the same sensor (one sample line) and air toxics species by separate sample line. Methane and C_2H_6, and air toxics sensors are installed next to but separate from the AMN. Inlet lines for gases extend straight out from the window and the inlet of the sample lines are perpendicular to the surface of the vehicle. A Nafion™ dryer is placed inline, just inside the window, to help control the water content in the sampled air for CH_4 and C_2H_6. A Nafion™ dryer is not used for other gases, including air toxics.

To minimize particle loss, sample lines for $PM_{2.5}$ and BC are composed of copper and antistatic silicone conductive tubing with brass fittings. The sample lines for $PM_{2.5}$ and BC extend from the inlet mounting surface with its opening situated parallel to the car exterior and facing towards the front of the car. Separate sample lines are used for $PM_{2.5}$ and BC. No size discrimination for $PM_{2.5}$ occurs at the inlet. Size discrimination is based on optical scattering. A Nafion™ dryer is placed inline for each, just inside the window, to help control the water content in the sampled air for $PM_{2.5}$ and BC. The BC sample line includes a cyclone (sharp-cut 2.9 μm at 150 mL/min) in order to keep larger particles from depositing inside the sensor and deteriorating performance.

The $PM_{2.5}$ sensor reports particle counts in six size ranges. Aclima uses a standard mass conversion model assuming spherical particles and constant particle density across the different size ranges to convert particle counts to particle mass. The $PM_{2.5}$ and BC sensors are used as provided by the manufacturer and only flow rate is verified before use. Additional details on limitations of the PM sensor are discussed in Section 5.6.2 and in the accompanying

"Hyperlocal Ambient Concentration Estimate Validation and Quality Assurance Plan" document (Section 5.2).

Table 1: Aclima Mobile Monitoring Platform Sensor Operational Characteristics

Parameter	Measurement Method	Measurement Frequency	Operational Range*
O_3	UV Spectroscopy	2 sec	0-100 ppm
NO	Electrochemical	1 sec	0-20 ppm, adjustable
NO_2	Electrochemical	1 sec	0-2 ppm
CO	Electrochemical	1 sec	Adjustable
CO_2	Nondispersive Infrared	1 sec	0-5000 ppm
CH_4	Direct Absorption Spectroscopy	1 sec	1 ppb-1%
C_2H_6	Direct Absorption Spectroscopy	1 sec	1 ppb-1%
TVOC	Photoionization Detection	1 sec	1 ppb - 20 ppm
$PM_{2.5}$[1]	Optical Light Scattering	1 sec	0-300,000 counts/L
BC	Optical Absorption at 880 nm	1 sec	0-1 mg BC/m^3
Air Toxics (BTEX)	UV Differential Optical Absorption Spectroscopy	1 sec	0-500 ppb (variable by species)
Temperature	Diode voltage	NA	-40-85 ℃
Relative Humidity	Resistive	1 sec	0-100%
Pressure	Capacitance	NA	300-1100 hPa
Location (GPS)	Satellite GPS	1 sec	> 5 m
Time	Real-time clock (RTC) via Network Time Protocol	32.768 kHz oscillator	NA

Additional sensor performance metrics are included in Table 5.

* Manufacturers' specification

[1.] $PM_{2.5}$ is determined from particle number (PN) counts, where PN is measured in six size ranges between 0.3 and 2.5 μm with count concentration (c/L) converted to mass concentration assuming spherical particles and constant density across the six size ranges.

Sensors located within the AMN each record in-line temperature (T), pressure (P), percent relative humidity (RH%), and flow-rate at 1 Hz. These sensors are located close to the respective sensor so environmental conditions are known when the pollutant is measured. The BC and $PM_{2.5}$ sensors also include T, P, RH, and flow rate (at 1 Hz) as part of their measurement system, while the CH_4 and C_2H_6 sensor includes temperature, pressure, and absolute H2O concentration (at 1 Hz). The air toxics sensor includes T, P, and RH as part of its measurement system (at 1 Hz). These environmental measurements provide system state diagnostics or *status indicators* and are used in Aclima's sensor models to convert the sensor signal (e.g., voltage) into a physically meaningful concentration measurement ($\mu g/m^3$, ppb, ppm) and to ensure sensor measurements are maintained within their appropriate operating conditions. Because the T, P, and RH measurements are specifically designed to measure internal instrument conditions, they are not representative of ambient atmospheric conditions and are thus not provided as part of Aclima's data products.

To account for differences in sample transport time among instruments, an algorithm was developed to time-shift concentrations to maximize the temporal correlations between the individual pollutants on a daily basis (Apte et al. 2017, Supplemental Information). Transport time is the time from when the sample enters the inlet until an instrument response is observed. The degree of time-shifting depends on the response time of the instrument and its position on the sampling manifold. For example, in Apte et al. (2017), daily time shifts for individual instruments ranged from ~0.1 – 12 seconds, and were relatively stable during the course of the one year experiment. These time shifts represent when the inlets were collocated and were positioned in a forward-facing orientation, several inches above the roof line at the rear edge of the front window. In Aclima's current mobile platform, where the inlets are located at the rear passenger window, time shifts range from about 2-10 seconds and have been estimated based on sample tube lengths, inner diameter, and flow rate. Uncertainty around these estimates primarily comes from the uncertainty in flow rates, estimated at approximately ±5%.

The geographic position of each measurement, described using latitude and longitude, is provided by GPS (manufacturer specifications of ±5 m) and is measured at 1 Hz. Evaluation of GPS position during field operations when parked indicated an uncertainty of ±6 to ±12 m depending on location and averaging time of the data, with smaller uncertainty at longer averaging times. Vehicle speed is determined based on change in GPS location and provided at 1-Hz. All measurements are synchronized using Network Time Protocol (NTP), which synchronizes the internal computer to Coordinated Universal Time (UTC), and are reported to the nearest segment in UTC. Individual sensors are tracked throughout by project, serial number, parameter, model, and start and stop time of use.

The measurement method for each of the sensors employed in Aclima's mobile monitoring platform is listed in Table 1, along with the reporting frequency and operational range as reported by the manufacturer.

Table 2: Manufacturer Specifications for Reference Instruments and Gas Standards

Target Pollutant[1]	Detection Principle	Resolution[2]	Range	Response Time
Reference Instruments				
O_3	UV Absorption	0.1 ppb	0.001-2 ppm	4 sec
NO	Chemiluminescence	0.001 ppm	0-0.5 ppm to 0-100 ppm selectable	<1 sec
NO_2	Cavity attenuated phase shift spectroscopy	0.010 ppb	0-1 ppm	<10 sec
CO_2	Non-Dispersive Infrared (NDIR)	0.1 ppm	0-20,000 ppm	~1 sec
$PM_{2.5}$[3]	a) light scattering[4] b) beta attenuation	a) < 0.1 ug/m³ b) 1.0 µg/m³ (24 hr)	a) 0.1 - 10,000 µg/m³ b) 0 - 10,000 µg/m³	a) 10 s data rate b) 1 hour
Certified Reference Gas Standards				
Target Pollutant	Standard Gas	Span Values	Standard Gas Uncertainty	
CH_4	NOAA and NIST Certified Reference Gas Standards[5]	~2 ppm ~5 ppm ~50 ppm	NOAA standards: +/-0.2% NIST standards: +/-0.2% (at 2 and 5 ppm) or +/-2% at (50 ppm)	
C_2H_6	NOAA and NIST Certified Reference Gas Standards[5]	~ 1 ppb ~ 50 ppb ~ 5000 ppb	NOAA standards: +/-0.2% NIST standards: +/-0.2% (at 1 and 50 ppb) or +/-2% at (5000 ppb)	
TVOC	NIST Certified Reference Gas Standards[6]	zero and ~1 ppm after dilution	20.00 ppm ± 5%	
Benzene Toluene Ethyl Benzene Xylenes	NIST Certified Reference Gas Standards[7]	zero and variable (up to ~250 ppb)	5 ppm ± 10%	

[1] Reference methods are not currently used in the Mobile Calibration Laboratory for $PM_{2.5}$, BC, and air toxics instruments. These sensors are used as received by the manufacturer with ambient concentrations directly reported by the instrument.

[2] Resolution is the ability to distinguish between two measurements from each other and does not refer to the limit of detection or the ability to distinguish from the baseline single.

[3] Installation of a PM reference method in the Mobile Calibration Laboratory is under development.

[4] Light scattering provides particle counts across various size ranges that are converted into mass concentration ($\mu g/m^3$). Mass concentration is reported by the instrument.

[5] A combination of reference standards prepared by NOAA's Central Calibration Laboratory and working standards prepared by NIST are used for CH_4 and C_2H_6. The latter acts as a transfer standard. The uncertainties for the NIST working standards were characterized against the NOAA standards.

[6] Isobutylene in nitrogen.

[7] Blend of 4 gases (benzene, toluene, ethyl benzene, and xylenes). Exact isomeric mixture of xylenes not known in the blend.

2.1.2 Reference equipment and standards used for calibration and performance assessment

Aclima uses laboratory-grade reference equipment in Aclima's location-specific Mobile Calibration Laboratories and at Aclima's location-specific Ambient Calibration Laboratories to calibrate O_3, NO, NO_2, and CO_2 measured by the Aclima mobile platform. Instruments selected for use as reference for calibration are historically tested and widely-used detection methods accepted by the atmospheric science community as a high-quality measurement. For some sensors, NIST certified gas standards were used as a reference for calibration and evaluation. All reference instruments have well defined uncertainty over a wide range of conditions. Table 2 provides a summary of manufacturers' specifications or certifications for all reference instruments or gas standards. (Note that the reference equipment is not operated in a manner that meets the requirements of regulatory reference methods, e.g. mobile sampling uses 1-Hz data rates vs typical gas reference method requirements of a 60-sec averaging time.) NOAA (primary standard) and NIST (transfer standard) traceable gases are used to calibrate the combined CH_4 and C_2H_6 sensor (Section 5.2.4). Individual $PM_{2.5}$ and CO sensor performance (precision and bias) and sensor-versus-sensor comparability are quantified, under real world driving conditions, pre- and post-deployment by peer-to-peer collocation using Aclima's Mobile Calibration Laboratories (Section 5.2.2). Aclima has recently installed a reference instrument for PM at its primary Ambient Calibration Laboratory in San Francisco, CA. The PM reference instrument will serve as a benchmark for future PM calibrations. Volatile organic compounds (TVOC) undergo an initial calibration using a reference gas (Section 5.2.4) and are validated through the peer-to-peer collocation approach (Section 5.2.2). The air toxics instrument is calibrated using zero air and a multi-point span using a reference gas mixture containing benzene, toluene, ethyl benzene, and xylenes (Section 5.2.4). The BC sensor flow rate is calibrated using a certified flow calibrator provided by the sensor manufacturer specifically for use with the BC sensor. The BC sensor is not routinely collocated with a reference method or with other BC sensors as part of its standard calibration procedure. However, occasional collocation experiments have been performed in order to characterize general sensor performance (see Sections 5.2.3 and 5.5 for more details).

Table 3: Reference instruments[1], quality control, and corrective action

Description	Modalities	Frequency	Acceptable Bounds	Corrective Action[2]
Reference Instrument span using NIST traceable methods	O_3, NO, NO_2, CO_2	Monthly	± 5% (O_3, NO, NO_2); ±10 ppm (CO_2)	Recalibrate the instrument span
Reference Instrument zero checks NIST traceable methods	NO, NO_2, O_3, CO_2	Monthly	±0.5 ppb (O_3, NO_2); -0.3-1.3 ppb (NO)	Recalibrate the instrument offset
O_3 Generator[3] and Gas Calibrator	NO, NO_2, O_3	As needed		Manufacturer service and recalibration
NIST traceable gases (cylinder)[4]	CH_4, C_2H_6 TVOC[5] (isobutylene) NO CO_2 BTEX	As needed when cylinders reach <25% full or certification expires		Replacement with new cylinder
Zero Air Generator	NO, NO_2, O_3, BTEX	As needed		Manufacturer service
Flow Calibrator	BC	Approximately annually, as recommended by mfr		Recalibration performed by manufacturer

1. Reference instruments are not available at this time for use in the Mobile Calibration Laboratories for CO, air toxics, $PM_{2.5}$, and BC. Aclima uses a relative reference or the average of an ensemble of CO, TVOC, or $PM_{2.5}$ sensors of the same make and model as the test sensor as described in Section 5.2.2. BC and $PM_{2.5}$ are currently used as received by the manufacturer.

2. If reference instruments fail to meet acceptance criteria they are sent back to the manufacturer for recalibration or replacement. Instruments are sent back annually for recalibration.

3. The O_3 generator is not routinely sent to the manufacturer for recalibration/certification, but only when service is required and may result in a small source of systematic bias for O3 and NO2.

4. For CH_4 and C_2H_6, Aclima uses a combination of reference standards prepared by NOAA's Central Calibration Laboratory and working standards prepared by NIST. The latter acts as a transfer standard.

5. TVOC - isobutylene certified gas standard in cylinder at ~1 ppm for an initial calibration, which is followed by peer-to-peer collocation in an Aclima Mobile Calibration Laboratory where the slope and R^2 values for bias and precision are obtained from comparison to the relative reference, similar to PM, TVOC, and CO.

Since laboratory-grade reference equipment forms the basis for data quality, Aclima performs a suite of quality control checks that are integrated into our operating protocols to quantify uncertainty in the reference measurements to ensure they are within an acceptable range. Table 3 details the type of checks performed with reference instruments, the frequency at which they are performed, the values triggering corrective action, and the type of corrective action triggered. Internally developed processes and instrument manuals are used to conduct the checks.

3.0 Fleet Operations

This section focuses on driving operations – the planning and operational information about how Aclima maps with a mobile measurement fleet. Details regarding field and laboratory procedures for measurements and data verification and validation are provided in Sections 2.1 and 6.0.

Pre-planned driving route assignments are delivered as daily map-based drive plans directly to drivers using customized advanced mapping software. Daily driving routes consist of spatially-limited geographic areas associated with the staging area (hub) for each set of vehicles that monitor that area.

Drive plans are assigned as a series of connected road segments, or "Ways," and drivers are instructed to drive across all of the Ways (streets) in their assigned plan during their daily shift. This approach has been found to spatially randomize the mobile sampling artifacts and increase the number of visits within neighborhood-level geographic regions like census tracts despite targeting the same number of passes per segment.

Drive plans are automatically generated to assign areas and road segments randomly across time to achieve temporal balance of each road segment by time of day (24 hours), day of week (seven days per week), and seasonally. As the monitoring period progresses, the drive-planning algorithm assesses sample balance and prioritizes areas and road segments that are determined to be undersampled to maximize temporal and spatial representativeness. The current driving approach also helps to minimize potential bias associated with variations in meteorology and source emissions.

Drive planning based on progress towards coverage of road segments supports the air toxics data product, which aims to detect and localize large enhancements in TVOC concentration over short time periods with respect to a baseline. Time-resolved data from individual drives are analyzed for times when measured pollution concentration varies substantially from average local background levels. Detection of these short time-scale peaks on repeat visits is an indication of a persistent localized emissions source. The balanced sampling approach combined with multiple visits to the same road location supports the identification of persistent local sources of emissions

3.1 Mobile Fleet Maintenance and Diagnostics

Preventative maintenance tasks are implemented at various frequencies as specified by the manufacturer or based on past experience as outlined in internal SOPs. Aclima uses a multi-level approach to mobile platform maintenance to help ensure the successful collection of high quality data while minimizing data loss. A rigorous schedule of system checks and replacements has been established, which includes daily, weekly, and monthly system maintenance tasks. A separate AMN/sensor database links each unique instrument (by serial number) to allow for tracking preventative maintenance due dates for individual sensors and mobile platforms.

Aclima technicians perform a full diagnostic check on the system at each installation of an AMN or sensor into a vehicle and whenever any on-demand maintenance is required. Diagnostic checks are also performed immediately prior to uninstallation in order to verify that the integrity of the system, especially the sampling system, was maintained over the duration of the deployment. The following tasks are performed and logged as part of these diagnostic checks:
- gas sample line filter change
- BC sensor sample line cyclone cleaning
- flow checks using an external calibrated flow sensor, with acceptable thresholds specified in SOPs
- PM sample line zero check
- full inspection of all sample line connections
- sample line cleaning using forced air
- checks on integrity of the mounting hardware
- power system and communications checks

Multiple diagnostic tests are also conducted by the drivers on-site with the vehicles and remotely by Aclima staff or when the vehicles report back to one of Aclima's location-specific Ambient Calibration Laboratories. On-site, drivers perform the following on a daily basis:
- visual inspection of the system, including sampling line connections
- PM zero checks (using a filter on the PM sample line),
- visual inspection of sampling lines (e.g., presence of water or disconnection)
- data connectivity checks
- cleaning the BC sensor sample line cyclone
- other system parameters, including flow system status diagnostics

Drivers are trained and guided by Aclima staff to make simple repairs, such as change a filter, or replace or repair a sampling line or leak. Remotely, Aclima staff check various diagnostic variables and make adjustments to account for drift (see section 5.3).

If more advanced maintenance is required, the AMN or other sensor device ($PM_{2.5}$, BC, CH_4 and C_2H_6, or Air Toxics sensor) is brought back to one of Aclima's location specific Ambient Calibration Laboratories for recalibration or swapped out in the field with a different sensor. Examples of repairs that may occur at Aclima include replacement of in-line filters for the gas sensors,replacement of PM sheath air filters used to keep internal components clean, and replacement of consumable materials, such as batteries and other filters.
Planned routine car maintenance occurs when cars report to one of the location-specific Ambient Calibration Laboratories or in the field by drivers (e.g., oil change) during field operations as required or as time permits. Planned preventative car maintenance is tracked in a separate database from the sensor tracking system. If issues occur outside of planned preventative maintenance periods (sensors or vehicle), a system has been established, including inclusion of relevant information in the vehicle tracking database that helps to ensure a quick resolution with minimal loss of data during deployment.

Aclima drive managers and technical staff follow procedures developed to interact with a driver when safety, car, communication, or measurement problems are detected. Based on those procedures, Aclima trains drivers on safe vehicle operations, device operation, route navigation, basic device troubleshooting, and protocols for coordinating with Aclima fleet operations staff at headquarters to address any anomalies. The procedures are designed to maximize the collection of valid data while ensuring driver safety. These procedures continuously evolve with lessons learned in the field.

4.0 Measurement Quality Objectives

Measurement Quality Objectives are quantitative criteria used to assess data quality and ensure that the DQOs are achieved. Aclima's MQOs include setting acceptable quantifiable criteria at the device level for the DQIs for precision, bias, and completeness. Per USEPA (2002) guidance, the term "accuracy" is not used since it represents both precision and bias. Aclima's MQOs define the performance and acceptance criteria that calibrated and verified sensors must meet before being installed into the Aclima Mobile Platform. MQOs are derived through Aclima calibrations and quality control checks (Section 5.0; Table 4). Data from each sensor is evaluated during and after collection to determine the validity of the 1-Hz data. Data verification includes two multilevel steps, which may lead to data revision and final validated data as described in Section 6.0.

Multiple in-line measurements of T, P, RH, and flow rate are recorded in conjunction with measurements from each sensor. These data provide system state diagnostics or Status Indicators, which are applied to sensor models along with calibration results and to verify that sensors remain within their normal physical operating limits of T, P, RH, flow rate, and concentration. If limits are exceeded, data are flagged at two levels and either excluded from additional aggregation and analysis, maintained but considered suspect (associated with a larger uncertainty), or investigated further (Section 5 .3 and 6.0). Consistent exceedance of

Status Indicator limits suggests potential issues with the system within the mobile platform that may impact the validity of the pollutant measurements. It may also indicate an issue with one or more of the state diagnostic sensors, which is important to ensure appropriate data for sensor models. In either case, tracking these measurements allows preventative maintenance before sensor or other system failure that might impact the measurement, minimizing loss of data. System state diagnostic limits are based on manufacturer specifications as well as internal evaluations of sensor performance.

Table 4: MQO - Aclima Mobile Node Acceptance Metrics

Pollutant	Acceptance Metric vs. Reference Instrument	
	Precision (± value) (Section 4.1.1)	Bias (Section 4.1.1)
O_3	3 ppb	3 ppb
NO	10 ppb	10 ppb
NO_2[1]	5 ppb	6 ppb
CO_2	11 ppm	12 ppm
CH_4[2]	10 ppb	50 ppb
C_2H_6[2]	10 ppb	15 ppb
BTEX[3]	LoD (see Table 6)	10%
	Acceptance Metric using Other Approaches	
	Precision based on R^2	Bias based on Slope
CO	>0.8	0.75 - 1.25
TVOC	>0.75	0.5 - 2.0
PM2.5[4]	>0.8	0.75 - 1.25
BC[5]	Successfully pass internal flow calibration procedure	

[1] NO_2 uncertainty is a combination of Ox (NO_2+O_3) and O_3 uncertainty since NO_2 is measured by difference between a sensor that measures Ox and a sensor that measures only O_3. Generally, under the conditions of our data collection, the Ox uncertainty dominates.

[2] See Section 5.2.4.

[3] BTEX = Benzene, Toluene, Ethyl Benzene, and Xylenes. See Section 5.2.4.

[4] See Section 2.1, Table 1, Footnote 1 regarding conversion of counts to mass concentration.

[5] The BC sensor is used as delivered from the manufacturer.

Other supporting data include GPS that provide high precision latitude and longitude position, and car speed. Sensor and system state data are used to create single pass mean values, which are used to create the high spatial resolution baseline aggregates. These aggregates are then used to develop spatial maps for Aclima's current mapping products.

Device-level MQOs or acceptance criteria are provided in Table 4; precision and bias are described in Section 4.1.1. These criteria were arrived at after extensive characterization over different phases of development from prototype to V2 version of the AMN and other sensors outside the AMN (BC, CH_4 and C_2H_6, and air toxics). These criteria must be met before a sensor is included in the Aclima mobile platform, and thus represent the upper limit (worst case) of what is expected from real-world (on-road) performance for a sensor at the start of deployment. Actual performance, which is characterized in Section 5.5, is typically better than these MQO criteria, but can also degrade over time due to sensor drift.

4.1.1 Precision and Bias

Precision and bias for O_3, NO, NO_2, and CO_2 are calculated based on the methodology used for meeting regulatory quality assurance requirements (Federal Register, 2022). Precision and bias are calculated using ambient, on-road concentration data measured by reference instruments that are collocated with sensors in Aclima's Mobile Calibration Laboratories. Inlets of the reference instruments and sensor are placed in close proximity to each other so both are sampling from similar air parcels. A slightly modified approach is used as the primary comparison statistic for precision and bias, where the difference between the sensor measurement concentration and the reference or target measurement (Equation 1) is used instead of the percent difference (Federal Register, 2022).

$$d_i = (\text{sensor - target}) \qquad \text{(Equation 1)}$$

Precision is expressed as the upper bound (at 90% confidence) of the standard deviation of the difference (d_i) measured over a specified time period (e.g., Quarterly Baseline).

Bias is expressed as the upper bound of the mean absolute value of the difference (d_i) over the same time period as given in Equation 2.

$$\left|AB\right| = 1/n(\sum_{i=1}^{n} |d_i|) + t0.95_{,n-1} (AS/\sqrt{n}) \qquad \text{(Equation 2)}$$

where, n is the number of single point checks being aggregated; $t0.95_{,n-1}$ is the 95th quantile of a t-distribution with n–1 degrees of freedom. AS is the standard deviation of the absolute value of the d_i's.

For CO, TVOC, and $PM_{2.5}$, precision and bias are based on regression analysis, where the test sensor (x-axis) is collocated in the Mobile Calibration Laboratory with an ensemble of sensors of the same type (make and model; y-axis). The average of the ensemble of sensors is referred to

Patented & Patents Pending │ Trade Secret & Proprietary Business Information

as a relative reference. While the regression is usually linear, it also may be based on a nonlinear function depending on the pollutant and sensor. Precision is given by the coefficient of determination (R^2) and the bias is based on the slope of the regression line.

The reference for CH_4 and C_2H_6 is based on a two point (~ambient and high concentration span) calibration using NIST traceable calibration gases (training) with a subsequent middle span check (testing), which must pass the acceptance criteria in Table 4 prior to deployment. The same three reference gases are used for periodic calibration checks. The NIST CH_4 and C_2H_6 gas mixture is a working or transfer standard verified by comparison to a gold standard NOAA gas mixture.

The UV-DOAS Air Toxics instrument undergoes a zero and multi-point calibration check using a NIST traceable calibration gas (containing benzene, toluene, ethyl benzene, and xylene) diluted with zero air. Precision is calculated as the standard deviation of the measurement while sampling zero air. Each compound measured by the instrument has a slightly different Limit of Detection (discussed in Table 6, Section 5); if the precision from a zero air check is greater than the LoD for a given species, the instrument does not meet the acceptance criteria. The bias is calculated based on the slope of the multi-point calibration. If the slope is more than 10% different from 1.0, the acceptance criteria is not met.

The BC sensor is used as provided by the manufacturer. Manufacturer specifications that relate to sensor performance include: flow repeatability ±0.4% in the range of 0-1 L/min; flow accuracy ± 5% at the range of 0-1 L/min. Performance of the sensor has been more directly evaluated through various one-off experiments, including peer-to-peer comparisons and collocations with regulatory measurements. Details on performance are provided in Section 5.5.

Quality control procedures for calibration and quality control checks for gases and particles are described in Section 5.2.

4.1.2 Data Completeness

Data completeness is calculated as the number of final verified 1-Hz data points collected relative to the number of data points expected, based on the amount of time a vehicle is actively driving on the road and the reporting frequency of each instrument (e.g., most gases are 1 Hz, ozone is 0.5 Hz). Aclima's completeness goal for 1-Hz data collection is >95%, with one exception. The completeness goal for the CO_2 sensor is >80% because its internal process results in sporadic drops of 1-second data points. Overall, a lower percent completeness can result, for example, from missing data (not measured because of power or network failure, vehicle or driving issues, calibration, etc) or data flagged as invalid (data excluded from further analysis during the data verification and validation steps, Section 6.0).

5.0 Quality Control Procedures

At Aclima, quality control begins with initial assembly of the AMN and initial testing of all sensors that will be installed into the mobile platform (Section 5.1) and extends through to final data verification (Section 6.0). Aclima staff meet weekly to assess vehicle and sensor performance, data quality, and status of the mapping effort. Automatic Status Indicator flags, which flag data outside specified Status Indicator limits, are established in the database before Aclima staff manually review data. Problems or irregularities are logged, investigated, and suitable corrective actions implemented as soon as possible. This section outlines the procedures implemented by Aclima as part of the quality control efforts. These processes underpin the technical success of the measurements and mapping campaign and help to guarantee Aclima's data are of the highest quality possible, providing critical information to effectively address broader community-level objectives and outputs.

5.1 Evaluation during AMN production

Detailed testing of sensors embedded in the AMN occurs during production and after the AMN is fully built. Checks include verifying voltages, communication, programming, flow rate, leak checks, and a series of visual inspections that include, for example, verifying labels, batteries are installed, and all fasteners are present and tightened. Once AMN production is complete, the checks are repeated on the entire AMN. Sensors external to the AMN also undergo detailed testing. After final verification of the AMN's operation, the AMN and other sensors that are not built into the AMN ($PM_{2.5}$, BC, CH_4 and C_2H_6, and air toxics) are sent for calibration and or verification (pollutant dependent) (Section 5.2) at one of Aclima's location-specific Ambient Calibration Laboratories or location-specific Mobile Calibration Laboratories. All checks are tracked and results are logged into a tracking system designed for this purpose.

In support of AMN production and integration of other sensors into the mobile platform, a series of SOPs have been developed that provide detailed instructions for AMN assembly and installation of the AMN and external AMN sensors into the mobile platform.

5.2 Pre-deployment calibration

A completed AMN along with any external sensor units installed in Aclima's mobile monitoring platform undergo initial calibration for O_3, CO, CO_2, NO, NO_2, CH_4 and C_2H_6, and/or performance verification for $PM_{2.5}$ prior to deployment. Currently, the BC and air toxics sensors are used as provided by the manufacturer. The sensors installed in the mobile platform are tested and/or calibrated by one of three approaches, depending on the pollutant: 1) collocation with reference instruments in one of Aclima's location-specific Mobile Calibration Laboratories for O_3, NO, NO_2, CO_2; 2) peer-to-peer collocation in one of Aclima's location-specific Mobile Calibration Laboratories, between an ensemble of sensors of the same make and model as the test sensor

for $PM_{2.5}$, CO, and TVOC (TVOC after initial span and zero with reference gas in an Ambient Calibration Laboratory); and 3) calibration using NIST traceable calibration gases in one of Aclima's location-specific Ambient Calibration Laboratories for CH_4 and C_2H_6. These types of calibrations occur annually (at minimum) for each device.

With each of these pre-deployment calibration/verification processes, the performance metrics are quantified from the "test" time period and compared against Aclima's Measurement Quality Objectives or acceptance criteria (Table 4). If a sensor does not meet the stated acceptance criteria, then individual sensors undergo a thorough evaluation to determine the root cause of the issue, replacing the affected sensor if necessary. Regardless of the cause of the problem, sensors must then undergo a repeat calibration and/or verification. Acceptance criteria for all sensors within an AMN as well as those not built into the AMN must be achieved before a sensor is installed into a vehicle for deployment.

These calibration approaches are used to derive and test sensor model input parameters prior to deployment. Sensor models are used to convert the sensor signal (e.g., voltage) into a physically meaningful measurement ($\mu g/m^3$, ppb, ppm). The sensor model may also include T, RH, pressure, and other pollutants that may interfere or be related to the measurement of interest.

Where possible, deployment of Mobile Calibration Laboratories in the same areas where data collection is occuring in order to calibrate under similar environmental conditions as where the AMNs will be deployed. However, this is not always possible and assumptions must be made about performance of the AMNs under different conditions. As part of our data validation processes, we use regulatory reference station data to support mobile-to-stationary comparisons to evaluate and correct for bias in our measurements that may result from differences in environmental conditions from those experienced during calibration, which is outlined in Section 5.0 of the Hyperlocal Ambient Concentration Estimate Validation and Quality Assurance document. This is particularly important for PM2.5 because chemical composition and size distribution of the particles often vary between geographic regions, and both affect the measurement reported by the sensor.

5.2.1 Collocation with Reference Instruments (O_3, NO, NO_2, and CO_2)

Pre-deployment calibration of the O_3, NO, NO_2, and CO_2 sensors is achieved by collocating sensors with reference methods in one of Aclima's location-specific Mobile Calibration Laboratories. This ensures that these sensors are calibrated over a large dynamic range of analyte concentrations and measured under atmospherically relevant on-road conditions, including variations in pollutant gases and particle concentrations, and varying meteorological conditions. In the mobile calibration laboratory, sensor and reference instrument inlets are situated near each other, essentially measuring pollutants and greenhouse gases from the same air volume. The collocation period for sensor calibration and verification is one week

(calibration approaches 1 and 2) and includes a calibration "training" period (4 days) to derive input parameters for the sensor calibration models, followed by a "testing" period (typically 3 days) to verify the input parameters and quantify sensor performance. The performance metrics derived during "testing" are considered to be representative of likely sensor performance when sensors are deployed for Baseline data collection. Regression analysis is applied to determine the gain (slope) and offset (intercept) for use in the sensor calibration model. While the regression is usually linear, it also may be based on a nonlinear function depending on the pollutant and sensor. Collocation ensures that all sensors are providing similar measurements within the expected uncertainty of the measurement and after the calibration model is applied it allows for an estimate of precision and bias relative to the reference method. If sensors meet acceptance criteria, they are installed and used in Aclima's mobile platform. If not, the calibration process is repeated.

The reference instruments used for the calibration of the O_3, NO, NO_2, and CO_2 sensors (Table 2) undergo monthly zero and span checks to confirm that the calibration is stable. Periodically or as needed, for example, if span and zero checks are outside performance specifications (Table 3), the zero and span of the sensor are changed based on manufacturer protocol or they are periodically (yearly unless warranted) returned to the vendor for recalibration.

5.2.2 Peer-to-peer collocation ($PM_{2.5}$, CO, and TVOC)

$PM_{2.5}$, CO, and TVOC (TVOC after initial span and zero with reference gas in an Ambient Calibration Laboratory, see section 5.2.4 for details) sensor performance and sensor-versus-sensor compatibility are quantified pre-deployment by collocating sensors in one of Aclima's location-specific Mobile Calibration Laboratories. During collocation periods, each test sensor is compared to an ensemble of other collocated sensors of the same make and model (minimum of 3, although >10 are possible). Inlets are in close proximity so they are sampling from a similar air parcel. The comparison is analogous to the calibration "testing" period for the sensor-versus-reference collocation, but instead of using a reference instrument, the average of the ensemble of collocated sensors is used as a relative reference for comparison to the test sensor.

$PM_{2.5}$ measurements are used as provided and only the flow rate is evaluated. If the flow rate of a $PM_{2.5}$ sensor does not fall within the expected error threshold of ±5%, the flow rate is calibrated according to manufacturer protocols. No other adjustments are applied for the $PM_{2.5}$ sensors – if a $PM_{2.5}$ sensor does not meet the acceptance criteria it is replaced. Peer-to-peer collocation ensures that all sensors are providing similar measurements within the expected uncertainty of the measurement as well as precision and bias. A discussion of accounting for systematic bias that may result from treating $PM_{2.5}$ and CO in this way is included in the accompanying "Hyperlocal Ambient Concentration Estimate Quality Assurance System" document.

5.2.3. Black Carbon

The BC sensor uses a well accepted detection principle (aethalometry) and as a result the sensor factory calibrations are used as is. As needed or annually, the BC sensor is sent back to the manufacturer for calibration and maintenance. The flow rate of each BC sensor is routinely evaluated, and if it does not fall within the expected error threshold of ±2%, the flow rate is calibrated according to manufacturer recommendations. Currently, BC sensors are not routinely placed in the Mobile Calibration Laboratory for peer-to-peer comparison as part of our standard operating procedures, although periodic collocations in the Ambient Calibration Laboratory have been conducted to confirm agreement among the BC sensors (see Section 5.5 for details). The sensor is operated according to manufacture specifications to ensure no significant impacts from loading effects. No additional processing is performed to smooth the signal.

5.2.4 Reference Gas Calibrations (CH_4 and C_2H_6, TVOC, and Air Toxics)

The CH_4 and C_2H_6 sensor calibrations are performed using NIST traceable reference gases of accurately known CH_4 and C_2H_6 concentrations. The CH_4 and C_2H_6 sensor offset and gain are adjusted if needed using a low target at approximately ambient concentrations of 2 ppm CH_4 + 1 ppb C_2H_6 and a high target at approximately 50 ppm CH_4 + 5 ppm C_2H_6, and then tested using a mid-range target at 5 ppm CH_4 + 50 ppb C_2H_6. If, after calibration, the difference between the test tank concentration and that reported by the sensor instrument is less than ±50 ppb CH_4 or ±15 ppb C_2H_6 or if the standard deviation of the 1 second data during the low concentration reference measurement is less than 10 ppb, then the acceptance criteria (Table 4) are met and the sensor can be deployed. If not met, then the calibration process is repeated as needed or the sensor is replaced. The current best estimate for total uncertainty of the reference gas concentrations are the larger of 25 ppb or 1% for CH_4 and 5 ppb or 0.2% for C_2H_6. This is based on total uncertainty of our gold standard reference gases obtained from NOAAs Central Calibration Laboratory and repeatability of characterization experiments of the NIST certified working or transfer standards.

TVOC sensors in the AMN undergo an initial pre-deployment calibration based on a single span using a NIST certified reference cylinder of isobutylene in the concentration range of ~1 ppm (range 0.68 to 1 ppm) and a zero check. The zero and span of the TVOC sensor are not checked in the field. The initial calibration is followed by peer-to-peer collocation in a Mobile Calibration Laboratory where the slope and R^2 values for bias and precision are estimated from the relative reference, similar to $PM_{2.5}$ and CO.

The UV-DOAS Air Toxics instrument undergoes a zero and multi-point calibration check using a NIST traceable calibration gas (containing benzene, toluene, ethyl benzene, and xylene) diluted with zero air. The process occurs on an ongoing basis over the course of the deployment. The Air Toxics vehicles typically operate out of one of Aclima's calibration laboratories and can receive calibration checks more frequently (up to weekly) than devices in the full collection

fleet. If during one of these calibration checks, either the precision or bias acceptance criteria (see Table 4) are not met, the instrument is serviced according to the user manual and with guidance from the manufacturer.

5.3 Measurement verification during deployment

During deployments, Aclima engineers and technical staff are notified of potential sensor performance issues based on the following: 1) notifications from drivers or hardware technicians as they perform their routine system checks; 2) as a result of automated flagging within Aclima's data processing pipeline; and 3) during regular (weekly) device-level reviews by data review analysts.

Automated status indicator flags signal when measurements are collected outside the environmental or physical specifications of the sensors, such as values greater than or less than the acceptable concentration range of the sensor or negative values below the uncertainty of the detection limit, data reported during quality control checks (e.g., zero and span checks), or periods where the flow rate is out of the optimal range. In many cases, status indicator flags are part of normal operation of certain sensors, for example when temperature is changing rapidly at warm-up. In other cases, excessive data flagging can be an indication of a hardware issue that requires a corrective action (i.e. low flow rates due to a failing pump or blocked flow path). In addition to these automated flags, manual data review is conducted on a weekly basis over all deployed sensors. During these weekly reviews, up to six weeks of previously collected measurements can be examined at one time to allow for observation of trends, e.g., drift or erratic behavior, or other behaviors that are not currently automatically flagged. Analysts have access to visualizations of time series data at different temporal aggregation scales of multiple devices deployed in the same geographic area to visually scan for anomalous behavior. Correlations between different pollutants are also visualized, which can provide useful context for understanding whether a particular sensor is behaving abnormally (CO vs CO_2, for example). Diagnostic variables such as temperature, pressure, relative humidity, and flow are also visualized.

For deployments where there are active NAAQS sites established by regulators in the mapping area, data review can rely on data from these sites for contextualizing large scale trends in different air pollutants over time. This is a useful tool for understanding whether features in Aclima's sensor data are anomalous or consistent with trends detected by the regulatory network. It is often possible to detect systematic bias across the network or identify individual sensors that have drifted substantially between their pre- and post-deployment calibration checks (Section 5.4). There are a number of different factors that can impact these comparisons that must be taken into account, including distance between mobile and stationary measurements, road type, site type (i.e. near-road vs neighborhood scale vs background site), and temporal aggregation. In some cases the type of measurement method

used at the site can be important, for example different FEM methods for $PM_{2.5}$ can be biased in different ways in relation to the FRM gravitational measurements.

In addition, customized alerts can be set by the technical team to alert for specific conditions that may indicate atypical behavior for a given sensor that needs to be resolved. For example, prolonged periods of elevated CO_2 may indicate possible sampling from inside the vehicle. This can alert the team to potential issues to trigger an intervention in a more efficient way than the standard weekly data review. These customized alerts do not automatically flag data, but trigger a manual review by an analyst who makes the final decision as to whether data needs to be reviewed or not.

When issues are detected, they are documented and tracked through an internal ticketing system. Each ticket is tied to a specific vehicle, sensor serial number, and AMN ID and contains a description of the symptom, which sensors are impacted, recommendation for intervention, description of how the issue was resolved, and the date resolved. If any manual omissions of data are required as a result of any issue, this is noted in the ticket in a manner that provides traceability in Aclima's database back to the reason for omission.

Some issues can be addressed remotely (i.e. firmware update or settings modification), but most issues rely on a physical intervention by either the driver or a trained hardware technician. Once identified, issues are resolved as soon as possible. When a notification is received, Aclima engineers and technical staff work with drivers to perform simpler repairs in the field (Section 3.2), such as a loose wire or blocked tubing. If fixing the issue in the field is not possible or if in-field service impacts performance, the sensor (AMN if the sensor is within the AMN) is returned to one of Aclima's location-specific Ambient Calibration Laboratories for service and repair or swapped for a newly calibrated sensor (or AMN). The returned device is fixed, recalibrated, and re-deployed when needed.

Issues tracked through this system are limited to those that impact data quality. The drivers and fleet operations team are responsible for routine maintenance of the vehicle itself, including monitoring and resolving any issues.

5.4 Post-deployment recalibration

After field deployment, mobile platforms are returned to one of Aclima's location-specific Ambient Calibration Laboratories. Each sensor is recalibrated using the same process as described for pre-deployment calibration (Section 5.2). The same acceptance criteria are applied (Table 4). If sensor performance does not meet the acceptance criteria, the data are adjusted based on the pre- and post-deployment calibrations and how the sensor calibration may have changed. This post-deployment calibration check occurs every time a node or the other sensors (CH_4/C_2H_6 and BC) are removed from the vehicle, and, at minimum, annually.

The procedure for handling calibration drift between pre- and post-deployment calibrations varies depending on the data product. For the Ambient Concentration Estimates, drift in both the offset and gain is important and can impact the bias in the data product. For the air toxics product, which is based on detecting short-lived enhancements (on the order of seconds) above the baseline, any drift in the baseline is taken into account by the difference calculation used to estimate the enhancement (or delta) value. Therefore, the air toxics data product is, by nature, only sensitive to drift in the gain of the measurement and not drift in the offset. This is discussed further in Section 5.5.

5.4.1 Drift correction: Ambient Concentration Estimates Data Product

If the acceptance criteria are not met and a calibration adjustment is required, Aclima staff review the pre- and post-deployment calibrations/verifications with other collocated data available during deployment. Collocated data may include data from stationary regulatory sites or sensors in other Aclima deployed mobile platforms. This evaluation is used to determine whether the sensor calibration change occurred as a step change or drifted over time. Once determined, appropriate calibration adjustments are applied. Certain sensor types are more prone to drift over time (e.g., CO_2 and CH_4) and a linear calibration function is applied based on the pre- and post-deployment calibrations, whereas other sensors (e.g., NO_2) have been found to occasionally experience fast step-changes in calibration during deployment. Once the date and time of the step change is identified, the sensor model parameters are recalculated after that period, applied, and revised concentrations are updated for the period after the step change. Where there is no obvious shift or if other collocated data are not available, the drift is assumed to be linear over time between the pre- and post- deployment calibrations.

Once determined, appropriate calibration adjustments are applied. If the newly applied calibration values allow the sensor to meet the acceptance criteria (Table 4), then, at this point of the review process, the data undergo a final review. If not, the issue is investigated in more detail and data from that sensor may be excluded from further analysis. All changes are tracked as a metric in a built-for-purpose database that documents adjustments identified after post-deployment calibration. This table contains calibration records by device. A history of calibration values and their corresponding date ranges are stored as metadata with the 1-Hz data as well as a record of the value used for each time stamp.

5.4.2 Drift correction: Air Toxics Data Product

For the Air Toxics instrument, calibrations are generally performed at a higher frequency (typically weekly) than for the sensors in the larger collection fleet. On these timescales, drift in the gain of the instrument is small. If the slope of the multi-point calibration drifts substantially, it likely indicates that the instrument has malfunctioned and will require manual omission of data for other reasons. The TVOC sensor may drift outside of the acceptable range based on the MQOs in Table 4; however, since the TVOC sensor is only partially quantitative, given its wide range of sensitivities to different VOC species (over 2 orders of magnitude differences), the TVOC sensor calibration values are typically not adjusted after the fact. It is assumed that

other uncertainties are much higher than the sensitivity of the sensor. If the precision criteria fails during a post-deployment calibration check, indicating a failed sensor, then manual omissions are applied starting at the likely time of failure.

5.4.3 Systematic Bias: Ambient Concentration Estimates Data Product

Systematic bias can be defined as the difference between Aclima's ambient concentration measurements and measurement that can be considered a suitable representation of the true concentration. This bias is always a possibility when comparing two different measurement networks that may use different primary standards, detection methods, and levels of data quality. Sensors in the platform that are not routinely calibrated using a reference instrument (CO, PM2.5, and Black Carbon) may be particularly prone to systematic bias. We use regulatory reference station data to perform mobile-to-stationary comparisons to evaluate and, if necessary, to correct for bias in our measurements. More detail on this process can be found in Section 5.0 of the Hyperlocal Ambient Concentration Estimate Validation and Quality Assurance document.

5.5 In-field sensor performance

The performance of all Aclima sensors that have been deployed have been evaluated for sensor bias, precision, limit of detection, and drift using different approaches (i.e. performance metric) depending on the relevant data products associated with each sensor type. This analysis, performed over many devices over 3 years of calibration events, indicates typical performance expected for each sensor while deployed in vehicles and mapping.

The Ambient Concentration data product aims to describe typical ambient concentrations at a given location (segment or group of adjacent segments, e.g., within a census tract). For this data product, it is important to characterize sensor bias and drift at typical ambient concentrations and to estimate precision and limits of detection at relevant temporal scales (e.g., seconds to minutes).

The relevant approaches are different for the Air Toxics data products than for the ambient concentration data product. The Air Toxics products aim to detect and localize large enhancements (with respect to the baseline) in pollutant concentrations over short time periods (typically as short as a few seconds). Sensor bias is important only as it relates to measurements of the magnitude of these enhancements above the ambient baseline rather than in the absolute magnitude of the ambient baseline. Since this product relies on detecting concentration differences from the same sensor, bias and drift in the sensor gain are more important than bias and drift in the sensor offset. Precision for these pollutants is characterized at 1-Hz defining the lower limit at which a fast enhancement (~seconds) can reliably be discerned from random noise. Baseline instability and sensor drift can impact precision and bias at longer timescales (minutes to hours and up) but can be largely ignored for the sensors exclusively used in the air toxics data products.

The following tables contain results for the DQIs for each sensor, representing the device-level uncertainty contributions to our two data products (Tables 5 and 6). The DQIs used to describe sensor performance are: bias, drift, precision, and limit of detection. The metrics and experiments used to quantify these indicators vary by how they are calculated, sensor, and data product, and these details are given below for each data product. The following provides a general description of each of the DQIs as applied to Aclima's two current data products:

- Ambient Measurement Bias: Typical errors in accuracy for sensors immediately following a calibration event and evaluated for pre-deployment calibration checks.

- Ambient Measurement Drift: Typical errors in accuracy for sensors at the end of their deployment relative to pre-deployment evaluations.

- Air Toxics Product Measurement Drift:
    - For TVOCs: Typical errors in accuracy at full scale (as a percent) at the end of their deployment relative to pre-deployment evaluations
    - For Air Toxics Instrument (BTEX): Standard deviation of the derived linear slope from repeat multi-point calibrations, performed periodically over the course of a deployment.

- Ambient Measurement Precision: Random differences in agreement between two sensors of the same type at typical ambient concentrations, defined at specific time aggregations and evaluated for pre- and post-deployments.

- Limit of detection
    - Ambient concentration data product: Random differences in agreement between two collocated sensors of the same type evaluated at ambient background concentrations (as defined by the lowest 15th percentile concentrations and below) and calculated as the standard deviation (2σ) of differences between the two sensors. Note: These calculations are derived from ambient data and not from traditional laboratory evaluations with zero air. For pollutants that typically do not reach zero concentration in the atmosphere (e.g. CO_2, CH_4, CO, O_3), these values should be interpreted as the minimum detectable values above typical background concentrations.
    - Air toxics data product: Minimum detectable peak above background at a 1-Hz sampling frequency determined as the standard deviation (3σ) of reported values while sampling a constant concentration reference gas.

Table 5: Ambient concentration data product - Typical In-field Performance[1]

Pollutant	Bias[2]	Drift	Precision			Limit of Detection (2σ)		
			1 s	8 s	170 s	1 s	8 s	170 s
NO (ppb)[3]	19%	22%	9.7	4.0	1.0	15.6	9.6	2.2
NO_2 (ppb)	2.4	4.6	4	3.1	2.1	6.4	5.2	3.4
O_3 (ppb)	1.5	1.5	1.8	1.8	1.2	3	2.8	1.8
CO (ppm)	0.03	0.03	0.028	0.026	0.023	0.028	0.02	0.012
CO_2 (ppm)	5.5	18.1	3.4	3.3	3	2.4	2.2	2
CH_4 (ppb)	21	41	37	34	27	68	74	60
C_2H_6 (ppb)[4]	2	4	2.5	2.4	2.3	5.2	5.0	2.8
$PM_{2.5}$ (ug/m3)	0.6	0.6	1.4	0.9	0.5	1.6	1.0	0.4
BC (ug/m3)	0.23	TBD	1.7	0.6	0.26	2.8	1.2	0.72

1. These values are our best estimates to-date of "typical" sensor performance and may change in the future as additional data are evaluated. Typical is defined as the median value across all sensors tested, and is not meant to be a performance specification.

2. This bias does not account for systematic bias that can arise from differences in sensor response when mapping in different environmental conditions from those occurring during calibration. See Section 5.2 in the Hyperlocal Ambient Concentration Estimate Validation and Quality Assurance System document.

3. For the NO sensor: as a result of how data is processed from this sensor, bias and drift are only expected to manifest in the gain of the sensor, so values are represented as a percent based on errors in the slope vs reference.

4. See description of limitations of the ethane sensor in Section 5.6.1

The temporal aggregation used to characterize precision and limit of detection impacts the calculated value and is specified for each case. As is typically seen, lower values (indicating higher precision and improved LOD) are observed at longer integration times. For the ambient concentration data product, the relevant time scales for aggregation are at 1 second (corresponding to the sensor sampling rate), 8 seconds (which is the typical single pass duration for 100 m segments), and 170 seconds (which is the typical number of seconds aggregated into a segment with 20 passes). The precision and LOD estimates at 170 seconds, therefore is an approximation of the contribution of random sensor precision to the overall uncertainty in our segment level aggregations, but does not include other sources of uncertainty such as drift or sensor-to-sensor variability across all 20 passes, which are collected non-contiguously over the duration of the Ambient Concentration Estimate collection period. As described in the Hyperlocal Ambient Concentration Estimate Validation and Quality Assurance System document, each segment in the Ambient Concentration Estimate receives a

confidence interval that better describes the true precision uncertainty, which includes a combination of measurement uncertainty and atmospheric variability over time.

Table 6:   Air toxics product In-field Performance

Pollutant	Drift	Limit of Detection (3σ)
TVOC (ppb)	32%	0.6
Benzene (ppb)	10%[1]	11
Toluene (ppb)	10%[1]	18
Ethyl Benzene (ppb)	10%[1]	24
m-Xylene (ppb)	10%[1]	12

1. Air Toxics instrument (UV-DOAS) drift determination is still in process: manufacturer's specification is 10%

5.6 Important Limitations for Specific Sensors

5.6.1 Ethane (C_2H_6)

Ethane measurements can exhibit significant baseline instability on hourly time scales and the degree of drift can vary quite a bit between individual sensors. The effects of this baseline drift are not represented in the reported "typical" precision or accuracy metrics (Section 5.5) . These artifacts are significant relative to ambient concentrations of ethane, thus, Aclima does not recommend using C_2H_6 in the typical use cases for the ambient concentration data product. The ethane data is useful where CH_4 enhancements are high (~above the reported precision limits) in order to determine whether the source of methane is biogenic or thermogenic in nature.

5.6.2 PM2.5

The PM2.5 sensor operates based on sizing and counting particles in six distinct channels between 0.3 and 2.5 μm in diameter and converting to mass using a sensor model that includes assumptions about particle density, optical properties, and shape. These parameters are known to vary depending on the size distribution and chemical composition of the aerosol, resulting in differing optimal sensor model parameters that are location-specific. As part of deployments in new geographies, we evaluate any potential systematic bias as a whole, rather than at the individual sensor level, using comparisons to local regulatory monitors where available. The accompanying "Hyperlocal Ambient Concentration Estimate Quality Assurance System" document discusses how this is handled in more detail.

5.6.3 Nitric Oxide (NO)

The NO sensor is prone to baseline drift, which we account for by performing a baseline adjustment as part of data processing. As a result, the final measurements can obscure some of

the larger spatial and longer temporal trends in ambient NO concentrations. Additionally, the detection limit for the NO sensor is high (at 1 second and 8 second measurement frequency), relative to typical ambient concentrations under clean conditions. Therefore, the NO measurements primarily indicate locations and times when NO is present at high concentrations and close to sources, resulting in sharp spikes in the time series on timescales of seconds to minutes.

5.6.4 Nitrogen Dioxide (NO_2)

The NO_2 sensor requires the use of the O_3 sensor in deriving an NO_2 measurement because the sensor is sensitive to both NO_2 and O_3. As a result, bias in either the O_3 sensor or the NO_2 sensor can impact the derived NO_2 measurement and the magnitude of error can vary as the relative concentrations of the two pollutants vary. The uncertainty values reported in Table 5 reflect the uncertainty from both sensors and are a realistic description of uncertainty during deployments resulting from this source of error.

5.6.5 Total Volatile Organic Compounds (TVOC)

The TVOC sensor operates on the principle of photo-ionization, which permits detection of most VOCs but does not report information on the specific compounds detected. While the TVOC sensor is sensitive to a wide range of VOCs, the sensor has variable sensitivity to different classes of VOCs spanning multiple orders of magnitude. The sensor is calibrated using isobutylene, so the units reported actually represent isobutylene-equivalent concentrations. The response factors for many compounds relative to isobutylene, can vary significantly; for example, benzene has a response factor of about 2x that of isobutylene while ethylene has a response of about 0.1x that of isobutylene.

The TVOC sensor is also prone to baseline drift on various timescales. However, rather than attempting to approximate ambient concentrations the air toxics data product uses enhancements (on the order of seconds to minutes) in the signal with respect to the baseline to detect localized hotspots. For this reason, baseline drift is not a critical aspect of the sensor's performance for its intended use case.

The acceptance criteria for r^2 and slope are more tolerant for the VOC sensors compared to CO and PM. In part this is due to baseline instability between different sensors, but there are also likely differences in the response factors to different VOCs from sensor to sensor. As a result, when sampling real-world mixtures of ambient VOCs, the agreement from sensor to sensor can vary depending on the exact mixture, and therefore the sensor-to-sensor variability during deployments can be large (-50% to +100%) in some cases. Typical sensor agreement is on the order of +/-30% (as shown in Table 6).

5.6.5 Air Toxics Instrument (BTEX)

The Air Toxics instrument operates on the principle of UV differential optical absorption spectroscopy. The primary limitation is that certain molecules that absorb strongly in the UV

spectrum can interfere with the detection of more weakly absorbing molecules. The most commonly encountered example of this during normal on-road mobile sampling is from nitric oxide (NO). As a result, when NO is present, it likely inhibits the measured enhancement of certain BTEX species.

Baseline drift can also be significant for this instrument, especially during warm up and during periods of unstable temperature in the vehicle. However, as with the TVOC sensor, the air toxics data product uses enhancements (on the order of seconds to minutes) in the signal with respect to the baseline to detect localized hotspots, rather than attempting to approximate ambient concentrations. For this reason, baseline drift is not a critical aspect of this sensor's performance for the intended use case.

6.0 Data Stages

After raw data from Aclima's sensors are initially ingested, they go through a data processing pipeline that ultimately transforms the collected data into our final data products. While the raw data will never be directly modified or removed in the database, modifications are often made to historical data at various stages throughout the processing pipeline as part of our standard QA operations in order to ensure the highest possible data quality in our final products.

In order to communicate to the user when the underlying 1 Hz data (0.5 Hz for Ozone) are final and will not undergo further changes, Aclima uses a simple data stage system that identifies each individual 1 Hz data point by one of two possible stages: (1) Preliminary (data are likely to change) and (2) Verified (data should be considered final). The data stage for each 1 Hz measurement is propagated downstream into the data products that aggregate the 1 Hz data. For this reason, different data products can be constructed from a combination of preliminary and verified data. Each data product that aggregates this 1 Hz data remains in the Preliminary stage until 100% of the 1 Hz data used is labeled as verified. Note that the data stages only apply to processed data and not the raw data as raw data in the ingestion database are never modified or removed.

Preliminary data are available within 24 hours of collection and have gone through a number of data processing steps. The first step is the conversion from raw measurements into calibrated physical quantities representing pollutant concentrations. Calibration input parameters derived from the most recent pre-deployment calibration are used in this stage. Next, the data go through an automated status indicator flagging step based on predetermined thresholds of concentrations or associated diagnostic variables designed to identify and remove as much poor-quality data as possible without requiring manual intervention. Status indicator flags indicate two basic categories: yellow or red. A yellow flag indicates data that are considered useful but likely have a slightly higher uncertainty, such as if the measured concentration range is outside the optimum concentration range of the sensor. Currently, the only instance where a yellow flag is applied to Aclima data is when the PM sensor is measuring above the

manufacturer-specific range. A red flag indicates that data are not valid and are excluded from further analysis but are maintained in the raw data archive. Red flags are applied as a result of standard sensor operations (e.g. during zero checks of the PM sensor), during sensor warm-up periods, or if there was an obvious problem where the integrity of the flow system (versus some slight drift in the pump speed, i.e., within 15% of nominal).

In the Preliminary stage, data are routinely modified and should be considered likely to change. These modifications typically fall into two categories: 1) data are removed as a result of manual omissions stemming from issues found during the routine data review process, or 2) calibration input parameter updates as a result of sensor drift as detected during post-deployment calibration checks. Occasionally, changes to status indicator thresholds or modifications to the underlying sensor calibration model may occur as well, but this is less common and not part of routine operations.

When sensors undergo post-deployment calibration checks, this triggers a manual review of all data collected by that sensor between its pre- and post-deployment calibration. This manual review is similar to the routine weekly mid-deployment review (Section 5.3), but the analyst views the entire deployment at once, rather than only the most recent week's data. Additionally, the analyst reviews additional metadata associated with the deployment, including records of all issues detected during the deployment, any relevant findings from post-deployment servicing that might impact data quality, post-deployment calibration results, and any drift corrections applied.

The reviewer's responsibility is to make sure that all required manual omissions and calibration adjustments have been identified and applied. Once this has been ensured, the reviewer marks that sensor as Verified in a tracking system. Each verification mark is applied to a specific sensor and vehicle with start and end dates associated with the pre- and post-calibration events, respectively. Once this occurs, each 1 Hz data point between the start and end dates from that sensor is marked as Verified in the database. Once data have been marked as Verified they are not expected to undergo any further changes. Data products that use verified 1 Hz data may undergo revisions, changing how the underlying 1 Hz data are processed and aggregated, and these changes are tracked and communicated by software versioning, however, the underlying verified data used as inputs to the data products will not change.

7.0 Data Management

Raw mobile sensor measurements are recorded each 1-second (2-seconds for O_3) along with an associated date, time, and location of each measurement as well as other critical diagnostic parameters at 1-second, including sensor specific T, RH, P, latitude and longitude (GPS), and flow rate. The latter parameters are used to verify sensor operations and include data needed for the sensor calibration models or sensor verification (T, P, RH).

Data from the mobile platform is transmitted every minute through integrated communications software to Aclima's cloud-based backend. Transmission is primarily through either an internal LTE module or by WiFi. During times when LTE connectivity is poor, data messages are queued on the device and sent when connectivity returns. The AMN contains enough flash storage to store up to a few weeks of 1-second data, including data obtained by the BC, CH_4, C_2H_6, and air toxics sensors as well as diagnostic and platform variables. All raw sensor readings, location, and diagnostic data are packaged into coded messages with an associated network timestamp based on a real-time clock linked to the AMN through a NTP. On a nightly basis, the data is further processed and the data products are stored in the cloud-based data store in addition to a permanent archive of the raw data. Raw data are never removed from this initial backend data store. All appropriate diagnostic variables and automatically-generated data quality flags for subsequent data review stages and application processing are included in the cloud-based storage. Metadata including sensor and AMN identifiers, units of measurement, and other relevant information is also associated with each measurement. Sensors are tracked throughout deployment by serial number, parameter, model, and start and stop time of operation. These results are stored in the Google Cloud Platform. This platform allows for storage and subsequent access to the large data files through various protocols, such as BigQuery.

8.0 Citations

Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., and Hamburg, S. 2017. High-resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. Environ. Sci. Technol. 51, 6999–7008. https://doi.org/10.1021/acs.est.7b00891.

Chambliss, S.E., Preble, C.V., Caubel, J.J., Cados, T., Messier, K.P., Alvarez, R.A., LaFranchi, B., Lunden, M., Marshall, J.D., Szpiro, A.A., Kirchstetter, T.W. and Apte, J.S. 2020. Comparison of Mobile and Fixed-Site Black Carbon Measurements for High-Resolution Urban Pollution Mapping. Environ. Sci. Technol. 54, 13, 7848–7857. https://doi.org/10.1021/acs.est.0c01409.

EPA. 2002. Guidance for Quality Assurance Project Plans, EPA QA/G-5. U.S. Environmental Protection Agency, Washington, DC. https://www.epa.gov/sites/production/files/2015-06/documents/g5-final.pdf.

Federal Register. 2022. 40 CFR, Part 58, Appendix A, 4. Calculations for Data Quality Assessment. U.S. Environmental Protection Agency, Washington, DC. https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-58

Guan, Y., Johnson, M.C., Katzfull, M., Mannshardt, E., Messier, K.P., Reich, B.J., and Song, J.J. 2019. Fine-Scale Spatiotemporal Air Pollution Analysis Using Mobile Monitors on Google Street View Vehicles. Journal of the American Statistical Association. 115(531), 1111-1124. https://doi.org/10.1080/01621459.2019.1665526.

Messier, K.P., Chambliss, S.E., Gani, S., Alvarez, R., Brauer, M., Choi, J.J., Hamburg, S.P., Kerckhoffs, J., Lafranchi, B., Lunden, M.M., Marshall, J.D., Portier, C.J., Roy, A., Szpiro, A.A., Vermeulen, R.C.H., and Apte, J.S. 2018. Mapping Air Pollution with Google Street View Cars: Efficient Approaches with Mobile Monitoring and Land Use Regression. Environ. Sci. Technol. 52(21), 12563– 12572. https://doi.org/10.1021/acs.est.8b03395.

Seinfeld, J.H. and Pandis, S.N. (2016). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 3rd ed. John Wiley & Sons, Inc, Hoboken, NJ. https://www.wiley.com/en-us/Atmospheric+Chemistry+and+Physics:+From+Air+Pollution+to+Climate+Change,+3rd+Edition-p-9781118947401.