

Quality assurance of air quality data

Priscilla Adong
19th-September-2023



Why is this necessary;

Improve and increase of confidence levels in data and analyses

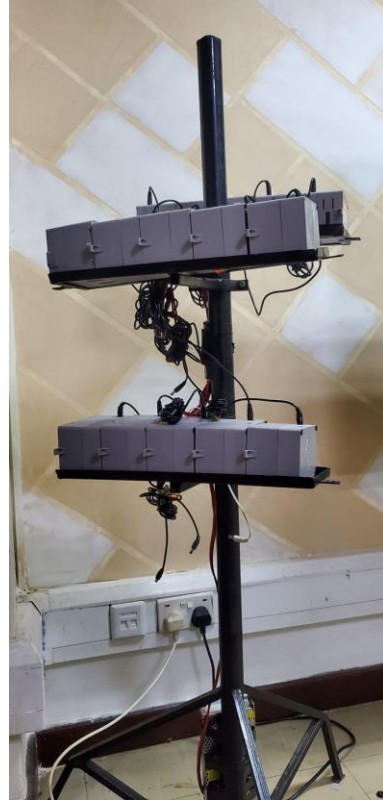
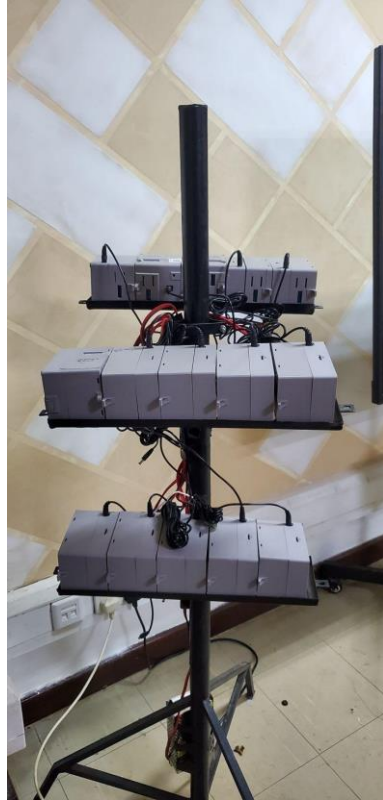
- Maintain accuracy
- Ensure consistency of measurements
- Minimise measurement errors
- Ensure data reliability



MAKERERE UNIVERSITY

Quality assurance Procedures

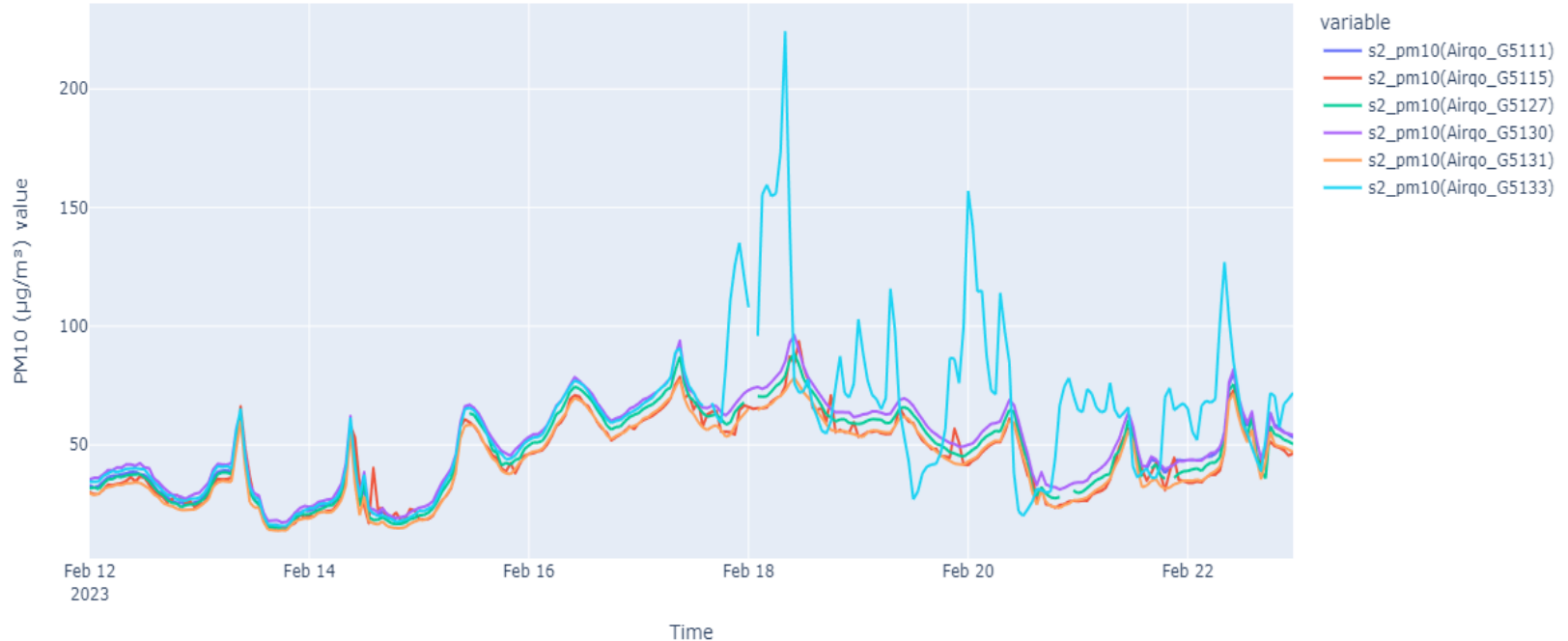
In Lab Colocation for AirQo monitors



- Inter Device **Performance evaluation** in lab setting.
- Evaluation of performance against organisation threshold
- Sensor1, Sensor2, internal and external temperature and humidity sensor performance review.

In Lab Colocation for LCS

PM10 Sensor 2 for the devices

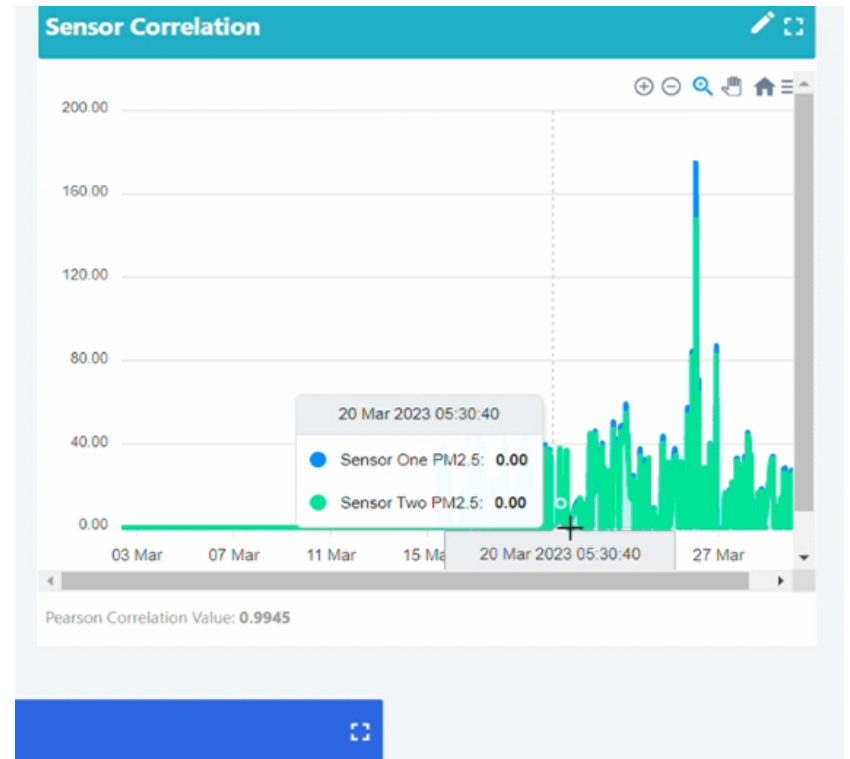


Field Colocation for LCS

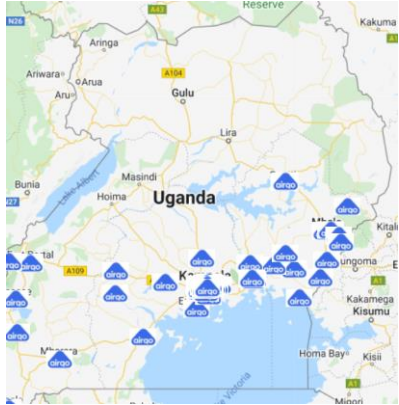


- Inter Device **Performance evaluation** in field setting.
- Sensor1, Sensor2, internal
- External temperature and humidity sensor performance review.
- Performance evaluation VS Met One BAM1022

Device monitoring and maintenance

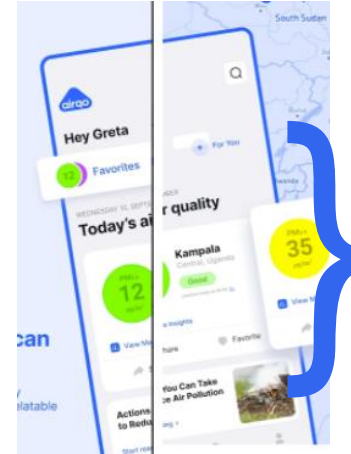


Overview of the Air quality data pipeline



Physical network

1. Custom sensor network monitoring
2. Data (pre)processing and storage
3. AI-powered calibration model
4. Analytics modelling: spatial and temporal modeling
5. Air quality data access: digital platforms and API



Public data access
through custom
digital platforms
and API

Data Sources and Collection Methods

- AirQo monitors
 - streamed via GSM network to the analytics platform
- Reference monitors (AirQo)
 - Streamed via GSM network to the analytics platform (Involves using the BAM logger)
 - Collected via flash disk
- Reference monitors (partners)
 - Streamed via GSM network to the analytics platform (UNEP)
 - Downloaded via CSV from partner websites (e.g. <https://www.airnow.gov>)

Data Sources and Collection Methods

- Low-cost monitors (partners)
 - Data integrated into the AirQo analytics platform (e.g Clarity devices installed by KCCA)
- Weather data
 - Streamed from Tahmo
 - Accessed via CSV from Met office
- Satellite data (potential)
 - Can be downloaded via various platforms including Google Earth Engine, official websites or code

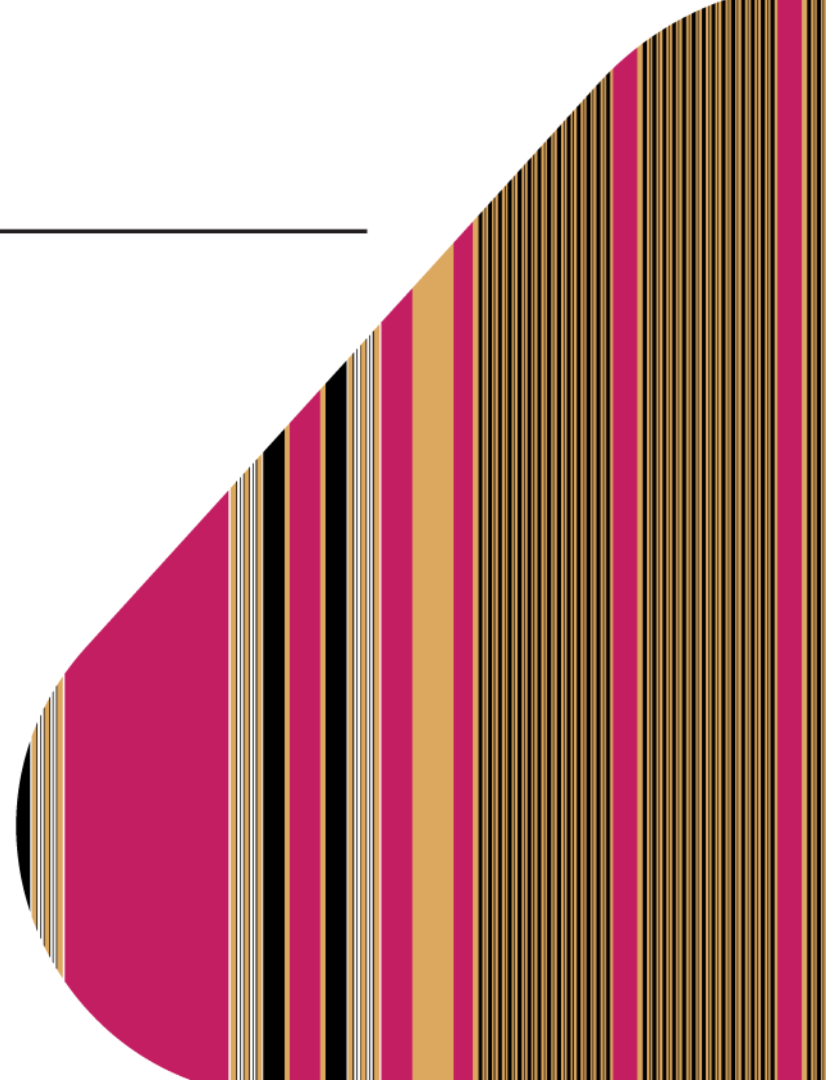
Preprocessing: data quality metrics

- Accuracy
 - Correctness of the data
- Completeness
 - Affects the outcomes & insights of data analysis
- Timeliness
 - Important for near realtime applications e.g mobile APP



MAKERERE UNIVERSITY

Field Calibration



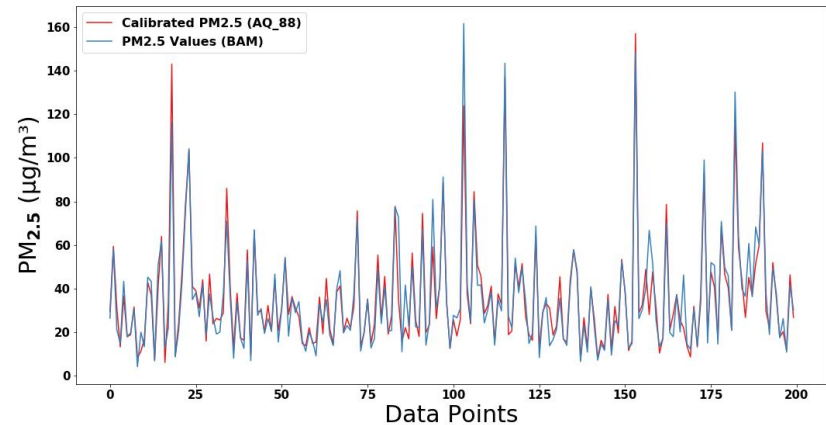
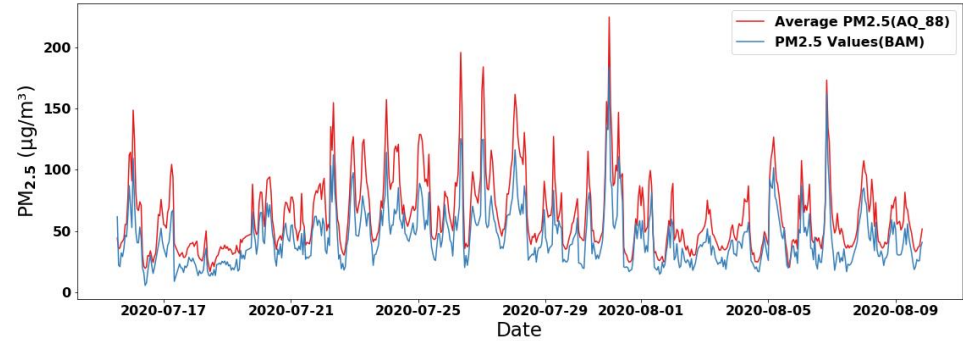
Field calibration infrastructure



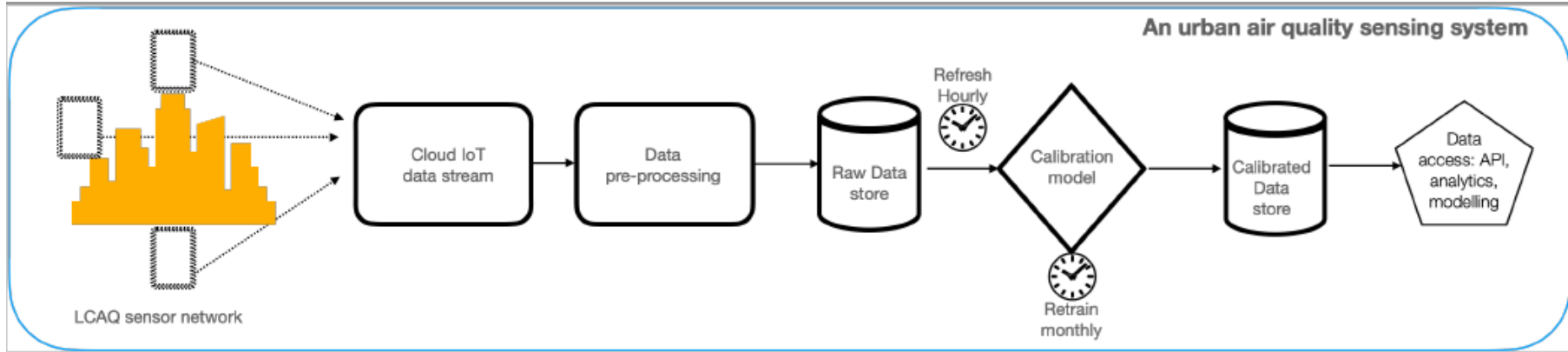
- Several calibration methods, we opted to use ML
- 2 permanent colocation sites in Kampala with BAM 1022 and AirQo monitors
- Co-location data used develop $PM_{2.5}$ and PM_{10} ML-based calibration methods

Performance against reference monitors

- Correlation values reaching up to 0.97 after calibration
- Cross-unit and cross site validation.
 - Uganda
 - Addis Ababa
 - Nairobi (Inprogress)
 - AirQo and Purple air devices
- Calibration applied to the entire network

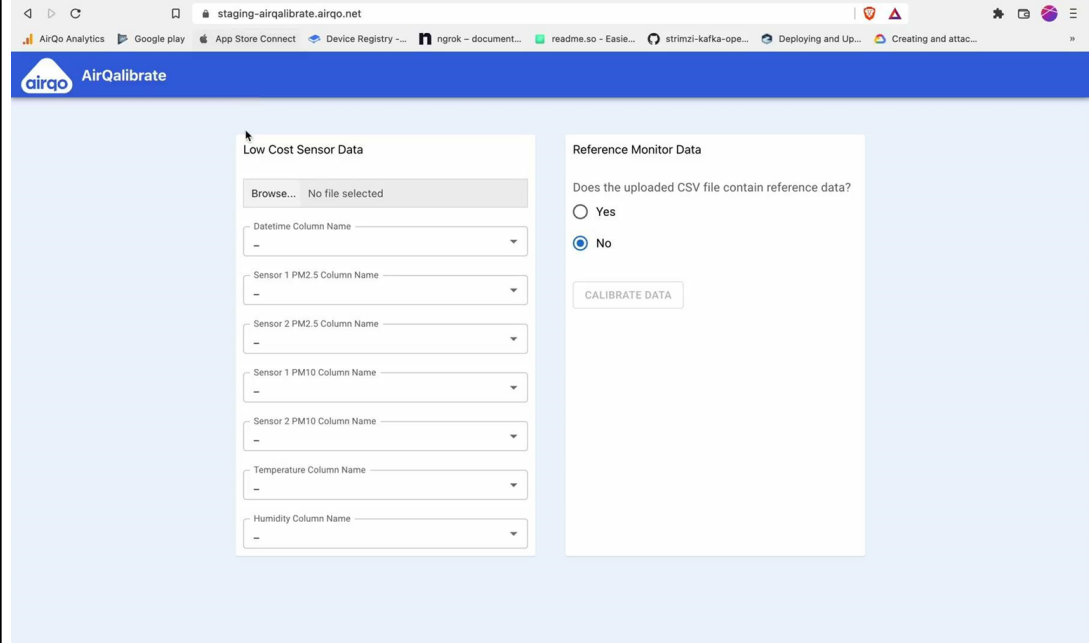


Field calibration pipeline



Calibration as a usable product

- A number of low-cost sensor initiatives emerging from the continent
- Data quality assurance still a challenge for individual networks
- Opportunity to leverage a data platform for quality assurance and data access



The screenshot shows the AirQalibrate web application interface. The browser address bar displays "staging-airqalibrate.airqo.net". The page header includes the AirQo logo and the text "AirQalibrate". The main content area is divided into two panels:

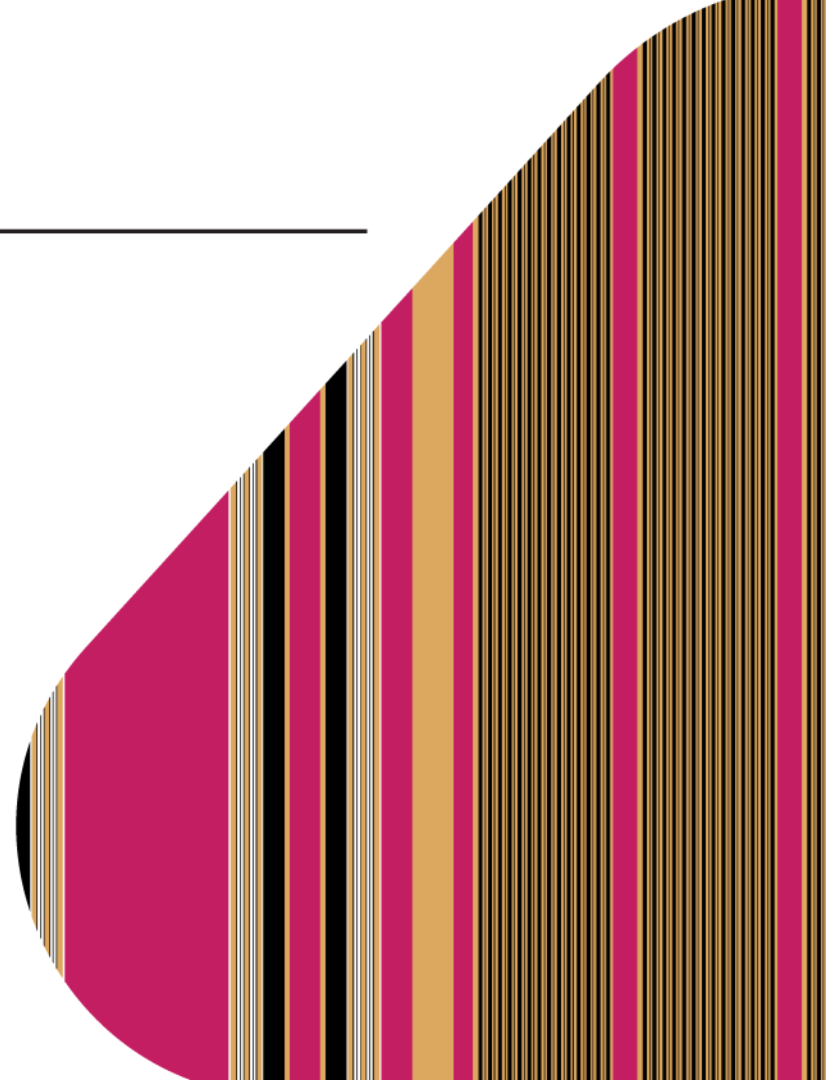
- Low Cost Sensor Data:** This panel contains a file upload section with a "Browse..." button and the text "No file selected". Below this are several dropdown menus for mapping columns: "Datetime Column Name", "Sensor 1 PM2.5 Column Name", "Sensor 2 PM2.5 Column Name", "Sensor 1 PM10 Column Name", "Sensor 2 PM10 Column Name", "Temperature Column Name", and "Humidity Column Name". Each dropdown menu currently shows a hyphen "-" as the selected option.
- Reference Monitor Data:** This panel asks "Does the uploaded CSV file contain reference data?". It has two radio button options: "Yes" and "No". The "No" option is selected. Below the radio buttons is a "CALIBRATE DATA" button.

<https://airqalibrate.airqo.net/>



MAKERERE UNIVERSITY

Lessons & Challenges



- Quality assurance is a multi step process
 - Right from production of devices
- Limited access to reference monitors in other cities still a major challenge (Almost none monitor PM_{10})
- Intermittent internet & power supply affects data completeness
- Meteorology data access was a main hindrance for calibration in several locations across Africa
 - We are improving AirQo devices to provide ambient temp & humidity

Data Integration

- Difference in temporal resolutions (minute, hour and daily intervals)
- Difference in spatial resolutions - effect on Macro analysis
- Automated processes vs manual processes
- The need for standardised formats and protocols for data integration.



MAKERERE UNIVERSITY

@AirQoProject
@Priscillaa_says

<https://www.airqo.africa>





**THANK
YOU**