This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

A Distributed Low-Cost Pollution Monitoring Platform

Thomas Becnel, Kyle Tingey, Jonathan Whitaker, Tofigh Sayahi, Katrina Le, Pascal Goffin, Anthony Butterfield, Kerry Kelly, Pierre-Emmanuel Gaillardon
University of Utah
Salt Lake City, Utah, USA
{thomas.becnel, pierre-emmanuel.gaillardon}@utah.edu

***Abstract*—Personal exposure to heightened levels of fine airborne *Particulate Matter* (PM) has been linked to numerous adverse health effects in sensitive groups. However, researchers investigating these correlations are struggling to find spatio-temporal datasets sufficient for study. Current airborne PM monitoring solutions are highly accurate, but expensive. Therefore they are not feasible candidates for spatially dense deployments, and cannot be used to analyze the effects of exposure to pollution microclimates. In this work, we present a low-cost pollution monitoring station that operates as a single node in a wireless network. Each node periodically collects airborne pollution and supporting meteorological data and uploads measurements to a central, open-source database. A total of 50 nodes were deployed across a large metropolitan area (roughly 100 km^2) over a six-month campaign. Experimental results show good correlation ($R^2 = 0.88$) between devices co-located with the Federal Equivalent Methods, which have accuracy traceable to the National Institute of Standards and Technology. By applying linear corrections derived from *in situ* field measurements to each PM sensor, we were able to demonstrate a 1.8× decrease in root-mean-squared error over the raw measurements.**

***Index Terms*—IoT, Low-Cost Pollution Monitoring, Particulate Matter, $PM_{2.5}$, Wireless Sensor Network**

I. Introduction

According to the *World Health Organization* (WHO), more than 80% of people living in urban areas are exposed to air quality levels that exceed the WHO limit [1]. These airborne *Particulate Matter* (PM) concentrations, most notably PM smaller than 2.5µm in diameter ($PM_{2.5}$), have been linked to numerous adverse health effects, such as increased incidents of cardiac arrhythmia, asthma, lung cancer, heart disease, and mortality [2], [3], [4], [5]. The WHO estimates that in 2012 approximately 3.7 million people died as a result of exposure to ambient air pollution [6].

The profound link between airborne $PM_{2.5}$ exposure and personal health deterioration has led researchers to thoroughly investigate ambient air pollution, especially in the form of urban microclimates. Current analyses of urban air pollution primarily rely on measurements from highly accurate air quality monitoring stations (AQMs) [7]. These standardized sources produce reliable measurements, but have several drawbacks: 1) AQMs are typically large or bulky, and in some cases require high-voltage power or peripherals such as heating/cooling, and 2) AQMs can be prohibitively expensive to deploy and maintain. These drawbacks can severely limit the number of installations in a given metropolitan region, and as a result current AQM solutions alone are incapable of producing the resolution required to evaluate the fine-grained formation and propagation of ambient airborne pollution in urban environments [8]. Introducing low-cost pollution sensor networks into these regions can help interpolate the spatial gaps between contiguous AQMs, and ultimately offer insight into the spatial and temporal diversity of airborne pollution [9].

Increased accessibility of cheap, yet powerful WiFi-enabled *Microcontroller Units* (MCUs) have given rise to many low-cost pollution monitoring platforms [10] [11]. The implementation of such solutions have since seen a great response from the community, further expanding the market and allowing many institutions and small, non-profit organizations to successfully target specific applications, such as personal exposure or large-scale monitoring networks [12] [13]. A popular and successful technique for dissemination of these low-cost sensor networks is through community engagement [14]. Residents can host a device at little to no cost to them, which greatly reduces the burden from the research groups and ultimately allows for the successful deployment of these large-scale networks. However, there are several challenges presented by relying on such networks for research-grade data. First is the issue of maintenance and upkeep of the devices. While WiFi-enabled embedded systems have proven to be reliable, they can be far from intuitive to troubleshoot, due to the lack of a user interface and standardized tools. As a result the researchers must ensure a quick and clear line of communication with the hosts to keep the networks functioning properly. The second, and more predominant issue, is the quality of data the integrated low-cost sensors produce. Device-level variations during production can lead to non-uniform correlations with the reliable AQMs, making interpretation of data a particular challenge.

In this paper, we propose a low-cost pollution monitoring station that operates as a single node in an interconnected large-scale sensor network, to which it contributes a temporally dense dataset of air pollution concentrations and supporting meteorological data.

More specifically, we make the following major contributions in this paper:

- We propose a novel, low-cost embedded platform, referred to hereon as the AirU Pollution Monitor, or

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

AirU, that measures ambient PM_1, $PM_{2.5}$, and PM_{10} (airborne particulate matter with diameter less than 1μm, 2.5μm, and 10μm respectively) concentrations (measured in $\mu g/m^3$), temperature, relative humidity, light intensity, carbon monoxide, nitrogen oxide, and is equipped with a battery-assisted real-time clock, a *Global Positioning System* (GPS), and a micro-SD card.

- We demonstrate the usability of the AirU through a deployment campaign of 50 AirUs across the Salt Lake Valley, spanning November 1, 2018 to April 30, 2019. This time period was chosen to coincide with elevated airborne pollution due to winter inversions. These devices are hosted by volunteers in the local community. They collect air quality and meteorological data and periodically upload it to our server, where it is stored in a database and made available to the public.
- We verify the quality of the data by analyzing a subset of AirUs co-located with the high-quality *Department of Air Quality* (DAQ) references. Results of the deployment show that the AirU Pollution Monitors correlate well with their co-located *Federal Equivalent Method* (FEM, discussed in Section II) ($R^2 = 0.88$), and an error improvement of 1.84× of the calibrated data over the uncalibrated data.

The remainder of this paper is organized as follows: Section II discusses the issues of large-scale, real-time data acquisition, as well as the current state of low-cost IoT sensor networks. Section III covers the design of the AirU Pollution Monitor and network, Section IV discusses calibration of the various sensors, Section V describes the field experiments conducted, and discusses the results of the aforementioned experiments. Finally Section VI concludes the work presented here.

II. Background

In this section we discuss the current state of air quality monitoring. We discuss the study region (Salt Lake City, UT) and the current pollution monitoring techniques used by the region. We then describe state-of-the-art low-cost alternatives to the current pollution monitoring techniques.

A. Air Quality Monitors in the Study Region

In this work, we compare measurements from the AirU network against a type of *Air Quality Monitor* (AQM) with very high accuracy and tight tolerance, known as a *Federal Equivalent Method* (FEM). The *Environmental Protection Agency* (EPA) defines an FEM as "a method of sampling and analyzing the ambient air for an air pollutant that has been designated as an equivalent method" [15]. The EPA requirements for a commercial particle counter to be classified as an FEM are stringent and rigorous. FEMs must be manufactured in an ISO 9001-registered facility, and all FEMs must be tested and meet the requirements for comparability to a *Federal Reference Method* (FRM) [16]. FRM data is referred to as the "gold standard" of the ambient pollution (PM_1, $PM_{2.5}$, and PM_{10}) concentrations, and is a 24-hour integrated sample collected from the ambient air and analyzed gravimetrically [16]. There are two FEMs in the Salt Lake Valley, where this research was conducted, located at the two DAQ sites (Hawthorne Elementary School, and Rose Park Elementary School), which we refer to throughout the study when analyzing the integrity of the AirU network measurements. Both FEM Pollution Monitors are ThermoFisher 5030i SHARP Monitors, which use a combination of light-scattering nephelometry and beta attenuation technology to produce PM measurements. We will refer to these two FEMs as *DAQ: Hawthorne* and *DAQ: Rose Park* throughout this paper.

B. Salt Lake County as a Test Bed

Salt Lake County, Utah, periodically experiences some of the worst air quality in the United States [17]. The area is surrounded by mountains, and undergoes heavy atmospheric inversions during the winter months. This has the effect of trapping the pollution on the valley floor for days or weeks at a time, and prompting concern of personal exposure health risks in the community. The Greater Salt Lake area also has a variety of geographical and urban factors, such as a large lake and suburban, industrial, and dense metropolitan regions that impact pollution microclimates.

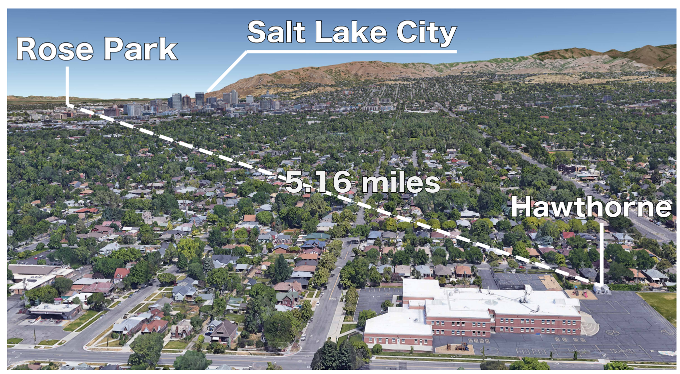

Fig. 1: The locations of the two FEMs maintained by the Department of Air Quality in the immediate Salt Lake Valley.

Figure 1 shows the location of the two FEMs located in the Salt Lake Valley, which are separated by a distance of over 5 miles. The elevation change between the two stations is less than 100 feet, and the path between them does not intersect any interesting geographical or urban features, such as downtown Salt Lake City. Figure 2 shows a sample of hourly measurements taken by the two stations during one of the more predominant periods of heightened ambient PM levels in the valley. A short analysis of Figure 2 shows large measurement disparities between the two locations, as well as abrupt temporal fluctuations. In some cases, measurements differ by as much as 27 $\mu g/m^3$, which lends proof to the existence of urban microclimates at these sources. The EPA defines $PM_{2.5}$ as harmful to human health if a daily average exceeds 35 $\mu g/m^3$ [18]. Figure 2 demonstrates that an individual's personal exposure risk may then be misclassified if they are not located near one of the FEMs. To this end, there is a need for interpolated data between the high-quality stations to more accurately represent airborne pollution events. These factors make the Salt Lake Valley an ideal test bed for the initial deployment of the AirU Pollution Monitoring Network.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

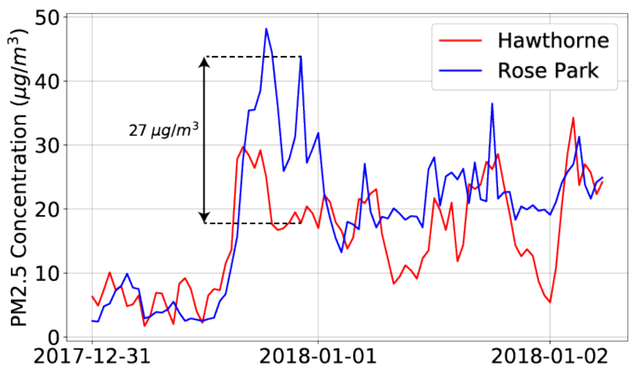

Fig. 2: $PM_{2.5}$ concentrations for the two DAQ FEM PM monitors located at Hawthorne Elementary School and Salt Lake Center for Science Education (Rose Park), during an elevated pollution event in the Salt Lake Valley. Note that measurements can differ by up to 27 $\mu g/m^3$. $R^2 = 0.635$

C. Existing Low-Cost Pollution Monitoring Strategies and Networks

Low-cost alternatives to the highly complex AQMs have been proposed, and demonstrate low variability, high correlation, and good repeatability with the high-quality AQMs [19]. For example, Mead *et al.* demonstrate a low-cost network of sensors containing *carbon monoxide* (CO), *nitrogen monoxide* (NO), and *nitrogen dioxide* (NO_2) measurement instruments that show low-noise and high-linearity results, as well as low gain attenuation after a 12-month period [20]. Bell *et al.* show that spatially heterogeneous ambient pollution levels may arise from regional PM sources, therefore resulting in a misclassification of personal exposure levels [21]. The group investigated the spatial relationship of $PM_{2.5}$ chemical constituent concentrations for the period 1999 to 2007 for 480 monitors in the United States for seven chemical constituents of $PM_{2.5}$ and found an inverse relationship between sensor distance and measurement correlation. Zikova *et al.* recently showed that low-cost particle counters correlate well with one another, and are useful in proving heterogeneity in regional PM levels [22]. A network of 25 sensors was deployed in Rochester, New York for two seasons during the winter months of October-April. The sensor network could be used to identify high-concentration areas or provide very detailed monitoring within a smaller area. Hu *et al.* have demonstrated the use of low-cost mobile monitors paired with smartphones to crowd-source urban air pollution from the community in order to more accurately estimate personal air pollution inhalation dosage [13]. Many groups have also demonstrated the feasibility of custom purpose-built network nodes at driving down the price point while retaining the accuracy of previous low-cost monitoring studies. Ali *et al.* demonstrate a General Packet Radio Service-enabled sensor network of mobile sensors that relay CO, NO_2, and *sulfur dioxide* (SO_2) measurements, along with GPS location and time stamp, to a central server [23]. The online interface makes it easy for users to access and view the data. Vineeth *et al.* demonstrate a mobile monitoring system that can monitor vehicular pollution in real-time along with the information about the area being monitored.[24]. The group show that the low-cost mobile device is capable of detailed analysis of ambient pollution constituents, such as PM and CO. The device secures the data by sending measurements to a central server via WiFi.

There also exist a number of commercially available low-cost pollution monitoring solutions, such as Dylos and PurpleAir. Dylos sensors are marketed as indoor pollution monitors to help track personal exposure inside one's home. Real-time data can be viewed via a web browser, or directly from the monitor itself. The approach of PurpleAir involves connecting WiFi-enabled pollution monitors to the exterior of residential homes to track outdoor pollution levels. PurpleAir has been selling network-connected pollution monitors in northern Utah for several years and have approximately 100 sensors currently deployed in the Salt Lake Valley. The data collected by the sensors is hosted on a database open to the public.

The aforementioned systems have helped to make substantial advancements in the field of urban pollution research, but fail to address all the issues to make a low-cost pollution monitoring network a feasible device for data collection, such as individual sensor calibration and verification prior to deployment, use of only low-cost components, a small, lightweight form-factor, and ease-of-use.

Our specific criteria for a feasible low-cost pollution monitor are as follows: 1) the device must incorporate a low-cost, calibrated PM sensor, 2) temperature and humidity sensor, 3) wireless connectivity, 4) small, lightweight, attractive form-factor, and 5) ease of deployment and maintenance. Due to the incomplete design criteria of previous work and the need for spatially and temporally dense urban pollution datasets, this paper presents a low-cost pollution monitor that acts as a single node in a sensor network and fits the above criteria. Accuracy and validity of the acquired data is ensured by testing every particle sensor in a laboratory setting and assigning a unique linear correction to the raw PM data. In addition, every AirU is equipped with temperature and humidity sensors, due to the relationship between elevated humidity and PM measurements [25]. While Salt Lake City has consistently low humidity and is therefore not affected by this relationship, the inclusion of these sensors allows the AirU to operate in many different environments in future deployments. Using the Salt Lake Valley as a test bed, we have added to the existing infrastructure of approximately 100 wireless sensor nodes set in place by PurpleAir.

III. Design Considerations of the AirU Pollution Monitor

In this section we discuss the rationale into the hardware and software design of the AirU Pollution Monitor and the AirU network. We discuss the on-board hardware/sensors, data transport and retention protocol, server/database architecture, electrical characteristics, cost, and storage capabilities.

A. General Design Considerations of the AirU Pollution Monitor

We first formalize the requirements for an AirU Pollution Monitor. A device must contain: 1) a low-cost particle counter

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

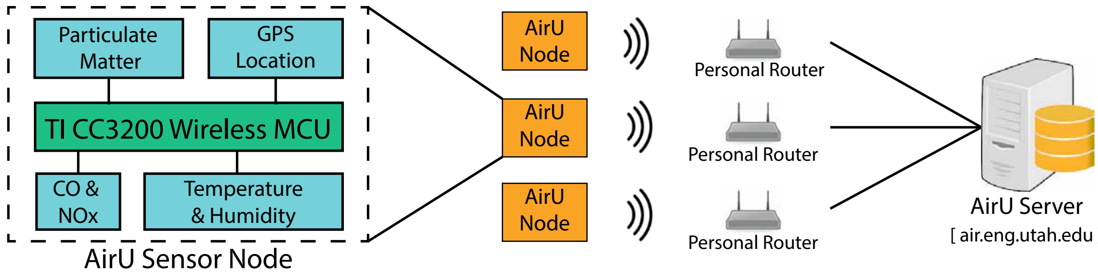

Fig. 3: System diagram of a single AirU node and the proposed interconnected network.

(<\$20), 2) temperature and humidity sensors, 3) WiFi connectivity, and 4) ease of deployment and maintenance. An optical light-scattering particle counter was chosen because they are a popular option for measuring PM concentrations at micron/sub-micron diameters in real-time, and are readily available at a low price point [26], [27]. The AirU network was targeted as a large-scale urban deployment, so WiFi communication was the most practical solution, as opposed to a *radio frequency* (RF) protocol such as LoRa or Zigbee. WiFi-enabled MCUs are much less expensive than these alternative RF solutions, and almost every household has a wireless access point. A new device can be easily connected to a network in minutes, and a robust set of protocols keeps the device connected to the Internet and pushing data as long as the host maintains a valid home network connection. Figure 3 shows the proposed system diagram for a single AirU device and its wireless connection in the AirU network.

Every AirU is given a unique serial number, denoted as *SAxxx*. We will consistently use this notation to refer to the AirU monitors for the remainder of this text.

B. AirU Pollution Monitor Hardware Design

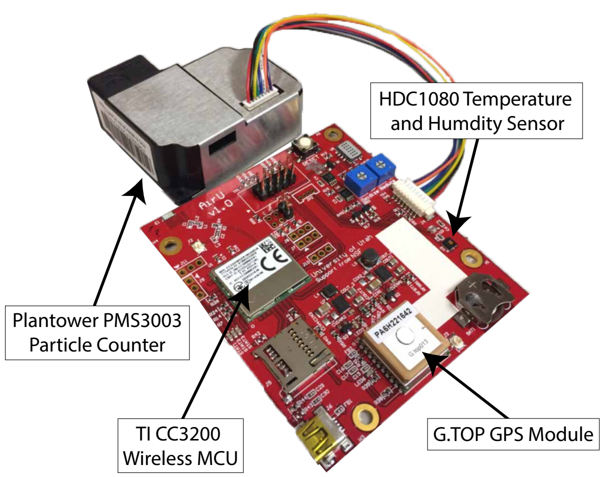

Fig. 4: AirU Pollution Monitor with relevant hardware labeled.

Figure 4 shows the block diagram containing the relevant sensors on the AirU Pollution Monitor. The PM sensor is the Plantower PMS3003. The PMS3003 samples at roughly 1Hz, and categorizes measurements into three bins: PM_1, $PM_{2.5}$, and PM_{10}. The sensor uses a photodiode and photosensor pair and the light-scattering principle to obtain particle counts, and translates these into $\mu g/m^3$ concentrations. The data is transferred via the *Universal Asynchronous Receiver/Transmitter* (UART) protocol to the host MCU. Every sample from the PMS3003 for the current acquisition period (typically one minute) is aggregated, and single averages for PM_1, $PM_{2.5}$, and PM_{10} are sent in a packet during the upload stage.

The temperature and humidity sensor is a Texas Instruments HDC1080. It is a low power, high accuracy digital sensor that communicates via an *Inter-Integrated Circuit* (I^2C) interface. The HDC1080 was calibrated to help correct manufacturer variability, and is discussed further in Section IV-C. Raw measurements are uploaded to the database during the upload stage, and the linear correction is provided online separately. The on-board humidity sensor was also calibrated to account for device-level fluctuations of the sensors.

A GPS module was incorporated into the system to allow for mobile acquisition, as well as to remove the burden from the host of manually entering the global position of the device. The GPS also provides elevation of the device, a crucial component to consider when evaluating atmospheric inversion climates, such as Salt Lake County. The GPS module on the AirU Pollution Monitor is the G.TOP FGPMMOPA6H, which is capable of measuring speed and velocity as well as spatial location. The G.TOP FGPMMOPA6H also includes a *real-time clock* (RTC) and backup battery connection, so the current time and ephemeral GPS data will not be lost if the AirU device loses power. The GPS module can keep time for up to four years on battery power from the 3V coin cell battery and RTC. Standardized packets containing time and GPS measurements, and known as *National Marine Electronics Association* (NMEA) sentences, are transmitted via UART to the AirU MCU at a frequency of roughly 1Hz.

The AirU Pollution Monitor incorporates several other hardware devices to assist in the study and acquisition of the PM data. A 4GB micro-SD card is included in the AirU monitor, and can store roughly one year of data. All packets uploaded to the online database are also stored on the micro-SD card. In the event that the AirU monitor becomes disconnected from the host WiFi network the data will continue to be logged, and when the device regains an Internet connection the logged data is uploaded to the database. This procedure is described in depth in III-D. The station also includes an ambient light sensor and a SGX SensorTech MiCS-4514 carbon monoxide and nitrogen oxide sensor. The use of the gas and ambient light sensors is out of the scope of this work, and will not be discussed here. The MCU used by the AirU is the Texas

Instruments CC3200 Wireless MCU with ARM Cortex-M4.

C. Network Considerations

Locally gathered data is transferred via the MQTT protocol from the AirU device to the server, where the Mosquitto MQTT Broker is maintained. MQTT was chosen as the communication protocol because it is an extremely lightweight messaging protocol and offers simple machine-to-machine communications through a centralized node (the MQTT Broker) through a standard Internet connection. AirU devices periodically send data packets, which contain the device's MAC address, *Coordinated Universal Time* (UTC) time, global coordinates in decimal-degrees, temperature, humidity, PM_1, $PM_{2.5}$, PM_{10}, and number of seconds since the last system reboot. This data is then transferred via Telegraf to an InfluxDB database, where it is made publicly available. Data stored on the InfluxDB instance can easily be retrieved at any time using the command-line InfluxDB tool, via the Chronograf web interface, or the InfluxDB library written for JavaScript, C#, Java, Python, and many other languages.

The MQTT communication structure is advantageous in that it allows clients (the AirU devices) to subscribe to, and listen to, an arbitrary number of topics. This provides a simple and efficient way to communicate with multiple AirU devices at once in order to update and maintain separate firmware versions. The simple messaging protocol also allows us to include a number of helpful remote debugging tools, such as instigating file (HTML and firmware patches) downloads from the secure server.

The CC3200 MCU hosts an internal HTTP server, which is the primary method for the user to communicate with the device. The user can access a web page hosted on the AirU, allowing them to connect the device to their local network, and view some settings such as MAC address and local IP address.

The AirU Pollution Monitor is designed to be an all-inclusive airborne particulate monitor, and therefore includes several sensors that have not yet been calibrated and implemented into the system. Since new features and sensor corrections are constantly being applied to the system, we have enabled the platform with over-the-air firmware updates capabilities. The binaries are hosted on a secure server, and a simple MQTT message allows a device to download a new binary from the server and execute it, without any interaction from the AirU's host.

D. Offline Data Upload Procedure

The inclusion of a GPS module, RTC, and micro-SD card allows the AirU to collect pollution data while offline, and upload this data when the device regains an Internet connection. A file is maintained on the micro-SD card which tracks the Internet connectivity status of the AirU. When the device regains an Internet connection, a software task is triggered. The task first retrieves the timestamp that the device lost connectivity, then sends the corresponding data files to the server via HTTP. Each data file stored on the micro-SD card contains pollution measurements for a 24-hour period, in 1-minute increments, with the filename *YYYY-MM-DD.csv* (year-month-day). The task iteratively uploads each daily data file contained in the down-time period. In the event that the AirU loses Internet connectivity multiple times a day, some duplicate data will be sent to the server. This is done intentionally to keep the constrained AirU from using resources to parse these data files, and instead we put this load onto the server. The AirU devices are currently limited to two uploads a day to minimize the traffic.

The server hosts an HTTP route dedicated to receiving and handling these incoming data files from the AirU devices. When a CSV file is received, each row of data is compared against the database and inserted if missing. In this way duplicate data points are discarded and the database maintains a data point for every minute the AirU is powered on.

This functionality also allows the AirUs to operate as mobile pollution monitors with no additional requirements. The AirU can be powered through a 5V battery pack and is not constrained to a WiFi access point. The on-board GPS will track the location of the device, and when the AirU regains an Internet connection, the offline data will be uploaded.

E. Power, Cost, and Storage Characteristics

Here, we describe the electrical characteristics of the AirU Pollution Monitor, which is of particular interest when the AirU is used as a mobile station. Power measurements described here were derived from samples taken at 1 KHz and averaged over a 60-second duration. Measurements were acquired using a Rigol DP832 programmable power supply and Fluke 117 multimeter. The AirU is powered by a 5V supply, and draws 108 mA while fully active. While in a light-sleep state (CC3200 asleep, sensors disabled), the AirU draws 2 mA. After a wake-up event, the PMS3003 particle counter takes 10 seconds to calibrate before it produces valid measurements. For a standard 60-second sampling period, the AirU spends 40 seconds in light-sleep, 10 seconds calibrating, and 10 seconds collecting measurements. The temperature and humidity sensors only require several milliseconds to calibrate after becoming active. The average power consumption of the AirU for a 60-second period is then:

$$P = [2mA * \frac{40s}{60s} + 108mA * \frac{20s}{60s}] * 5V = 187mW \qquad (1)$$

Additionally, an AirU costs less than \$200 for hardware, fabrication, assembly, and housing. The AirU contains a 4GB micro-SD card, on which it stores the collected air quality and meteorological data. Every data packet is less than 250 bytes. For a standard 60-second sampling period, this results in over 30 years of potential data storage. The large micro-SD was chosen because the AirU is capable of sampling periods as low as two seconds, and logs a number of different messages and statistics from the device onto the micro-SD card.

IV. Calibration of AirU On-board Sensors

Due to the innate manufacturing variability of the sensors on the AirU Pollution Monitors, a calibration of the on-board sensors was performed in order to reduce measurement error.

A generic linear calibration was derived via an *in situ* field calibration technique in order to properly calibrate the sensors to the correct aerosol makeup of the Salt Lake airshed.

A. Calibration Methodology

A subset of eight AirU monitors were deployed at the DAQ sites in the Salt Lake Valley (four co-located with DAQ: Hawthorne, four co-located with DAQ: Rose Park) over a deployment campaign ranging from November 1, 2018 to April 30, 2019 (six months). During this time, the AirU monitors collected $PM_{2.5}$, temperature, and humidity at one-minute intervals. The data was aggregated into one-hour averages in order to properly align with the one-hour FEM sampling rates, and then compared against the FEMs. FEM data was downloaded directly from the DAQ database and was used without any alterations or filtering. To derive a $PM_{2.5}$ calibration, a subset of co-located sensors were used to train simple linear regression models, while the remaining co-located sensors were to used for verification of the model. Over the six-month period, we did not see a significant change in the gain and offset calibration parameters, and therefore believe a single linear calibration is sufficient to properly characterize the sensor measurements.

Results of the calibration are demonstrated using Bland-Altman plots, also known as mean-difference plots. The x-axis is the mean of the two measurement techniques, while the y-axis is the difference between the two measurement techniques. More concisely, the Cartesian coordinates of a given sample S are given by:

$$S(x,y) = \left(\frac{S_1 + S_2}{2}, S_1 - S_2\right) \tag{2}$$

where S_1 and S_2 refer to samples from the two measurement techniques. When plotted, these points give insight to the relationship between the two methods that is not made clear with a scatter plot and correlation alone [28]. Three horizontal lines are drawn on the plot: the mean-difference, and upper and lower *limits-of-acceptance* (LoA). The mean-difference, denoted $\bar{d}$, is the average of the difference between the two methods:

$$\bar{d} = \frac{1}{N}\sum_{n=0}^{N}(\hat{x}_n - x_n) \tag{3}$$

where N is the number of samples, $\hat{x}$ is the predicted value (FEM measurements), and x is the dependent variable (AirU measurements). The LoAs, also known as the 95% LoAs, are approximately two standard deviations on either side of the mean-difference, and help define the amount that the two methods may disagree.

The *root-mean-squared error* (RMSE) is a popular error metric for comparing two methods, and is used frequently throughout this text. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=0}^{N}(\hat{x}_n - x_n)^2} \tag{4}$$

where N is the total number of samples, $\hat{x}$ is the predicted value, and x is the dependent variable.

B. Calibration of Plantower PMS3003 PM Sensor

A dataset of over two million PM measurements was used to derive a generic linear calibration. The AirU datasets were first cleaned by removing outliers greater than 1000 $\mu g/m^3$. Values that exceeded this threshold were removed because they have been attributed to malfunctioning particle counters, due to debris in the chamber or a loose wire powering the PMS3003 fan and laser [30]. These measurements were then averaged into hourly measurements to coincide with the hourly measurements from the DAQ PM sensors (ThermoFisher 5030i SHARP Monitor at both Hawthorne and Rose Park sites). Three PMS3003 particle counters from each DAQ site (six sensors total) were randomly chosen for the training datasets. AirU sensors located at the DAQ Hawthorne site were SA127, SA008, and SA137. AirU sensors located at the DAQ Rose Park site were SA079, SA072, and SA134. Linear least-squares regressions were fitted to each training sensor, using measurements from the ThermoFisher 5030i SHARP Monitor located at each DAQ site as the FEM reference. Table I shows the gain and offset calculated for each of these sensors. The gains and offsets were averaged, as shown in Table I, to produce a generic linear calibration, which was then verified against the remaining two test sensors, one sensor located at each DAQ site.

TABLE I: $PM_{2.5}$ gain and offset from linear least-squares regression for the training sensors located at the DAQ sites.

	DAQ: Hawthorne			DAQ: Rose Park			
	SA008	SA127	SA137	SA072	SA079	SA134	Mean
Gain	1.89	1.22	1.43	1.47	1.67	1.61	1.55
Offset	-4.03	-2.96	-3.17	-3.91	-4.53	-4.02	-3.77

The results of this calibration procedure are shown in Table II, and supplemented by the Bland-Altman plots shown in Figure 5. We first note that all the PMS3003 sensors showed high correlation with the corresponding DAQ references, as shown in Table II, as well as very high correlation amongst themselves, as shown in Figure 9. The test sensors (SA016, SA131) saw 1.77× and 1.84× reductions in RMSE after calibration, respectively. The calibration also greatly tightened the 95% LoA for both test sensors, which demonstrates that the calibration successfully reduced the spread of measurements, especially at high PM concentrations. This can be seen by the decrease in the 95% LoA metrics in Table II, as well as the calibrated 95% LoA lines in Figure 5.

C. Calibration of the Texas Instruments HDC1080 Temperature & Humidity Sensor

Similar data processing was done to perform calibrations of the temperature and humidity sensors. Two AirU temperature sensors from each DAQ site (four sensors total) were randomly chosen for the training datasets. AirU sensors located at the DAQ Hawthorne site were SA127 and SA016. AirU sensors located at the DAQ Rose Park site were SA079 and SA072. Linear least-squares regressions were fitted to each training sensor, using measurements from the MET One 597A Temperature Sensor located at each DAQ site as the reference. The humidity sensors were calibrated in an identical fashion.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

TABLE II: PMS3003 *in situ* field calibration statistics for $PM_{2.5}$ measurements.

The applied calibration equation was the inverse of the equation derived in Table I: $y^{-1} = 0.65x + 2.43$. This equation was applied to each sensor measurement. Bland-Altman Plots were created for the raw and calibrated datasets for each sensor and compared against the corresponding DAQ FEM dataset. Mean-difference and Limits-of-Agreement were extracted from the plots and are presented in this table. We also present the root-mean-squared error, derived from the PMS3003 and FEM.

Metric ($\mu g/m^3$)	SA016	SA131	SA127	SA008	SA137	SA079	SA072	SA134
Correlation (R^2)	0.85	0.90	0.85	0.87	0.87	0.90	0.87	0.87
RMSE (Raw)	7.91	6.41	5.95	10.77	6.87	7.68	6.89	8.02
RMSE (Cal)	4.46	3.49	4.07	5.59	3.95	3.85	3.90	4.27
95% LoA (Raw)	15.79	12.80	11.64	20.79	13.75	15.30	13.78	15.97
95% LoA (Cal)	8.92	6.89	7.85	10.75	7.90	7.70	7.73	8.54
Mean Diff (Raw)	-0.49	0.38	1.26	-2.82	-0.13	-0.63	0.27	-0.73
Mean Diff (Cal)	-0.03	0.56	1.09	-1.53	0.20	-0.09	0.48	-0.16

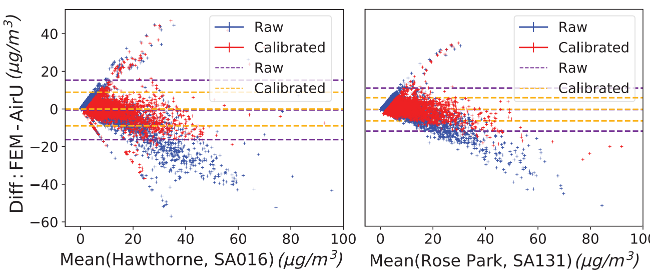

Fig. 5: Bland-Altman plots for PM test sensors SA016, located at DAQ: Hawthorne, and SA131, located at DAQ: Rose Park. Both test sensors saw a large reduction in measurement spread at high $PM_{2.5}$ concentrations, as well as tightening of the LoA bounds, after calibration.

Tables III and IV show the gain and offset calculated for each of these sensors, for temperature and humidity, respectively. The gains and offsets were averaged (also shown in Tables III and IV) to produce a generic linear calibration, which was then verified against the remaining two test sensors, one sensor located at each DAQ site.

TABLE III: HDC1080 temperature sensor calibration coefficients. Sensors were co-located at the DAQ sites, inside the AirU enclosures.

	DAQ: Hawthorne		DAQ: Rose Park		
	SA016	SA127	SA072	SA079	Mean
Gain	0.99	1.07	0.89	0.97	0.98
Offset	10.50	10.85	12.62	11.64	11.4

TABLE IV: HDC1080 humidity sensor calibration coefficients. Sensors were co-located at the DAQ sites, inside the AirU enclosures.

	DAQ: Hawthorne		DAQ: Rose Park		
	SA016	SA127	SA072	SA079	Mean
Gain	1.03	0.97	1.02	1.27	1.07
Offset	25.89	28.56	28.88	15.04	24.59

The results of these calibrations are shown in Table V, and supplemented by the Bland-Altman Plots shown in Figure 6. From Table V and Figure 6, we see that the application of a generic linear calibration to the temperature sensors was very successful. The test sensors (SA008, SA134) saw 2.2× and 3.3× reduction, respectively, in RMSE. We also note that the one-sided measurement error was largely corrected, which can be seen by the reduction in the mean difference. The HDC1080 temperature sensors are very noisy, which can be seen in the large spread in Figure 6, and unfortunately cannot be controlled with linear calibration alone. However, all the co-located temperature sensors have a high correlation ($R^2 > 0.85$) with the corresponding DAQ sensors. By looking at the linear calibration ($y = 0.98 + 11.4$), we see that most of the discrepancy can be attributed to the sensor drift, meaning the measurements are affected mostly by a constant offset. We believe this is largely due to the plastic enclosure, which only has several small holes at the bottom. With limited air circulation, the AirU is quick to heat, and slow to cool, which makes the temperature inside the enclosure lag behind the ambient temperature during cooling events.

TABLE V: *In situ* field temperature calibration of DAQ co-located sensors. Sensors SA008 and SA134 were used as test data, while SA127, SA016, SA079, and SA072 were used as the training data. *Raw* refers to uncalibrated data, and *Cal* refers to calibrated data.

Metric ($^\circ C$)	SA008	SA134	SA127	SA016	SA079	SA072
Correlation (R^2)	0.80	0.83	0.94	0.74	0.78	0.75
RMSE (Raw)	10.30	12.39	10.61	12.45	12.58	12.32
RMSE (Cal)	4.75	3.73	2.08	5.34	3.90	4.54
95% LoA (Raw)	8.50	7.24	3.72	10.47	7.56	8.92
95% LoA (Cal)	8.67	7.38	3.78	10.68	7.67	9.07
Mean Diff (Raw)	-9.38	-11.85	-10.45	-11.30	-12.00	-11.49
Mean Diff (Cal)	1.94	-0.56	0.86	-0.01	-0.71	-0.20

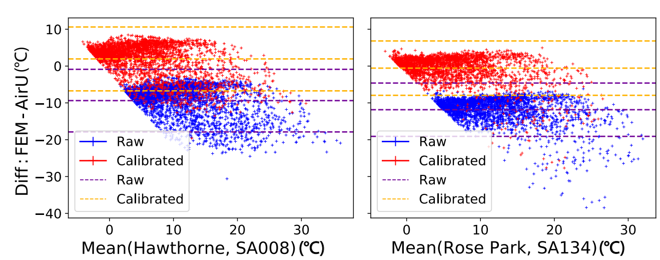

Fig. 6: Bland-Altman Plots for temperature test sensors (SA008 and SA134), located at DAQ: Hawthorne and DAQ: Rose Park, respectively. Both sensors saw a large reduction in RMSE after calibration, as well as a decrease in the 95% LoA bound. The calibration also greatly reduced the mean-difference of the measurements.

The HDC1080 humidity sensors were calibrated in the same fashion described for the temperature sensors. From Table VI, we see relatively good correlation between the AirU sensors and the DAQ FEMs. Analysis of the Bland-Altman

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

Plots shown in Figure 7 reveal that the sensors systematically experience a large deviation from the mean. We believe this may be due to several factors, such as extreme temperatures (triggering freezing, evaporation on the exposed sensor face), and contamination (such as dust, PM, or debris). The effects of the enclosure tend to amplify these issues, which points to the cause of the large 95% LoAs, which were not affected by the calibration. However, Table VI also shows large reductions in RMSE compared to uncalibrated datasets (2.56× reduction for SA008, 2.51× reduction for SA134), which shows that a linear calibration was successful in improving measurement accuracy.

TABLE VI: *In situ* field humidity calibration of DAQ co-located sensors. Sensors SA008 and SA134 were used as test data, while SA127, SA016, SA079, and SA072 were used as the training data. *Raw* refers to uncalibrated data, and *Cal* refers to calibrated data.

Metric ($\%RH$)	SA008	SA134	SA127	SA016	SA079	SA072
Correlation (R^2)	0.66	0.60	0.76	0.68	0.67	0.78
RMSE (Raw)	30.17	32.02	26.04	29.18	28.19	23.82
RMSE (Cal)	11.77	12.77	10.83	11.61	11.79	11.27
95% LoA (Raw)	23.42	24.94	20.93	23.09	23.36	19.96
95% LoA (Cal)	23.44	24.93	20.80	23.22	23.46	19.88
Mean Diff (Raw)	27.81	29.49	23.85	26.80	25.65	21.63
Mean Diff (Cal)	1.03	2.76	-3.05	-0.02	-1.19	-5.32

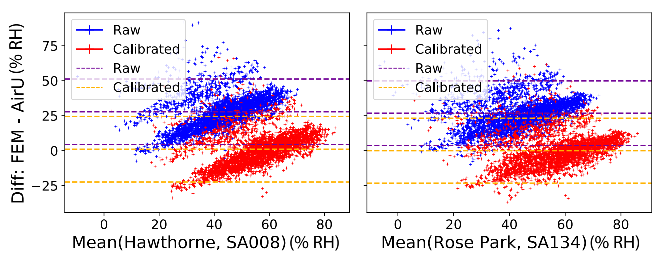

Fig. 7: Bland-Altman Plots for humidity test sensors (SA008 and SA134), located at DAQ: Hawthorne and DAQ: Rose Park, respectively. Both sensors saw a large reduction in RMSE after calibration, but did not see a decrease in the 95% LoA bounds. This can mainly be attributed to the large calibration offset and unity gain.

V. Deployment of the AirU Network

1) Deployment Methodology: The AirU monitors were designed, manufactured, and fabricated in Salt Lake City. Each board was programmed with the initial firmware and loaded with the appropriate files prior to deployment. All boards were housed in a custom weather-resistant plastic enclosure designed at the University of Utah. The initial deployment of the AirU Pollution Monitors used Salt Lake county as a test bed. Locations were largely chosen based on geography in the county in order to maintain a good spatial distribution. In total, 50 AirU Pollution Monitors were deployed and are currently active in Salt Lake County. AirU pollution monitors were shipped to participating volunteers along with instructions to connect the device to WiFi. When a user first powers on their AirU, it configures itself as a wireless access point, broadcasts its SSID, and starts an HTTP server. The user then connects their smartphone, tablet, or laptop to the AirUs WiFi network and navigates to the AirU landing page. Here, the user enters their home WiFi credentials and submits the form to the AirU. The AirU then connects itself to the user's home WiFi network. Navigating back to the AirU landing page through their home WiFi network, the user is then able to register their device. Name, contact information, and device MAC address are sent to our server and stored in a MongoDB instance. When the device obtains a valid GPS location, it is automatically added to the interactive map, which can be viewed at https://aqandu.org. In Section V-A, we show the deployed AirU locations across Salt Lake County in Figure 10.

Each pollution monitor samples temperature, humidity, PM_1, $PM_{2.5}$, and PM_{10} data. The PM measurements are taken once a second, and a running total and count is kept. When a data packet is constructed to upload to the server, the PM accumulations are averaged, and this is the measurement sent. Data is sent to the server in one minute intervals. The temperature and humidity sensors are only sampled immediately prior to a data packet being uploaded. GPS coordinates are also queried only once prior to uploading the data. Each DAQ site (Rose Park and Hawthorne) have four co-located AirU monitor alongside the FEMs. The AirU monitors were deployed during December, 2017 and are still active as of June, 2019.

A. Deployment Results

Here we describe the results of the AirU deployment campaign, which includes details on the reliability of the monitors, as well as integrity of the data they produce, and how calibration affects the raw measurements. In total, 50 AirU Pollution Monitors were deployed across Salt Lake County, Utah. Once connected to the host's WiFi network, the devices proved to be very reliable over the course of the campaign. Figure 8 shows the percentage of data packets successfully transmitted to the database by the AirU monitors. All 50 AirU monitors transmitted at least 84% of the packets, and 26 of the 50 monitors successfully transmitted 95% or more of the data packets. This high transmission rate can be largely attributed to the collection of data packets while the AirU is offline, and the transmission of these packets when the AirU regains an Internet connection. The missing packets can be attributed to power outages/issues at the host's home, or time spent offline while being repaired.

Figures 10 and 11 provide good visual representations of the quality of the $PM_{2.5}$ dataset collected over the course of this deployment campaign by the distributed AirU network. Figure 10 show $PM_{2.5}$ concentrations for the deployed network at 19:00 on December 8, 2018, during an elevated pollution event. Circles indicate the locations of AirU monitors, and stars indicate locations of the DAQ FEMs. Size and color of the markers correspond to the pollution concentration recorded by that device. The image on the left shows raw, uncalibrated measurements, while the figure on the right shows calibrated measurements, using the linear correction derived in Table I. In Figure 10a., the AirU monitors are systematically overestimating the $PM_{2.5}$ concentrations, which can be seen

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

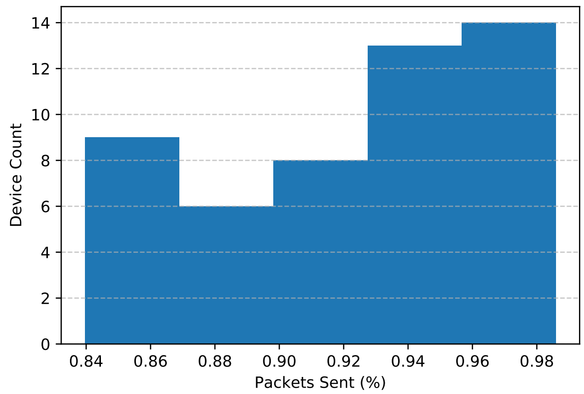

Fig. 8: Percent of data packets sent by all 50 AirU monitors over the 6-month deployment campaign.

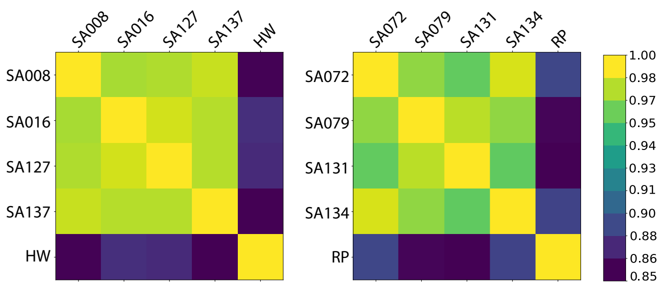

Fig. 9: PMS3003 sensor correlation (r) for AirU monitors located at the DAQ sites. DAQ: Hawthorne is denoted as HW, and DAQ:Rose Park is denoted as RP. Sensors show good correlation with the corresponding DAQ sensors, and very high correlation amongst themselves.

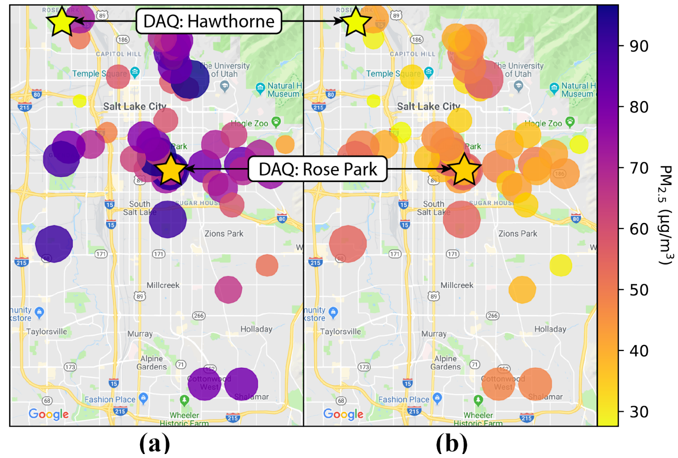

Fig. 10: Pollution concentrations recorded by the AirU network during an elevated pollution event on December 08, 2018. Each circle represents the location of an AirU Pollution Monitor. The size and color of each circle is proportional to the $PM_{2.5}$ concentration. Stars represent the DAQ FEMs. Star colors also represent PM concentration. (a) shows raw $PM_{2.5}$ measurements from the network, and (b) shows calibrated measurements. This figure visually demonstrates the large discrepancy between FEM data and uncalibrated AirU data, and how a single generic calibration derived from field measurements is able to drastically decrease the error between the two measurement techniques.

by the disparity in color between the DAQ FEMs (stars) and AirUs (circles). After applying the linear correction, the AirUs show much more reasonable measurements.

Figure 11 shows the correlation between all 50 PM sensors over the course of the deployment campaign. The figure was created by ordering the AirU monitors by cardinal distance from the Rose Park FEM, then plotting the Pearson correlation between all sensors, using a dataset which spanned from November 1, 2018 to April 30, 2019. The average correlation is $R^2 = 0.72$ for these 50 sensors. Weaker correlations can generally be attributed to larger distances, as depicted by the color gradient of the plot, with higher correlations towards the top-left, and lower correlations along the bottom and right sides. We can also see that one sensor (row/column 26 of this figure) is most likely defective, because of the lack of correlation with all other sensors.

Other studies using the Plantower PMS family of optical particle counters found similar results that we describe here. In a one-year field evaluation of the Plantower PMS5003 and PMS7003, Bulot *et al.* found correlation of $R^2 = 0.98$, and RMSE of 2.2 between 12 PMS sensors. On average, the 12 sensors reported Pearson Correlations of 0.86 with the background reference station, and an average of 6.7 RMSE [29]. Sayahi *et. al.* performed a similar evaluation of four Plantower PMS sensors, over the course of 320 days, which were co-located with one tapered element oscillating microbalance (TEOM), and one gravimetric *Federal Reference Method* (FRM) [30]. The sensors showed good correlation with both the hourly TEOM ($R^2 > 0.87$) and 24-hour FRM ($R^2 > 0.88$). They also found similar calibration gains (1.8) that we report here. Badura *et. al.* compared a Plantower PMS7003 to a TEOM 1400a analyser [31], and found excellent daily correlations ($R^2 = 0.92$). These studies help reiterate the integrity and reliability of the PMS3003 sensor used by the AirU Pollution Monitor.

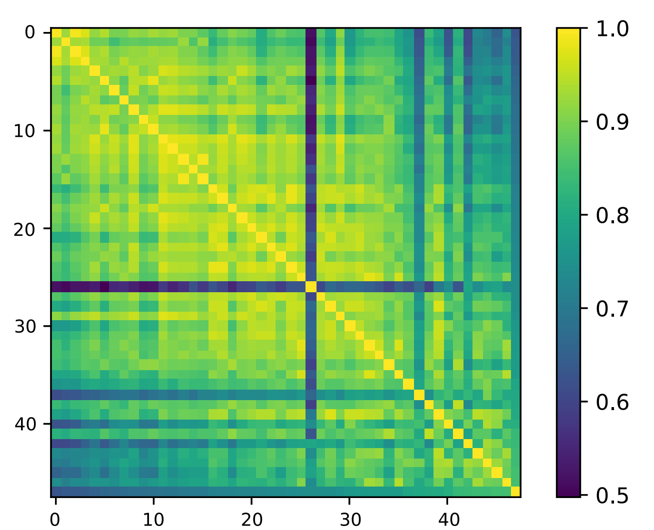

Fig. 11: PM sensor correlation over the length of the campaign. Average correlation is $R^2 = 0.72$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

VI. Conclusion

In this work we have demonstrated the design of a low-cost air pollution monitoring station (dubbed the AirU Pollution Monitor) that can provide reproducible, dense air quality and supporting meteorological measurements. These WiFi-enabled devices gather airborne particulate matter (PM_1, $PM_{2.5}$, and PM_{10}) concentrations, temperature, humidity, and global positioning data and periodically upload it to a server. The sensors are hosted and maintained by participating members of the community, which not only supplements community engagement, but eases a time-consuming burden of maintaining hundreds of sensors by a small team across a large geographical region.

Through a six-month deployment campaign of 50 sensor across a 100 km^2 metropolitan area, we demonstrate a spatially and temporally dense representation of pollution concentrations in urban environments, thus aiding significantly in the study of personal exposure to ambient pollution. The low-cost optical particle counters housed on the AirU are calibrated to account for any manufacturing irregularities. The derived linear correction was effective in correcting the linear regression of the best-fit line, and resulted in a 1.84$\times$ root-mean-squared error improvement in regards to their colocated Environmental Protection Agency-approved Federal Equivalent Method (FEM) standards, which are maintained by the Department of Air Quality. The AirU monitors also include temperature and humidity sensors, which also saw a benefit from linear calibration.

The AirU is capable of real-time monitoring, as well as automatically uploading measurements after loss of Internet connection, which are stored on a micro-SD card. As a result, intermittent Internet connections on the host's personal network will not affect the continuity and density of the data collected. The monitors are outfitted with GPS modules, and therefore may act as mobile pollution monitoring stations, collecting air quality and meteorological data and automatically uploading to the online database upon regaining Internet access. In this work we have shown that a great reduction in size and cost from an FEM standard does not necessarily mean a dramatic decrease in measurement accuracy, and thereby demonstrate that a scalable, low-cost PM sensor network is a reliable and efficient means for collecting spatially and temporally dense heterogeneous pollution data across large metropolitan areas, and ultimately greatly aids in the analysis of urban pollution microclimates.

Acknowledgments

The authors would like to thank Miriah Meyer, Ross Whitaker, Iman Adib, and Kara Douville. This work was funded under the National Science Foundation CPS grant 1646408. Pierre-Emmanuel Gaillardon and Kerry Kelly have financial interest in the company Tetrad: Sensor Network Solutions, LLC, which manufactures air quality sensing solutions and provides engineering services.

References

[1] World Health Organization. (2017). WHO Global Urban Ambient Air Pollution Database (update 2016). [online]. Available: http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/, Accessed on: Dec. 22, 2017

[2] American Lung Association, "State of the Air 2015," Tech. Rep. American Lung Association, 2015.

[3] J. Lepeule, F. Laden, D. Dockery, and J. Schwartz, "Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard six cities study from 1974 to 2009," *Environmental Health Perspectives*, vol. 120, no.7, pp. 965970, 2012.

[4] A. Peters, E. Liu, R. L. Verrier, J. Schwartz, D. R. Gold, M. Mittleman, J. Baliff, J. A. Oh, G. Allen, K. Monahan, and D. W. Dockery, "Air pollution and incidence of cardiac arrhythmia." *Epidemiology*, vol. 11 no. 1, 2000.

[5] A. Pope, R. T. Burnett, M. C. Turner, A. Cohen, D. Krewski, M. Jerrett, S. M. Gapstur, and M. J. Thun. "Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: Shape of the exposure-response relationships," *Environmental Health Perspectives*, vol. 119, no. 11, pp. 16161621, 2011.

[6] World Health Organization, "Burden of disease from Ambient Air Pollution for 2012," Tech. Rep. World Health Organization, 2014.

[7] US-EPA, DRAFT roadmap for next generation air monitoring, U.S.Environ. Protect. Agency, Washington, DC, USA, Tech. Rep., 2013.

[8] N. Zikova, M. Masiol, D. Chalupa, D. Rich, A. Ferro, and P. Hopke, Estimating Hourly Concentrations of PM2.5 across a Metropolitan Area Using Low-Cost Particle Monitors, Sensors, vol. 17, no. 8, p. 1922, 2017.

[9] K. B. Shaban, A. Kadri, and E. Rezk, Urban Air Pollution Monitoring System With Forecasting Models, IEEE Sensors Journal, vol. 16, no. 8, pp. 25982606, Apr. 2016.

[10] H. Zhi-gang and C. Cai-hui, "The Application of Zigbee Based Wireless Sensor Network and GIS in the Air Pollution Monitoring", 2009 International Conference on Environmental Science and Information Application Technology, 2009.

[11] Polidori, A., Feenstra, B., Papapostolou, V. and Zhang, H. (2018). "Field Evaluation of Low-Cost Air Quality Sensors," [online] Aqmd.gov. Available at: http://www.aqmd.gov/docs/default-source/aq-spec/protocols/sensors-field-testing-protocol.pdf [Accessed 13 Apr. 2018].

[12] A. Northcross, R. Edwards, M. Johnson, Z. Wang, K. Zhu, T. Allen and K. Smith, "A low-cost particle counter as a realtime fine-particle mass monitor", Environ. Sci.: Processes Impacts, vol. 15, no. 2, pp. 433-439, 2013.

[13] K. Hu, Y. Wang, A. Rahman and V. Sivaraman, "Personalising pollution exposure estimates using wearable activity sensors", *IEEE Ninth International Conference on Intelligent Sensors*, Sensor Networks and Information Processing (ISSNIP), 2014.

[14] S. V. Brussel and H. Huyse, Citizen science on speed? Realising the triple objective of scientific rigour, policy influence and deep citizen engagement in a large-scale citizen science project on ambient air quality in Antwerp, Journal of Environmental Planning and Management, vol. 62, no. 3, pp. 534551, 2018.

[15] "US Code of Federal Regulations, Title 40, Part 50 (40CFR§50.1)", Choice Reviews Online, vol. 50, no. 02, pp. 50-0634-50-0634, 2012.

[16] C. Noble, R. Vanderpool, T. Peters, F. McElroy, D. Gemmill and R. Wiener, "Federal Reference and Equivalent Methods for Measuring Fine Particulate Matter", Aerosol Science and Technology, vol. 34, no. 5, pp. 457-464, 2001.

[17] US EPA. (2018). Download Daily Data — US EPA. [online] Available at: https://www.epa.gov/outdoor-air-quality-data/download-daily-data [Accessed 14 Apr. 2018].

[18] United States Environmental Protection Agency, NAAQS Table, EPA, 20-Dec-2016. [Online]. Available: https://www.epa.gov/criteria-air-pollutants/naaqs-table. [Accessed: 19-Jun-2019].

[19] T. Sayahi, D. Kaufman, T. Becnel, K. Kaur, A. Butterfield, S. Collingwood, Y. Zhang, P-E. Gaillardon, K. Kelly, "Development of a Calibration Chamber to Evaluate the Performance of Low-cost Particulate Matter Sensors", *Environmental Pollution*, 2019.

[20] M. Mead, O. Popoola, G. Stewart, P. Landshoff, M. Calleja, M. Hayes, J. Baldovi, M. McLeod, T. Hodgson, J. Dicks, A. Lewis, J. Cohen, R. Baron, J. Saffell and R. Jones, "The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks", *Atmospheric Environment*, vol. 70, pp. 186-203, 2013.

[21] M. Bell, K. Ebisu and R. Peng, "Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research", *Journal of Exposure Science & Environmental Epidemiology*, vol. 21, no. 4, pp. 372-384, 2010.

[22] N. Zikova, M. Masiol, D. Chalupa, D. Rich, A. Ferro and P. Hopke, "Estimating Hourly Concentrations of PM2.5 across a Metropolitan Area Using Low-Cost Particle Monitors", *IEEE Sensors Journals*, vol. 17, no. 8, p. 1922, 2017.

[23] A. Al-Ali, I. Zualkernan and F. Aloul, "A Mobile GPRS-Sensors Array for Air Pollution Monitoring", *IEEE Sensors Journal*, vol. 10, no. 10, pp. 1666-1671, 2010.

[24] K. Vineeth, B. Vamshi and V. Mittal, "Identifying high pollution level regions through a terrestrial mobile monitoring system", *IEEE Region 10th Symposium (TENSYMP)*, 2017.

[25] Y. Wang, J. Li, H. Jing, Q. Zhang, J. Jiang, and P. Biswas, Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement, Aerosol Science and Technology, vol. 49, no. 11, pp. 10631077, 2015.

[26] A. Northcross, R. Edwards, M. Johnson, Z. Wang, K. Zhu, T. Allen and K. Smith, "A low-cost particle counter as a realtime fine-particle mass monitor", *Environ. Sci.: Processes Impacts*, vol. 15, no. 2, pp. 433-439, 2013.

[27] C. Clausen, A. Han, M. Kristensen and A. Bentien, "Optical sensor technology for simultaneous measurement of particle speed and concentration of micro sized particles", *IEEE Sensors Journal*, 2013.

[28] J. M. Bland and D. Altman, Statistical Methods For Assessing Agreement Between Two Methods Of Clinical Measurement, The Lancet, vol. 327, no. 8476, pp. 307310, 1986.

[29] F. M. J. Bulot, S. J. Johnston, P. J. Basford, N. H. C. Easton, M. Apetroaie-Cristea, G. L. Foster, A. K. R. Morris, S. J. Cox, and M. Loxham, Long-term field comparison of multiple low-cost particulate matter sensors in an outdoor urban environment, Scientific Reports, vol. 9, no. 1, 2019.

[30] T. Sayahi, A. Butterfield, and K. Kelly, Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors, Environmental Pollution, vol. 245, pp. 932940, 2019.

[31] M. Badura, P. Batog, A. Drzeniecka-Osiadacz, and P. Modzel, Evaluation of Low-Cost Sensors for Ambient PM2.5 Monitoring, Journal of Sensors, vol. 2018, pp. 116, 2018.

Thomas Becnel Thomas Becnel is working towards his Ph.D. in the Laboratory for NanoIntegrated Systems, led by Pierre-Emmanuel Gaillardon, in the department of Electrical and Computer Engineering, University of Utah, Salt Lake City, UT, USA. He received the electrical engineering degree from the University of Utah, and the M.Sc. degree in computer engineering from the University of Utah in 2018. His areas of research involve the design of large-scale sensor networks, low-noise capacitive CMOS sensors, and advanced low-power communication techniques. He plans to graduate with a Doctorate of Philosophy in the field of computer engineering in 2021.

Kyle Tingey Kyle Tingey obtained a Bachelor of Science in Computer Engineering from the University of Utah in 2017. His areas of specialty include custom environmental sensors, implantable medical devices, and wireless communications. Kyle co-founded Tetrad Sensor Network Solutions to translate the research he advanced during his undergraduate studies. He successfully developed devices and conducted experiments that demonstrated a network of hyperlocal air quality sensors at low cost. Kyle continues to improve others lives by developing innovative neurostimulators and other medical devices. His future goals are to expand Tetrads product line and furthering his education.

Jonathan Whitaker earned his B.S. degree in Computer Engineering from the University of Utah. Jonathan's research interests include real-time sensor networks, highly real-time embedded systems and software, as well as formal verification techniques. He has published papers regarding formal verification techniques for both the International Conference on Software Engineering (ICSE) and the Computer Aided Verification (CAV) Conference. In addition to his academic involvements, Jonathan is a large contributor to open source projects for embedded technologies and DIY projects spanning various domains.

Tofigh Sayahi Tofigh Sayahi is a Chemical Engineering PhD candidate at the University of Utah, Salt Lake City, Utah, USA. His main research focuses on laboratory and ambient calibration of low-cost air quality sensors. He holds a bachelors degree from Petroleum University of Technology, Ahvaz, Iran, and a masters degree from Ferdowsi University of Mashhad, Mashhad, Iran, both in Chemical Engineering. His research interests include the design of laboratory chamber to calibrate low-cost sensors, field evaluation of sensor networks, pool boiling heat transfer of nanofluids, and waste water treatment.

Katrina Lê Katrina Mỹ Quyên Lê is an aspiring engineer and paid intern at the Chemical Engineering Department at the University of Utah where she has worked on teaching module development and outreach as well as sensor calibration within the AQ&U Project for over two years. She is currently a freshmen in chemical engineering at the University of Utah, where she plans to work towards her Ph.D. in chemical engineering.

Pascal Goffin Pascal Goffin received his PhD in Computer Science from Université Paris-Saclay in France in 2016. During his PhD he worked for the Aviz visualization group at Inria. He holds a Masters degree in Computer Science from ETH Zurich in Switzerland. His interest span information visualization, text visualization, and human computer interaction. His current research focuses on how to support the communication of air quality in urban environments to citizens. He also builds tools to assist the exploration of urban air quality data.

Kerry Kelly Dr. Kelly is an Assistant Professor in Chemical Engineering at the University of Utah specializing in the links between energy, air quality and human health. She received her B.S. in Chemical Engineering from Purdue University, a M.S. in Environmental Engineering from the University of North Carolina-Chapel Hill, and her PhD in Environmental Engineering from the University of Utah. Her research is motivated by local and regional air-quality challenges. Dr. Kelly served 8 years on Utahs Air Quality Board, and she currently chairs Utahs Air Quality Policy Board. Her research currently includes projects to develop the next-generation of low-cost particulate matter sensors, to develop real-time estimates of particulate matter concentration and uncertainty, and to help engage high-school and middle-school students as citizen scientists. She was recently awarded the UCAIR person of the year by the governor for her work.

Anthony Butterfield Anthony Butterfield is an associate professor (lecturing) at the University of Utah. His research interests center around STEM community outreach, citizen scientist efforts, retention of underrepresented groups, and project base learning, particularly as applied to first-year students. He has been awarded the GLBT Educator Award from NOGLSTP and AIChEs 2017 Award for Innovation in Chemical Engineering Education. He is a member of the ASEE Chemical Engineering Division's board and AIChE's Societal Impact Operating Counsel.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2019.2941374, IEEE Internet of Things Journal

Pierre-Emmanuel Gaillardon Pierre-Emmanuel Gaillardon (S10M11SM16) received the Electrical Engineering degree from the cole Suprieure de Chimie Physique lectronique de Lyon, Lyon, France, in 2008, the M.Sc. degree in electrical engineering from the Institut National des Sciences Appliques de Lyon, Lyon, France, 2008, and the Ph.D. degree in electrical engineering from the CEA-LETI, Grenoble, France and the University of Lyon, Lyon, France, in 2011. He is currently an Assistant Professor with the Electrical and Computer Engineering Department and an Adjunct Professor with the School of Computing, The University of Utah, Salt Lake City, UT, USA, where he leads the Laboratory for NanoIntegrated Systems. Prior to joining The University of Utah, he was a Research Associate with the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, within the Laboratory of Integrated Systems (under the supervision of Prof. De Micheli), and a Visiting Research Associate with Stanford University, Palo Alto, CA, USA. Previously, he was a Research Assistant with the CEA-LETI. His research interests include the development of reconfigurable logic architectures and digital circuits exploiting emerging device technologies and novel EDA techniques. He is the recipient of the C-Innov 2011 Best Thesis Award and the Nanoarch 2012 Best Paper Award. He is an Associate Editor for the IEEE Transactions on Nanotechnology. He is a TPC member for many conferences, including DATE1518, DAC1618, and is a Reviewer for several journals and funding agencies. He is a Topic Cochair Emerging Technologies for Future Memories for DATE1718.