Understanding “Atmosome”, the Personal Atmospheric Exposome: Comprehensive Approach

Hari Bhimaraju; Nitish Nag, PhD; Vaibhav Pandey, PhD; Ramesh Jain, PhD
Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA, United States

Corresponding Author:
Hari Bhimaraju
Donald Bren School of Information and Computer Sciences
University of California
Donald Bren Hall, 6210
Irvine, CA, 92697
United States
Phone: 1 949 824 7427
Fax: 1 949 824 3976
Email: hari.bhimaraju@columbia.edu

Abstract

Background: Modern environmental health research extensively focuses on outdoor air pollutants and their effects on public health. However, research on monitoring and enhancing individual indoor air quality is lacking. The field of exposomics encompasses the totality of human environmental exposures and its effects on health. A subset of this exposome deals with atmospheric exposure, termed the “atmosome.” The atmosome plays a pivotal role in health and has significant effects on DNA, metabolism, skin integrity, and lung health.

Objective: The aim of this work is to develop a low-cost, comprehensive measurement system for collecting and analyzing atmosomic factors. The research explores the significance of the atmosome in personalized and preventive care for public health.

Methods: An internet of things microcontroller-based system is introduced and demonstrated. The system collects real-time indoor air quality data and posts it to the cloud for immediate access.

Results: The experimental results yield air quality measurements with an accuracy of 90% when compared with precalibrated commercial devices and demonstrate a direct correlation between lifestyle and air quality.

Conclusions: Quantifying the individual atmosome is a monumental step in advancing personalized health, medical research, and epidemiological research. The research explores the significance of the atmosome in personalized and preventive care for public health. Researchers will now have the data to model the impact of indoor air pollutants on the health of individuals, communities, and specific demographics, leading to novel approaches for predicting and preventing diseases.

(JMIR Biomed Eng 2021;6(4):e28920) doi: 10.2196/28920

KEYWORDS
exposome; exposomics; personal health; indoor air quality; health state estimation; health informatics; public health policy; epidemiology; embedded systems; internet of things

Introduction

At any moment in time, health is affected by various internal and external factors, such as the genome, microbiome, and exposome. The exposome consists of everything an individual is exposed to across his or her lifespan [1]. It considers lifestyle, occupation, socioeconomic factors, and the environmental conditions in which people live to develop an in-depth understanding of how an individual’s surroundings impact his/her health. The atmospheric exposome, a subset of the complete exposome and which is presented in this work, focuses on the health effects from the air that people breathe.

The term “atmosome” was coined to describe the atmospheric subset of an individual’s exposome. Common indoor air pollutants include PM$_{2.5}$ (particulate matter with diameter of $\leq 2.5$ μm), PM$_{10}$ (particulate matter with diameter of $\leq 10$ μm),...
carbon dioxide (CO₂), nitrogen dioxide (NO₂), carbon monoxide (CO), volatile organic compounds (VOCs), ozone (O₃), liquid petroleum gas (LPG), natural gas (NG), formaldehyde (HCHO), and biological contaminants such as bacteria and fungi. Measuring the quality of indoor air can provide insights into the potential adverse effects of poor air quality and preventative measures to keep the exposure cleaner. A cleaner atmosphere, in turn, has a positive impact on health and well-being. Thus, there is a need for a portable, real-time, multichannel air measurement system for enabling data-driven analytics and research [2].

A number of studies have been conducted by several federal, state, and local agencies that monitored, collected, and stored outdoor air quality data in the Environmental Protection Agency’s (EPA) Air Quality System database [3,4]. The Environmental Defense Fund in collaboration with Google Earth [5], the World Health Organization Global Urban Ambient Air Pollution Database [6], the World Air Quality Historical Database [7], and many other organizations generated air quality maps. These data are used for various modeling studies, to review policy implementation plans, and to generate reports for the Congress (US) [5]. However, these agencies have overlooked similar quantitative studies of indoor air quality (IAQ).

Previous studies have shown that indoor air is much more polluted than outdoor air and represents a major public health challenge especially in developing countries [8]. Many studies focused on specific or limited indoor air pollutants such as VOCs, CO₂, PM₂.₅, PM₁₀, and halogen flame-retardants [9-12]. Further, though potentially unexpected, there exist a myriad of well-defined sources of indoor air contamination and, correspondingly, numerous contaminants [13]. To improve air quality and minimize pollution-related disease and mortality, the atmosphere must be defined, measured, and analyzed to mitigate adverse environmental conditions and improve health outcomes. Thus, the motivation for this work is to use the atmosphere to further personalize an estimation of individual health from multimodal data [14,15].

According to the EPA, indoor air pollution is one of the top 5 environmental risks to public health. Annually, 9 out of every 10 people breathe air containing high levels of pollutants at some point [16,17]. Air pollutants can be in the form of pet dander, mold, dustmites, CO, radon, pests, lead, and secondhand smoke [17]. Americans spend approximately 90% of their time indoors, where the concentrations of some airborne pollutants are 2-5 times higher than those outdoors [13]. This poor IAQ can cause various infections, lung cancer, and chronic lung diseases, including asthma [18]. It can also contribute to the development of atherosclerosis, a root cause of many cardiovascular diseases [19]. In 2020, a long-term study of over 63 million US adults indicated a surprising correlation between PM₂.₅ and hospitalizations for severe neurological diseases [20].

According to the State of Global Air 2020 [21], nearly 500,000 newborns died in 2019 in their first month of life due to exposure to all types of air pollutants described previously. Household inhalation of mold spores and infant pulmonary hemorrhage are found to be linked in some studies [22]. Air pollution even impacts children while they are in their mothers’ womb [23,24], with the effect of air pollution on pregnant women and their fetuses comparable to smoking tobacco [21]. People often assume that indoor spaces are safe from outdoor air pollution, but this is inaccurate. Therefore, IAQ is a significant threat to public health [25]. The Program Needs for Indoor Environments Research (PNIER) document details EPA’s research needs for the indoor environment and recommends that the EPA and other governmental and private sector agencies and organizations address this issue [26].

The environmental research field of IAQ is nascent. Nonetheless, several researchers have recently demonstrated portable gas detection systems using various sensor technologies covering a limited set of analytes. For example, MQ sensors were embedded in a system where VOCs were detected [27]. MQ sensors are well-known to exhibit acceptable selectivity, but low sensitivity. Interestingly, the researchers implemented an artificial neural network and dramatically improved the sensitivity to gas concentrations at single-digit parts per million (ppm). However, the supported analyte set is far too small given the contaminants of concern, as reported by the EPA. A related research effort embedded photoionization detectors (PIDs) into a portable system for the detection of isobutylene, ethanol, propanol, and acetone [28]. PID sensors are well-known to exhibit high sensitivity (ie, on the order of parts per billion [ppb]), but lower selectivity than MQ sensors. Further, that work demonstrated only a small set of detected analytes. Other recent efforts include a portable system with embedded gas chromatography PID sensors [29]. This system detects benzene, toluene, and xylene. The researchers also employed an elegant algorithm using various quantification parameters (eg, pumping time, temperature) and calibration curves to optimize selectivity. This system also requires a pumping time up to 90 seconds and an analysis time of 10 minutes. Nonetheless, as expected, very high sensitivity was achieved, but again the analyte set is small and the system cost is substantially higher than an implementation with low-cost devices, such as MQ sensors. Last, none of these research efforts considered management of the data in an actionable manner. The systems reported were to demonstrate selectivity and sensitivity, the 2 most critical metrics for gas sensors.

This recent research focused on utilizing more common sensor technologies for portable applications. However, more exotic sensors have been developed recently in portable gas detection systems that have been demonstrated for indoor and outdoor use. For example, in [30], a mobile microscopy system (coined as the c-Air device) is presented and utilizes microscopy as a sensing technique and includes machine learning algorithms to increase accuracy. Further, it includes a mobile software app for data display. The device requires a sample of 6.5 L of air and an analysis time of 30 seconds. Also, similar to related studies, the c-Air device supports only a limited set of analytes, including total suspended particulates, PM₁₀, and PM₂.₅. Nonetheless, very good results are achieved and it is an advancement that the system is linked to a mobile app, though nothing actionable is reported.
In [31], a so-called portable cyber-physical system is presented for gas detection and is embodied in 2 distinct architectures, a stationary and portable device, each using well-established electronics related to this work, including an Arduino microcontroller and Raspberry Pi system-on-chip (SoC), corresponding to each embodiment. The 2 systems also use MQ sensors. However, the only sensors supported by the system are an MQ-4 (a methane sensor) and an MQ-8 (a hydrogen sensor). Further, although the system supports connectivity to the internet, the functionality is to merely upload the data to cloud storage. No manipulation or presentation of the data is reported.

Basic IAQ commercial-off-the-shelf (COTS) products also exist. Xiaomi, for example, offers a countertop product which detects PM$_{2.5}$, total VOC (tVOC), CO$_2$, temperature, and humidity [32]. It includes a touch screen and Wi-Fi connectivity. The company also offers handheld products that measure individual analytes, such as PM$_{2.5}$ [33]. Further, Xiaomi offers air purifiers that can be controlled by a mobile phone app and are marketed under the brand Mi. There exist many other related COTS products for IAQ monitoring. Although interesting, these products measure a very small set of analytes and offer minimal, if any, actionable information based on the data collected.

In contrast to both the recent research and commercial work, this paper introduces a patent pending and low-cost embedded system called the Atmosome Measurement System (AMS) [34], which can reliably, accurately, and instantaneously monitor and measure significantly more indoor air pollutants such as PM$_{2.5}$, PM$_{10}$, CO$_2$, NO$_x$, CO, VOCs, O$_3$, LPG, NG, equivalent CO$_2$ (eCO$_2$), hydrogen as well as environmental parameters including temperature, humidity, pressure, and altitude. This set of analytes was chosen as it represents the primary sources of indoor air pollution as well as the leading causes of adverse impact on human respiratory health, according to the EPA. Further, the presented work is implemented as a scalable and low-cost embedded system utilizing COTS electronics including an Arduino microcontroller, Raspberry Pi SoC, MQ gas sensors, and simple environmental sensors. MQ sensors were selected for low cost, high selectivity, and sufficient sensitivity. Besides, in contrast to previous studies [10,12], AMS is built with the goal of providing a cloud-based infrastructure that stores, analyzes, and presents insights into IAQ and trends that correlate with personal lifestyles. It displays historic and real-time data from multiple sensors in a user-friendly web application, enabling users to interpret their data, and recommends environmental changes to improve personal atmosphere conditions.

Therefore, the development of a system that can evaluate IAQ by using multiple analytes, process and visualize pollutant data, recommend remediation steps, and be built at an affordable price point is the foundation of this research. The system can be configured with a variety of optional customizations including the frequency at which the users would like to monitor their air quality; their geographic location details including the zip code, city, state, and country; the indoor space details such as home, office, or car; the location within the space such as kitchen, bedroom, garage; and the activity details such as cleaning, cooking, routine. Further, AMS supports representational state transfer (REST) application programming interface (API) to download data for further exploration and analytics. AMS also includes an option for the users to anonymously share their data to further indoor air pollution research. This opens the possibility of developing a public indoor air quality database while maintaining user confidentiality allowing for extended research on indoor air quality, its impacts, and health policy modeling.

Taken together, AMS can be a useful tool for improving public health outcomes as it can provide the necessary data that people need to manage their IAQ in a cost-effective and convenient manner. Moreover, the data can display different atmospheres between 2 neighbors or within neighborhoods of different socioeconomic classes, which can be useful for public health officials or policy researchers that work toward enhancing the health of citizens.

Comparisons of AMS with various COTS products are available in Multimedia Appendix 1 [35-41]. AMS stood out in both number of analytes and cost compared with the nearest COTS product, Aeroqual, that covers multiple analytes [35]. Comparisons are also made with recent studies and those details are available in Multimedia Appendix 2 [42-47]. Aspects such as data sampling duration and pollutant streams in AMS are found to be much more extensive than similar indoor air quality assessments in a college campus [44] and homes in a temperate region of the United States [46]. Comparisons are not made with the cited research because the analyte sets are dramatically smaller than those supported by AMS and there is no actionable interpretation of the data in the studies, because they focus primarily on sensor selectivity and sensitivity, as described previously. Further, few of those efforts included internet connectivity and none included any actionable information based on the data collected.

**Methods**

**Study Approach and Design**

An experimental approach combining internet of things (IoT) hardware and software development was used to measure air pollutants and air quality metrics. The researchers used AMS to nonintrusively monitor air quality through daily indoor life. AMS provided visuals and recorded trends that could indirectly indicate the relationship between lifestyle and observed pollutant values.

In this work, the researchers used AMS to collect air quality data indoors at home (in Cupertino, California; South Lake Tahoe, California; and Hyderabad, Telangana, India) as well as during local commute in the United States and during an inflight journey from the United States to India. These locations were chosen to monitor and evaluate the performance of AMS in environments associated with distinct indoor air quality profiles and climates. The device, similar to a thermostat placed in a room, is completely noninvasive. To initiate readings, it is powered on with a USB cable (or a wall socket) and AMS software is launched by the user. The studies were conducted, intermittently, between January 2020 and January 2021.

After initial calibration, to ensure the continued accuracy of the measurements made by AMS on an ongoing basis, the sensors

https://biomedeng.jmir.org/2021/4/e28920
were recalibrated every 3 months. Measurements were repeated over similar activities in several different timeframes to better analyze and predict possible relationships between activities and associated pollutants. To control for bias, measurements were also carried out across several recalibrations.

**System Architecture**

AMS comprises 3 distinct modules that collect, analyze and store, and display data. As shown in Figure 1, these 3 interconnected modules include the data collection system (DCS), data analysis and storage engine (DASE), and user interface (UI). The DCS is a sensor system that collects the user’s environmental data in user-defined intervals. The data gathered are then analyzed and stored by the system’s DASE, and the air quality metrics obtained are displayed by the system’s web-based UI.

**Materials and Software**

Materials used to build the DCS and the software used to create DASE and UI are discussed in the following sections.

**Data Collection System**

**Overview**

The DCS is an IoT air monitoring sensor system for acquiring an individual’s unique geospatial data to track air quality. It includes 17 environmental sensors that measure 22 different air pollutant data streams, an Arduino Mega, a Raspberry Pi, and a power source. Referring back to Figure 1, the Arduino Mega is a microcontroller that captures sensor data and the Raspberry Pi is an SoC which supports Wi-Fi and enables an interface to the cloud.

**Sensors**

The DCS monitors environmental conditions such as temperature, pressure, altitude, humidity, and various analytes including PM$_{2.5}$, PM$_{10}$, CO, O$_3$, CO$_2$, eCO$_2$, tVOCs, LPG, methane, hydrogen, flammable gases, aromatic compounds, hydrogen sulfide, ammonia, nitrogen oxide, NG, and HCHO as shown in Table 1. Metal oxide semiconductor (MOS) gas sensors were used to detect the different air pollutants. MOS-based sensors detect the concentration of various kinds of gases by acting as a chemiresistor, where a change in resistance of the metal oxide occurs due to the adsorption of specific gases. These sensors, and specifically MQ series sensors, are ideally suited for low-cost and low-power applications in indoor environments. Selectivity to certain gases is dependent on the specific sensor model, which is indicated numerically. Some MQ sensors are sensitive to multiple gases; for example, both MQ-5 and MQ-6 measure LPG, but MQ-6 exhibits higher selectivity and sensitivity to LPG and is calibrated for that particular gas. As individual sensors are calibrated for their specific gases, they are less selective to other gases. A total of 17 different kinds of sensors were used to detect and measure the level of pollutants as well as environmental parameters (Table 1).
Table 1. List of sensors, including target gas analytes and parameters as well as sensor type, in the data collection system implementation. The system supports 22 data streams from 17 different sensors.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Analyte/parameter</th>
<th>Sensor type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS811/BME280</td>
<td>Total volatile organic compounds, equivalent carbon dioxide, temperature, humidity, pressure, altitude</td>
<td>Environmental combo sensor</td>
</tr>
<tr>
<td>CO2</td>
<td>Carbon dioxide</td>
<td>Nondispersive infrared</td>
</tr>
<tr>
<td>PM2.5</td>
<td>Particulate matter 2.5</td>
<td>Optical, infrared-emitting diode</td>
</tr>
<tr>
<td>MQ2</td>
<td>Smoke/particulate matter 10</td>
<td>MOS(^a)</td>
</tr>
<tr>
<td>MQ4</td>
<td>Methane</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ6</td>
<td>Liquefied petroleum gas</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ7</td>
<td>Carbon monoxide</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ131</td>
<td>Ozone</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ3</td>
<td>Alcohol (ethyl alcohol)</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ5</td>
<td>Natural gas</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ8</td>
<td>Hydrogen</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ9</td>
<td>Flammable gases</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ135</td>
<td>Aromatic compounds</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ136</td>
<td>Hydrogen sulfide</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>MQ137</td>
<td>Ammonia</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>NO(_x)</td>
<td>Nitrogen oxides</td>
<td>MOS gas sensor</td>
</tr>
<tr>
<td>HCHO</td>
<td>Formaldehyde</td>
<td>MOS gas sensor</td>
</tr>
</tbody>
</table>

\(^a\)MOS: metal oxide semiconductor.

Realized AMS

Figure 2 presents the complete realized AMS including sensors and the Raspberry Pi platform, which are securely mounted to an FR4 printed circuit board (PCB) and an Arduino Mega 2560 microcontroller reverse mounted to the same. The Arduino Mega microcontroller is programmed to capture data from the sensors. All of these system components and associated positions on the PCB are labeled clearly in Figure 2.

Figure 2. A photograph of the populated PCB with the Arduino (reverse mounted) and 17 sensors listed in Table 1. CO: carbon monoxide; CO\(_2\): carbon dioxide; H\(_2\)S: hydrogen sulfide; HCHO: formaldehyde; LPG: liquid petroleum gas; NG: natural gas; NH\(_3\): ammonia; PCB: printed circuit board; PM\(_{2.5}\): particulate matter with diameter of 2.5 \(\mu\)m or less; PM\(_{10}\): particulate matter with diameter of 10 \(\mu\)m or less; VOC: volatile organic compound.
IoT

The sensor PCB is connected using a USB cable linked to a computer (either a Raspberry Pi or a laptop that runs the DCS software). The computer receives sensor data from Arduino Mega microcontroller and posts it to the cloud (DASE) server or saves it in local memory in the absence of an internet connection. The PCB comes in 2 variations: a portable model that consists of 8 different sensors providing 13 data streams and powered by the USB; and a high-power model that has 17 sensors providing 22 data streams and which must be powered by a wall socket. Future work will aim to make the high-power model portable.

Data Storage and Analysis Engine

DASE runs in the cloud on Amazon Web Services (AWS) and stores the sensor data in a Postgres Database. AWS was selected for simplicity of implementation, low cost, and well-known user data security as shown in a recent case study [48]. Further, no personally identifiable information is collected for this study. The username and zip code are stored in the cloud and password protected. The software implementation is built using the Python Flask Framework and supports REST API to support receiving data from the sensor board at user location (DCS) and for sending data to the UI or alerts to the user. The source module layout and libraries are shown in Figure 3.

Figure 3. DASE (data analysis and storage engine) programming components and functional flow diagram.

The backend of the system is based on Python and Flask and has 3 distinct interconnected components that make up the Atmosome engine: the web application framework, the database model that receives data from the cloud database, and the numerical analysis libraries that operate on the data received from the user’s device and surface them to the user via the front end UI. To accomplish this workflow, DASE employs various Flask web framework components including flaskmigrate (to support modifications to existing DB tables), sqlalchemy (an object-relational mapper that enables reading from/writing to the DB python data objects without the need for using the DB’s SQL), flasksqlalchemy (to manage DB connections and sessions), flaskcors (to enable cross-domain REST API communication between the server and the UI), flaskmarshmallow (to handle python-json serialization/deserialization for REST API communication), numpy (Python’s multidimensional numerical analysis library), pandas (Python’s DataFrame support much like a database table...
in memory), and matplotlib (Python’s data visualization library). flaskproject.models defines the DB tables in python, flaskproject.config defines the web application configurations, for example, the type of database being used, and flaskproject.app is the code that handles all data computations.

**User Interface**

The UI is built as a progressive web application using the React Framework. This makes it available on any device with a web browser. Additionally, because it is a progressive web app, it automatically adjusts to the size of mobile platforms, and thereby presents a user experience similar to a native app. The source code is structured and modularized into pages and components within pages.

For example, Figure 4 illustrates the 6 javascript (.js) source files that render the various sections of the Atmosome dashboard. *header.js* is the header of the page, and displays the page title, user location, date, and time. *gauge.js* displays the AQI gauge which gives an “at-one-glance” state of the current indoor air conditions. *weatherParameters.js* displays the temperature, humidity, pressure, and altitude. Although weather metrics are not pollutants, they are a part of lifestyle conditions and are recorded along with information on pollutants. *pollutants.js* displays numeric and visual information about each of the pollutants. *averageExposure.js* displays the user’s average exposure to various pollutants in different intervals of time. *timeseries.js* displays the quarterly time series graphs of the user’s historic exposure to various pollutants. The UI is designed mainly to render data and does not store data or perform computations. It makes REST API calls to DASE to retrieve the data.

**Figure 4.** UI dashboard software and display components. The dashboard showcases the level of various pollutants (eg, PM$_{2.5}$, VOC, ozone, CO$_2$), weather metrics (eg, temperature, humidity, pressure, and altitude), Air Quality Index, average exposure levels over time, and quarterly temporal graphs reflecting the user’s historic exposure data. CO$_2$: carbon dioxide; PM$_{2.5}$: particulate matter with diameter of 2.5 μm or less; UI: user interface; VOC: volatile organic compound.

**Procedure**

In this section, the methods involved in calibrating the DCS device, collecting data from a setting, and posting data to DASE will be discussed. Moreover, the analyzed data can be viewed on UI screens for the given set up.
**Sensor Calibration**

The sensors in AMS’s DCS are first “burned-in,” meaning that they are placed in an environment with clean air and operated with active power for 48 hours. Next, AMS’s DCS is run through a meticulous calibration process to ensure accuracy. A flow diagram of the calibration algorithm is illustrated in Figure 5.

Figure 5. DCS calibration algorithm flow diagram. CO: carbon monoxide; CO\(_2\): carbon dioxide; DCS: data collection system; LPG: liquefied petroleum gas; NG: natural gas; O\(_3\): ozone; PM\(_{2.5}\): particulate matter with diameter of 2.5 μm or less; PM\(_{10}\): particulate matter with diameter of 10 μm or less; tVOC: total volatile organic compound.

To ensure the continued accuracy of the new system, the sensors were recalibrated every 3 months, and new training improves the accuracy of the data correction algorithms. Upon continued comparison of the DCS postcalibration values with precalibrated COTS devices, over 90% accuracy postcalibration was achieved when compared with precalibrated COTS devices. Alternative calibration approaches, such as calibrating outdoors against the values of the nearest outdoor weather station, were explored and similar accuracy was noted.

**Data Collection**

Once the user receives the portable DCS, they need to power it and connect it to a Wi-Fi network, if available. Then they need to launch the application. Once launched, the user can change any of their default settings, or retain the defaults, and initiate collection of air quality data.

Although AMS is intended to measure IAQ, part of the calibration routine was performed outdoors. This decision was made to compare results easily with the nearest, trusted outdoor air pollution station. Further, the environmental differences between the indoor and outdoor settings were not notable enough to introduce a substantial error. Certainly, an entirely indoor calibration routine would yield higher sensor accuracy, but precise indoor calibration techniques are complex and not suitable for consumer use. For example, precision routines to calibrate gas sensors are typically performed in the presence of a high concentration source of each analyte. Such an approach was deemed unrealistic and prohibitively expensive.

The researcher (HB) used AMS to monitor air quality through daily life at home, such as in the kitchen, living room, bedroom, home office, and during commute. The researcher also measured air quality during air travel to India, and at home in India and the United States.

**Data Transmission**

Local Storage and Transmission to Cloud

In the presence of Wi-Fi, as the DCS program collects air quality data from the sensors at the frequency specified by the user, it concurrently transmits the data to the cloud and requires no user interaction. In the absence of Wi-Fi, the DCS stores the data in local memory. Once a Wi-Fi connection is available, the user can connect the DCS to the Wi-Fi, and select an icon in the UI that indicates “Upload AMS data to the Cloud.” This loads the data to the cloud, and once complete, deletes the locally stored...
data, automatically removing the no-longer required data from local storage.

**Post/Receive Data to/From Cloud**

The DCS software wraps the readings from each of its sensors into JSON, a simple format, with a series of key value pairs, used to store and transmit data to DASE, which is running in the cloud. Once a JSON payload with the values of each pollutant, atmospheric data, and sampling location information has been assembled (a portion of which is shown in Figure 6) it is ready to be transmitted via the internet to DASE.

**Figure 6.** Data transmission: REST API POST data from DCS to DASE containing sensor types and values. API: application programming interface; DASE: data analysis and storage engine; DCS: data collection system; REST: representational state transfer.

The DCS software uses REST APIs to interact with DASE. REST APIs use HTTP requests to interact with a remote web server. The HTTP GET method is used to receive data from the server, and the HTTP POST method is used to send data to the remote server.

The DCS software sends this JSON payload by making a REST API POST request to DASE, which receives the data, and subsequently analyzes and stores them. Upon receiving an HTTP GET request from the UI to display data, DASE formats the information required by the UI into the JSON payload and sends it, as illustrated in Figure 7. The UI then displays these data.
Data Storage

DASE stores the received sensors data in its Postgres database. Postgres is a free database and is seamlessly integrated into DASE’s Python Flask framework to store and retrieve data using REST API calls. Currently, there is no limit on the amount of time the user’s history data are stored in DASE and there is no user action involved in this step.

AQI Calculation

General Equation

The Indoor Air Quality Index is calculated using the weighted mean formula. The contribution of each pollutant is multiplied by its weightage, whose calculation is explained in the next section, and divided by the sum of pollutant weightages. This can be represented by the formula:

$$\text{AQI} = \sum_{i=1}^{n} \frac{(W_i) \times (P_i)}{\sum_{i=1}^{n} (W_i)}$$

where $W_i$ is weightage of pollutant $i$ and $P_i$ is the reading of pollutant $i$.

Calculation of Pollutant Weightages

Weights for each of the pollutants have been calculated based on the concentrated means and their contribution to different AQI levels.

Individual AQIs (Table 2) and the breaking points for the concentration mean of different pollutants in a fixed cycle were used to arrive at the weights of each pollutant at the respective AQI levels. Weightage of each pollutant’s mean concentration at each AQI level was calculated by measuring the fractional contribution to the AQI.

Figure 7. JSON formatted REST API GET data from DASE to UI containing attributes stored for each pollutant and weather metric. API: application programming interface; DASE: data analysis and storage engine; REST: representational state transfer; UI: user interface.
Table 2. Individual Air Quality Indexes and the breaking points for the concentration mean of pollutants [49]a,b.

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>Sulfur dioxide 24 hours</th>
<th>Sulfur dioxide 1 hour</th>
<th>Nitrogen dioxide 24 hours</th>
<th>Nitrogen dioxide 1 hour</th>
<th>PM_{2.5}^c 24 hours</th>
<th>Carbon monoxide 24 hours</th>
<th>Carbon monoxide 1 hour</th>
<th>Ozone 1 hour</th>
<th>Ozone 8 hours</th>
<th>PM_{2.5}^c 24 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>150</td>
<td>40</td>
<td>100</td>
<td>50</td>
<td>2</td>
<td>5</td>
<td>160</td>
<td>100</td>
<td>35</td>
</tr>
<tr>
<td>100</td>
<td>150</td>
<td>500</td>
<td>80</td>
<td>200</td>
<td>150</td>
<td>4</td>
<td>10</td>
<td>200</td>
<td>160</td>
<td>75</td>
</tr>
<tr>
<td>150</td>
<td>475</td>
<td>650</td>
<td>180</td>
<td>700</td>
<td>250</td>
<td>14</td>
<td>35</td>
<td>300</td>
<td>215</td>
<td>115</td>
</tr>
<tr>
<td>200</td>
<td>800</td>
<td>800</td>
<td>280</td>
<td>1200</td>
<td>350</td>
<td>24</td>
<td>60</td>
<td>400</td>
<td>265</td>
<td>150</td>
</tr>
<tr>
<td>300</td>
<td>1600</td>
<td>f</td>
<td>565</td>
<td>2340</td>
<td>420</td>
<td>36</td>
<td>90</td>
<td>800</td>
<td>800</td>
<td>250</td>
</tr>
<tr>
<td>400</td>
<td>2100</td>
<td>f</td>
<td>750</td>
<td>3090</td>
<td>500</td>
<td>48</td>
<td>120</td>
<td>1000</td>
<td>g</td>
<td>350</td>
</tr>
<tr>
<td>500</td>
<td>2600</td>
<td>f</td>
<td>940</td>
<td>3840</td>
<td>600</td>
<td>60</td>
<td>150</td>
<td>1200</td>
<td>g</td>
<td>500</td>
</tr>
</tbody>
</table>

aData presented are mean values.
bSulfur dioxide (not collected by AMS) and nitrogen dioxide, which are primarily outside pollutants, are excluded from the indoor Air Quality Index calculation. Carbon dioxide and volatile organic compounds are much more common indoors and are more relevant and considered in the indoor Air Quality Index calculation.
cThe concentration means of 1-hour sulfur dioxide, nitrogen dioxide, and carbon monoxide just adapt to the real-time calculation for Indoor Air Quality Index, but the concentration means of 24-hour sulfur dioxide, nitrogen dioxide, and carbon monoxide were used to calculate for a whole day.
dPM_{10}: particulate matter with diameter of ≤10 μm.
ePM_{2.5}: particulate matter with diameter of ≤2.5 μm.
fThe concentration mean of 1-hour sulfur dioxide higher than 800 μg/m^3 is calculated with the concentration mean of 24-hour sulfur dioxide.
gThe concentration mean of 8-hour ozone higher than 800 μg/m^3 is calculated with the concentration mean of 1-hour ozone.

**VOC Information**

Values from Table 3 have been considered for the corresponding AQI windows, with the midpoint of concentration breaking points chosen as pollutant representation, which is used as the denominator, and the breaking point of the IAQI range as the numerator, similar to calculations performed for other pollutants.

Table 3. Individual AQIs and the breaking points for the concentration mean of VOCs and others [50].

<table>
<thead>
<tr>
<th>Level</th>
<th>AQI^a range</th>
<th>VOC^b concentration (BP_{LO}–BP_{HI})^c</th>
<th>CO^d concentration (BP_{LO}–BP_{HI})^d</th>
<th>PM^e concentration (BP_{LO}–BP_{HI})^e</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0-50</td>
<td>0-200</td>
<td>0-4.99</td>
<td>0-30</td>
<td>Good</td>
</tr>
<tr>
<td>B</td>
<td>51-100</td>
<td>201-350</td>
<td>5-9.99</td>
<td>31-90</td>
<td>Moderate</td>
</tr>
<tr>
<td>C</td>
<td>101-250</td>
<td>351-500</td>
<td>10-14.99</td>
<td>91-140</td>
<td>Unhealthy</td>
</tr>
<tr>
<td>D</td>
<td>251-400</td>
<td>501-757</td>
<td>15-2000</td>
<td>141-750</td>
<td>Very unhealthy</td>
</tr>
</tbody>
</table>

^aAQI: Air Quality Index.
^bVOC: volatile organic compound.
^cBP: breaking point (LO: low; HI: high).
^dCO: carbon monoxide.
^ePM: particulate matter.

**CO₂ Information**

Values from Table 4 have been considered for the corresponding AQI windows, with the midpoint of pollutant concentration calculated and chosen as pollutant representation, which is used as the denominator, and the breaking point of the IAQI range as the numerator, as is the case with others.
Table 4. Individual Air Quality Indexes and the pollutant concentration ranges of carbon dioxide and others [51].

<table>
<thead>
<tr>
<th>Carbon monoxide (ppm)</th>
<th>Carbon dioxide (ppm)</th>
<th>Hydrogen (ppm)</th>
<th>Ammonia (ppm)</th>
<th>Ethanol (ppm)</th>
<th>Hydrogen sulfide (ppm)</th>
<th>Toluene (ppm)</th>
<th>Oxygen (%)</th>
<th>Indoor Air Quality Index</th>
<th>Health effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-0.2</td>
<td>0-379</td>
<td>0-1</td>
<td>0-24</td>
<td>0-0.49</td>
<td>0-0.00033</td>
<td>0-0.0247</td>
<td>20.95</td>
<td>0-50</td>
<td>Good</td>
</tr>
<tr>
<td>0.21-2</td>
<td>380-450</td>
<td>1.1-2</td>
<td>25-30</td>
<td>0.5-10</td>
<td>0.00034-1.5</td>
<td>0.0248-0.6</td>
<td>19-20.9</td>
<td>51-100</td>
<td>Moderate</td>
</tr>
<tr>
<td>2.1-9</td>
<td>451-1000</td>
<td>2.1-3</td>
<td>31-50</td>
<td>11-49</td>
<td>1.6-5</td>
<td>0.7-1.6</td>
<td>15-19</td>
<td>101-150</td>
<td>Unhealthy for sensitive individuals</td>
</tr>
<tr>
<td>9.1-15.4</td>
<td>1001-5000</td>
<td>3.1-5</td>
<td>51-100</td>
<td>50-100</td>
<td>6-20</td>
<td>1.7-9.8</td>
<td>12-15</td>
<td>151-200</td>
<td>Unhealthy</td>
</tr>
<tr>
<td>15.5-30.4</td>
<td>5001-30,000</td>
<td>5.1-8</td>
<td>101-400</td>
<td>101-700</td>
<td>21-50</td>
<td>9.9-12.2</td>
<td>10-12</td>
<td>201-300</td>
<td>Very unhealthy</td>
</tr>
<tr>
<td>30.5-50.4</td>
<td>30,001-40,000</td>
<td>8.1-10</td>
<td>401-500</td>
<td>701-1000</td>
<td>51-100</td>
<td>12.3-100</td>
<td>&lt;10</td>
<td>301-400</td>
<td>Hazardous</td>
</tr>
</tbody>
</table>

The final weightage factor of each pollutant was calculated by taking the arithmetic mean as shown in Tables 5-10.

**Pollutant Weightage Formula**

\[ W_p = \left( \sum_{i=1}^{n} \frac{\text{AQI}_i}{\text{PCM}_i} \right) / n \]

where \( W_p \) is the weightage of pollutant; \( \text{AQI}_i \) is the individual AQI pollution level; and \( \text{PCM}_i \) is the corresponding concentration threshold mean of a pollutant.

**Weightage Worksheet Tables by Pollutant**

Table 5. Weightage calculation for PM\(_{10}\)\(^{a}\).

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>PM(_{10}) 24-hour mean</th>
<th>PM(_{10}) weightage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>50</td>
<td>1</td>
</tr>
<tr>
<td>100</td>
<td>150</td>
<td>0.666666667</td>
</tr>
<tr>
<td>150</td>
<td>250</td>
<td>0.6</td>
</tr>
<tr>
<td>200</td>
<td>350</td>
<td>0.571428571</td>
</tr>
<tr>
<td>300</td>
<td>420</td>
<td>0.714285714</td>
</tr>
<tr>
<td>400</td>
<td>500</td>
<td>0.8</td>
</tr>
<tr>
<td>500</td>
<td>600</td>
<td>0.833333333</td>
</tr>
<tr>
<td>Final weightage</td>
<td></td>
<td>0.740816327</td>
</tr>
</tbody>
</table>

\(^{a}\)PM\(_{10}\): particulate matter with diameter of \(\leq 10\) \(\mu\)m.

Table 6. Weightage calculation for CO\(^{a}\).

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>CO 1-hour mean</th>
<th>CO weightage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>150</td>
<td>35</td>
<td>4.28571429</td>
</tr>
<tr>
<td>200</td>
<td>60</td>
<td>3.33333333</td>
</tr>
<tr>
<td>300</td>
<td>90</td>
<td>3.33333333</td>
</tr>
<tr>
<td>400</td>
<td>120</td>
<td>3.33333333</td>
</tr>
<tr>
<td>500</td>
<td>150</td>
<td>3.33333333</td>
</tr>
<tr>
<td>Final weightage</td>
<td></td>
<td>5.37414966</td>
</tr>
</tbody>
</table>

\(^{a}\)CO: carbon monoxide.
### Table 7. Weightage calculation for O$_3^a$.

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>O$_3$ 1-hour mean</th>
<th>O$_3$ weightage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>160</td>
<td>0.3125</td>
</tr>
<tr>
<td>100</td>
<td>200</td>
<td>0.5</td>
</tr>
<tr>
<td>150</td>
<td>300</td>
<td>0.5</td>
</tr>
<tr>
<td>200</td>
<td>400</td>
<td>0.5</td>
</tr>
<tr>
<td>300</td>
<td>800</td>
<td>0.375</td>
</tr>
<tr>
<td>400</td>
<td>1000</td>
<td>0.4</td>
</tr>
<tr>
<td>500</td>
<td>1200</td>
<td>0.416666667</td>
</tr>
</tbody>
</table>

Final weightage 0.429166667

*O$_3$: ozone.

### Table 8. Weightage calculation for PM$_{2.5}^a$.

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>PM$_{2.5}$ 24-hour mean</th>
<th>PM$_{2.5}$ weightage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>35</td>
<td>1.428571429</td>
</tr>
<tr>
<td>100</td>
<td>75</td>
<td>1.333333333</td>
</tr>
<tr>
<td>150</td>
<td>115</td>
<td>1.304347826</td>
</tr>
<tr>
<td>200</td>
<td>150</td>
<td>1.333333333</td>
</tr>
<tr>
<td>300</td>
<td>250</td>
<td>1.2</td>
</tr>
<tr>
<td>400</td>
<td>350</td>
<td>1.142857143</td>
</tr>
<tr>
<td>500</td>
<td>500</td>
<td>1</td>
</tr>
</tbody>
</table>

Final weightage 1.248920438

*PM$_{2.5}$: particulate matter with diameter of ≤2.5 μm.

### Table 9. Weightage calculation for tVOC$^a$.

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>tVOC mean</th>
<th>tVOC weightage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>100</td>
<td>0.5</td>
</tr>
<tr>
<td>100</td>
<td>275</td>
<td>0.363636364</td>
</tr>
<tr>
<td>250</td>
<td>425</td>
<td>0.588235294</td>
</tr>
<tr>
<td>400</td>
<td>628</td>
<td>0.636942675</td>
</tr>
</tbody>
</table>

Final weightage 0.522203583

*tVOC: total volatile organic compound.
Table 10. Weightage calculation for CO$_2$$^a$.

<table>
<thead>
<tr>
<th>Indoor Air Quality Index</th>
<th>CO$_2$ mean</th>
<th>CO$_2$ weightage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>341</td>
<td>0.146627566</td>
</tr>
<tr>
<td>100</td>
<td>747</td>
<td>0.133868809</td>
</tr>
<tr>
<td>150</td>
<td>1305</td>
<td>0.114942529</td>
</tr>
<tr>
<td>200</td>
<td>5400</td>
<td>0.037037037</td>
</tr>
<tr>
<td>300</td>
<td>31,500</td>
<td>0.00952381</td>
</tr>
<tr>
<td>400</td>
<td>63,000</td>
<td>0.006349206</td>
</tr>
<tr>
<td>Final weightage</td>
<td></td>
<td>0.074724826</td>
</tr>
</tbody>
</table>

$^a$CO$_2$: carbon dioxide.

Data Analysis
The data analysis flowchart is depicted in Figure 8. Analysis of the collected data from the DCS is automatically executed in the background and is transparent to the user. Upon receiving a request from the UI, or through the background user alert process, the database is queried to retrieve data. The data are validated against thresholds predetermined by environmental safety limits [52,53]. If the user is registered to receive alerts and the values exceed safe thresholds, an alert is sent to the user. Additionally, the values and corresponding qualitative metrics of each pollutant are determined. Finally, AQI is computed and the data are returned.
The authors considered including an adaptive threshold for each analyte to account for spatial, temporal, and environmental variations. However, AMS is recalibrated in each new usage location and is intended for indoor use. Thus, environmental variations are small and no substantial error between AMS and COTS sensors was observed when analyzing the data. Nonetheless, it is well-known that MQ sensors exhibit high temperature coefficients and sensitivity to humidity [54]. Drift of these sensors is low and ongoing recalibration every 3 months is more than sufficient to address such drift [54]. Considering these factors, the ranges at which the AQI considered are absolute, but the unit itself is adjusted to account for location. Further, the sample rate is considered sufficient for the application. In the data that follow, it does not appear that the sensors are undersampled, so temporal variation is not considered in this embodiment. Nonetheless, one of the goals for the future work is to automate the calibration process and also allow customization of the calibration interval according to the user’s choice.
Data Alerts

If a user opts to be alerted when pollutants are above the optimal threshold, DASE sends an email and SMS text message alerts. This could be invaluable in preventing accidents or calling for emergency services in the case of an NG leakage or similar emergency.

Data Display

Current UI Module and Plans for the Next Version

The UI module displays the air quality information to users. The user accesses the UI at the atmosome website [55]. There is a dropdown menu to select the user’s zip code.

The next version of the atmosome UI will include a login screen for the user, instead of the current zip code selection. The various parts of the UI displayed to the user are described below.

**Hourly Dashboard**

The dashboard (Figure 9) presents the user with an “at-a-glance” state of their indoor air quality using a gauge that reflects AQI. The gauge conforms to the conventions of the US AQI gauge. The AQI is computed based on the pollutants that show the most variability due to user lifestyle, including PM$_{2.5}$, tVOC, CO$_2$, O$_3$, CO, and PM$_{10}$. The past hourly averages sample of 10 pollutant analytes is presented visually and quantitatively. Thresholds for each analyte are predetermined by environmental safety limits. Relative assessments of levels are color coded and reported as GOOD, MODERATE, POOR, or BAD. Four environmental parameters, including temperature, pressure, humidity, and altitude, are also shown quantitatively.

![Atmosome UI](image)

**Figure 9.** Representative example of hourly dashboard of user’s atmosome data displayed by AMS UI. The values of 4 environmental parameters are also quantitatively presented below the AQI gauge. AMS: Atmosome Measurement System; AQI: Air Quality Index; PM$_{10}$: particulate matter with diameter of 10 $\mu$m or less; UI: user interface.

**Recommendations**

Each pollutant in the dashboard is associated with detailed information about acceptable thresholds and specific suggestions on how to manage it to be within healthy limits. This is shown in Figures 10 and 11.
Figure 10. Lifestyle recommendations to improve atmosphere quality, provided to the user in AMS UI. Clicking on the “Tips” link of each of the pollutants in the UI dashboard, as shown, opens up a popup box showing the quantitative thresholds associated with the pollutant and gives the user recommendations on managing it to maintain healthy levels. AMS: Atmosome Measurement System; ppb: parts per billion; UI: user interface; VOC: volatile organic compound.

Clicking on “Tips” shows the recommendations

Figure 11. Examples in the AMS UI that show the recommendations and thresholds for NG and PM2.5. AMS: Atmosome Measurement System; NG: natural gas; PM$_{2.5}$: particulate matter with diameter of 2.5 μm or less; UI: user interface.

<table>
<thead>
<tr>
<th>NG</th>
<th>PM2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Leave area of suspected leak as quickly as possible</td>
<td>• Avoid smoke, burning wood, candles, incense, etc.</td>
</tr>
<tr>
<td>• Warn others to stay out of the area</td>
<td>• Ventilate and use exhaust fans</td>
</tr>
<tr>
<td>• Call local utility preferably or 911</td>
<td>• Use exhaust fan while cooking</td>
</tr>
<tr>
<td>• Further reading: <a href="https://www.peoples-gas.com/all-about-gas/safety/smell/what-to-do.php">https://www.peoples-gas.com/all-about-gas/safety/smell/what-to-do.php</a></td>
<td>• Use HEPA filters to purify air</td>
</tr>
</tbody>
</table>

Further reading: [https://www.airnow.gov/air/index.cfm](https://www.airnow.gov/air/index.cfm)  
Extremely high levels of PM2.5:  
• Good: 0 - 15.4  
• Moderate: 15.4 - 40.4  
• Poor: 40.5 - 65.4  
• Unhealthy: 65.5 - 150.4  
• Very unhealthy: 150.5 - 250.4

Average Exposure

The cumulative average values of different pollutants indicate the overall exposure across various periods as shown in Figure 12. These data provide potential insights into the underlying causes of poor AQI, which could be correlated with specific living conditions.
Figure 12. An example of AMS UI showing the average exposure statistics for different pollutants across various periods, ranging from 8 hours to 1 week. AMS: Atmosome Measurement System; CO$_2$: carbon dioxide; O$_3$: ozone; ppb: parts per billion; ppm: parts per million; UI: user interface; VOC: volatile organic compound.

<table>
<thead>
<tr>
<th></th>
<th>VOC</th>
<th>CO$_2$</th>
<th>O$_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>8 Hours</strong></td>
<td>907.35 ppb</td>
<td>585.89 ppm</td>
<td>0.11 ppm</td>
</tr>
<tr>
<td><strong>24 Hours</strong></td>
<td>395.95 ppb</td>
<td>733.02 ppm</td>
<td>0.1 ppm</td>
</tr>
<tr>
<td><strong>1 Week</strong></td>
<td>315.76 ppb</td>
<td>715.63 ppm</td>
<td>0.09 ppm</td>
</tr>
</tbody>
</table>

(a) VOCs exposure  (b) CO$_2$ exposure  (c) O$_3$ exposure

Quarterly Temporal Graphs
The UI time-series data display quarterly data from the different pollutant data streams (Figure 13). This enables users to visualize trends over time and gain deeper insights into which pollutants are affecting their air quality most substantially. For example, tVOC measurements are noted to be highly variable due to indoor sources resulting from occupant lifestyle, including exposure to cosmetics, cleaning products, room refreshers, cooking fumes, and more.

Figure 13. An example of AMS UI showing graphs of the user’s quarterly trends of tVOC, CO$_2$, O$_3$, and humidity data streams. AMS: Atmosome Measurement System; CO$_2$: carbon dioxide; O$_3$: ozone; tVOC: total volatile organic compound; UI: user interface.

Pollutants Information Page
For each category of pollutant, upon selecting its name in the UI, the user is taken to a new page that contains a brief description of the pollutant; associated health risks at different concentrations; and examples of how such data can be collected, graphed, and studied further are shown. This provides users further insights into each of the pollutants and the possibility of enabling additional research. An example description page is shown in Figure 14 for PM$_{2.5}$.
Figure 14. An example of AMS UI individual pollutant page that is displayed on a separate web page for every pollutant and can be found by selecting the pollutant name in the dashboard. These pages provide the user with greater detail about each pollutant. This includes information such as more details about the pollutant, common sources of the pollutant, possible research and data analysis that have been/could be done on the pollutant using the corresponding AMS sensor, and more. This figure is an example of the PM2.5 pollutant page. Besides providing more details about PM2.5, it shows how the AMS PM2.5 sensor and extensive statistical analysis of its data was used for new, internationally published research on low-cost enhancement of facial mask filtration. AMS: Atmosome Measurement System; PM2.5: particulate matter with diameter of 2.5 μm or less; UI: user interface.

**Particulate Matter (PM2.5)**

Particulate matter refers to mixtures of microscopic solid and liquid particles suspended in air. There are two types of particulate matter that are most relevant to air pollution: PM10 and PM2.5. PM10 refers to particles that are between 2.5 and 10 microns; some examples of these include dust, pollen, and particles of mold. PM2.5 consists of fine particles that are 2.5 microns in diameter or less; fuel combustion, cigarette smoke, aerosols, and more can form them. Particulate matter is a health risk because it is small enough to be inhaled and deposits itself in airways of the human body.

The AMS’s PM 2.5 sensor can also be used to spin off new research as done by the same researchers in “Low-cost enhancement of facial mask filtration to prevent transmission of COVID-19” published in the BBJ Journal.

A system as shown below was developed to compare the efficacies of different facial masks, based on their abilities to filter PM 2.5 particles and the data was analyzed using a nested two-way ANOVA model, besides many other statistical techniques.

**REST API**

An extended REST API is also available for advanced users and developers interested in conducting further research or data analysis using the DCS measurements. The REST API enables users to download their data in a .csv format. These data can be mined to gain deep insights into the dynamics of the various indoor air pollutants across time, address extreme or alarming conditions by taking appropriate corrective actions, and exploring possible connections between air pollution and various health conditions.

**Results**

**Study Purpose**

The purpose of this work was to measure various pollutants and other air quality metrics that affect individual environmental atmosomes. The following section presents the results of a variety of air quality metrics in selected environments and relates them with the conditions in their atmosome.

**Temperature and Humidity**

Figures 15 and 16 show the quantity and value of readings of relative humidity and temperature, respectively, taken at 3 different indoor locations: a home in Cupertino, California (shown in blue); a home in Hyderabad, Telangana, India (shown in orange); and an airplane economy cabin during a nonstop flight of 17 hours (shown in green). The low humidity readings from the airplane correlate with the dryness and discomfort often experienced by airplane passengers and align with the United States Centers for Disease Control and Prevention’s (CDC) air travel yellow book [56].
Relative Humidity (%): 30-50 marks the ideal range. Conditions within the airplane journey were low in humidity and conditions in Hyderabad went above the recommended range at times. The Cupertino home had ideal humidity values.

Temperature (°C): 20-27 marks a comfortable and healthy range. Conditions within the airplane journey and in Hyderabad were high in temperature. The Cupertino home had mostly ideal temperature values.

Carbon Dioxide

Figure 17 shows a graph of CO₂ measurements taken in the same environments as above. The readings in the home in the United States showed much higher indoor CO₂ levels than the readings in a more polluted area in India. Further analysis has revealed, however, that the closed windows and doors throughout the day during winter in the United States reduced ventilation and increased CO₂ concentration. Studies show that higher CO₂ exposure can cause drowsiness [57]. These data highlight the importance of ventilation during the winter. The readings on the lengthy airplane journey confirmed that the DCS readings and published values by the airline were within the range of each other. Figure 18 illustrates CO₂ readings in 2 rooms and shows the role AMS had in enabling the researcher to take corrective actions to improve indoor air quality.

CO₂ (ppm): 250-1000 is the safe range for typical indoor spaces with good air ventilation. Higher than 1000 leads to a range of adverse effects from drowsiness to headaches, nausea, and increased heart rate among other conditions. Conditions within the airplane journey and in Hyderabad were mostly ideal to moderate, while conditions in the Cupertino home were outside healthy limits. CO₂: carbon dioxide; ppm: parts per million.
Figure 18. Temporal data for CO$_2$ pollution in 2 different rooms in the Cupertino house: (A) Room 1; (B) Room 2. In both rooms, air conditions were initially unhealthy and then improved drastically. In Room 1, conditions became unhealthy again after some hours before improving once more. CO$_2$: carbon dioxide.

**Particulate Matter 2.5 µm (PM$_{2.5}$)**

Figure 19 depicts the variation in PM$_{2.5}$ during the course of a typical home activity (cooking food). The onset of cooking is associated with a sharp, transient increase in the levels of PM$_{2.5}$ followed by a sustained period of unhealthy PM$_{2.5}$ levels. In this particular scenario, the user responded to the elevated concentration of PM$_{2.5}$ by activating the exhaust fan, deactivating the stove, and opening all windows for improved cross-ventilation.

Figure 19. PM$_{2.5}$ (µg/m$^3$) readings in the kitchen while cooking food. Much of the time spent cooking was in very unhealthy air conditions. PM$_{2.5}$: particulate matter with diameter of 2.5 µm or less.

This set of mitigating actions lowered the PM$_{2.5}$ levels but unhealthy concentrations (defined by the horizontal orange and purple lines in Figure 19) persisted even after cooking ceased. These observational data highlight the power and utility of AMS in identifying and mitigating substantially unhealthy levels of indoor air quality. A longitudinal analysis of such data has the potential to offer rich insights into the interaction between indoor air pollutants and respiratory health outcomes, and aid data-driven health policy research.

**Volatile Organic Compounds and Ozone**

Raw data stored in DASE were downloaded in .csv format using its REST API for advanced users. The data were graphed in a python notebook using matplotlib. As shown in Figure 20, there is a substantial difference between Indian and American household air pollution levels in terms of tVOC measurements. The difference in results could be attributed partly to the substantially higher levels of air pollution in India compared with those in the United States. The current results are consistent with previous results suggesting vehicle exhaust as one of the leading sources of VOC-related pollution in India [58].
**Figure 20.** tVOCs (ppb): 0-200 is the safe range. Values above 2200 ppb are extremely unhealthy. Conditions in the airplane journey and in the Cupertino home were typically ideal, whereas conditions in Hyderabad were often very unhealthy. ppb: parts per billion; tVOC: total volatile organic compound.

**Figure 21.** Displays a strong correlation between tVOC and ground-level O₃ measurements. For the most part, spikes in tVOC result in increases in O₃ levels that cause nitrogen oxides to react with tVOC in the presence of sunlight to create ground-level O₃.

**Figure 21.** Ozone (ppb): Ozone levels from February to April 2020 at a home in the Sierra Mountains, El Dorado County, CA, USA. O₃: ozone; ppb: parts per billion; tVOC: total volatile organic compound.

**Figure 22.** Displays tVOC readings recorded during various daily activities such as house cleaning, cooking at night, and commuting by car. The tVOC values were 1549, 2008, and 2868 ppb, respectively, for the aforesaid activities, well above the accepted moderate range of 220-660 ppb (presented in Figure 8), and demonstrate the high levels of tVOC during innocuous common activities [59]. The results also show that tVOC readings impact the indoor AQI value.

**Figure 22.** tVOC readings during (A) home cleaning, (B) cooking, and (C) commuting by car. These statistics are displayed on AMS user interface. AMS: Atmosome Measurement System; tVOC: total volatile organic compound.
Studies carried out in 2003 by the California Air Resources Board show that cleaning products alone account for the release of 7.4 t of VOCs per day and that their various health effects include asthma attacks and eczema [60]. AMS can furnish necessary information to take precautions in such situations (e.g., a spike in tVOC levels). Homes with infants or expecting parents often undergo more cleaning than typical. AMS, however, suggests that a presumably “well-cleaned” house may, in fact, pose a heightened risk of childhood asthma to infants [60], delays in child language development [61], or prenatal exposure to the fetuses that may impact their postnatal growth [62].

In summary, these results show that IAQ may be at unhealthy levels while conducting typical daily activities. AMS is able to track air quality through its sensors and indicate the impact of harmful levels of indoor air pollutants such as \( \text{O}_3 \), tVOC, \( \text{CO}_2 \) levels, and the presence of PM\(_{2.5}\) inside homes.

**Discussion**

**Principal Findings**

This work has aimed to determine the performance of a low-cost AMS in indoor spaces in terms of gathering sensor data for a variety of air pollutants, transferring data reliably to an analysis engine in the cloud, displaying air quality monitoring results to the user, and sending alerts when the pollutants exceed safe thresholds. This system also provided a way for users to choose to share their IAQ data anonymously for further health research, and for users and researchers to retrieve data using REST APIs for further analysis and data analysis. This can be a foundation for building a public IAQ database across geographical regions.

The results establish that AMS is effective in analyzing multipollutant data streams in multiple settings, displaying AQI value, hourly visual and quantitative data of 18 pollutants, average exposure statistics of various pollutants, and the user’s quarterly pollutant exposure graphs. The dashboard also includes tips for each pollutant that display the GOOD, MODERATE, or POOR thresholds, enabling the user to better interpret the values and graphs shown in the UI and receive recommendations for keeping various pollutants in check with simple mitigation. The individual pollutant pages give the user further insights into the pollutant and show novel research in which AMS sensors and data collection have been employed. The system is also effective in sending alerts when pollutants exceed safe thresholds. The following sections discuss the findings of the work, which are corroborated by previous work and studies of a similar nature.

**Particulate Matter**

Particulate matter refers to mixtures of microscopic solid and liquid particles suspended in the air. There are 2 types of particulate matter that are most relevant to air pollution: PM\(_{10}\) and PM\(_{2.5}\). PM\(_{10}\) consists of particles that are between 2.5 and 10 \( \mu \text{m} \) (diameter); some examples of these include dust, pollen, and particles of mold. PM\(_{2.5}\) consists of fine particles that are 2.5 \( \mu \text{m} \) in diameter or less; some examples of these include fuel combustion, cigarette smoke, and aerosols. Particulate matter is a health risk because it is small enough to be inhaled and deposits itself in the airways of the human respiratory system. Smaller particles can lodge themselves deep in the lungs or enter the bloodstream. Even short-term exposure to PM\(_{2.5}\) has
been associated with worsening respiratory diseases and can lead to emergency care. Long-term exposure (ie, months to years) has been linked to premature death, especially in people with chronic conditions, and leads to reduced lung function in children [63]. The EPA set the maximum 24-hour exposure limit to PM$_{2.5}$ to 35 µg/m$^3$ and annual exposure limit to 12 µg/m$^3$ [64].

Quite interestingly, results of this work clearly indicate that when cooking is involved, the level of PM$_{2.5}$ increases well beyond recommended exposure limits and these levels remain high even after cooking has ceased. While it was observed that ventilation could mitigate levels to some extent, it is imperative that cooking methodologies are modified to reduce the emission of PM$_{2.5}$. A study that focused on emission of PM$_{2.5}$ from cooking in homes found that the levels were raised 20-40-fold to 160 µg/m$^3$ in the kitchen and 10-fold in the nearby living room to 60 µg/m$^3$ [65], clearly corroborating the results of this work. Measuring particulate matter in indoor air can direct researchers toward finding easily implementable corrective actions to reduce exposure levels.

**Carbon Dioxide**

CO$_2$ is considered to be a dominant air pollutant. Moderately high concentrations of CO$_2$ in indoor air can lead to drowsiness, fatigue, and headaches. Increasing amounts can cause dizziness and nausea [66]. In the experimental data, CO$_2$ levels increased dramatically in poorly ventilated rooms as shown in the temporal data in Figure 18. In both the rooms, air conditions were initially unhealthy and then improved drastically. In room 1, conditions became unhealthy again after some hours due to a room heater that did not circulate air. Taking advice furnished by AMS and adding a fan, along with regular ventilation, improved conditions again. These results corroborate the findings of a previous work in the Texas elementary schools that monitored CO$_2$ levels in 120 randomly selected classrooms in 2 school districts [67]. The simple process of improving ventilation and air circulation can reduce high concentrations of CO$_2$ and improve comfort.

**Volatile Organic Compounds**

The experimental results show that even seemingly routine daily activities can induce significant tVOC release. A variety of household items, such as candles, cooking fumes, room fresheners, cosmetics, cleaning products, and paints, emit VOCs. These are organic chemicals that are usually in gaseous form at room temperature and are photo-chemically active. Short-term exposure to VOCs can cause optic or respiratory irritation, headaches, memory lapses, and dizziness. Long-term exposure can cause nausea, fatigue, organ damage, and cancer [68]. Results shown in Figure 22 based on an AMS report of activities, such as cleaning, cooking, and commuting, corroborate the findings of previous work [69-72].

**Ozone**

Readings taken while cleaning, cooking, and commuting, as shown in Figure 22, highlighted a correlation between activities that lead to elevated VOC levels in the presence of sunlight and elevated O$_3$ levels. In support, studies in Los Angeles have found as many VOCs being emitted from household products as from vehicle exhaust pipes, which then react in the presence of sunlight to produce ground-level O$_3$ [73]. These results match the EPA’s report that mentions the adverse effects of synthetic chemicals and emissions caused by cars, power plants, and other industrial setups that react in the presence of sunlight to increase ozone levels [74]. The results of this work show that the same levels of VOCs do not elevate O$_3$ as much at night, as there is no sunlight to facilitate such a chemical reaction.

Ground-level O$_3$ can trigger a variety of health problems, including chest pain, coughing, throat irritation, and airway inflammation. It can also worsen bronchitis, emphysema, and asthma [74]. This can lead to the need for increased medical care [75]. In 2016, 90% of noncompliance to the national ambient air quality standards in the United States was due to O$_3$. Both short- and long-term exposures to O$_3$ at concentrations below the current regulatory standards are associated with increased mortality from respiratory and cardiovascular diseases [76].

**Temperature**

The DCS can measure both events induced by high temperature (eg, the formation of ground-level O$_3$ and low-temperature triggers [77]) and brown adipose tissue metabolism [78]. The direct impact of temperature reflected in a user’s heart rate data stream during their sleep was characterized. An increasing trend in ambient temperature even in the 68-80°F (20-27°C) range and most definitely beyond (>27°C; Figure 16) could be unhealthy. This observation is supported by a previous study highlighted in Figure 23, which shows data collected for 1 user over 424 nights using an iPhone with the Sleep Cycle app and a Garmin Fenix 5 Watch [79]. Clear increases in the user’s circulation flow (ie, reflected by resting heart rate) were observed and appear to be correlated with increases in ambient temperature.

**Humidity**

Lower humidity levels can cause skin dryness, corneal dryness, dry nasal passages, and sinusitis [80], while higher relative humidity levels can promote the growth of mold, bacteria, and viruses. According to previous studies, low humidity levels lead to increased aqueous tear evaporation [81,82]. The low humidity values in the flight shown in Figure 15 corroborated with the researcher’s discomfort due to dry eye symptoms during the flight.

**Significance**

Previous studies have shown that several air pollutants, including PM$_{2.5}$, PM$_{10}$, tVOC, CO, CO$_2$, and more, can be present at much higher quantities in indoor air than in the outdoors [83]. There has been an increasing concern among the scientific community about the connection between IAQ and the impact of personal health [84].

The study results are significant in several ways. Data access in AMS is both simple and comprehensive. Real-time AMS data can be obtained and analysis performed from any device through a progressive web application. The website also includes historic data logs and recommendations on how to improve the
user’s current air quality. Because it uses a cloud-based API, it opens the possibility to integrate new types of analytical graphs or recommendations into the system without the users needing to update their device. With all the air quality measurements collected and posted in tandem, there is an increased knowledge about the interconnections and impacts of pollutants on an individual, community, or specific demographic. Other researchers can also tap into the system to download data and include their own analyses for their unique studies.

Limitations
Because the DCS unit’s calibration is not performed on a large scale, the calibration training set is limited when developing data correction algorithms. Another limitation is that the data were collected only by the researcher, and not external participants, which indicates that the data set is small and does not provide insight into various indoor pollutants affecting various living situations. This limits the development of training sets for learning algorithms to generate optimal individual weights for each of the pollutants to calculate AQI and to generate customized optimal recommendations for managing the pollutants to maintain healthy indoor air quality. While the DCS unit can be USB powered and portable when providing 13 data streams, due to increased power consumption for the comprehensive 22 data streams model, it requires to be plugged into a wall socket, thereby limiting its portability.

Future Work
Future research could implement AMS on a large scale to safeguard public and personal health. Users can anonymously share their data to help researchers draw connections between indoor air pollution and public health and to further research in the development of new modeling techniques for public health, low-cost sensor data correction algorithms, and sensor modeling. The larger vision of AMS is to build a public database for indoor air, like EPA’s Air Quality System database [11,12] for ambient air. The REST APIs in the system can be used to access the open database populated by users willing to share their data.

AMS could include dynamic recalibration of the devices installed at various settings over the internet. With this auto-calibration, sensors would maintain the highest accuracy possible and users would not have to perform manual calibration themselves.

AMS can be researched for reduction of power consumption, while still providing the comprehensive set of all 22 data streams, thus maximizing its portability.

AMS can be used within households to compare indoor air quality levels between neighbors, or on a grander scale within communities. Large-scale distribution of AMS can lead to an expansion of preventative health approaches from an individual level to a community level and can provide valuable insights into public health.

Other possible enhancements include making the high-power model portable, and adding physical indicators, such as sound and light on AMS hardware, which can help indicate the hazardous levels of pollutants in a more noticeable way.

Cybernetic and navigational health approaches enable individuals to be in control of their health throughout their lives so that they have the necessary information to always maintain an ideal state of health [85,86]. Continued research in this field will focus on expanding into other health domains, improving quality metrics, and developing methods to combine atmosome data with other data streams to provide uniquely tailored lifestyle recommendations.

Last, AMS could implement more sophisticated software algorithms, such as an artificial neural network, as demonstrated in [27], and machine learning, as demonstrated in [30], and adaptive thresholding. In particular, the simplest algorithm to implement would be to account for environmental variation (e.g., temperature and humidity) using an adaptive threshold to improve sensor accuracy. Specifically, and referring to [54], an adaptive threshold or the computed gas concentration could be computed as a function of temperature and humidity. This would significantly increase the sensitivity of the MQ sensors. Other algorithms likely exist and could be explored with the intention of maintaining a low-cost and small sensor solution as opposed to transitioning to an alternative sensor.

Conclusions
In this work, the concept of an atmospheric exposome (atmosome) was presented and a low-cost approach to leverage multimodal sensors and cloud data storage in building a personal atmosome was proposed. This work shows that AMS offers both the concept and system to quantify and measure the atmosphere with continuous real-time data streams, which provide users and researchers access to pertinent air quality data through a cloud computing architecture. These data can be used by both members of the public and researchers. For example, AMS could be used to alert an elderly user with dementia about a gas leakage.

Further, AMS can find the effect of other atmospheric streams such as temperature and humidity on the user’s observed health and behavioral outcomes. This system monitors and analyses VOCs, which could help ensure that pregnant women breathe safe air, safeguarding not only their own health, but also that of their fetus(es). Particulate matter detection and warnings can help users to act on remediating the environment to avoid pulmonary complications in all age groups. Overall, the findings indicate that the multidimensional model of AMS is a step closer in considering a variety of pollutants and atmospheric characteristics that guide users toward a healthy lifestyle.

While the system can predict respiratory triggers, such as asthma attacks in vulnerable people, it is yet to be tethered to other health conditions (e.g., proneness of a patient with diabetes to blood glucose variations due to air quality). AMS could certainly provide insight into such correlations and dramatically improve personal health.
Acknowledgments
This research was conducted by HB under the supervision of NN and RJ. This research received no external funding.

Authors’ Contributions
Data from GoldenCheetah OpenData Project were retrieved by VP. HB and NN were responsible for study conceptualization and formal analysis. HB was also responsible for methodology, software implementation, data validation, investigation, procuring resources, writing—original draft preparation, and writing—review and editing. Both HB and VP took care of data curation. HB, NN, and VP handled data visualization. NN and RJ provided project supervision and administration. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Comparison of features and performance of COTS (Commercial-Off-The-Shelf) AQI sensor systems to AMS.

[PDF File (Adobe PDF File), 102 KB-Multimedia Appendix 1]

Multimedia Appendix 2
Summary of recent AQI studies as compared to the study presented in this work.

[PDF File (Adobe PDF File), 85 KB-Multimedia Appendix 2]

References


33. Mi PM 2.5 Detector - Your Own Eco-station. URL: https://xiaomi-mi.com/air-and-water-purifiers/xiaomi-mi-pm25-detector-white/ [accessed 2021-08-21]


37. Temtop M2AQMMP. 5 PM10 Detector HCHO Tester Data Export. Temtop. URL: https://tinyurl.com/5c8b4yv [accessed 2020-12-28]

38. BRWISSEN A18 Air Quality Monitor Indoor Pollution Tester for Co2 Dust Meter PM1.0 PM2.5 PM10 Particulate Matter Analyzer HCHO Formaldehyde TVOC Detector Data Record Storage Particle Counter. URL: https://www.amazon.com/gp/product/B08K7DTKN2/ [accessed 2020-12-28]

41. \. uHoo Indoor Air Quality Sensor - 9 in 1 Smart Air Monitor. URL: https://getuhoo.com/home [accessed 2021-01-18]
55. Atmosome Live Data. URL: https://www.atmosome.com [accessed 2021-11-10]
64. US Environmental Protection Agency. What are the Air Quality Standards for PM?. URL: https://tinyurl.com/3s4mcntt [accessed 2020-09-15]


79. Liversedge M. GoldenCheetah OpenData Project. URL: https://osf.io/6hfpz/ [accessed 2019-12-01]


**Abbreviations**

- **AMS**: Atmosphere Measurement System
- **ANN**: artificial neural network
- **API**: application programming interface
- **AQI**: Air Quality Index