

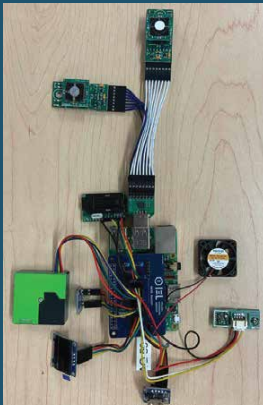



## Graphical Abstract

### Design, fabrication, and calibration of the Building EnVironment and Occupancy (BEVO) Beacon: a rapidly-deployable and affordable indoor environmental quality monitor

Hagen Fritz, Sepehr Bastami, Calvin Lin, Kingsley Nweye, Tung To, Lauren Chen, Dung Le, Angelina Ibarra, Wendy Zhang, June Young Park, William Waites, Mengjia Tang, Pawel Misztal, Atila Novoselac, Edison Thomaz, Kerry Kinney, Zoltan Nagy

#### Design, fabrication, and calibration of the Building EnVironment and Occupancy (BEVO) Beacon: a rapidly-deployable and affordable indoor environmental quality monitor



Design	Fabrication	Calibration and Testing
 <p>Provides data on 24 parameters including:</p> <ul style="list-style-type: none"><li>CO<sub>2</sub>: Carbon Dioxide</li><li>CO: Carbon Monoxide</li><li>PM: Particulate Matter</li><li>VOCs: Volatile Organic Compounds</li><li>T: Temperature</li></ul> <p>1-minute resolution</p> <p>local and cloud data storage</p>	<p>20 devices created</p>  <p>Designed:</p> <ul style="list-style-type: none"><li>Plywood housing</li><li>Printed circuit board</li><li>Software</li></ul>	<p><b>Simulated Home Environment:</b></p> <ul style="list-style-type: none"><li>CO<sub>2</sub></li><li>PM</li></ul>  <p><b>Environmental Chamber:</b></p> <ul style="list-style-type: none"><li>CO<sub>2</sub></li><li>CO</li><li>PM</li><li>TVOC</li><li>T</li></ul>  <p>Device-specific &gt; environment-averaged models</p> <p>Devices/models tested in 11 week deployment</p>

Fritz, H. et al, 2022



## Highlights

### **Design, fabrication, and calibration of the Building EnVironment and Occupancy (BEVO) Beacon: a rapidly-deployable and affordable indoor environmental quality monitor**

Hagen Fritz, Sepehr Bastami, Calvin Lin, Kingsley Nweye, Tung To, Lauren Chen, Dung Le, Angelina Ibarra, Wendy Zhang, June Young Park, William Waites, Mengjia Tang, Pawel Misztal, Atila Novoselac, Edison Thomaz, Kerry Kinney, Zoltan Nagy

- Developed open-source, consumer-grade indoor air quality monitor.
- Calibrated sensors using multiple techniques in controlled/uncontrolled environments.
- Deployed 20 devices in a field study to validate their use and robustness.
- Provide insights into design and calibration of sensors for successful applications.

# Design, fabrication, and calibration of the Building EnVironment and Occupancy (BEVO) Beacon: a rapidly-deployable and affordable indoor environmental quality monitor

Hagen Fritz<sup>a</sup>, Sepehr Bastami<sup>a</sup>, Calvin Lin<sup>a</sup>, Kingsley Nweye<sup>a</sup>, Tung To<sup>b</sup>, Lauren Chen<sup>c</sup>, Dung Le<sup>c</sup>, Angelina Ibarra<sup>b</sup>, Wendy Zhang<sup>a</sup>, June Young Park<sup>d</sup>, William Waites<sup>e</sup>, Mengjia Tang<sup>a</sup>, Pawel Misztal<sup>a</sup>, Atila Novoselac<sup>a</sup>, Edison Thomaz<sup>b</sup>, Kerry Kinney<sup>a</sup>, Zoltan Nagy<sup>a,\*</sup>

<sup>a</sup>Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX 78712

<sup>b</sup>Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712

<sup>c</sup>Department of Mechanical Engineering, University of Texas at Austin, Austin, TX 78712

<sup>d</sup>Department of Civil Engineering, University of Texas at Arlington, Arlington, TX, 76019

<sup>e</sup>Department of Computer and Information Sciences, University of Strathclyde, Glasgow, Scotland, G1 1XH

---

## Abstract

Indoor Air Quality (IAQ) monitoring is essential to assess occupant exposure to the wide range of pollutants present in indoor environments. Accurate research-grade monitors are often used to monitor IAQ but the expense and logistics associated with these devices often limits the temporal and spatial scale of monitoring efforts. More affordable consumer-grade sensors – frequently referred to as low-cost sensors – can provide insight into IAQ conditions across greater scales but their accuracy and calibration requirements need further evaluation. In this paper, we present the Building EnVironment and Occupancy (BEVO) Beacon. The BEVO Beacon is entirely open-source, including the software, hardware, and design schematics which are all provided on GitHub. We created 20 of these standalone, stationary devices which measure up to 24 parameters at a one-minute resolution of which we focus on carbon dioxide, carbon monoxide, total volatile organic compounds, temperature, and size-resolved particulate matter. We investigated the efficacy of two different calibration approaches – device-specific and environment-averaged – for these sensors as well as also provide an extensive discussion considerations for each of the sensors. Calibrated sensors performed well when compared to reference monitors or calibrated gas standards. The CO sensors yielded the best agreement ( $r^2=0.98-0.99$ ), followed by temperature ( $r^2=0.89-0.99$ ), CO<sub>2</sub> ( $r^2=0.62-0.99$ ), and PM<sub>2.5</sub> ( $r^2=0.13-0.91$ ). In all cases, the device-specific calibration approach yielded the most accurate results. We evaluated our devices through a successful 11-week field study where we monitored the IAQ in participants' bedrooms. The work we present on consumer-grade sensors adds to the existing literature by considering sensor-specific calibration techniques and analysis. The BEVO Beacon adds to the successful line of similarly developed devices by providing an open-source framework that researchers can readily adapt and modify to their own applications.

*Keywords:* field study, indoor air quality, low-cost sensors

---

## Acronyms

**BEVO Beacon** Building Occupancy and EnVironment Beacon.

**CGS** Consumer-Grade Sensors.

**CH<sub>4</sub>** Methane.

**CO** Carbon Monoxide.

**CO<sub>2</sub>** Carbon Dioxide.

**EC** Electrochemical.

**HCHO** Formaldehyde.

**I<sup>2</sup>C** Inter-Integrated Circuit.

**IAQ** Indoor Air Quality.

**LCS** Low-Cost Sensors.

**MOS** Metal Oxide Semiconductor.

**NDIR** Non-Dispersive Infrared.

---

\*Corresponding author

Email address: nagy@mail.utexas.edu (Zoltan Nagy)

**NO<sub>2</sub>** Nitrogen Dioxide.

**O<sub>3</sub>** Ozone.

**PCB** Printed Circuit Board.

**PM** Particulate Matter.

**PM<sub>2.5</sub>** Particulate Matter with aerodynamic diameters less than 2.5  $\mu\text{m}$ .

**RH** Relative Humidity.

**RPi** Raspberry Pi 3B+.

**RTC** Real-Time Clock.

**SO<sub>2</sub>** Sulfur Dioxide.

**T** Temperature.

**TVOC** Total Volatile Organic Compound.

**UART** Universal Asynchronous Receiver-Transmitter.

**VOC** Volatile Organic Compound.

**ZAG** Zero Air Gas.

## 1. Introduction

Indoor Air Quality (IAQ) is an issue of broad concern as both acute and chronic exposure to common indoor air pollutants can contribute adverse health effects. Poor IAQ can exacerbate or induce illnesses relating to the respiratory [1] and cardiovascular systems [2] in addition to negatively affecting occupant mood [3], productivity [4], and performance [5]. These effects are compounded by the fact that people spend more time indoors, especially in developed nations where occupants spend nearly 90% of their day inside [6] – 69% of which is spent in residences. The World Health Organization (WHO) reports that IAQ can be up to 5 times worse than ambient air pollution concentrations, and that nearly 3.8 million people die annually due to household air pollution, primarily in developing countries [7]. In addition to human health concerns, the need to balance human comfort and energy considerations [8] motivates the development of wide-spread and accurate IAQ monitoring tools which inform occupants and building managers about pollution events that require intervention.

Traditionally, IAQ has been measured with research-grade equipment that has undergone extensive calibration and certification. These sensors are very accurate, but the cost of the equipment, training needed to properly operate the instruments, and space requirements often make using these devices challenging and expensive, especially for large-scale deployments. However, within the last 10-15 years, technological advances have allowed for the mass production of affordable,

Consumer-Grade Sensors (CGS) designed to measure atmospheric particles and gases [9]. The rapid development of CGS has led to a paradigm shift in how researchers, companies, and government agencies are monitoring air pollution [10].

CGS – also referred to as Low-Cost Sensors (LCS) – do not have any universal definition but are typically thought to cost less than a few hundred US dollars. In a recent review paper, authors defined CGS as “anything costing less than the instrumentation cost required for demonstrating compliance with the air quality regulations can be termed as low-cost” [11]. The affordability of CGS allows users to create vast sensor networks that can monitor multiple locations and delineate spatiotemporal trends of specific pollutants, typically in near real time. Low maintenance requirements and ease/affordability of replacing damaged sensors make CGS ideal for these networks which can help to supplement sparse, pre-existing air pollution networks [12, 13]. Furthermore, the cost barrier to work with CGS is significantly lower which opens up possibilities for community-driven science [14, 15].

The simplicity of CGS tend to lead to issues with data accuracy and reliability. CGS used to measure gaseous pollutants are often based on Electrochemical (EC) or Metal Oxide Semiconductor (MOS) technology, both of which are typically sensitive to multiple compounds, require frequent calibration, and have short lifespans [14]. Other systems such as Non-Dispersive Infrared (NDIR) which are used for detecting pollutants like Particulate Matter (PM), are limited because they cannot directly measure the mass of particles – needing to assume a particle density – and typically cannot detect particles less than 0.3 $\mu\text{m}$  in diameter [16]. Another concern with CGS is the manufacturing process which can also lead to significant differences between device sensitivity resulting in issues with reproducibility and inter-sensor variability [17, 18]. Also, CGS typically do not have corrections for factors like temperature and relative humidity as is common in reference-grade monitors meaning many CGS are sensitive to changes in ambient conditions [19]. Lastly, CGS are more prone to sensor drift which can be exacerbated by the environment in which these sensors are located. While CGS do not provide the accuracy that reference-grade monitors do, not all applications of CGS for monitoring IAQ require high accuracy instrumentation [20].

### 1.1. Related Work

CGS are especially useful in the indoor environment, where adjacent rooms or even locations within the same space can have dramatically different pollutant concentrations. The body of research related to CGS for IAQ monitoring is growing and is attracting researchers from fields outside of environmental engineering [21, 22]. These devices vary in their layout (single unit or distributed) and number/types of indoor air pollutants that are measured. We highlight some of the single unit IAQ monitoring devices in Table 1 which are similar to the device we created and have been used in previous studies. A majority of these studies focus less on developing accurate devices and more on aspects such as communication protocols, data storage, online data processing, portability, and dashboard/application development (see references in [23]). However, many studies forego calibration which is vital for commercial and research applications. Authors in [23] indicate that only 10 of the 35 devices identified included some level of calibration with a reference monitor, pointing to calibration as a major criteria to include in future studies. These devices can be further improved if made open-source so that researchers and other end-users can adapt and refine devices to meet their needs.

### 1.2. Contribution

In this paper, we detail the development, calibration, and deployment of the BEVO Beacon: an affordable, entirely open-source IAQ monitor that can be rapidly developed and deployed to understand occupied indoor environments. Our aim when developing the BEVO Beacon was to provide a device that researchers with little knowledge in the domain of embedded systems engineering – defined as the software and hardware design of a microcontroller-based smart system – could easily replicate by providing all the necessary documentation to create their own. We also highlight a variety of techniques that can be used to calibrate these sensors depending on the availability of calibration environments, reference-grade monitors, and/or gas standards. In addition to the techniques, we provide evidence for the use of device-specific calibration models which account for sensor-to-sensor variability that aggregated models applied across multiple devices do not. Lastly, 20 devices were deployed as part of a large-scale field study to assess their performance and reliability. The intention of

this paper is to provide insight into the decisions, design process, calibration, and deployment of the BEVO Beacon so that researchers who opt to design their own IAQ monitoring device may learn from our experiences.

## 2. Materials and Methods

As part of our commitment to the open science movement [35], the hardware schematics, software, and assembly instructions for the BEVO Beacon are included alongside this publication. This and more detailed information can also be found on the projects GitHub repository: [https://github.com/intelligent-environments-lab/bevo\\_iaq](https://github.com/intelligent-environments-lab/bevo_iaq)

### 2.1. Hardware Selection

The primary processing unit of our device is a Raspberry Pi 3B+ (RPi), a credit card sized single-board computer that runs on a Linux-based operating system. We chose the RPi because it is easy to interface with, contains a large amount of accessible and easy-to-understand documentation, has built-in WiFi and Bluetooth capabilities, and can be programmed using Python – a popular, well-documented programming language. The RPi powers all components, reads data from each of the sensors, provides local storage, and can be configured to send the data to a cloud-based storage system when connected to WiFi.

The parameters measured and the specific sensors on the BEVO Beacon are detailed in Table 2. We opted for sensors that measure common indoor air pollutants, namely Carbon Dioxide (CO<sub>2</sub>), Particulate Matter with aerodynamic diameters less than 2.5  $\mu\text{m}$  (PM<sub>2.5</sub>), and Total Volatile Organic Compound (TVOC). CGS for CO<sub>2</sub> provide some of the most reliable measurements relative to other parameters and CO<sub>2</sub> can provide details on ventilation as well as occupancy. Monitoring PM<sub>2.5</sub> is important because this pollutant is health hazard and is ubiquitous in home environments as it is generated from sources such as cooking, candle/incense burning, smoking, pets, and nearby outdoor sources. TVOC sensors are typically less reliable than the CO<sub>2</sub> and PM<sub>2.5</sub> sensors since they are not standardized and are manufactured with varying sensitivities to different compounds. However, most CGS for TVOC are sensitive to activities like cooking, cleaning, and smoking indoors. On our device, we also include sensors for Carbon Monoxide (CO)

Table 1: Recently developed single-unit IAQ monitoring devices including parameters monitored, evaluation environments, calibration, and approximate cost – if provided otherwise we indicate Not Provided (NP). The final entry corresponds to our device, the Building Occupancy and EnVironment Beacon (BEVO Beacon).

Reference	Parameters Monitored	Calibration			Evaluation		Open-Source	Cost (USD)
		Field	Lab	Empirical	Field	Lab		
[24]	T, RH, CO <sub>2</sub> , PM, CO, NO, O <sub>3</sub> , SO <sub>2</sub> , TVOC			✓	✓			NP
[25]	T, RH, CO <sub>2</sub> , light				✓	✓	✓	NP
[26]	T, RH, CO <sub>2</sub> , PM				✓	✓		250
[27]	T, RH, CO <sub>2</sub> , PM			✓	✓			NP
[28]	T, RH, CO <sub>2</sub> , PM, light, noise				✓	✓		550
[29]	T, RH, CO <sub>2</sub> , PM, CO, TVOC, light, noise		✓		✓			200
[30]	T, RH, CO <sub>2</sub> , PM, CO, NO, NO <sub>2</sub> , noise		✓	✓	✓	✓		300
[31, 32]	T, RH, CO <sub>2</sub> , PM, CO, TVOC, HCHO, light, noise, air velocity		✓		✓	✓		250
[33]	T, RH, CO, NO <sub>2</sub> , C <sub>2</sub> H <sub>6</sub> OH, H <sub>2</sub> , NH <sub>3</sub> , CH <sub>4</sub> , C <sub>3</sub> H <sub>8</sub> , C <sub>4</sub> H <sub>10</sub>				✓			60
[34]	T, RH, CO <sub>2</sub> , PM, CO, NO, NO <sub>2</sub> , O <sub>3</sub> , CH <sub>4</sub>	✓	✓		✓	✓		NP
This Study	T, RH, CO <sub>2</sub> , PM, CO, NO <sub>2</sub> , TVOC, light	✓	✓		✓	✓	✓	350

and Nitrogen Dioxide ( $\text{NO}_2$ ) since they are implicated in many health issues associated with poor IAQ. CO measurements can be used to understand health affects associated with indoor combustion, primarily through the use of gas stoves.  $\text{NO}_2$  is typically associated with outdoor air quality and vehicle emissions, therefore providing information regarding proximity to major roadways and/or the “leakiness” of the building envelope. The CO and  $\text{NO}_2$  modules also include a temperature and Relative Humidity (RH) sensor. Lastly, we include an ambient light sensor on the device which can provide us with information such as when lights are switched on or off.

There are a variety of CGS on the market that monitor the pollutants of interest for this study. We opted for sensors that use Inter-Integrated Circuit ( $\text{I}^2\text{C}$ ) to communicate with the RPi. The RPi supports multiple communication protocols, but  $\text{I}^2\text{C}$  communication provides the ability to scale up the number of sensors more easily. However, at the time of development, there were no  $\text{NO}_2$  or CO sensors that used  $\text{I}^2\text{C}$ . The two sensors we used for  $\text{NO}_2$  and CO are manufactured by the same company and come with a Universal Asynchronous Receiver-Transmitter (UART) to USB-A adapters that were connected to two of the four USB-A ports available on the RPi. To ensure communication between the RPi and  $\text{I}^2\text{C}$  sensors, we created our own Printed Circuit Board (PCB) (Figure 1) which allowed us to connect multiple sensors and provide the necessary pull-up resistors on the Serial Clock (SCL) and Serial Data (SDA) lines.

Figure 2 shows each of the sensors listed in Table 2, the cooling fan, Real-Time Clock (RTC), PCB, and connections between each component and the RPi. Housing (Figure 3) was designed and cut from 0.25 inch plywood using precision laser cutting. The RPi and sensors are separated by 0.25 inch plywood insert cut with a two rectangular holes for wiring and connections for the two USB-connected sensors. The cooling fan is mounted on the inside and pulls air through openings in the housing to provide cooling to the processing unit on the RPi. The light and TVOC sensors are mounted on the outside of the top panel. The adapter boards for CO and  $\text{NO}_2$  sensors are mounted to one of the side panels with square holes cut so that top-mounted digital EC sensors are exposed directly to the surrounding air. The  $\text{CO}_2$  sensor is mounted to the inside wall of a smaller side panel with a hole cut to expose the inlet of the sensor. Lastly, the PM

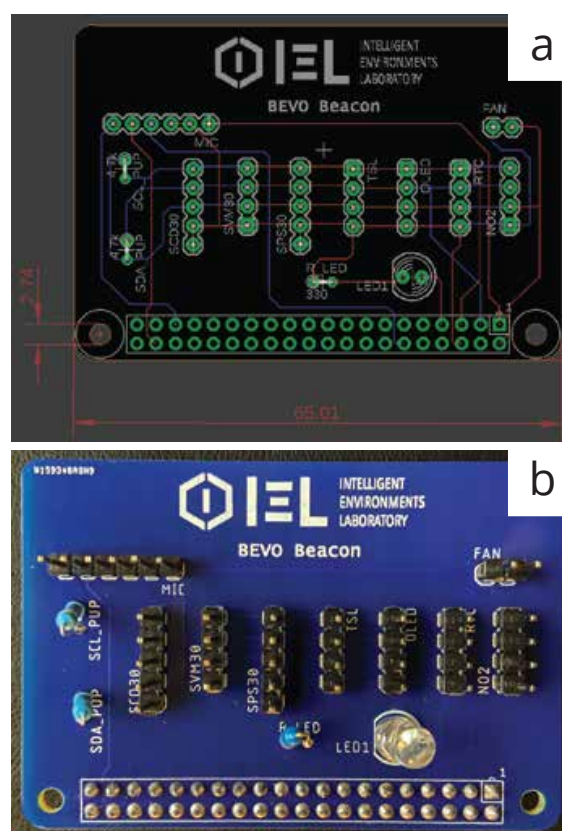


Figure 1: The PCB we developed for the BEVO Beacon: (a) digital schematic and (b) fully assembled.

Table 2: Manufacturer’s specifications provided for each sensor included in the BEVO Beacon. Manufacturer-reported accuracy is typically a function of Measured Value (MV).

Sensor Name	Variable(s) Measured	Measurement Accuracy	Measurement Range	Communication Protocol	Approximate Cost
Sensirion SCD30	CO <sub>2</sub> T	$\pm(30 + 3\% \text{ of MV})$ ppm $\pm(0.4 + 0.023 \times (MV - 25))^\circ\text{C}$	0 - 40,000 ppm -40 - 70°C	I <sup>2</sup> C	75 USD
Sensirion SPS30	PM <sub>1</sub> PM <sub>2.5</sub> PM <sub>4</sub> PM <sub>10</sub>	$\pm 10 \mu\text{g}/\text{m}^3$ ( $< 100 \mu\text{g}/\text{m}^3$ ) $\pm 10 \mu\text{g}/\text{m}^3$ ( $< 100 \mu\text{g}/\text{m}^3$ ) $\pm 25 \mu\text{g}/\text{m}^3$ ( $< 100 \mu\text{g}/\text{m}^3$ ) $\pm 25 \mu\text{g}/\text{m}^3$ ( $< 100 \mu\text{g}/\text{m}^3$ )	0 - 1,000 $\mu\text{g}/\text{m}^3$	I <sup>2</sup> C	75 USD
Sensirion SVM30	TVOC T	15% of MV $\pm 1^\circ\text{C}$	0 - 60,000 ppb 5 - 55°C	I <sup>2</sup> C	50 USD
SPEC DGS CO	CO T	15% of MV $\pm 0.4^\circ\text{C}$	0 - 1000 ppm -10 - 85°C	UART	75 USD
SPEC DGS NO <sub>2</sub>	NO <sub>2</sub> T	15% of MV $\pm 0.4^\circ\text{C}$	0 - 5 ppm -10 - 85°C	UART	75 USD
Adafruit TSL2591	light	Not Provided	Not Provided	I <sup>2</sup> C	20 USD

sensor is secured to the bottom panel of the housing where the side of the sensor with the inlet is inserted through a small opening in the same side panel the CO<sub>2</sub> sensor is mounted to. Figure 4 shows the completed BEVO Beacon and highlights the IAQ sensors and their locations once assembled.

## 2.2. Software Design

There are three Python files that are run synchronously upon a successful boot of the RPi: *main.py*, *display.py*, and *connection.py*. The first script is the main program file that connects to and reads measurements from the sensors. The *display.py* file reads the most recently measured values and displays them on a small OLED screen mounted underneath the top panel (see Figure 4). Readings cycle every 3 seconds across 8 parameters. The time and parameters displayed can be customized for the given study if certain pollutants are more or less relevant. If available, measurements are first corrected using locally-stored calibration files. Otherwise, the raw measurements are shown. Lastly, the *connection.py* script checks for internet connectivity every 5 seconds and indicates the status with a LED.

For each of the sensors that use I<sup>2</sup>C communication, open-source software was readily accessible

and used. We developed our own software to read from the NO<sub>2</sub> and CO sensors. For each of the sensors, the software we used was developed entirely in Python. Figure 5 shows the process the device employs to measure from each of the sensors on the BEVO Beacon. Each sensor is enabled sequentially, scans the environment synchronously, stores the average measurement for each parameter from five scans on the RPi, and then each sensor is disabled. To ensure that measurements are taken each minute, we calculate the number of seconds until the next minute,  $t$ , based on the time logged at the “Enable Sensors” step and the device sleeps for  $t$  seconds. Data from each scan is appended to the same file until the next day, ultimately providing daily data files with 1-minute measurements from each sensor.

## 2.3. Device Setup and Operation

The device setup is outlined in the projects README file. Users simply install the latest version of Raspbian Lite on their RPi before following the steps on the project’s Github repository to get the software running. The devices are meant to work regardless of the components that the user wishes to install. Operating the BEVO Beacon simply requires the user to power the device from a



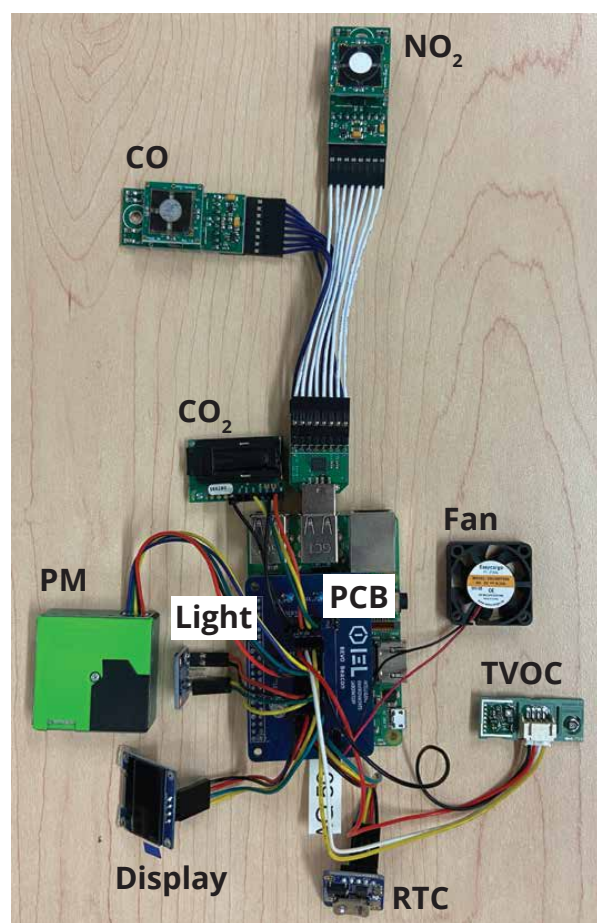


Figure 2: Sensors, other hardware, and connections on the BEVO Beacon

traditional power outlet. After the RPi is finished booting up, the three Python scripts outlined in Section 2.2 will begin running automatically. To address a timeout issue that occurs infrequently with the  $\text{PM}_{2.5}$  and  $\text{CO}_2$  sensors if the BEVO Beacon is operating for longer than a few days, we programmed in a daily reboot at midnight. The device collects data at a 1-minute resolution minimum which generates a daily CSV file of 450 KB maximum. The project and RPi operating system account for about 1 GB of space, meaning for a standard 8 GB memory card, the BEVO Beacon can log data for years without having to remove data files.

#### 2.4. Calibration

Each BEVO Beacon was calibrated in a variety of environments depending on the sensor. Details regarding the calibration process used in each environment are explained in the subsequent sections, but the general process is outlined in Figure 6. Details regarding the pollutants that were calibrated are provided in Table 3, indicating that each sensor on the BEVO Beacon was calibrated excluding the  $\text{NO}_2$  and light sensors. We did not calibrate the  $\text{NO}_2$  sensor because the manufacturer-specified resolution was 20 ppb. Previous studies indicate that household  $\text{NO}_2$  measurements are typically lower than 20 ppb [36, 37], meaning that our sensor would be unable to differentiate concentrations at these levels. With respect to the light sensor, it was only intended to determine when lights were on or off in a room and calibration was not required for this purpose.

Due to the variety of sensors included on the BEVO Beacon, we were provided an opportunity to address various concerns associated with calibration, namely:

1. How do calibration results vary between controlled and uncontrolled environments?
2. What methods are suitable to calibrate sensors?
3. What assumptions are safe to make?

For  $\text{PM}_{2.5}$  and  $\text{CO}_2$ , we were able to calibrate both sensors in controlled and uncontrolled environments against research-grade monitors. This process provides insight into how necessary a controlled environment is to properly calibrate these devices. Calibration in an uncontrolled setting is easier to conduct and more realistic, but introduces

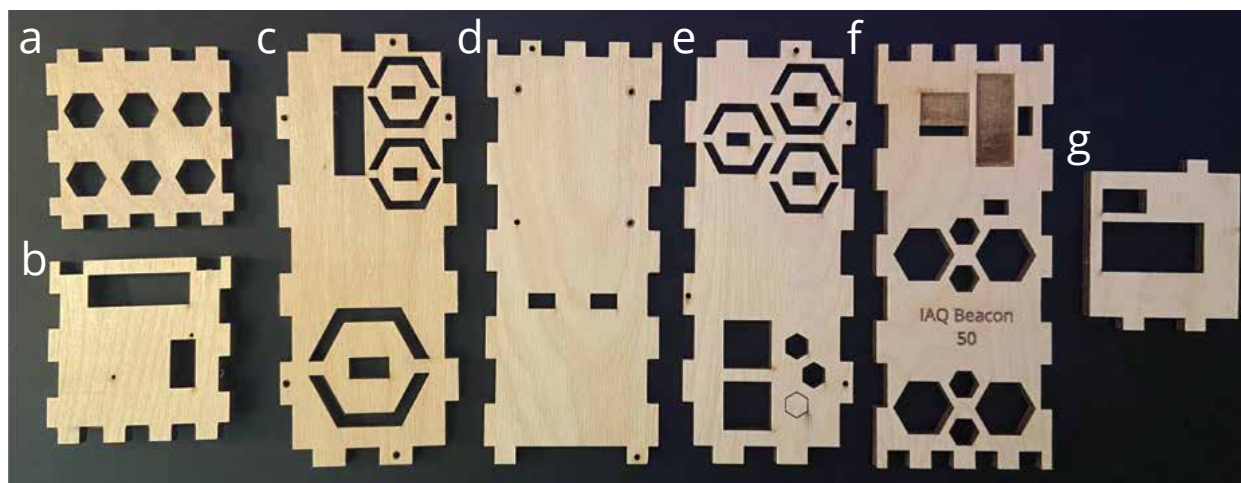


Figure 3: Plywood housing used to encase the BEVO Beacon: (a) small end panel, (b) small sensor panel, (c) left long panel, (d) bottom, (e) right long panel, (f) top, and (g) middle partition

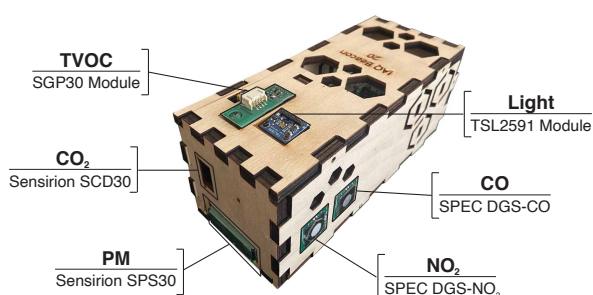


Figure 4: The BEVO Beacon and the location, type, and parameter measured by each of the IAQ sensors.

issues such as incomplete mixing which can lead to the development of incorrect calibration model parameters. However, discrepancies in model parameters between environments should be minor especially if similar concentrations and profiles are generated during calibration and multiple experiments are conducted for quality control.

Ideally, sensors should be calibrated by comparing measurements to a reference-grade monitor which we do for CO<sub>2</sub>, PM<sub>2.5</sub>, and temperature. However, a research-grade monitor is not always available. In this case, a gas standard can be used and diluted to various concentrations to conduct a step calibration. This process is what we enacted for the CO sensor. In other cases, a research-grade monitor is not feasible to use, especially for non-specific analytes such as TVOCs. The sensitivity of TVOC sensors to different compounds can vary significantly depending on the manufacturer and

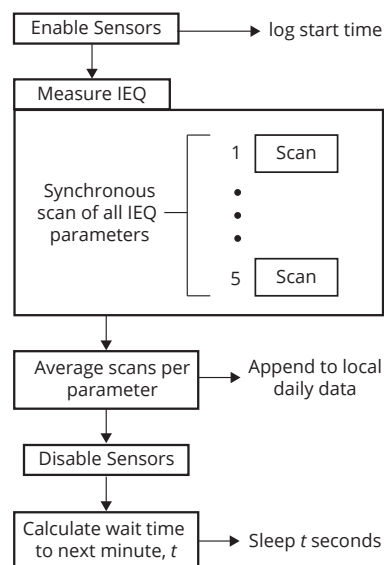


Figure 5: Software for measuring IAQ parameters from the BEVO Beacon executed in the *main.py* script.

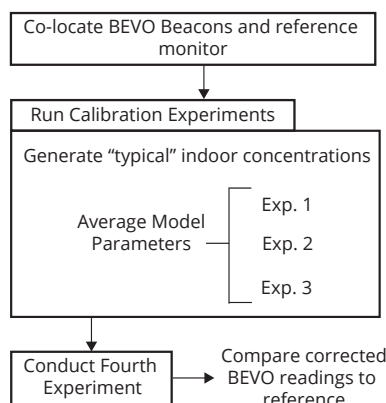


Figure 6: General calibration process used for all parameters except for CO.

these specifications are often not available. This issue makes calibrating these devices complicated because not all sensors respond identically to the same Volatile Organic Compound (VOC) standard. Thus, these TVOC sensors often provide a relative measure of TVOCs present in indoor air but cannot provide an absolute measurement. We opt to calibrate our TVOC sensors in each BEVO Beacon by co-locating them in the same environment and normalizing their response to the same concentration.

For each sensor we calibrate, we derive univariate linear models of the form  $y = b + mx$  where  $y$  is the corrected CGS reading,  $x$  is the raw CGS measurement, and  $b$  and  $m$  are intercept and slope parameters. In the context of our work,  $m$  is related to the sensitivity of the sensors while  $b$  provides insight into a sensor’s base offset. We evaluate sensor performance by considering the coefficient of determination,  $r^2$ , which measures the interrelation between variables and provides information on the model’s goodness-of-fit. Values for  $r^2$  range from 0 to 1 where 1 would indicate perfect agreement between the CGS and the ground-truth measurements made by the reference-grade instrument. Negative  $r^2$  values are possible and indicate that a horizontal line would be a better fit which, in our case, is likely an indication that the CGS and reference measurements are out of sync i.e. CGS measurements increase when reference values decrease or vice-versa.

#### 2.4.1. UTest House

The experimental home environment (UTest House at UT’s Pickle Research Center) shown in Figure 7 represents our uncontrolled environment which we used to calibrate the  $PM_{2.5}$  and  $CO_2$  sensors only. Each of the 20 devices were set up in the kitchen next to the research-grade monitors.  $CO_2$  was emitted into the center of the space through the use of a pressurized gas cylinder and we allowed the concentration to reach approximately 2000 ppm before closing the cylinder. In separate experiments, particles to calibrate the  $PM_{2.5}$  sensors were introduced on two occasions during a two-hour period using a hand-operated nebulizer containing ultra-fine particles (PTI Arizona Test Dust A1) with median diameter between 3 and 5  $\mu m$ . Prior to use, the particles used in the experiment were placed in a 10L incubator with a desiccant (anhydrous calcium sulfate, Drierite Company) to remove excess moisture. To induce mixing during the  $PM_{2.5}$  and  $CO_2$  experiments, two box fans (not pictured) were set up in the room.

We conducted three two-hour experiments to help monitor consistency and to ensure the pollution events we generated were sufficiently captured. We obtain values for  $b$  and  $m$  by averaging the values for each parameter over the three experiments. As a final check, we conducted a fourth experiment and applied the linear models to the raw data collected by the BEVO Beacon to assess model performance relative to the reference standard readings.

#### 2.4.2. Environmental Chamber

The 27  $m^3$  stainless steel chamber shown in Figure 8 represents the controlled environment which we used to further characterize the  $PM_{2.5}$  and  $CO_2$  sensors in addition to the TVOC and CO sensors. Rather than emitting  $CO_2$  using a pressurized cylinder, we generated  $CO_2$  by asking one researcher to occupy the space and breath normally while in a seated position for 30 minutes after an initial unoccupied period of 30 minutes. Human breath is a natural source for VOCs so the researcher also represented a source to calibrate the TVOC sensors with. Following this period, the occupant left and we monitored the  $CO_2$  concentration for 45 minutes with the chamber’s ventilation system deactivated. For the remaining time, the door to the chamber remained slightly ajar to allow for the  $CO_2$  concentration to gradually return to the background concentration. We controlled the

Table 3: Calibration details for each of the variables measured on the BEVO Beacon

Variable	Reference	Calibration Environment		Model
		UTest House	Laboratory	
CO <sub>2</sub>	LI-COR Model 6252	✓	✓	Linear
PM <sub>2.5</sub>	TSI Aerodynamic Particle Sizer Model 3321	✓	✓	Linear
TVOC	None		✓	Linear
NO <sub>2</sub>	None			None
CO	Background, Gas Standard		✓	Constant, Linear
T	Michell Instruments S8000		✓	Linear
RH	Michell Instruments S8000		✓	Linear
Light	None			None

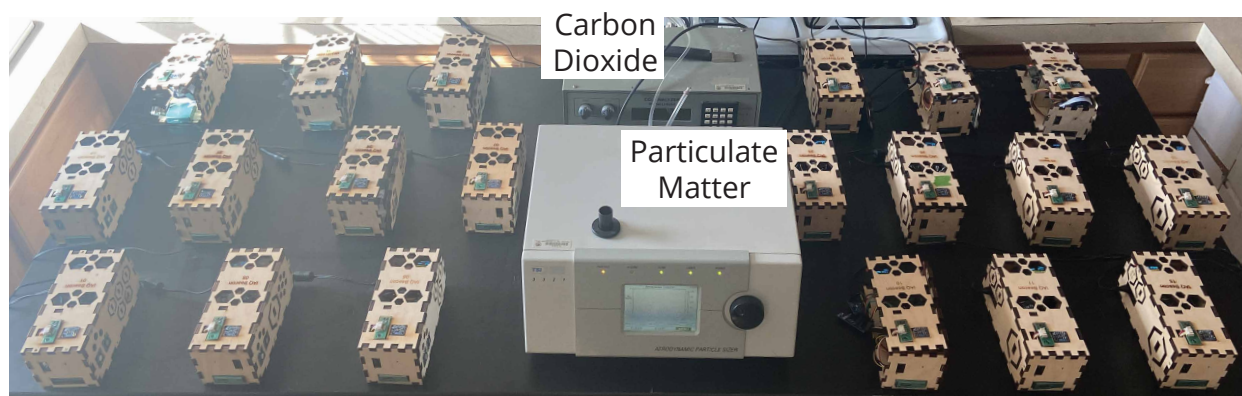


Figure 7: Calibration setup in the UTest House.

period of occupancy to ensure that the CO<sub>2</sub> concentration did not exceed 1500 ppm.

PM<sub>2.5</sub> was generated through the use of the same hand-operated nebulizer containing the same ultra-fine particles used in the UTest House environment (Section 2.4.1). We used two fans (not pictured) to provide mixing throughout the entire duration of the experiment. We controlled the particle injection so that the PM<sub>2.5</sub> concentration measured by the reference monitor did not exceed 50 µg/m<sup>3</sup> to ensure sensors would be calibrated for concentrations consistent with the indoor environment [38, 39].

For TVOC sensors, we do not have a reference monitor to compare measurements. Instead, we create an average curve from TVOC measurements made by each CGS at each timestamp. The average curve is then used as a reference to compare against. This process, while not a true calibration, helps ensure that the TVOC sensors are measuring similar concentrations given the same input.

Again, we conducted three two-hour experiments for each of the four pollutants to generate linear models for each BEVO Beacon by averaging parameters over the three experiments. We then conducted a fourth experiment to assess the models. Models for PM<sub>2.5</sub> and CO<sub>2</sub> sensors were created by comparing CGS measurements to research-grade monitors while models for TVOC sensors were created by comparing to the average, reference curve.

#### 2.4.3. Incubator

We calibrated temperature sensors through the use of a retrofitted incubator equipped with a electric heater and used a chilled-mirror hygrometer (Michell Instruments S8000) as a reference monitor. Experiments were conducted in 5 separate batches of four devices, each lasting two-hours. We varied the temperature from room temperature ( 21°C) to 32°C throughout the course of the experiment, but did not vary the RH. We conducted one experiment per device and compared measurements from the CGS to the reference values from the same experiment to create linear models.

#### 2.4.4. Gas Standard

We used a 10 ppm CO gas standard to perform a step-calibration on the CO sensors and to test the assumption that the background concentration in the laboratory is zero. We diluted the gas standard with Zero Air Gas (ZAG) to achieve CO concentrations of 0, 1, 2, and 4 ppm – typical indoor CO concentrations. We ensured the same 1 L/min

flowrate was achieved by varying the ratio of standard to ZAG. Batches of three BEVO Beacons were placed in a 5L chamber and allowed to run for 24 hours with only ZAG supplied. After this period, we ran each step in the calibration span for 2 hours. For each span, we only consider the middle 60 minutes which is obtained by removing the initial and last 30 minutes. Then we calculate the average CO over the 60-minute period and compare this value to calibration standard concentration at the current step. This method provides 4 data points from each step which we use to fit a linear model to correct CO readings made by each BEVO Beacon.

#### 2.5. Field Study

To assess the performance of the BEVO Beacons, 20 devices were deployed as part of a larger study of home environments from June 15th, 2020 to September 1st, 2020. We asked participants to place devices in their bedroom at approximately 1 meter above the ground and out of direct sunlight if possible.

### 3. Results

#### 3.1. Calibration

The following sections highlight the calibration results for each of the IAQ sensors on the BEVO Beacon. When presenting the results, we limit the number of devices we show since multiple devices share similar characteristics. However, the Appendix contains calibration results for all devices where data are available.

We develop device-specific models for each of the sensors we calibrate, defined as:

**Device-Specific** linear model parameters  $m$  and  $b$  which are unique for each sensor on each device

Since we calibrate the CO<sub>2</sub> and PM<sub>2.5</sub> sensors in two different environments, we can also create and compare environment-averaged models:

**Environment-Averaged** linear model parameters  $m$  and  $b$  which have been calculated by averaging across all devices calibrated in a specific environment and are applied to a given sensor for all devices

We refer to these two model types frequently and highlight their differences when presenting results for the CO<sub>2</sub> and PM<sub>2.5</sub> results, specifically.



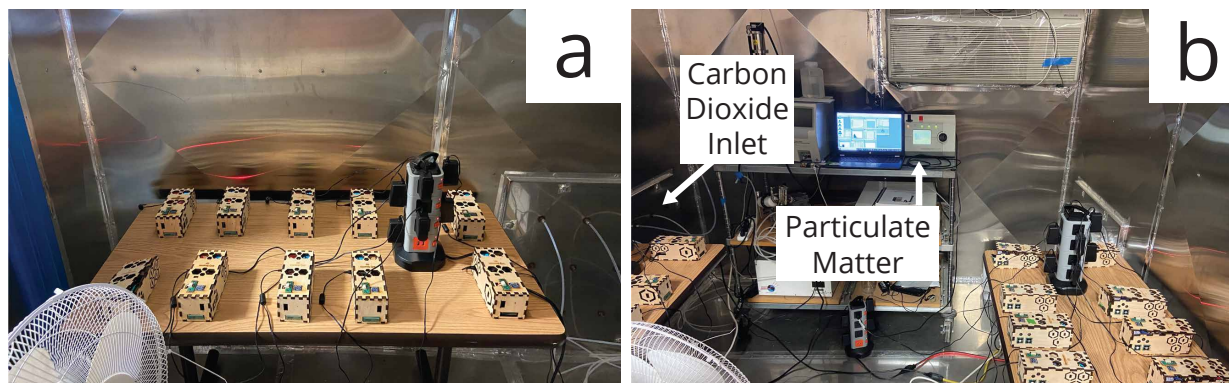


Figure 8: Calibration setup in the environmental chamber where (a) shows one of two tables holding 10 BEVO Beacons and (b) shows the  $PM_{2.5}$  monitor and inlet for the  $CO_2$  monitor

### 3.1.1. Carbon Dioxide

Figure 9a shows a performance summary of the device-specific linear correction models from  $CO_2$  calibration conducted in the UTest House. Eight devices have  $r^2 > 0.995$  with only one device having  $r^2 < 0.94$ . Device 24 has the lowest  $r^2$  of 0.619, but the device is still able to capture the general trend in  $CO_2$  concentration. Five devices exhibit more variation, or “noise”, in their measurements especially after the peak concentration is reached, but still exhibit excellent agreement with the reference monitor. The range of averaged  $m$  values for the correction models is 0.64 to 1.17. Values for  $b$  vary between 166.87 ppm to 435.17 ppm, indicating that the  $CO_2$  sensors used on the BEVO Beacon tend to underpredict the true concentration. Table A.1 contains all parameter values for each BEVO Beacon.

Figures 9b and 9c compare the  $CO_2$  outputs between device-specific and environment-averaged models used to correct  $CO_2$  measurements. Figure 9b shows the time series  $CO_2$  measurements made by all devices from both model types compared to the reference monitor. The figure includes a threshold around the reference measurement consistent with the resolution of the CGS for  $CO_2$  (see Table 2). Ideally, all curves should fall within this range, regardless of the calibration model. Errors between both model outputs and the reference tend to be smaller at lower  $CO_2$  concentrations and exhibit greater variation at elevated levels. The majority of device-specific model outputs are difficult to discern since they are contained within a narrow range near the reference while multiple environment-averaged

curves are easy to discern both above and below the reference line. The difference in these errors is illustrated in Figure 9c. The vast majority of errors between device-specific outputs and reference measurements are within the tolerance of the  $CO_2$  CGS while errors between environment-averaged corrections for 11 of 20 devices are outside this tolerance entirely.

Figure 10a highlights the results from the device-specific linear models for  $CO_2$  sensors from calibration conducted in the environmental chamber. Again, the CGS outputs corrected by the device-specific models have an excellent agreement with measurements made by the reference monitor – 16 of 20 devices have  $r^2 > 0.99$ , four of which have  $r^2 > 0.997$ . The remaining four devices tended to overpredict concentrations, especially at higher concentrations. Device 34 performed the worst with  $r^2 = 0.822$ . The sensor on this device is the only one with a  $m < 1$  and exhibits the greatest variation across experiments compared to all other devices. The remaining devices have averaged  $m$  coefficients in the range of 1.05 to 1.43 while there is more variability exhibited in averaged  $b$  values (see Table A.2).

Figure 10b shows the traces of  $CO_2$  measurements from each BEVO Beacon corrected by the device-specific and environment-averaged models from calibration experiments conducted in the environmental chamber. Outputs from the environment-averaged model exhibit greater variation, especially at elevated concentrations measured during the middle of the experiment. Again, the traces for the device-specific outputs are hard to

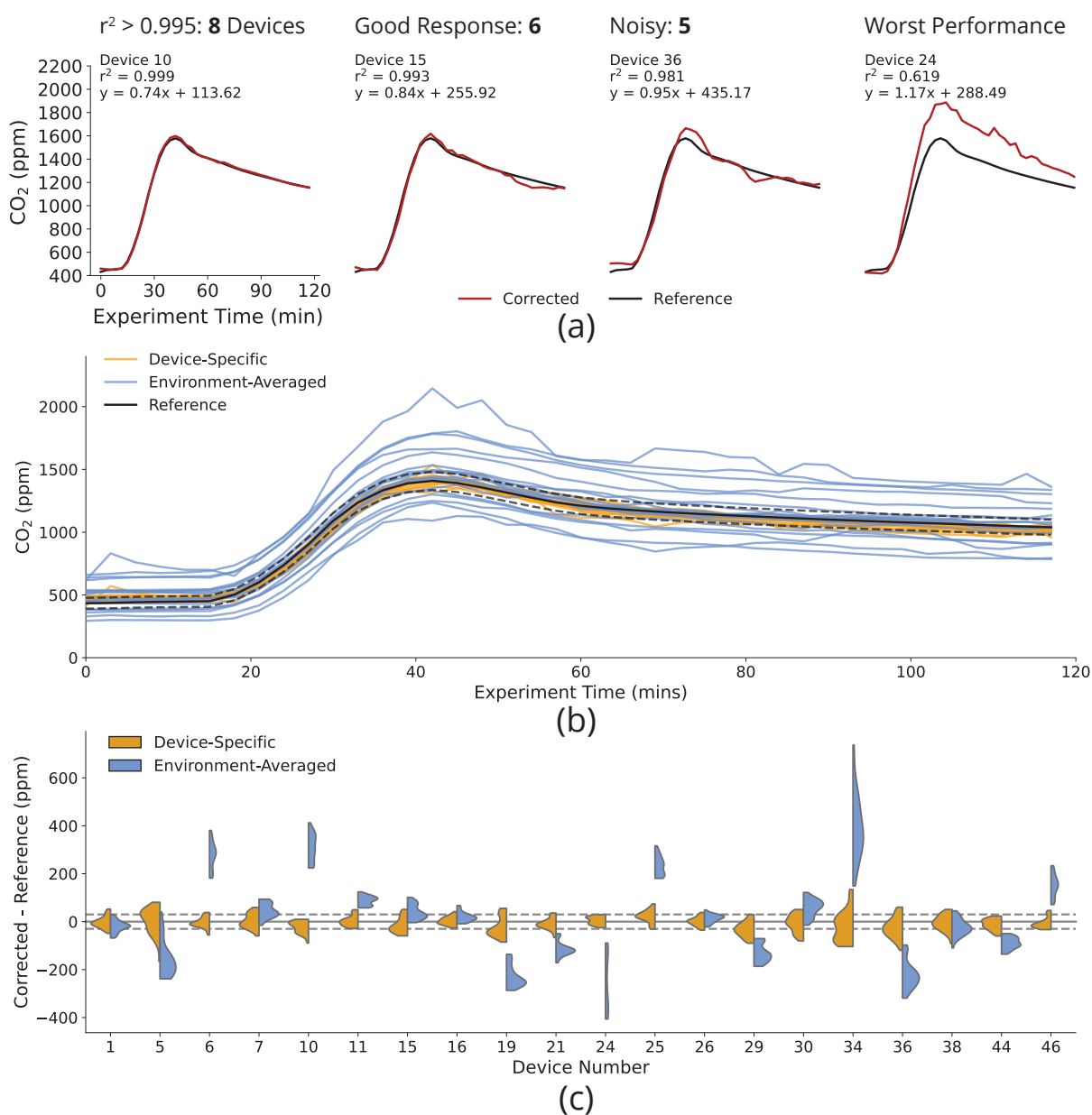


Figure 9: Performance summary of CO<sub>2</sub> linear regression models averaged from three experiments conducted in the UTest House. Data shown are from a fourth experiment where models are applied to devices' measurements and compared to the reference. Panel (a) shows typical responses for each of the 20 devices corrected by device-specific models while Panel (b) illustrates each CGS output when corrected with device-specific and environment-averaged models while Panel (c) highlights the distribution of errors between these models and the reference. Dashed lines around the reference in (b) and (c) correspond to the CGS resolution (see Table 2).

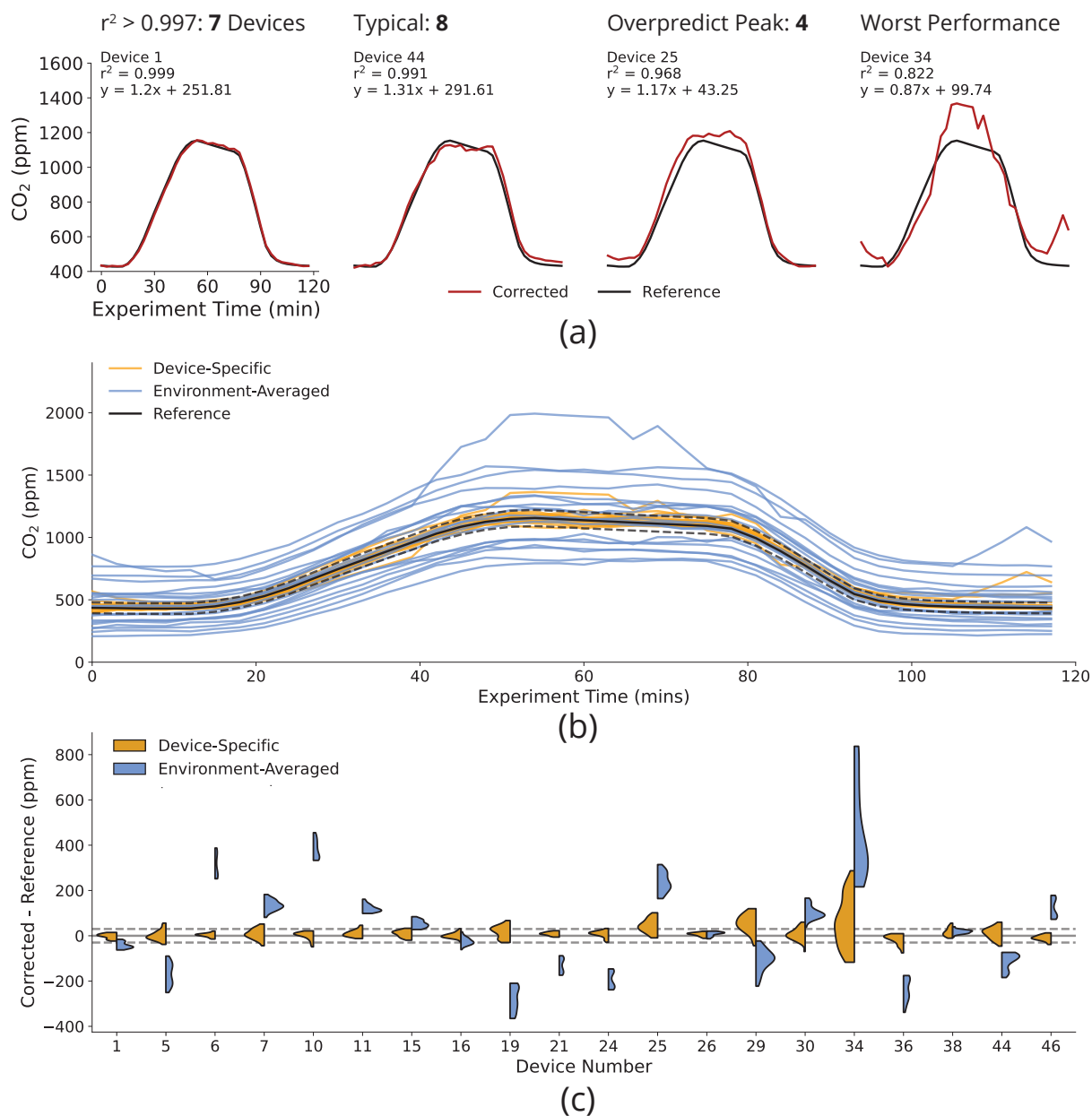


Figure 10: Performance summary of CO<sub>2</sub> linear regression models averaged from three experiments conducted in the environmental chamber. Data shown are from a fourth experiment where models are applied to devices' measurements and compared to the reference. Panel (a) shows typical responses for each of the 20 devices corrected by device-specific models while Panel (b) illustrates each CGS output when corrected with device-specific and environment-averaged models while Panel (c) highlights the distribution of errors between these models and the reference. Dashed lines around the reference in (b) and (c) correspond to the CGS resolution (see Table 2).



discern since many of them are between the tolerance range around the reference curve. The distribution of errors between reference and model outputs are given in Figure 10c. The majority of device-specific errors are contained within the resolution for the CGS for CO<sub>2</sub> with the exception of Devices 25, 29, and 34. However, the errors from these devices corrected by the device-specific models are still smaller than those corrected by the environment-averaged model. There are 14 out of 20 devices with all errors outside the tolerance range when CGS measurements are corrected by the environment-averaged model.

### 3.1.2. Particulate Matter

Figure 11a illustrates the performance of the calibration models for the PM<sub>2.5</sub> sensors derived from experiments conducted in the UTest House. Models typically exhibit poor performance with low, and sometimes negative,  $r^2$  values, ranging from -0.585 to 0.571. Measurements from the reference instrument indicate two clear events when PM<sub>2.5</sub> was generated. However, there are six devices that do not detect any signal from the first event and 3 devices that detect a significant third event around minute 90. In either case, these issues are the cause of the poor performance of these calibration models. However, there are some consistencies amongst the model parameters, namely that  $b$  values are all negative ranging from  $-3.1 \mu\text{g}/\text{m}^3$  to  $-15.1 \mu\text{g}/\text{m}^3$  and all  $m$  values are positive between 1.62 and 3.65. These values seem to indicate that these CGS tend to underpredict concentrations and are less sensitive to increases in concentration. However,  $m > 1$  could be the model compensating for the negative  $b$  values.

Figure 11b highlights the individual outputs of the device-specific and environment-averaged PM<sub>2.5</sub> models from experiments conducted in the UTest House. Thresholds around the reference line correspond to the resolution of the CGS for PM<sub>2.5</sub> (see Table 2) and represent a tolerance that all CGS measurements should be within. Nearly every measurement from each CGS is within this limit and distributions of the errors between devices and the reference are shown in Figure 11c. There is less of a difference between models for PM<sub>2.5</sub> than with CO<sub>2</sub> outputs. For devices such as 5 and 25, the model output are significantly different, with the environment-averaged model performing better than a device-specific model. The opposite is true for Devices 6 and 11, where the device-specific mod-

els are more appropriate. While results indicate that most measurements are within the tolerance we define, these sensors are still best suited to detect large variation in PM<sub>2.5</sub> concentrations which is evident by considering the  $r^2$ . The tolerance is based on a range of  $\pm 10 \mu\text{g}/\text{m}^3$  which is large considering most measurements from our field study are between 0 - 40  $\mu\text{g}/\text{m}^3$  (see Section 3.2).

Figure 12a highlights the performances of the calibrated PM<sub>2.5</sub> sensors from experiments conducted in the environmental chamber. Values for  $r^2$  still exhibit a wide range, between -0.131 to 0.913, but are better than results from the experimental test-house environment. Many of the sensors were able to accurately detect the injection of particles at the beginning of the experiment, but performance typically deteriorated at lower concentrations. Poor performance at lower concentrations is likely due to overfitting at higher concentrations. Many of the  $m$  parameters are high which means that any small perturbations in concentration will be amplified resulting in the cyclic behavior exhibited by most sensors after minute 60.  $b$  values are large and negative, which is again likely compensation for the large, positive  $m$  values. Yet, parameter values from each of the experiments conducted in the laboratory chamber exhibit less variability than in the UTest House which is expected since the laboratory exemplifies a more controlled environment. However, there is still considerable variability in the  $b$  values, some of which are larger than the measurement resolution of the CGS.

Individual traces for the device-specific and environment averaged model outputs for each device are shown in Figure 12b, corresponding to calibration experiments conducted in the environmental chamber. In general, the device-specific models performed better, with fewer traces outside the tolerance limits. However, both model outputs tend to underpredict the reference at both low and high concentrations, which is evident in Figure 12c. The majority of error distributions are skewed toward negative values, indicating reference concentrations are higher on average. Only a few devices – namely Devices 6, 11, and 38 – have significantly different error distributions when comparing the device-specific and environment-averaged models. The range of errors is larger for PM<sub>2.5</sub> calibration conducted in the chamber compared to the experiments in the UTest House, but this is likely because concentration profiles are much different.

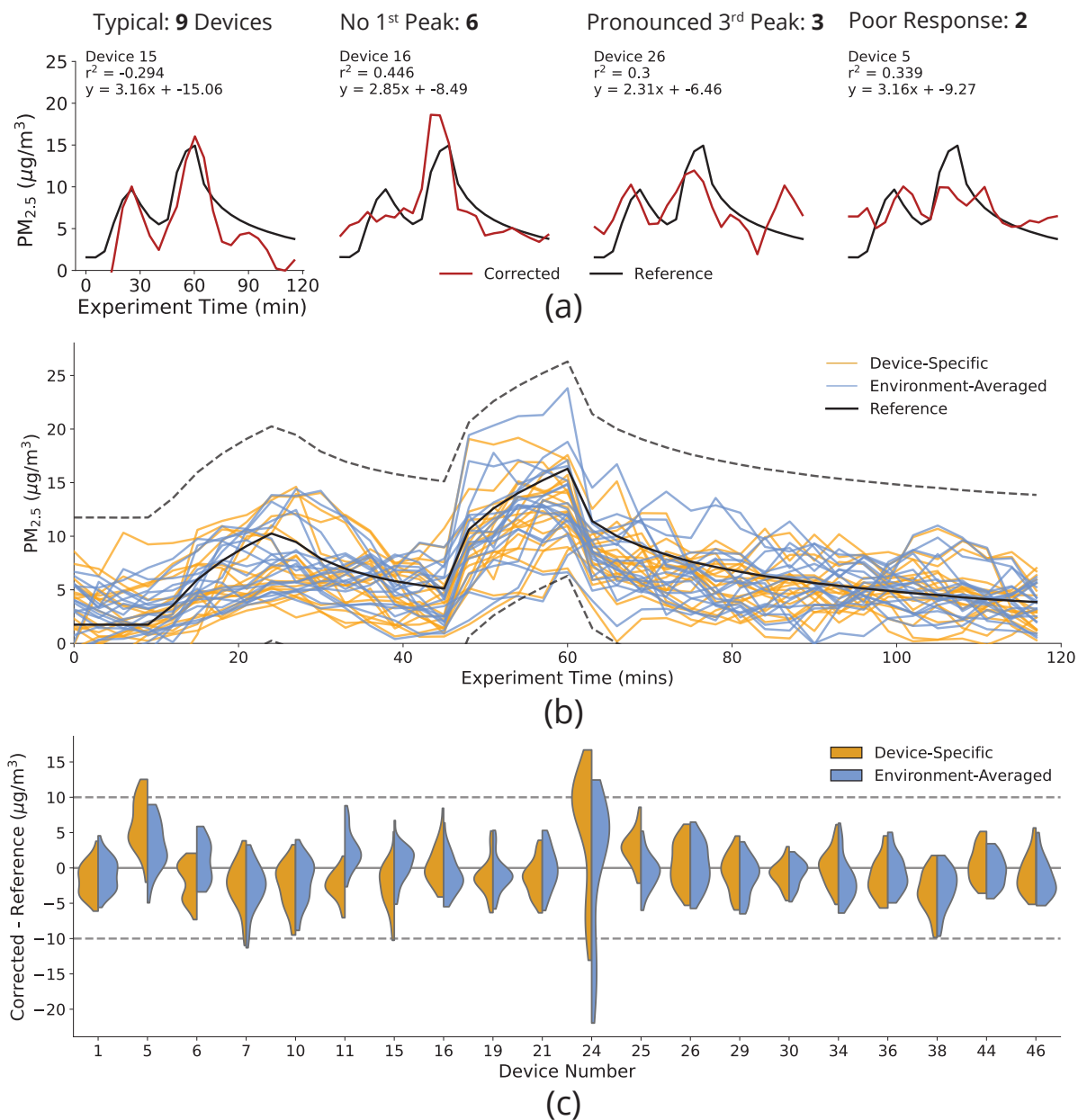


Figure 11: Performance summary of  $PM_{2.5}$  linear regression models averaged from three experiments conducted in the UTest House. Data shown are from a fourth experiment where models are applied to devices' measurements and compared to the reference. Panel (a) shows typical responses for each of the 20 devices corrected by device-specific models. Panel (b) illustrates each CGS output when corrected with device-specific and environment-averaged models while Panel (c) highlights the distribution of errors between these models and the reference. Dashed lines around the reference in (b) and (c) correspond to the CGS resolution (see Table 2).

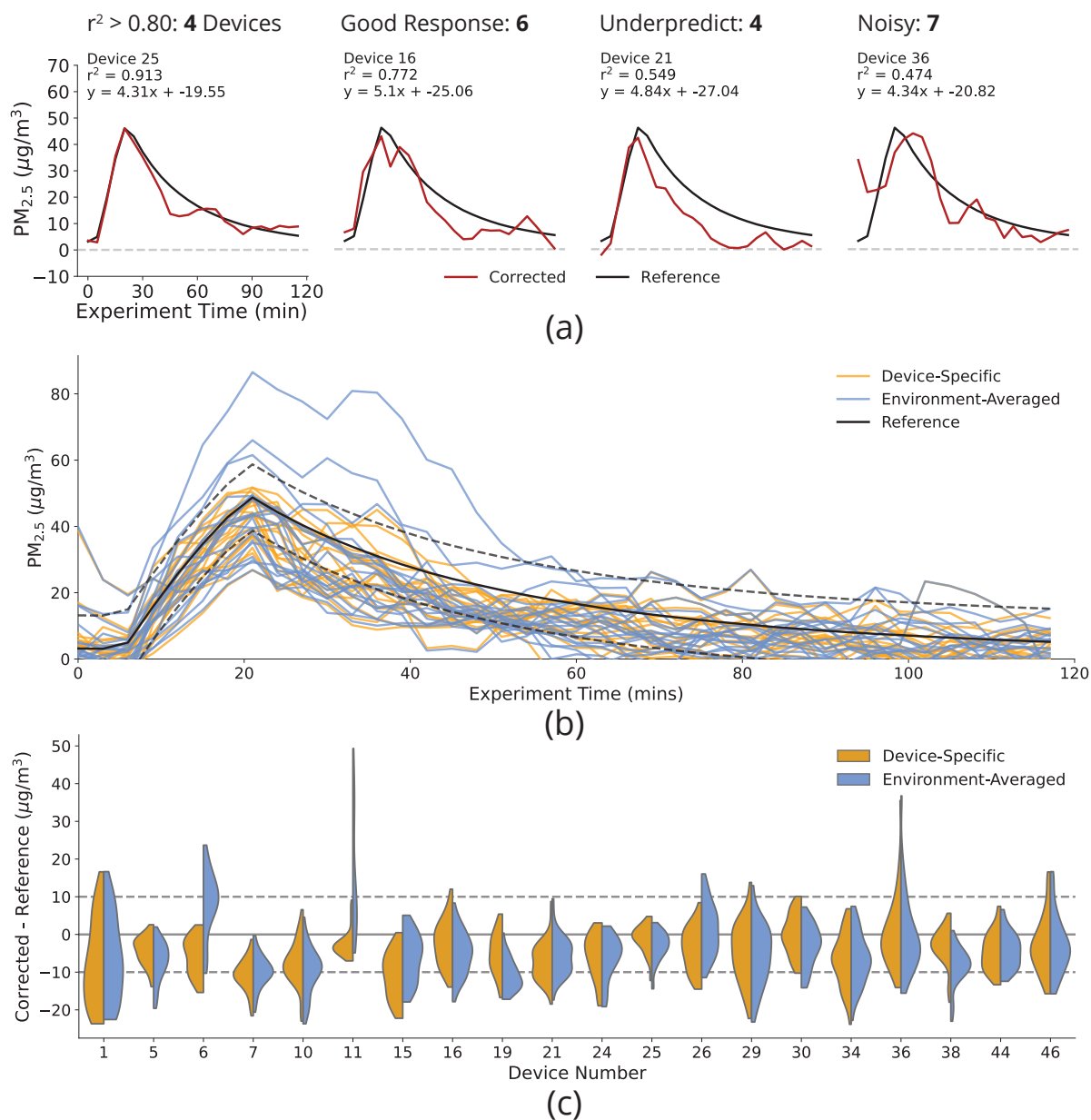


Figure 12: Performance summary of PM<sub>2.5</sub> linear regression models averaged from three experiments conducted in the environmental chamber. Data shown are from a fourth experiment where models are applied to devices' measurements and compared to the reference. Panel (a) shows typical responses for each of the 20 devices corrected by device-specific models. Panel (b) illustrates each CGS output when corrected with device-specific and environment-averaged models while Panel (c) highlights the distribution of errors between these models and the reference. Dashed lines around the reference in (b) and (c) correspond to the CGS resolution (see Table 2).

### 3.1.3. Total Volatile Organic Compounds

The results from calibrating the TVOC sensors in the laboratory chamber are shown in Figure 13. The black reference line indicates the average concentration calculated from all devices. Devices 34 and 44 had the worst performing models with  $r^2 = 0.873$  and  $r^2 = 0.585$ , respectively, but the remaining devices all had  $r^2 > 0.95$ . The strength of the models is surprising given the wide range of  $m$  and  $b$  values from the three calibration experiments (see Table A.5 for more details). Typically, TVOC sensors tended to underpredict the peak concentration while 3 devices overpredicted peak concentrations. While the averaged  $b$  parameters appear to exhibit a wide range of values, the variation is relatively small given the high TVOC concentrations that were measured during experiments. The averaged  $m$  coefficient ranged from 0.46 to 1.63 which indicates that perhaps issues with manufacturing cause sensors to be more or less sensitive to the same VOCs.

### 3.1.4. Carbon Monoxide

Figure 14 shows the results from calibrating the CO sensor against the gas standard. The original data used to derive the linear models for Devices 5, 11, 16, and 24 was overwritten and lost while the CO sensors on Devices 19, 38, and 46 were non-responsive during experiments. Of the remaining 13 devices, the CO sensors respond similarly and the linear models all have  $r^2 > 0.98$ . As mentioned in Section 2.4.4, the parameters are determined by only considering measurements over the span of 60 minutes for each step after allowing 30 minutes for the sensors to acclimate to the new concentration. This process helps to ensure that we are calibrating to the correct value since some BEVO Beacons have a slower response rate to the change in concentration which is evident by inspecting measurements made directly after the CO concentration increased from 2 ppm to 4 ppm.

### 3.1.5. Temperature

Table 4 shows the linear model parameters for temperature sensors on each of the BEVO Beacons. We group the devices that were calibrated in the same experiment together. Only the last group (Devices 11 and 15), show similarities in the model parameters although we only have two devices to compare. For the remaining groups, there do not appear to be clear patterns amongst the  $b$  nor  $m$  values. However,  $b$  values appear to be distributed

Table 4: Linear model parameters for each of the temperature sensors from one experiment conducted in the small laboratory chamber.

Device	$b$	$m$	$r^2$
1	-33.28	1.81	0.94
24	-1.39	1.00	0.99
34	-0.16	0.96	0.99
36	-19.41	1.70	0.91
44	-18.80	1.73	0.92
46	-2.20	1.07	0.98
6	-16.79	1.53	0.94
10	-4.06	1.18	0.89
26	-17.39	1.55	0.93
30	-12.39	1.46	0.95
7	-3.27	1.09	0.90
16	-4.18	1.12	0.92
19	-4.86	1.21	0.94
21	-5.50	1.15	0.94
5	-13.03	1.43	0.96
25	-2.73	1.00	0.99
29	-15.23	1.53	0.95
38	-1.53	1.00	0.99
11	-32.08	2.03	0.92
15	-35.41	2.18	0.94

into three distinct groups based on decreasing values: -0.16 to -5.50, -12.39 to -19.41, and -32.08 to -35.41. The corresponding  $m$  values follow an opposite pattern where devices with lower  $b$  values have larger  $m$  coefficients. Devices with  $b$  close to 0 have  $m$  values closer to 1. Despite the differences in model parameters, the correlation coefficients indicate excellent agreement with the reference monitor ranging from 0.89 to 0.99. Devices 24, 25, and 38 all have  $m = 1$  and the highest correlation coefficients, indicating a constant model could be appropriate.

### 3.2. Field Study

Table 5 summarizes all measurements made by the 20 devices during the field study conducted in the summer of 2020. For all parameters listed, the mean and median values are generally similar indicating that the distributions are likely Gaussian. Distributions of each IAQ parameter, ordered by increasing mean value, made by each BEVO Beacon are shown in Figure 15. All but Device 6 measured more than 41 day's worth of data for each sensor with 7 devices recording more than 70 day's worth of data for each sensor. Device 6 measured

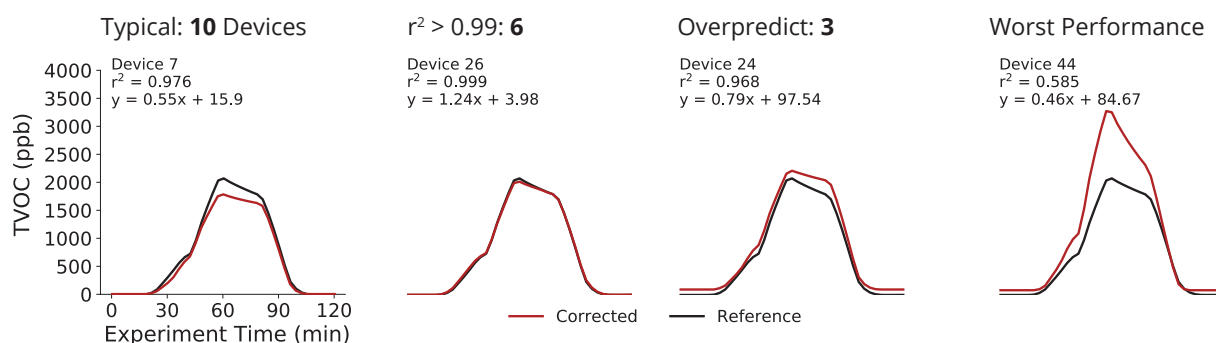


Figure 13: Summary of the performance of TVOC linear regression models averaged from the three experiments conducted in the environmental chamber. Data shown are from a fourth experiment where models are applied to devices’ measurements and compared to the reference monitor. We highlight results from three devices that represent typical performance, excellent performance, and devices with a tendency to over-predict peak concentrations in addition to showing the worst performing model as measured by the  $r^2$ .

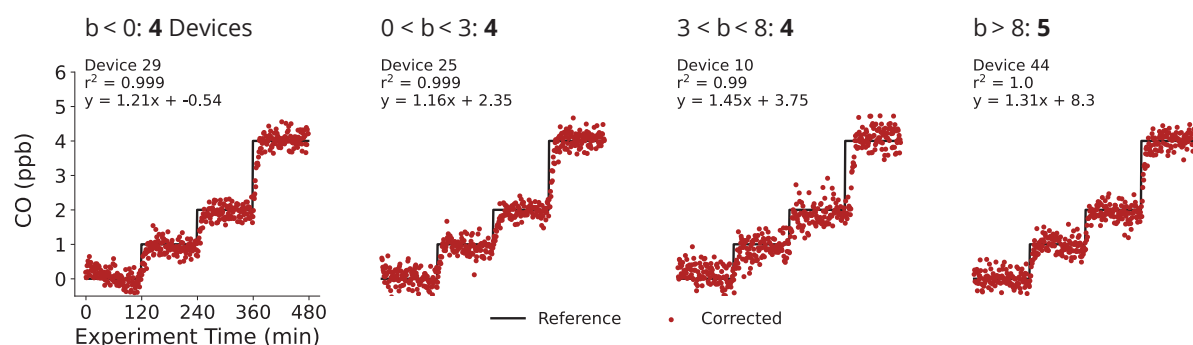


Figure 14: Summary of the CO linear models we generated from the gas standard calibration arranged by the  $x_0$  parameter. Three sensors malfunctioned prior to calibration so we only summarize 17 devices.

Table 5: Summary of calibrated measurements made by the BEVO Beacons during the field study. For  $PM_{2.5}$ , TVOC, CO, and light sensors there were many instances of non-detect (ND) indicating that concentrations were approximately zero.

Variable	Mean	Median	Min	25%	75%	95%
CO <sub>2</sub> (ppm)	1050.3	970.7	236.6	759.5	1268.5	1786.9
PM <sub>2.5</sub> ( $\mu\text{g}/\text{m}^3$ )	13.3	11.6	ND	6.0	17.6	33.3
TVOC (ppb)	242.4	215.5	ND	113.7	338.9	552.6
CO (ppm)	3.5	2.7	ND	1.2	4.5	10.4
Light (lux)	24.1	2.0	ND	ND	11.3	53.4
T ( $^{\circ}\text{C}$ )	25.8	26.3	15.4	24.5	27.7	29.8
RH (%)	42.0	41.3	23.4	38.3	44.5	51.8

the minimum of just over 9 day’s worth of data for each sensor.

CO<sub>2</sub> distributions are generally normal and unimodal, peaking around 1000 ppm with the exception of Devices 34, 46, and 11 which indicate clear bimodal distributions. Bimodal distributions are likely an indication of periods of low and high occupancy or ventilation. Measurements on Device 11 are significantly higher than any device and exhibit a wide range of values indicating a poorly ventilated space. CO<sub>2</sub> measurements made by Devices 44, 25, and 1 indicate the opposite – a narrow range of low concentrations – despite measuring over a similar range of days. Device 16 measured the minimum of 9.5 days while Device 30 measured 77.6 days.

The PM<sub>2.5</sub> distributions shown in Figure 15b are skewed right, characterized by a few episodes of high concentrations. These events are likely due to activities like vacuuming or cooking that tend to generate a large amount of PM. Device 44 is the only BEVO Beacon that shows a clear bimodal distribution, which indicates more instances of these episodes. There are only a few devices that measure concentrations as low as the detection limit while the majority measure minimum concentrations of approximately 5  $\mu\text{g}/\text{m}^3$ . Device 16 again measured the minimum amount of PM<sub>2.5</sub> data points while Device 25 measured the maximum of 77.1 days.

The Temperature (T) measurements, averaged between the CO and NO<sub>2</sub> modules (if available) are shown in Figure 15c. Generally, each device measures over a narrow range of values within a participant’s environment. T distributions exhibit multimodal behavior which is a consequence of the resolution of the T sensor. Devices 21, 7, and 15 have the greatest range of T measurements – over the range of approximately 10°C while Devices 5 and 44 show minor variations in T. Device 6 did not record any T measurements while Device 30 measured the maximum amount of data at 77.6 days.

Summary statistics for TVOC measurements in Table 5 highlight that the aggregate measurements are approximately normally distributed, and many of the distributions of TVOC in Figure 15d confirm this observation. The majority of TVOC concentrations measured by each device are within the same range of values – up to approximately 500 ppb. The devices with greater mean concentrations are characterized by a few episodes of high concentration measurements, reminiscent of the PM<sub>2.5</sub> distributions. Only Device 26 operates over a narrow measurement range of approximately 200 ppb, while

the remaining devices measure significantly larger ranges of concentrations. Many of these distributions are unimodal with a few devices such as 11, 16, and 34 that exhibit two identifiable modes. Device 6 measured the minimum amount of TVOC – 9.5 day’s worth of data – while and Device 10 measured the maximum of 76.7 days.

Figure 15e highlights the CO concentrations measured during the field study. There is a clear distinction between low CO households and more polluted environments starting with Device 21. Distributions of measurements with means less than this device are characterized by measuring below the detection limit for nearly the entire study period with a few instances of high concentrations that are likely measured when the device was powered on for the first time. The remaining devices also measure a few episodes of high concentrations, but their mean concentrations are significantly higher which perhaps indicates that the home uses a gas-powered stove. Devices 21, 15, 5, 7, and 1 all have higher, but safe concentrations of measured CO. Devices with greater mean concentrations – starting with Device 36 – measure concentrations that might induce minor health effects like headaches or sensory irritation. The last device, Device 34, measures even higher concentrations of CO that are likely causing the occupant some level of discomfort. In terms of the number of data points measured, Device 16 measured the minimum of 9.5 days while Device 25 measured all possible CO data points, reporting 78 days of data.

Lastly, Figure 15f indicates that light levels were uncharacteristically low for the majority of the study. Only Device 38 measured light levels greater than 10 lux consistently. There are no records to indicate the location of the device or if the light sensor was obscured in any way. We expected light levels to be higher and more consistent with results from [40] which used the same sensors.

## 4. Discussion

### 4.1. Comparison to Similar Devices

CGS for indoor air quality monitoring are being used successfully and we highlight many of these studies in Table 1. All devices we include in Table 1 measure T and RH, and each device, except the one developed in [33], measure CO<sub>2</sub> and PM. Sensors for CO are also common, included on seven of the ten devices. A few of the less commonly measured

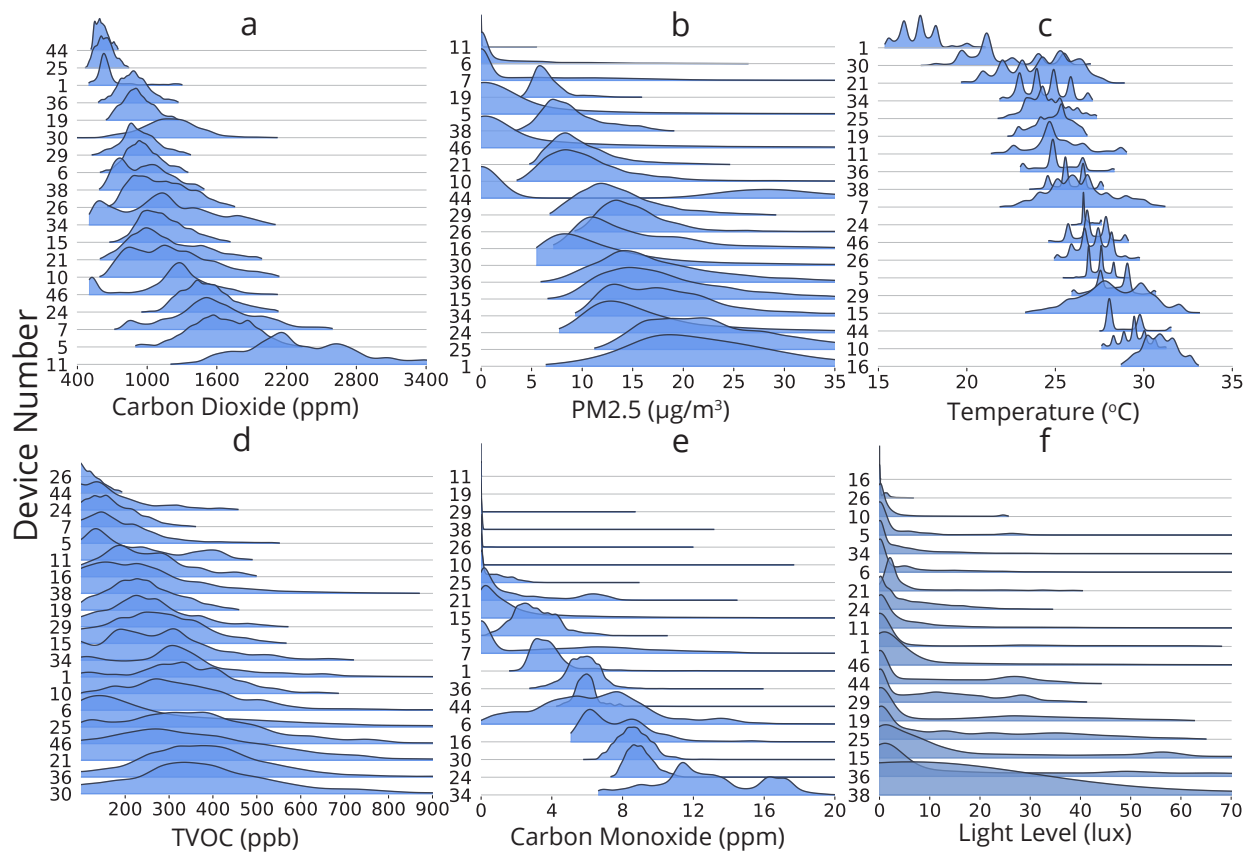


Figure 15: Distributions of measured IAQ values from each sensor on the BEVO Beacons ordered by increasing mean concentration.

parameters include oxides of nitrogen (5), Ozone ( $O_3$ ) (2), and Sulfur Dioxide ( $SO_2$ ) (1). Rather than measure TVOC concentrations, a few devices opt to measure specific VOCs like Formaldehyde (HCHO) [31] and Methane ( $CH_4$ ) [33, 34]. The device in [33] measures a wide variety of VOCs and includes empirical corrections to account for cross-sensitivities.

Our study is one of two studies we identified that mentions calibration in both field and laboratory settings. All studies evaluate their devices in the field through IAQ monitoring campaigns, but do not necessarily provide information regarding calibration [26, 28]. In [24], the device uses a characteristic table with coefficients to empirically correct readings while the researchers in [27] develop calibration coefficients by using information provided in the sensors' datasheets along with a correction for T and RH. The remaining related studies in Table 1 apply more traditional calibration techniques, which we discuss in more detail in Section 4.2.

The BEVO Beacon shares many similarities with recent devices in regard to measured parameters, approximate costs, and the calibration processes. However, a few key components distinguish our device from others: (1) complete open-source availability, (2) option to customize the type and quantity of parameters measured with minimal effort thus reducing the overall cost, and (3) ease of development and use. We were able to create, calibrate, and deploy 20 devices. Furthermore, the CO and  $NO_2$  sensors are developed by a company that currently produces four other modules that can easily be switched in with no changes needed to the software. The remaining CGS use I<sup>2</sup>C protocol, which RPs can support tens of I<sup>2</sup>C connections. Therefore, researchers can include more devices by wiring them in series with the existing connections to suit their research needs. Once created, the BEVO Beacons can begin to collect data immediately once powered on. Data can be easily accessed from the device and used immediately for analysis. So far these devices have been used multiple times by students for class projects and more formal research [41] inquiries including studies analyzing the relationship between sleep and IAQ [42] and ventilation estimation [43].

The BEVO Beacon is meant to provide a framework for other researchers to reduce the upfront time needed to develop a similar device. While our field study was successful, the BEVO Beacons can still be improved. Unforeseen software and hard-

ware issues caused significant data loss for some devices. Sensors also require routine calibration and newer, more robust sensors are being developed which could help ameliorate the accuracy and array of IAQ parameters monitored. Our hope is that by providing open-source documentation, collaborators can build from our design, modify for their own purposes, and provide contributions to the project to improve our device.

#### 4.2. Consumer-Grade Sensor Calibration

The calibration model parameters that we developed and used to correct the measurements from our CGS are unique to our devices and are likely not applicable to those created by other researchers. Furthermore, these parameters will need to be recalculated since sensors – both CGS and reference-grade – are prone to drift over time [19]. The primary purpose of our extensive calibration efforts was to provide an overview of techniques that could be used depending on the availability of calibration environments, reference-grade monitors, and/or gas standards, which we discuss further in the subsequent paragraphs. Calibration of CGS is a necessary step since the manufacturing processes – for both the individual sensors modules and the device – can induce minor differences in sensitivity which causes sensors manufactured by the same company to respond differently even under similar IAQ conditions [34]. The wide range of linear model parameters we derive illustrates this point since parameters would be more consistent if sensors were manufactured with similar sensitivities. In addition, manufacturer calibration might have been conducted under conditions which are different than those researchers intend to monitor. Ideally, CGS should be calibrated under similar conditions to those they will be deployed to [23]. Furthermore, calibration should occur after the device has been fully built in case the hardware configuration causes sensor modules to receive more or less power which might alter how the CGS converts voltage to concentration.

The variety of IAQ sensors on the BEVO Beacon provided an opportunity to explore various calibration techniques: (1) co-locating CGS with research-grade monitors, (2) exposing CGS to known pollutant concentrations, and (3) correcting sensors relative to each other to control for issues with batch quality in the absence of a research-grade monitor. A vast majority of studies calibrate CGS in field



or lab environments by co-locating with research-grade monitors and providing a pollutant signal [44, 45] – similar to our process for CO<sub>2</sub>, PM<sub>2.5</sub>, and T sensors. A more systematic approach involves generating pollution either by diluting gas standards with ZAG to target specific concentrations – like we did with CO sensors – or creating known mixtures of compounds. The former method is useful for pollutants like CO and NO<sub>2</sub> [46] which are challenging to generate safely while the latter can be used for PM and TVOCs sensors which respond differently depending on the mixture [47, 48, 49, 50]. In either case, the span of possible concentrations is limited and thus models are fit with only a few data points. In the absence of a research-grade monitor and the ability to generate known concentrations, researchers can calibrate sensors relative to each other by defining a reference curve based on the aggregated measurements from all devices being tested [51]. We used this process for TVOC sensors and, while not a true calibration, this method helps correct for issues with batch quality.

For any of calibration method, one must consider the concentration profile generated during the experiment. Ideally, CGS should be calibrated to concentrations that are likely to occur in their planned location. Calibrating outside this range is unnecessary and might introduce bias since the models that are generated might be sensitive to extreme measurements. CGS are known to perform differently at disparate concentration ranges [52], which is evident in many CGS datasheets. This behavior is evident when comparing the error distribution between CGS outputs and the reference results from the PM<sub>2.5</sub> calibration in the UTest House (Figure 11c) and environmental chamber (12). Errors are smaller and span a more limited range in the UTest House primarily because PM<sub>2.5</sub> concentrations measured by the reference are lower than in the environmental chamber, which has a peak concentration more than twice that in the UTest House. Furthermore, the shape of the concentration profile can affect the final model since CGS might be more or less sensitive to increases and/or decreases in pollutant concentrations. With the exception of the CO sensor, we introduce pollution on one or more occasions during calibration experiments. The CO<sub>2</sub> sensor (Figures 9 and 10) captures the concentration profiles exceptionally well. PM<sub>2.5</sub> sensors (Figures 11 and 12) capture the increase in concentration well but tend to underpredict more

lengthly decreases, and the TVOC sensors (Figure 13) tend to underpredict peak concentrations.

The environment in which sensors are calibrated is important since models can vary significantly between controlled and more realistic environments [46]. Results from our study confirm this finding for the CO<sub>2</sub> and PM<sub>2.5</sub> sensors. No BEVO Beacon had similar device-specific models between environments, even when considering CO<sub>2</sub> sensors which had excellent agreement with the research-grade monitors in both environments. Device-specific parameters for each sensor from each environment are provided in the Appendix. Many researchers recommend field tests over laboratory calibration because sensors are exposed to more realistic environmental conditions [53]. Despite this suggestion, many of the devices listed in Table 1 are calibrated in laboratory environments and evaluated in the field. Laboratory settings are favored because controlled settings allow researchers to remove confounding variables like mixing-conditions and T/RH variation. Temperature and RH are two environmental parameters that are known to affect CGS, most notably PM<sub>2.5</sub> [54] and MOS sensors [55]. Many of the studies that include T and RH in their calibration models are for CGS operating in ambient conditions (see references within [45]) which have more varied conditions [56] compared to indoor environments which is why we do not include T or RH in our models. Heat and cold tests performed on MOS sensors in [30] indicated that temperatures between 15.6°C and 23.8°C – consistent with indoor temperatures – did not noticeably affect sensor readings. However, researchers in [34] include T and RH corrections for each of their sensors using data gathered from field evaluations.

For some pollutants, researchers often make the assumption that the background concentration is zero or some other constant value. In this case, researchers only need to correct readings by subtracting a constant value. Initially, we tested this assumption on the CO sensor, placing devices in the environmental chamber and assuming a background concentration of 0 ppm. However, these experiments indicated appreciable variations in CO concentration, measuring CO up to 4 ppm during various experiments. For some environments, such as a typical laboratory setting, this assumption is generally safe to make. However, researchers should be careful, especially for CGS that monitor gaseous pollutants and can be sensitive to other compounds.

Beyond the calibration setup and procedure, re-

searchers are then faced with the challenging decision of which model they should use to correct the CGS readings. These models can be as simple as ordinary least-squares regression [57, 58] to highly complex neural networks [59]. As discussed earlier, these models can incorporate covariates such as T and RH in addition to other gaseous compounds that present appreciable cross-sensitivities. In this study, we opt for univariate, least-squares linear regression models because of their simplicity, explainability, and ubiquity in the related literature [45]. Related device presented in [32, 30, 34] use linear models to correct their sensors' readings while [29] uses constant offsets derived from experiments with ZAG.

Researchers must also make the decision as to whether device-specific or environment-averaged parameters should be used. Our study indicates the former is more appropriate which has been corroborated by other studies on CGS [60] including a related study [30]. Researchers in [32] calibrate 100 devices but do not indicate whether they apply device-specific or averaged models while devices in [34] are calibrated frequently, including an on-line calibration system which implies device-specific models are used. The CO<sub>2</sub> model results in Figures 9 and 10 – specifically Panels (b) and (c) – highlight how important device-specific models are. Measurement errors between the reference and device-specific models span a far narrower range and are typically centered around zero while errors from environment-averaged models vary considerably across devices. However, when considering PM<sub>2.5</sub> models in Figures 11 and 12 the difference is not as stark which implies that for some sensors, an averaged model can be appropriate. This approach simplifies the calibration process considerably especially if researchers plan to develop many devices. Rather than calibrate each device individually, researchers could simply calibrate a few devices, generate a model, and apply it across all devices.

## 5. Conclusion

In this study we presented the Building Occupancy and EnVironment Beacon – an all-in-one IAQ monitor that leverages multiple, consumer-grade sensor modules. Our goal was to create a device that researchers with limited knowledge in embedded systems could replicate and customize to their research efforts. The BEVO Beacon uses

six sensors to measure 19 different parameters including CO<sub>2</sub>, PM number and mass concentrations, TVOCs, CO, T, RH, and light. Data are measured at a one-minute resolution and stored locally on the device but can be accessed by researchers remotely if configured to WiFi.

We present extensive results regarding calibration options for each of the primary IAQ sensors. Linear models for PM<sub>2.5</sub> and CO<sub>2</sub> sensors were developed from a controlled laboratory setting and a home testing environment by comparing measurements to research-grade monitors. Models for both IAQ parameters were different depending on the environment, but in both settings, we were able to correct CO<sub>2</sub> readings to achieve  $r^2 > 0.9$  for many devices. Models for PM<sub>2.5</sub> varied widely and performance was significantly worse. We also compared the performance of PM<sub>2.5</sub> and CO<sub>2</sub> models with device-specific parameters to models with parameters averaged across all beacons in each environment. We found that device-specific models contained parameters that varied significantly and were more appropriate than a single model applied to multiple devices. TVOC sensors were calibrated relative to each other to ensure each sensor was reading similarly. CO sensors were calibrated through a controlled step calibration in a 5L chamber, achieving  $r^2 > 0.98$  for all devices. T was calibrated against a reference monitor by comparing temperatures spanning from room temperature to 10°C degrees warmer in retrofitted incubator. Models were able to achieve  $r^2 > 0.89$  across all devices, but parameters varied considerably and we only considered a narrow T range.

To understand the performance of our devices, we deployed 20 BEVO Beacons in a field study for 11 weeks during the summer of 2020 to student participants living in home and apartment dwellings. Measurements from each device indicate that the majority of the measured IAQ parameters were within typical ranges for indoor environments. Devices recorded a mean of 62 days with a minimum of 9 and maximum of 78 days – the entire study period. The quality and quantity of measurements we recovered from the field study underscore the ability of CGS to provide insight into the IAQ from a large number of spaces and highlight their ability to gather data over extended periods of time.

The variety and affordability of consumer-grade IAQ sensors means that researchers can now develop and tailor devices to their specific needs. For example, our contribution provides a robust base

hardware setup for the field implementation of occupancy detection [61], occupant-centric building controls [62], and post occupancy evaluation studies [63]. While these sensors have accuracy limitations, they can help to answer research questions, especially where large variation in IAQ parameters is expected. In addition, companies and third-party users provide extensive documentation, reducing the burden of developing these sensors for researchers and the public alike which helps grow community science efforts toward measuring and understanding air quality. Creating devices like the BEVO Beacon is still challenging, but is likely to get easier as researchers and communities get more involved and sensor technology matures, leading to an increase in the quantity and quality of air quality measurements.

## 6. Acknowledgements

This work was supported by Whole Communities—Whole Health, a research grand challenge at the University of Texas at Austin.

## References

- [1] B. Lévesque, V. Huppé, M. Dubé, R. Fachehoun, Impact of indoor air quality on respiratory health: results of a local survey on housing environment, *Public health* 163 (2018) 76–79.
- [2] H.-C. Chuang, K.-F. Ho, L.-Y. Lin, T.-Y. Chang, G.-B. Hong, C.-M. Ma, I.-J. Liu, K.-J. Chuang, Long-term indoor air conditioner filtration and cardiovascular health: a randomized crossover intervention study, *Environment international* 106 (2017) 91–96.
- [3] N. Fiedler, K. Kelly-McNeil, P. Ohman-Strickland, J. Zhang, J. Ottenweller, H. M. Kipen, Negative affect and chemical intolerance as risk factors for building-related symptoms: a controlled exposure study, *Psychosomatic medicine* 70 (2) (2008) 254–262.
- [4] I. Mujan, A. S. Anđelković, V. Munćan, M. Kljajić, D. Ružić, Influence of indoor environmental quality on human health and productivity—a review, *Journal of cleaner production* 217 (2019) 646–657.
- [5] O. A. Seppänen, W. Fisk, Some quantitative relations between indoor environmental quality and work performance or health, *Hvac&R Research* 12 (4) (2006) 957–973.
- [6] N. E. Klepeis, W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, W. H. Engelmann, The national human activity pattern survey (nhaps): a resource for assessing exposure to environmental pollutants, *Journal of Exposure Science & Environmental Epidemiology* 11 (3) (2001) 231–252.
- [7] A. Prüss-Üstün, J. Wolf, C. Corvalán, R. Bos, M. Neira, Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks, World Health Organization, 2016.
- [8] J. Y. Park, Z. Nagy, Comprehensive analysis of the relationship between thermal comfort and building control research—a data-driven literature review, *Renewable and Sustainable Energy Reviews* 82 (2018) 2664–2679.
- [9] A. L. Clements, W. G. Griswold, A. Rs, J. E. Johnston, M. M. Herting, J. Thorson, A. Collier-Oxandale, M. Hannigan, Low-cost air quality monitoring tools: from research to practice (a workshop summary), *Sensors* 17 (11) (2017) 2478.
- [10] E. G. Snyder, T. H. Watkins, P. A. Solomon, E. D. Thoma, R. W. Williams, G. S. Hagler, D. Shelow, D. A. Hindin, V. J. Kilaru, P. W. Preuss, The changing paradigm of air pollution monitoring, *Environmental science & technology* 47 (20) (2013) 11369–11377.
- [11] A. C. Rai, P. Kumar, F. Pilla, A. N. Skouloudis, S. Di Sabatino, C. Ratti, A. Yasar, D. Rickerby, End-user perspective of low-cost sensors for outdoor air pollution monitoring, *Science of The Total Environment* 607 (2017) 691–705.
- [12] N. Castell, M. Viana, M. C. Minguillón, C. Guerreiro, X. Querol, Real-world application of new sensor technologies for air quality monitoring, *ETC/ACM Technical Paper* 16 (2013) 34.
- [13] P. M. Santos, J. G. Rodrigues, S. B. Cruz, T. Lourenço, P. M. d’Orey, Y. Luis, C. Rocha, S. Sousa, S. Crisóstomo, C. Queirós, et al., Portolivinglab: An IoT-based sensing platform for smart cities, *IEEE Internet of Things Journal* 5 (2) (2018) 523–532.
- [14] R. M. White, I. Paprotny, F. Doering, W. E. Cascio, P. A. Solomon, L. A. Gundel, et al., Sensors and ‘apps’ for community-based atmospheric monitoring, *EM Air Waste Manag. Assoc. Mag. Environ. Manag* 5 (2012) 36–40.
- [15] J. E. Thompson, Crowd-sourced air quality studies: A review of the literature & portable sensors, *Trends in Environmental Analytical Chemistry* 11 (2016) 23–34.
- [16] K. A. Koehler, T. M. Peters, New methods for personal exposure monitoring for airborne particles, *Current environmental health reports* 2 (4) (2015) 399–411.
- [17] L. Zhang, F.-C. Tian, X.-W. Peng, X. Yin, A rapid discreteness correction scheme for reproducibility enhancement among a batch of mos gas sensors, *Sensors and Actuators A: Physical* 205 (2014) 170–176.
- [18] P. J. Peterson, A. Aujla, K. H. Grant, A. G. Brundle, M. R. Thompson, J. Vande Hey, R. J. Leigh, Practical use of metal oxide semiconductor gas sensors for measuring nitrogen dioxide and ozone in urban environments, *Sensors* 17 (7) (2017) 1653.
- [19] L. Morawska, P. K. Thai, X. Liu, A. Asumadu-Sakyi, G. Ayoko, A. Bartonova, A. Bedini, F. Chai, B. Christensen, M. Dunbabin, et al., Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?, *Environment international* 116 (2018) 286–299.
- [20] R. Williams, V. Kilaru, E. Snyder, A. Kaufman, T. Dye, A. Rutter, A. Russell, H. Hafner, *Air sensor guidebook*, US Environmental Protection Agency (2014).
- [21] J. Saini, M. Dutta, G. Marques, A comprehensive review on indoor air quality monitoring systems for enhanced public health, *Sustainable Environment Research* 30 (1) (2020) 6.
- [22] J. Saini, M. Dutta, G. Marques, Indoor air quality monitoring systems based on internet of things: A systematic review, *International journal of environmental research and public health* 17 (14) (2020) 4942.

- [23] H. Chojer, P. Branco, F. Martins, M. Alvim-Ferraz, S. Sousa, Development of low-cost indoor air quality monitoring devices: Recent advancements, *Science of The Total Environment* 727 (2020) 138385.
- [24] J.-Y. Kim, C.-H. Chu, S.-M. Shin, Issaq: An integrated sensing systems for real-time indoor air quality monitoring, *IEEE Sensors Journal* 14 (12) (2014) 4230–4244. doi:doi.org/10.1109/JSEN.2014.2359832.
- [25] A. S. Ali, Z. Zanzinger, D. Debose, B. Stephens, Open source building science sensors (osbss): A low-cost arduino-based platform for long-term indoor environmental data collection, *Building and Environment* 100 (2016) 114–126. doi:doi.org/10.1016/j.buildenv.2016.02.010.
- [26] Y. Wang, M. Boulic, R. Phipps, C. Chitty, A. Moses, R. Weyers, J. Jang-Jaccard, G. Olivares, A. Ponder-Sutton, C. Cunningham, Integrating open-source technologies to build a school indoor air quality monitoring box (skomobo), in: 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), IEEE, 2017, pp. 216–223. doi:doi.org/10.1109/APWCConCSE.2017.00046.
- [27] Z. Idrees, Z. Zou, L. Zheng, Edge computing based iot architecture for low cost air pollution monitoring systems: a comprehensive system analysis, design considerations & development, *Sensors* 18 (9) (2018) 3021. doi:doi.org/10.3390/s18093021.
- [28] A. Carre, T. Williamson, Design and validation of a low cost indoor environment quality data logger, *Energy and Buildings* 158 (2018) 1751–1761. doi:doi.org/10.1016/j.enbuild.2017.11.051.
- [29] A. Tiele, S. Esfahani, J. Covington, Design and development of a low-cost, portable monitoring device for indoor environment quality, *Journal of Sensors* 2018 (2018). doi:doi.org/10.1155/2018/5353816.
- [30] S. E. Gillooly, Y. Zhou, J. Vallarino, M. T. Chu, D. R. Michanowicz, J. I. Levy, G. Adamkiewicz, Development of an in-home, real-time air pollutant sensor platform and implications for community use, *Environmental Pollution* 244 (2019) 440–450. doi:doi.org/10.1016/j.envpol.2018.10.064.
- [31] T. Parkinson, A. Parkinson, R. de Dear, Continuous ieq monitoring system: Context and development, *Building and Environment* 149 (2019) 15–25. doi:doi.org/10.1016/j.buildenv.2018.12.010.
- [32] T. Parkinson, A. Parkinson, R. de Dear, Continuous ieq monitoring system: Performance specifications and thermal comfort classification, *Building and Environment* 149 (2019) 241–252. doi:doi.org/10.1016/j.buildenv.2018.12.016.
- [33] G. Marques, R. Pitarma, A cost-effective air quality supervision solution for enhanced living environments through the internet of things, *Electronics* 8 (2) (2019) 170. doi:doi.org/10.3390/electronics8020170.
- [34] C. Buehler, F. Xiong, M. L. Zamora, K. M. Skog, J. Kohrman-Glaser, S. Colton, M. McNamara, K. Ryan, C. Redlich, M. Bartos, et al., Stationary and portable multipollutant monitors for high-spatiotemporal-resolution air quality studies including online calibration, *Atmospheric Measurement Techniques* 14 (2) (2021) 995–1013. doi:doi.org/10.5194/amt-14-995-2021.
- [35] T. H. Mboa Nkoudou, D. Bild, F. Grey, H. Toivonen, J. Serrano, J. P. Maestre, B. Paz, Global open science hardware roadmap: making open science hardware ubiquitous by 2025 (2018).
- [36] C. Kornartit, R. Sokhi, M. Burton, K. Ravindra, Activity pattern and personal exposure to nitrogen dioxide in indoor and outdoor microenvironments, *Environment international* 36 (1) (2010) 36–45. doi:doi.org/10.1016/j.envint.2009.09.004.
- [37] K. Belanger, T. R. Holford, J. F. Gent, M. E. Hill, J. M. Kezik, B. P. Leaderer, Household levels of nitrogen dioxide and pediatric asthma severity, *Epidemiology (Cambridge, Mass.)* 24 (2) (2013) 320. doi:doi.org/10.1097/EDE.0b013e318280e2ac.
- [38] L. Morawska, C. He, J. Hitchins, K. Mengersen, D. Gilbert, Characteristics of particle number and mass concentrations in residential houses in brisbane, australia, *Atmospheric Environment* 37 (30) (2003) 4195–4203. doi:doi.org/10.1016/S1352-2310(03)00566-1.
- [39] C. Liao, M. Delghust, J. Laverge, Association between indoor air quality and sleep quality, in: 33rd Annual Meeting of the Associated-Professional-Sleep-Societies (SLEEP), Vol. 42, Oxford Univ Press Inc, 2019.
- [40] J. Y. Park, T. Dougherty, H. Fritz, Z. Nagy, Lightlearn: An adaptive and occupant centered controller for lighting based on reinforcement learning, *Building and Environment* 147 (2019) 397–414.
- [41] C. Wu, H. Fritz, S. Bastami, J. P. Maestre, E. Thomaz, C. Julien, D. M. Castelli, K. de Barbaro, S. K. Bearman, G. M. Harari, et al., Multi-modal data collection for measuring health, behavior, and living environment of large-scale participant cohorts, *GigaScience* 10 (6) (2021) giab044.
- [42] H. Fritz, K. Kinney, C. Wu, D. M. Schnyer, Z. Nagy, Data fusion of mobile and environmental sensing devices to understand the effect of the indoor environment on measured and self-reported sleep quality, *Building and Environment* (2022) 108835. doi:doi.org/10.1016/j.buildenv.2022.108835.
- [43] H. Fritz, K. Kinney, D. Schnyer, Z. Nagy, Ventilation and indoor air quality in residential bedrooms., *Whole Communities-Whole Health-Published Research* (2021).
- [44] B. Maag, Z. Zhou, L. Thiele, A survey on sensor calibration in air pollution monitoring deployments, *IEEE Internet of Things Journal* 5 (6) (2018) 4857–4870. doi:doi.org/10.1109/JIOT.2018.2853660.
- [45] F. Karagulian, M. Barbieri, A. Kotsev, L. Spinelle, M. Gerboles, F. Lagler, N. Redon, S. Crunaire, A. Borowiak, Review of the performance of low-cost sensors for air quality monitoring, *Atmosphere* 10 (9) (2019) 506. doi:doi.org/10.3390/atmos10090506.
- [46] N. Castell, F. R. Dauge, P. Schneider, M. Vogt, U. Lerner, B. Fishbain, D. Broday, A. Bartonova, Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?, *Environment international* 99 (2017) 293–302. doi:doi.org/10.1016/j.envint.2016.12.007.
- [47] Y. Wang, J. Li, H. Jing, Q. Zhang, J. Jiang, P. Biswas, Laboratory evaluation and calibration of three low-cost particle sensors for particulate matter measurement, *Aerosol Science and Technology* 49 (11) (2015) 1063–1077. doi:doi.org/10.1080/02786826.2015.1100710.
- [48] P. J. Dacunto, N. E. Klepeis, K.-C. Cheng, V. Acevedo-Bolton, R.-T. Jiang, J. L. Repace, W. R. Ott, L. M. Hildemann, Determining pm 2.5 calibration curves for a low-cost particle monitor: common indoor residential aerosols, *Environmental Science: Processes & Impacts* 17 (11) (2015) 1959–1966.

- doi:doi.org/10.1039/C5EM00365B.
- [49] S. Sousan, K. Koehler, L. Hallett, T. M. Peters, Evaluation of consumer monitors to measure particulate matter, *Journal of aerosol science* 107 (2017) 123–133. doi:doi.org/10.1016/j.jaerosci.2017.02.013.
- [50] A. M. Collier-Oxandale, J. Thorson, H. Halliday, J. Milford, M. Hannigan, Understanding the ability of low-cost mox sensors to quantify ambient vocs, *Atmospheric Measurement Techniques* 12 (3) (2019) 1441–1460. doi:doi.org/10.5194/amt-12-1441-2019.
- [51] Y. Cheng, X. Li, Z. Li, S. Jiang, Y. Li, J. Jia, X. Jiang, Aircloud: a cloud-based air-quality monitoring system for everyone, in: *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, 2014, pp. 251–265. doi:doi.org/10.1145/2668332.2668346.
- [52] T. Zheng, M. H. Bergin, K. K. Johnson, S. N. Tripathi, S. Shirodkar, M. S. Landis, R. Sutaria, D. E. Carlson, Field evaluation of low-cost particulate matter sensors in high-and low-concentration environments, *Atmospheric Measurement Techniques* 11 (8) (2018) 4823–4846.
- [53] R. Piedrahita, Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, R. P. Dick, Q. Lv, M. Hannigan, et al., The next generation of low-cost personal air quality sensors for quantitative exposure monitoring, *Atmospheric Measurement Techniques* 7 (10) (2014) 3325–3336. doi:doi.org/10.5194/amt-7-3325-2014.
- [54] A. Di Antonio, O. A. Popoola, B. Ouyang, J. Saffell, R. L. Jones, Developing a relative humidity correction for low-cost sensors measuring ambient particulate matter, *Sensors* 18 (9) (2018) 2790.
- [55] C. Wang, L. Yin, L. Zhang, D. Xiang, R. Gao, Metal oxide gas sensors: sensitivity and influencing factors, *Sensors* 10 (3) (2010) 2088–2106.
- [56] E. S. Cross, L. R. Williams, D. K. Lewis, G. R. Magoon, T. B. Onasch, M. L. Kaminsky, D. R. Worsnop, J. T. Jayne, Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements, *Atmospheric Measurement Techniques* 10 (9) (2017) 3575–3588.
- [57] L. Spinelle, M. Gerboles, M. G. Villani, M. Alexandre, F. Bonavitaola, Field calibration of a cluster of low-cost available sensors for air quality monitoring. part a: Ozone and nitrogen dioxide, *Sensors and Actuators B: Chemical* 215 (2015) 249–257. doi:doi.org/10.1016/j.snb.2015.03.031.
- [58] L. Spinelle, M. Gerboles, M. G. Villani, M. Alexandre, F. Bonavitaola, Field calibration of a cluster of low-cost available sensors for air quality monitoring. part a: Ozone and nitrogen dioxide, *Sensors and Actuators B: Chemical* 215 (2015) 249–257. doi:doi.org/10.1016/j.snb.2016.07.036.
- [59] E. Esposito, S. De Vito, M. Salvato, V. Bright, R. L. Jones, O. Popoola, Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems, *Sensors and Actuators B: Chemical* 231 (2016) 701–713. doi:doi.org/10.1016/j.snb.2016.03.038.
- [60] B. Mijling, Q. Jiang, D. de Jonge, S. Bocconi, Practical field calibration of electrochemical no2 sensors for urban air quality applications (2017).
- [61] Q. Huang, K. Kieffer, An intelligent internet of things (iot) sensor system for building environmental monitoring, *Journal of mobile multimedia* (2019) 29–50.
- [62] J. Y. Park, M. M. Ouf, B. Gunay, Y. Peng, W. O’Brien,

- M. B. Kjærsgaard, Z. Nagy, A critical review of field implementations of occupant-centric building controls, *Building and Environment* 165 (2019) 106351.
- [63] P. Li, T. M. Froese, G. Brager, Post-occupancy evaluation: State-of-the-art analysis and state-of-the-practice review, *Building and Environment* 133 (2018) 187–202.

## Appendix A. Calibration Results

The following figures highlight the performance of all BEVO Beacons for each of the IAQ sensors from the calibration experiments.

### Appendix A.1. Carbon Dioxide

Figure A.1 shows the performance of CO<sub>2</sub> sensors on the BEVO Beacons from the calibration experiments conducted in the mock home environment.

Table A.2 illustrates consistency amongst the linear regression parameters across all three experiments for the majority of devices. Only Devices 5 and 24 show significantly different results, both from Experiment 2, which explains why these two devices are the only ones with averaged  $y > 1$ . Excluding this experiment would result in  $m$  values more consistent with the other devices. Excluding Experiment 2 would decrease performance of Device 5, but increase the performance for Device 24.

Figure A.2 shows the performance of all CO<sub>2</sub> sensors on the BEVO Beacons from the calibration experiments conducted in the 27 m<sup>3</sup> laboratory chamber.

Table A.2 indicates  $b$  values from the first two experiments are similarly low, but increase by a factor of 6 in the final experiment. Device 10 was the only device with a negative, averaged  $b$  coefficient while the remaining values ranged from 15.76 to 448.01. These results again indicate that the CO<sub>2</sub> sensors we use tend to underpredict baseline concentrations.

### Appendix A.2. Particulate Matter

Figure A.3 shows the performance of the corrected PM<sub>2.5</sub> sensors on all BEVO Beacons from calibration experiments conducted in the mock home environment.

Table A.3 shows the model parameters for each of the three experiments. In some cases, sensors were not responsive for a majority of the experiment and therefore, we do not report parameters for these devices. Generally, devices with more negative  $b$  values, have larger  $m$ . This behavior is even evident by examining single experiments per device.

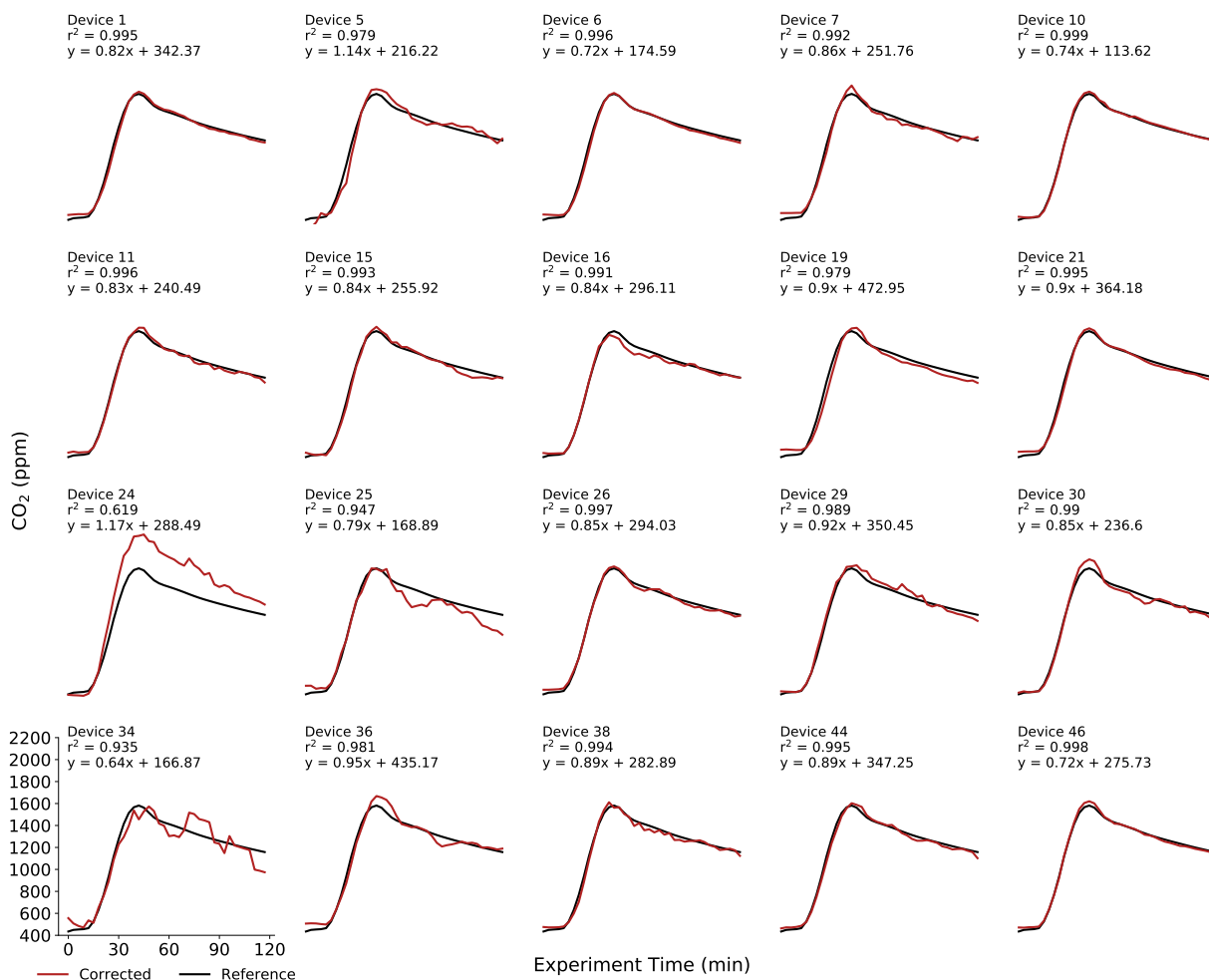


Figure A.1: Results of the CO<sub>2</sub> linear regression models averaged from the three experiments conducted in the experimental home environment. Data shown are from a fourth experiment where the models are applied to the devices' measurements and compared to the reference monitor. The  $r^2$  value corresponds to the goodness of fit from this fourth experiment.

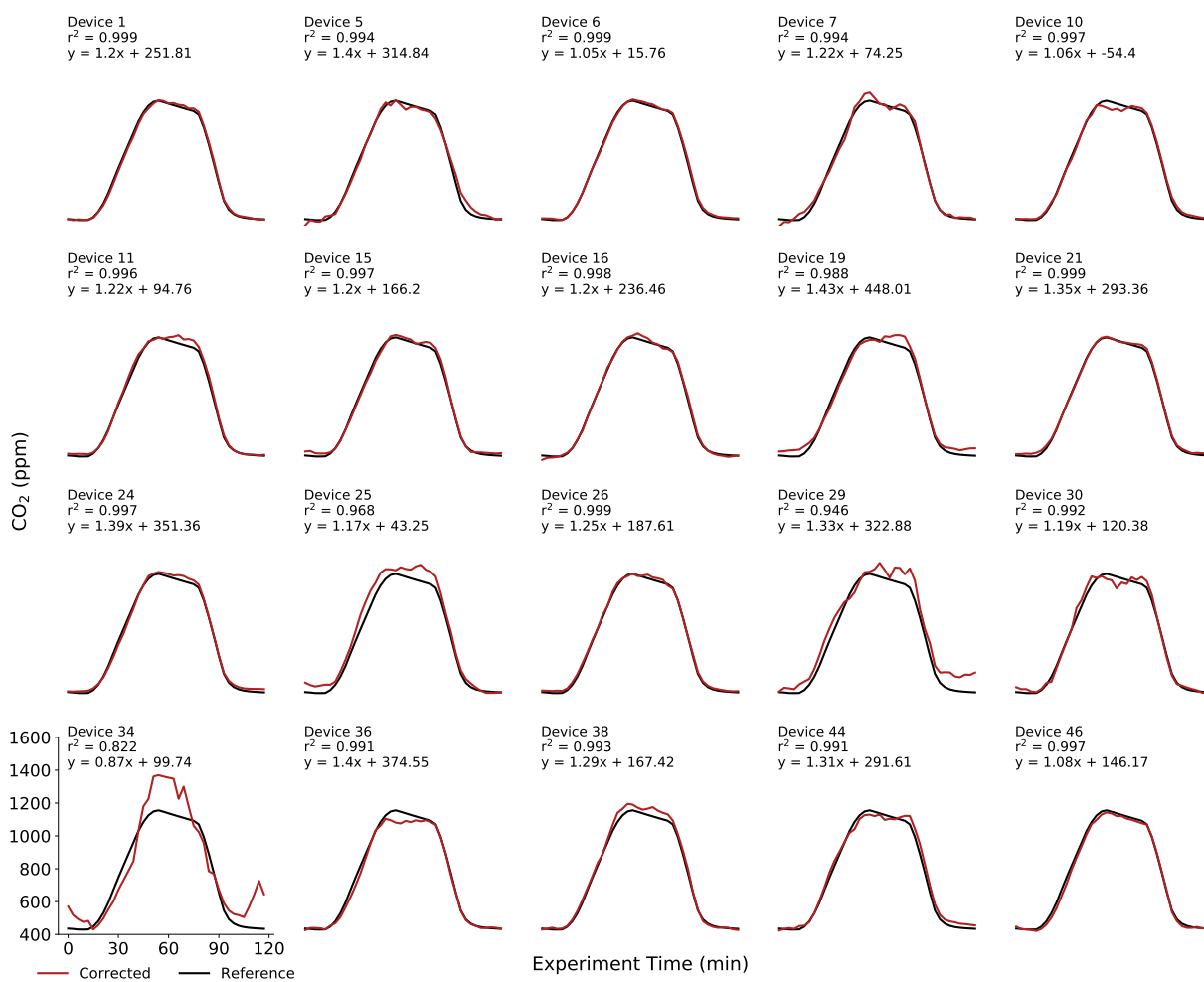


Figure A.2: Results of the CO<sub>2</sub> linear regression models averaged from the three experiments conducted in the laboratory chamber. Data shown are from a fourth experiment where the models are applied to the devices' measurements and compared to the reference monitor. The  $r^2$  value corresponds to the goodness of fit from this fourth experiment.

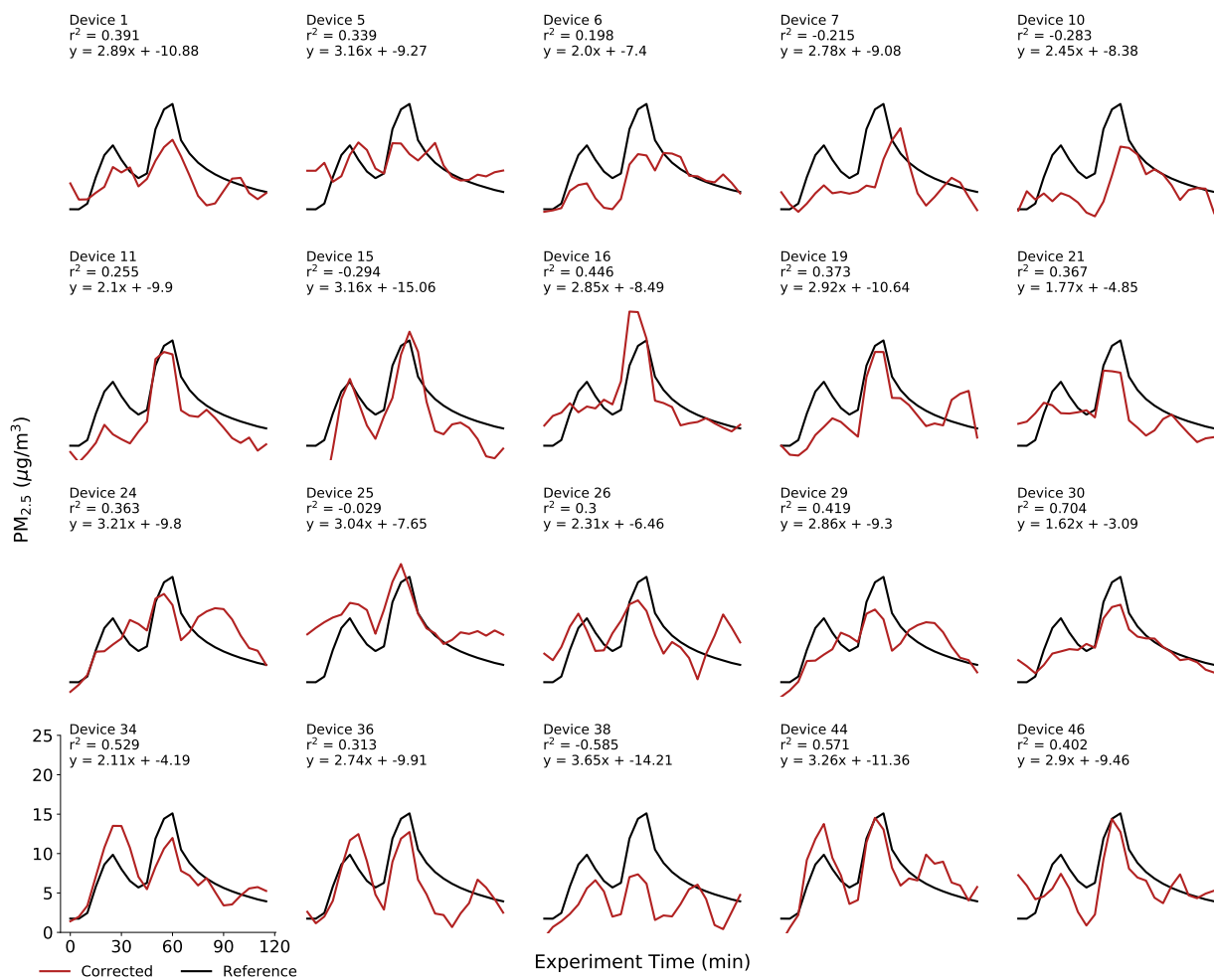


Figure A.3: Results of the PM<sub>2.5</sub> linear regression models averaged from the three experiments conducted in the experimental home environment. Data shown are from a fourth experiment where the models are applied to the devices' measurements and compared to the reference monitor.



Table A.1: Linear model parameters for each CO<sub>2</sub> sensor from the three calibration experiments conducted in the home environment.

Device	<i>b</i>			<i>m</i>		
	Experiment			Experiment		
	1	2	3	1	2	3
1	338.2	365.4	323.4	0.8	0.8	0.8
5	291.5	84.4	272.8	1.1	1.5	0.8
6	169.3	198.3	156.1	0.7	0.7	0.7
7	234.0	291.8	229.6	0.9	0.8	0.9
10	119.9	98.7	122.2	0.7	0.7	0.7
11	237.5	278.9	205.1	0.8	0.8	0.9
15	278.5	251.8	237.5	0.8	0.8	0.9
16	268.3	332.9	287.1	0.9	0.8	0.8
19	485.4	458.5	474.9	0.9	0.9	0.9
21	359.5	380.9	352.2	0.9	0.9	0.9
24	453.0	120.2	292.3	0.9	1.7	0.9
25	157.7	212.7	136.3	0.8	0.7	0.8
26	274.1	334.5	273.5	0.9	0.8	0.9
29	359.5	328.2	363.7	0.9	0.9	0.9
30	265.7	227.3	216.8	0.8	0.8	0.9
34	187.3	110.6	202.7	0.6	0.7	0.6
36	432.5	467.0	406.0	0.9	0.9	1.0
38	285.7	294.7	268.3	0.9	0.9	0.9
44	350.4	333.5	358.0	0.9	0.9	0.9
46	270.7	314.8	241.8	0.7	0.7	0.8

Table A.2: Linear model parameters for each of CO<sub>2</sub> sensors from the three experiments conducted in the laboratory chamber.

Device	<i>b</i>			<i>m</i>		
	Experiment			Experiment		
	1	2	3	1	2	3
1	256.8	234.9	263.7	1.2	1.2	1.2
5	309.6	322.9	312.0	1.5	1.3	1.4
6	24.2	5.7	17.4	1.0	1.1	1.0
7	85.0	79.4	58.4	1.2	1.2	1.2
10	-30.4	-63.8	-69.0	1.0	1.1	1.1
11	100.3	98.3	85.8	1.2	1.2	1.2
15	183.8	159.5	155.3	1.2	1.2	1.2
16	238.3	240.1	231.0	1.2	1.2	1.2
19	464.1	443.2	436.7	1.4	1.4	1.4
21	297.5	291.8	290.8	1.4	1.4	1.3
24	357.2	354.0	342.9	1.4	1.4	1.4
25	57.4	22.1	50.2	1.2	1.1	1.2
26	205.5	180.7	176.7	1.2	1.3	1.2
29	343.8	317.1	307.7	1.3	1.3	1.4
30	99.4	129.3	132.4	1.2	1.2	1.2
34	31.0	40.7	227.5	0.9	1.0	0.7
36	382.6	371.9	369.2	1.4	1.4	1.4
38	174.3	166.2	161.7	1.3	1.3	1.3
44	299.4	290.8	284.6	1.3	1.3	1.3
46	150.9	144.0	143.6	1.1	1.1	1.1

Table A.3: Linear model parameters for each of PM<sub>2.5</sub> sensors from the three experiments conducted in the home environment.

Device	<i>b</i>			<i>m</i>		
	Experiment			Experiment		
	1	2	3	1	2	3
1	-11.7	-	-10.1	3.3	-	2.4
5	-10.2	-11.1	-6.5	3.3	4.2	2.0
6	-4.9	-8.3	-8.9	1.9	2.5	1.6
7	-6.6	-8.0	-12.6	2.8	3.0	2.6
10	-5.3	-	-11.5	2.4	-	2.5
11	-4.3	-15.1	-10.3	1.6	2.9	1.7
15	-3.5	-21.6	-20.0	1.7	4.5	3.2
16	-3.3	-7.0	-15.2	1.9	3.6	3.1
19	-4.7	-	-16.6	2.1	-	3.8
21	-3.1	-	-6.6	1.7	-	1.8
24	-9.8	-	-	3.2	-	-
25	-10.9	-4.4	-	3.1	2.9	-
26	-4.6	-9.1	-5.7	2.0	3.5	1.5
29	-5.2	-9.4	-13.2	2.2	3.7	2.7
30	-5.1	-	-1.1	2.1	-	1.1
34	-1.5	-4.4	-6.6	1.6	3.1	1.6
36	-3.3	-15.5	-10.9	1.8	4.1	2.4
38	-8.6	-19.5	-14.5	3.2	4.7	3.1
44	-6.0	-18.7	-9.4	2.2	5.3	2.3
46	-5.1	-10.7	-12.5	2.3	3.9	2.6

For example, Device 44 has an *b* value of -6.0 in the first experiment with a corresponding *m* of 2.2. In the second experiment, the *b* drops to -18.7 and the *m* jumps to 5.3. This set of experiments along with others in Table A.3 highlight that parameters varied widely from experiment to experiment despite ensuring experiments were conducted in a similar fashion. The values for *b* range from -21.6 (Device 15, Experiment 2) to -1.1 (Device 30, Experiment 3) while values for *m* range from 1.1 (Device 30, Experiment 3) to 5.3 (Device 44, Experiment 2). The variability in the parameters would explain the models poor ability to correct raw CGS measurements to the reference.

Figure A.4 shows the performance of PM<sub>2.5</sub> sensors on the BEVO Beacons calibrated in the 27 m<sup>3</sup> chamber environment.

Table A.4 displays all parameter values for each of the PM<sub>2.5</sub> sensors from each calibration experiment conducted in the laboratory chamber. The *b* and *m* parameters from Experiment 1 are generally lower than the subsequent experiments, but overall there are no clear patterns in parameter values over the set of experiments.

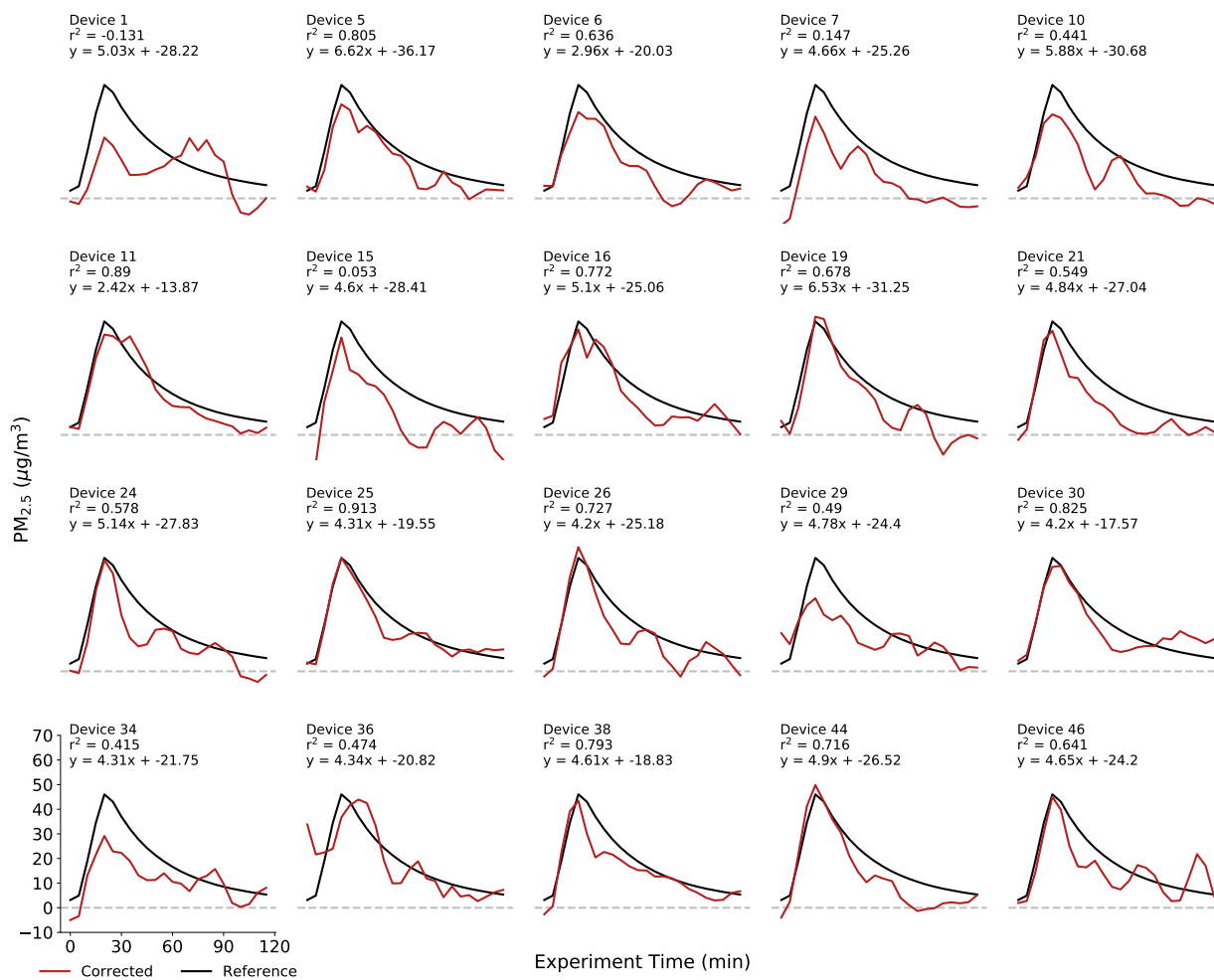


Figure A.4: Results of the PM<sub>2.5</sub> linear regression models averaged from the three experiments conducted in the laboratory chamber. Data shown are from a fourth experiment where the models are applied to the devices' measurements and compared to the reference monitor.

Table A.4: Linear model parameters for each of PM<sub>2.5</sub> sensors from the three experiments conducted in the environmental chamber.

Device	<i>b</i>			<i>m</i>		
	Experiment			Experiment		
	1	2	3	1	2	3
1	-26.6	-34.5	-23.6	5.4	5.8	3.9
5	-18.7	-42.2	-47.6	5.1	8.1	6.7
6	-7.0	-21.9	-31.2	2.6	3.1	3.1
7	-16.0	-28.6	-31.1	4.7	4.7	4.6
10	-20.7	-29.6	-41.7	5.6	6.3	5.7
11	-4.1	-18.9	-18.6	2.0	2.8	2.4
15	-13.4	-36.0	-35.8	3.8	5.4	4.6
16	-20.3	-20.6	-34.2	5.5	4.8	5.0
19	-23.1	-29.1	-41.5	6.4	6.4	6.9
21	-20.4	-23.5	-37.2	4.8	4.5	5.2
24	-20.1	-29.1	-34.3	4.8	5.7	4.9
25	-15.6	-17.6	-25.5	4.5	4.4	4.1
26	-18.4	-26.4	-30.8	4.0	4.6	4.0
29	-10.7	-25.6	-36.9	4.2	5.5	4.7
30	-8.7	-24.5	-19.5	4.1	4.8	3.7
34	-10.3	-20.0	-35.0	4.0	4.3	4.5
36	-16.7	-22.8	-23.0	4.3	4.8	3.9
38	-10.1	-19.5	-26.9	4.6	4.7	4.5
44	-16.6	-31.7	-31.3	4.8	5.7	4.3
46	-26.2	-16.4	-30.0	4.9	4.4	4.7

Table A.5: Linear model parameters for each of TVOC sensors from the three experiments conducted in the environmental chamber.

Device	<i>b</i>			<i>m</i>		
	Experiment			Experiment		
	1	2	3	1	2	3
1	502.3	120.0	-2.5	0.90	1.40	1.07
5	-3.9	56.4	-1.5	1.11	1.09	1.04
6	12.8	-51.9	0.2	0.84	0.98	1.00
7	56.2	-12.6	4.1	0.40	0.43	0.80
10	-74.2	-1.7	0.2	1.37	1.32	1.12
11	-24.5	-9.4	-1.8	1.05	1.11	1.01
15	1.3	-116.0	-0.2	0.85	1.01	1.01
16	24.5	-96.4	2.6	0.70	0.90	0.97
19	-13.3	49.7	-2.0	1.22	1.21	1.09
21	-23.1	-22.5	1.5	1.38	1.46	1.17
24	65.0	225.3	2.3	0.80	0.70	0.88
25	4.0	-80.5	-0.5	1.18	1.22	1.05
26	-16.1	28.2	-0.1	1.37	1.26	1.09
29	-38.9	-73.2	-1.0	1.48	1.46	1.10
30	-62.4	215.5	-3.7	1.51	1.15	0.99
34	-64.9	192.9	0.2	1.92	1.50	1.03
36	-3.6	-35.4	0.9	0.92	1.03	0.99
38	-57.7	-20.6	-1.0	1.43	1.44	1.11
44	73.5	176.9	3.6	0.39	0.34	0.67
46	-117.8	45.3	-2.0	1.99	1.80	1.10

### Appendix A.3. Total Volatile Organic Compounds

Figure A.5 shows the results of the calibrated TVOC sensors from experiments conducted in the laboratory chamber.

Table A.5 highlights the wide range of *m* and *b* values for the TVOC sensors from the three calibration experiments, both for a given device and across devices. While the *m* values tended to be similar across experiments with the exception of a few devices, the *b* coefficient varied considerably. Ten devices had *b* coefficients with oppositely signed values in at least one experiment. Of note are Devices 1, 30, and 34 which have differences in *b* coefficients of more than 200 ppb between Experiments 1 and 2. In Experiment 3, *b* coefficients were considerably lower and exhibited much less variability, ranging from -3.69 to 4.06.

### Appendix A.4. Carbon Monoxide

The model parameters were derived from a single experiment with the gas standard (Table A.6). We group devices that were calibrated in the same experiment to understand if the model parameters we derived were dependent on the experiment. Based

on the variety of model parameters from each batch, there does not appear to be any influence of the experiment group on model parameters. The *b* values range from -2.07 to 11.39. The *m* values are all greater than 1 and exhibit less variation, ranging from 1.10 to 1.80. Unlike previous models for other CGS, the *m* term does not seem to compensate for low *b* values. While Device 16 had the lowest *b* and the largest *m*, this trend does not extend to Devices 7, 26, and 29 which also had negative *b* parameters.

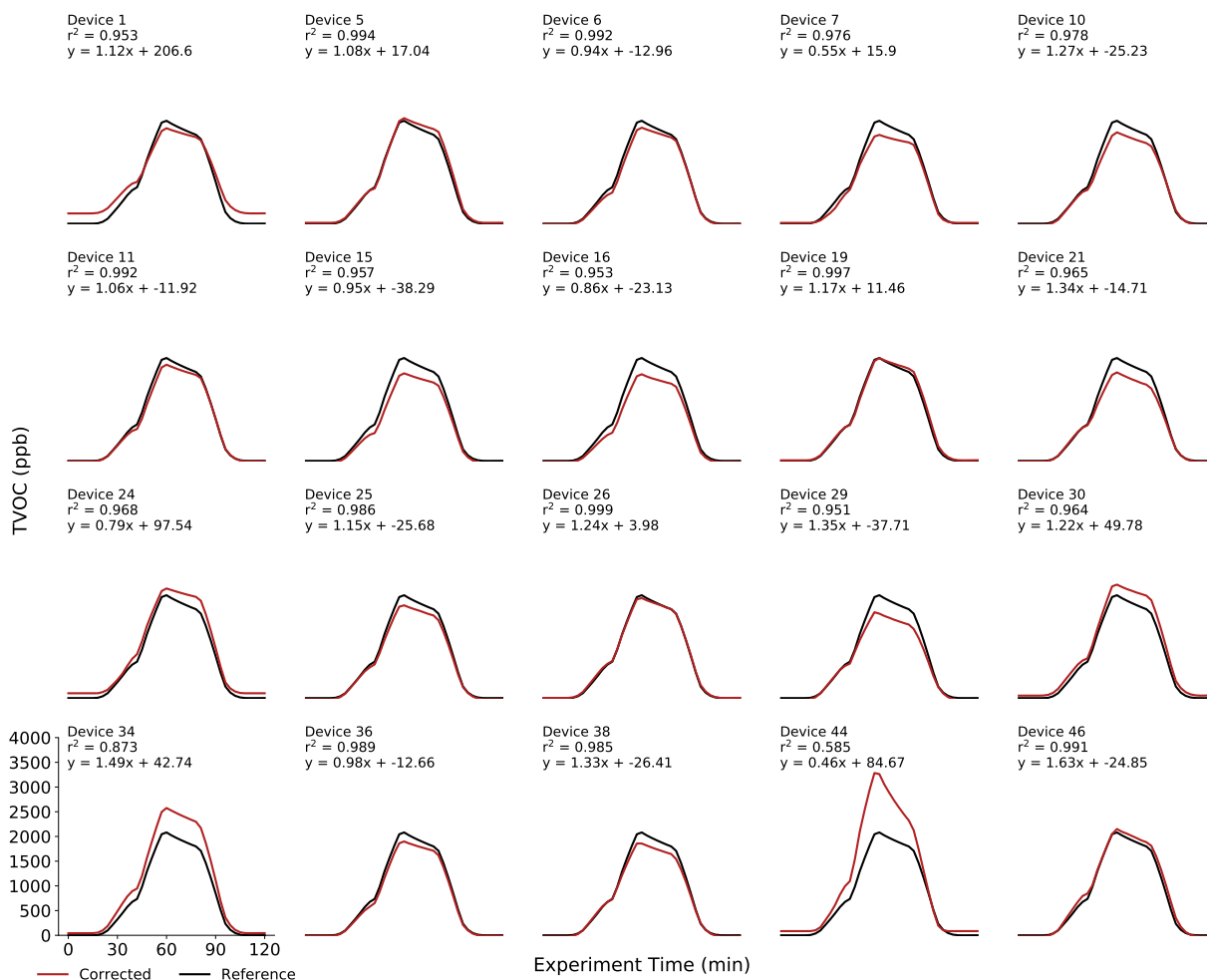


Figure A.5: Results of the TVOC linear regression models averaged from three experiments conducted in the laboratory chamber. Data shown are from a fourth experiment where the models are applied to the devices' measurements and compared to the reference monitor. The  $r^2$  value corresponds to the goodness of fit from this fourth experiment.

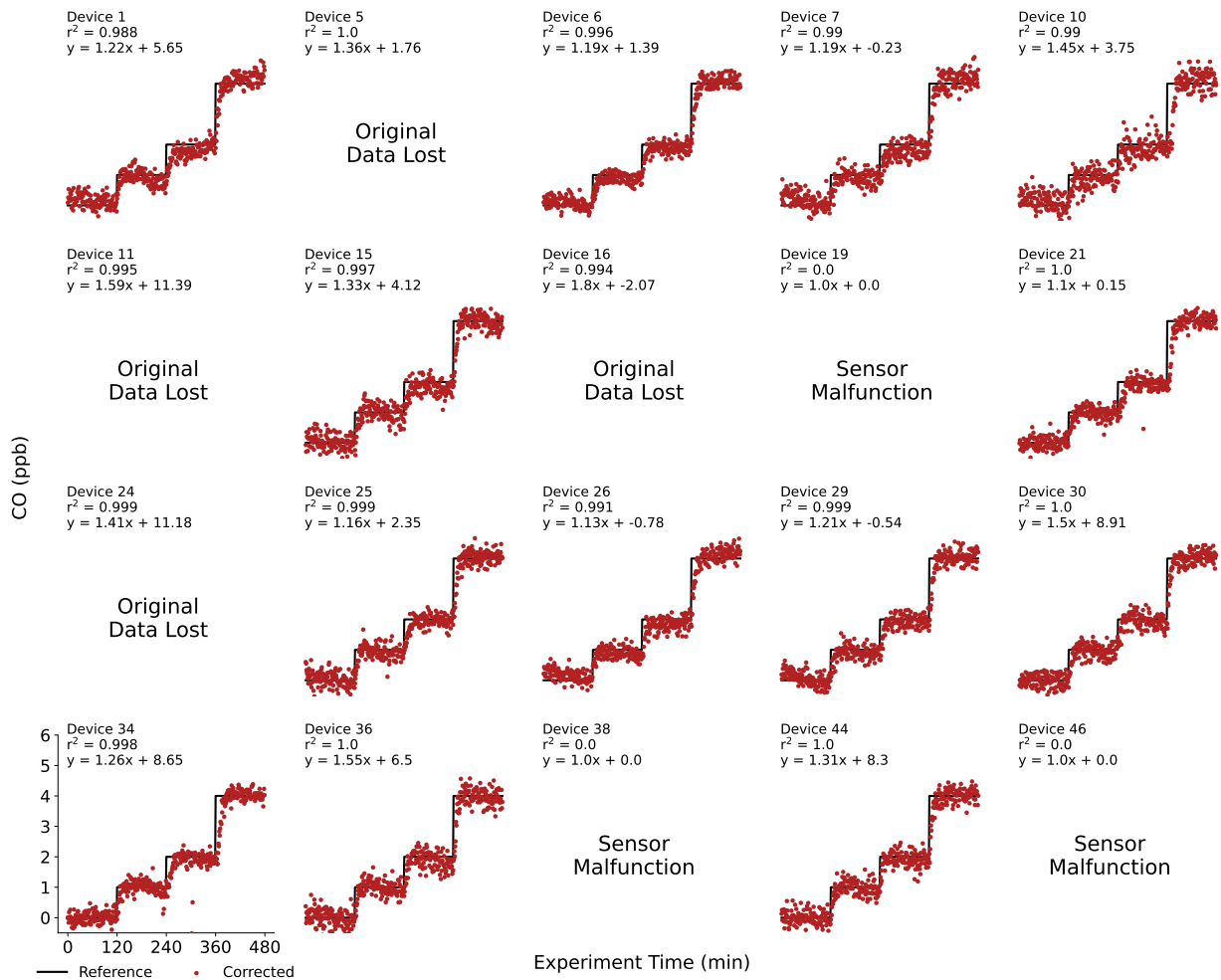


Figure A.6: Results of the CO linear models from the gas standard calibration.

Table A.6: Linear model parameters for each of CO sensors from the three experiments conducted against the gas standard.

Device	$b$	$m$	$r^2$
1	5.65	1.22	0.99
7	-0.23	1.19	0.99
26	-0.78	1.13	0.99
15	4.12	1.33	1.00
24	11.18	1.41	1.00
36	6.50	1.55	1.00
6	1.39	1.19	1.00
25	2.35	1.16	1.00
29	-0.54	1.21	1.00
21	0.15	1.10	1.00
34	8.65	1.26	1.00
38	-	-	-
5	1.76	1.36	1.00
30	8.91	1.50	1.00
44	8.30	1.31	1.00
10	3.75	1.45	0.99
46	-	-	-
11	11.39	1.59	1.00
16	-2.07	1.80	0.99
19	-	-	-