

Performance and Environmental Correction of a Low-Cost  
NDIR CO<sub>2</sub> Sensor

Cory R. Martin

A scholarly paper in partial fulfillment of the requirements for the  
degree of Master of Science

April 2016

Department of Atmospheric and Oceanic Science

University of Maryland

College Park, Maryland

Advisor: Dr. Ning Zeng

# Table of Contents

Abstract .....	ii
Acknowledgements .....	iii
List of Figures .....	iv
List of Symbols .....	v
1. Introduction .....	1
2. Instruments and methods .....	3
2.1 Noise reduction with temporal averaging .....	6
3. Controlled lab experiment .....	8
4. Ambient evaluation chamber .....	11
5. Calibration and environmental correction .....	14
5.1 Successive regression method .....	14
5.2 Multivariate linear regression method .....	17
5.3 Discussion .....	19
6. Conclusions .....	20
References .....	21

## **Abstract**

Non-dispersive infrared (NDIR) sensors are a low-cost way to observe carbon dioxide concentrations in ambient air, but their specified accuracy and precision are often not sufficient for scientific applications. We evaluated several such low-cost CO<sub>2</sub> sensors and selected one, the SenseAir K30, which has a manufacturer specified accuracy of 30 ppm. We tested several K30 sensors in lab and ambient air conditions. After noise reduction, calibration and environmental correction for atmospheric pressure, temperature and humidity, the root mean square (RMS) difference between K30 and a laser cavity-ringdown spectroscopy (CRDS) CO<sub>2</sub> analyzer Los Gatos FGGA was typically reduced to below 5 ppm. Thus, this sensor has the potential to compliment traditional higher accuracy but high cost instruments in certain CO<sub>2</sub> monitoring applications.

## **Acknowledgements**

First and foremost, I would like to thank Professor Ning Zeng for all of his guidance and support in preparation for this work as well as my academic career in general. Additionally, I would like to acknowledge Professor Russell Dickerson, Dr. Xinrong Ren, Bari Turpie, and Kristy Weber, for their contributions to the analysis described in this paper. I would also like to thank Dr. James Whetstone, and all others at NIST including Dr. Subhomoy Ghosh, Dr. Anna Karion, Dr. Israel Lopez-Coto, Dr. Kimberly Mueller, Dr. Kuldeep Prasad, and Dr. Tamae Wong, for help as well as NIST's Greenhouse Gas and Climate Science Measurements program for partial funding of this work. Finally, I would like to thank all of my friends and family, as I would not be where I am today without their decades of encouragement and help.

## List of Figures

1. CO <sub>2</sub> calibrations of the LGR during a month-long period .....	4
2. Picture of complete sensor package .....	5
3. Allan variance analysis of K30 sensor .....	7
4. Laboratory experiment setup schematic .....	8
5. 24 hours of measurements from two K30s and LGR instrument .....	9
6. Continuous time series data from ambient evaluation chamber .....	12
7. Calibration curve of one K30 sensor .....	15
8. Successive regression time series plot .....	16
9. Successive regression difference plots .....	17
10. Multivariate regression time series and difference plots .....	18

## List of Symbols

CO<sub>2</sub> – carbon dioxide  
ppm – parts per million  
NDIR – non-dispersive infrared  
Hz – hertz  
°C – degrees Celsius  
hPa – hectopascals  
CRDS – cavity ring-down spectroscopy  
K30 – Senseair K30 sensor  
LGR – Los Gatos Research Instrument  
CH<sub>4</sub> – methane  
NIST – National Institute of Standards and Technology  
mL – milliliters  
IR – infrared  
RPi – Raspberry Pi mini computer  
cm – centimeters  
UART - Universal Asynchronous Receiver/Transmitter serial bus  
i<sup>2</sup>c - Inter-Integrated Circuit serial bus  
SD – secure digital card  
HVAC – Heating Ventilation and Air Conditioning  
UTC – Universal Time Coordinate  
AC – alternating current  
CO<sub>2</sub> dry – dry carbon dioxide  
RMSE – root mean square error  
 $y$  – dependent variable in regression  
 $a_n$  – coefficient for  $n$ th independent variable in regression  
 $x_n$  –  $n$ th independent variable in regression  
 $b$  – regression offset  
 $\varepsilon_n$  – residual of  $n$ th regression step  
R<sup>2</sup> – coefficient of determination  
P – pressure  
T – temperature  
q – water vapor

## 1 Introduction

Carbon dioxide (CO<sub>2</sub>) is a major greenhouse gas, with fundamental importance to Earth's climate. Since measurements started at the Mauna Loa Observatory in the 1950s, (Keeling et al., 2005) the global mean concentrations of CO<sub>2</sub> has steadily risen from preindustrial levels of approximately 280 parts per million (ppm), to today's level exceeding 400 ppm. These observations, both from flask samples, as well as state-of-the-art continuous measurement instruments, have a typical accuracy of ~0.1 ppm. Flasks require observers to collect the samples, and then need to be transported to a lab for analysis, which costs significant amounts of time and money. Towers do not suffer from these continuous costs, but do have some maintenance costs associated with them, plus their high initial price of installation. Because of these limitations of both funding and manpower, the amount of carbon dioxide observations are relatively sparse, especially compared to other climatological variables such as temperature and precipitation. Recent research efforts have focused more locally and on the use of networks of observing sites that use instrumented towers similar to what is used for global monitoring, but applied to the urban environment (Briber et al., 2013; Kort et al., 2013; Turnbull et al., 2014). Data from these sparse tower sites are then used to create inversions to estimate the total greenhouse gas flux from the urban area in question. However, due to the cost of these networks being comparable to ones at the global scale, the observation towers are still sited at a relatively low density. Thus, to better constrain the inversion estimates, there is a need for more spatial density in the data collected in these campaigns.

Recently, however, small, low-cost sensors, some that observe trace gases, particulate matter, as well as traditional meteorological variables, using various technologies and accuracy have

become commercially available. Evaluation and implementation of these sensors is quite promising (Eugster and Kling, 2012; Holstius et al., 2014; Piedrahita et al., 2014; Young et al., 2014; Wang et al., 2015). Many of these instruments are based on electrochemical reactions to observe the concentrations of trace gases. With the advent of widely available and low cost mid-IR light sources and detectors, a small group of non-dispersive infrared (NDIR) CO<sub>2</sub> sensors have also become commercially available. They are designed for use in a number of applications including ventilation control, agricultural and industrial applications, and including them in stand-alone commercial products. Additionally, with the high volume of possible applications, these small NDIR CO<sub>2</sub> sensors are affordably priced on the order of \$100 to \$200 per sensor. Previous works have compared some of these NDIR CO<sub>2</sub> devices and concluded that after use of some type of calibration procedure, some of these devices can provide reasonably accurate measurements of ambient CO<sub>2</sub> concentrations (Hurst et al., 2011; Yasuda et al., 2012).

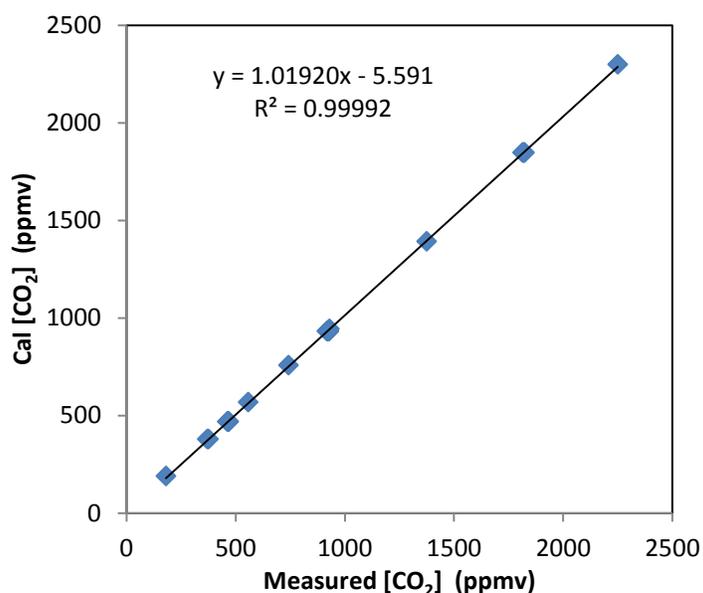
In this paper, we take one of these small NDIR devices and assess its accuracy (precision and drift) with and without environmental corrections. Section two describes the CO<sub>2</sub> sensor and its Allan variance, the other instruments included in our system, and our data collection and processing methodology. Section three shows the results of our initial laboratory experiment, which led to the initial ambient evaluation experiment described in Sect. 4. Finally, two methods are described to determine functional relationships and coefficient values to correct the observed values of the instrument for environmental variables.

## 2 Instruments and methods

To test the validity of using low-cost sensors for scientific applications, we implemented a sensor package consisting of various off the shelf components. The K30 sensor module (K30), manufactured by the Swedish company, SenseAir, is our NDIR CO<sub>2</sub> observing instrument. The K30 is a microprocessor-controlled device with on-board signal averaging, has a measurement range of 0 to 10,000 ppm, observation frequency of 0.5 Hz, and resolution of 1 ppm. The manufacturer's stated accuracy of the K30 sensor is given as  $\pm 30 \text{ ppm} \pm 3\%$  of reading (SenseAir, 2007). We initially evaluated additional NDIR sensors before selecting the K30, including the COZIR ambient sensor and GE Telaire T6615, having accuracies described as being  $\pm 50 \text{ ppm} \pm 3\%$  and  $\pm 75 \text{ ppm}$  respectively (Gas Sensing Solutions, 2014; General Electric, 2011). The K30 was chosen not only because of its highest manufacturer's specified accuracy, but also because of observed reliability and consistency with observations. In addition to CO<sub>2</sub>, temperature and pressure readings are recorded using a breakout board purchased from Adafruit. This board features a Bosch Sensortec BMP180, which according to the manufacturer's datasheet has an average absolute accuracy of  $\pm 1 \text{ }^\circ\text{C}$  and  $\pm 0.12 \text{ hPa}$ , and an output resolution of  $0.1 \text{ }^\circ\text{C}$  and  $0.01 \text{ hPa}$ , respectively (Bosch Sensortec, 2013).

To compare the performance of the K30 to better performing research instrumentation, we used a greenhouse gas analyzer based on cavity ring-down spectroscopy (CRDS) as the control. The LGR-24A-FGGA fast greenhouse gas analyzer from Los Gatos Research (LGR) provides CO<sub>2</sub>, CH<sub>4</sub>, as well as water vapor mixing ratios at a frequency of 0.5 Hz and has an un-calibrated uncertainty of less than one percent (Los Gatos Research, 2013). The LGR was connected to a tee connection, to allow either ambient air or a calibration source (during calibrations) to be

sampled continuously by the analyzer at a flow rate of 400 standard mL min<sup>-1</sup>. Calibrations for CH<sub>4</sub> and CO<sub>2</sub> were conducted using several NIST certified standard mixtures every 1-2 days for a period of one month with volume mixing ratios ranging from 1,800 to 3,900 ppb for CH<sub>4</sub> and from 360 to 2,200 ppm for CO<sub>2</sub>. See Fig. 1 for a calibration curve for the LGR calculated during this period.

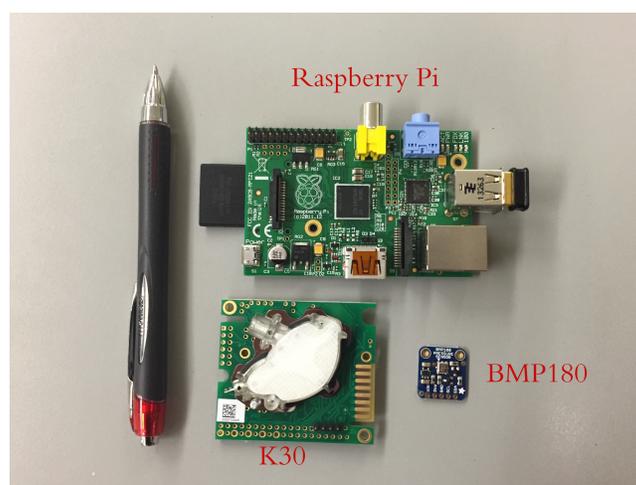


**Figure 1.** CO<sub>2</sub> calibrations of the LGR during a month-long period. Nineteen multi-point calibrations were conducted during this period with one calibration every 1 to 2 days.

It is important to note that there are differences in how CRDS works compared to non-dispersive infrared, most notably that the LGR and other instruments have a controlled cavity where pressure and temperature are kept nearly constant, removing potential environmental interference and the need for corrections, whereas the NDIR K30 works in the ambient environment without any mechanism for keeping temperature or pressure constant. To increase the effective path length, both the K30 and LGR use mirrors, but a CRDS system uses highly reflective mirrors that allow for an effective path length back and forth between the two that is many times larger

than the mirrors that reflect the IR signal inside the K30. Additionally, a CRDS instrument determines the concentration of a gas by how long it takes for the signal to degrade inside the cavity (the e-folding time), whereas a NDIR sensor merely measures the intensity of the signal received relative to the total intensity emitted.

For data collection, a Raspberry Pi (RPi) computer was used (Raspberry Pi Foundation, 2015). The RPi is a credit card sized (approximately 5 x 6 cm) computer running a full Linux distribution, allowing for easy customization and usability. The K30 was connected to the RPi over UART Serial, and the BMP180 over I<sup>2</sup>C. An image of the complete sensor package is shown in Fig. 2. Data is archived on the RPi and uploaded to a centralized data storage and processing server. The LGR collects and archives its own data, but an RPi was used here as well to collect the data from the LGR over a local area network and transfer it to the same centralized server. The added computational power of a Raspberry Pi over traditional data loggers allowed us to archive two levels of data, the raw data collected every two seconds, and one-minute averages.

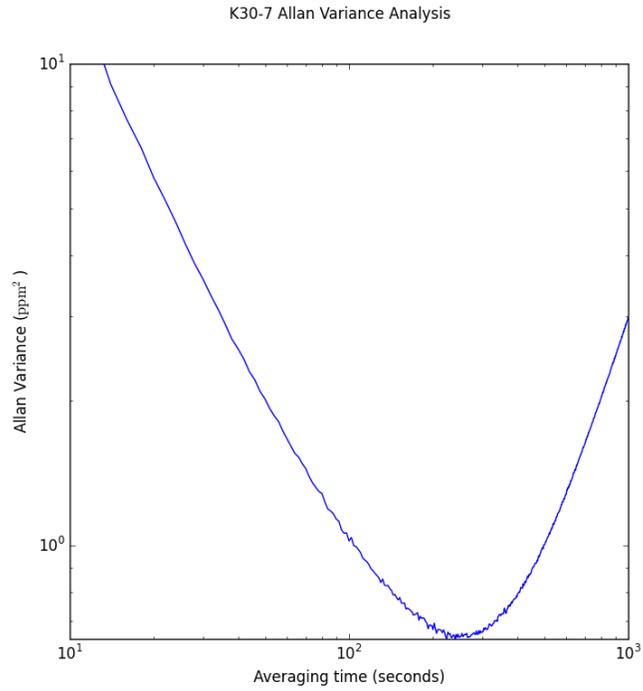


**Figure 2.** Picture of a Raspberry Pi computer (top), a SenseAir K30 CO<sub>2</sub> sensor (bottom center), a Bosch BMP180 temperature and pressure sensor (bottom right), and a ballpoint pen for size reference.

Archiving and comparing multiple datasets proved to be an interesting challenge, so steps were taken to ensure that each compared value is at the same observed time. Because of various complications, the data collection time of each K30 sensor package and the LGR are asynchronous. Additionally, we observed that power issues can corrupt parts of the plain text data files stored on the RPi's SD card with random characters. Thus, a post-processing procedure was developed that filtered extraneous characters, then each dataset is synchronized based off recorded time stamps and averaged over selected time periods. These new datasets can then be compared without worrying about missing data points or a slight time delay.

## **2.1 Noise reduction with temporal averaging**

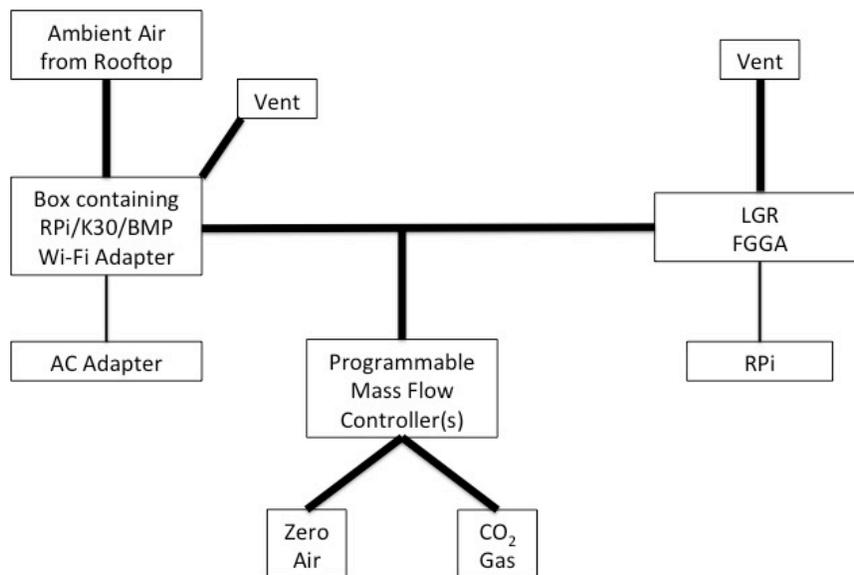
Like many similar sensors, a single measurement can be very noisy. Such can usually be reduced by temporal averaging. Allan variance (Allan, 1966) is a measure of the time-averaged stability between consecutive measurements or observations, often applied to clocks and oscillators. Here, Allan variance analysis was used to determine the optimum averaging interval for a dataset to minimize noise by sacrificing some temporal resolution. Figure 3 shows the Allan variance for one K30's raw two-second data. The original two-second data is somewhat noisy, but averaging even to ten seconds drops the variance significantly. The optimum averaging time, when the Allan variance is at a minimum (Langridge et al., 2008), is around four minutes. This means that with averaging periods below four minutes, noise may still exist, and longer than four minutes will result in loss of real signal. For the subsequent analysis, we chose an averaging time of one minute, as the Allan variance is only slightly higher than for four minutes, and it is relatively straightforward to work with data with a frequency of one minute compared to data every four minutes.



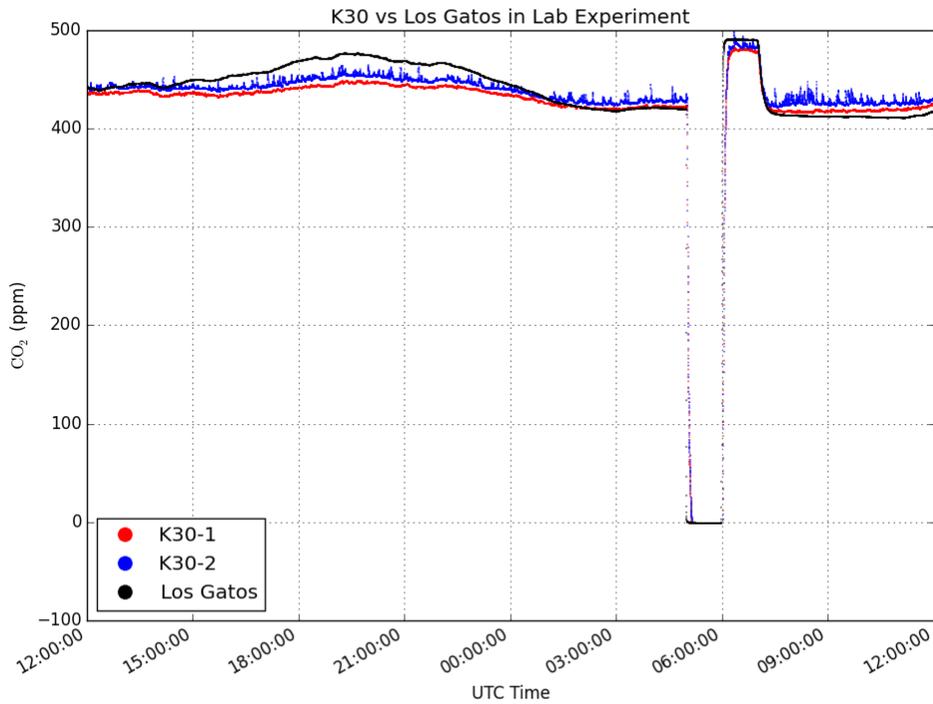
**Figure 3.** Allan variance analysis for a K30 CO<sub>2</sub> sensor without any signal averaging. Averaging times between 2 and 1,000 seconds are shown.

### 3 Controlled lab experiment

In one analysis, four K30s were installed in a semi-airtight container in a laboratory. The container was placed inline upstream of the Los Gatos Fast Greenhouse Gas Analyzer, and the LGR's included pump was used to pull ambient outdoor air from an inlet into the container holding the K30s and into the LGR. Additionally, a Dasibi (Model 5008) calibrator was attached to this system, allowing for two NIST traceable calibration gas standards (ultra pure zero air, as well as a CO<sub>2</sub>/CH<sub>4</sub> mixture near ambient levels) to be introduced to the sensors and the LGR simultaneously. The temperature was kept relatively constant for the K30s, as the laboratory, located in a classroom building, stays at room temperature through the building's HVAC system, but relative humidity may have varied as outside air was pulled into the sensor container. See Fig. 4 for a schematic of the instrument setup.



**Figure 4.** Schematic of the instrument setup used in the laboratory experiment described in Sect. 3. Ambient air is pumped through the box containing the K30 and associated equipment and into the LGR cavity. Between the LGR and the K30 box, a tee connection links a Dasibi calibrator acting as a mass flow controller into the system. The Dasibi calibrator was used to periodically introduce a known concentration of carbon dioxide (both zero air resulting in 0 ppm of CO<sub>2</sub> and a near ambient concentration of 490 ppm) for calibration and reference.



**Figure 5.** 24 hours of measurements of carbon dioxide mixing ratios from two K30 NDIR sensors and the Los Gatos Research Fast Greenhouse Gas Analyzer. During this time period, ambient air was introduced except for the two-step calibration period controlled by the mass flow controller starting at 05:00:00 UTC. Zero air was introduced into the system for an hour, and then a change to a mixture of zero air and carbon dioxide, which resulted in a carbon dioxide mixing ratio of around 490 ppm.

Figure 5 shows 24 hours of measurements of CO<sub>2</sub> from two K30s in the lab including a two-point calibration period introduced into the K30 container and to the LGR for one hour per gas sample starting at 05:00:00 UTC. During this time, the data from each K30 was collected every five seconds and then averaged over each minute, and the LGR's 0.5Hz data was also averaged over each minute. Because the container is not perfectly sealed, there is some slow diffusion from the air inside the room into the container, leading to some indoor influences in addition to the natural outdoor diurnal variation. For the calibration, first the synthetic zero air, containing only nitrogen and oxygen, was pumped into the system using the Dasibi calibrator, for an hour.

Then, the calibrator added CO<sub>2</sub> calibration gas into the zero air, and the newly mixed air with a diluted concentration of CO<sub>2</sub> was used to calibrate the LGR and K30s for an additional hour before returning to the ambient air. Both the K30s and the LGR reach equilibrium after a step change during the hour period, but this occurs much earlier for the LGR, because of the volume of the container holding the K30s being of sufficient volume that it takes longer for the calibration gas to completely fill the chamber, whereas the LGR cavity is of much smaller volume such that the step change is more quickly observed.

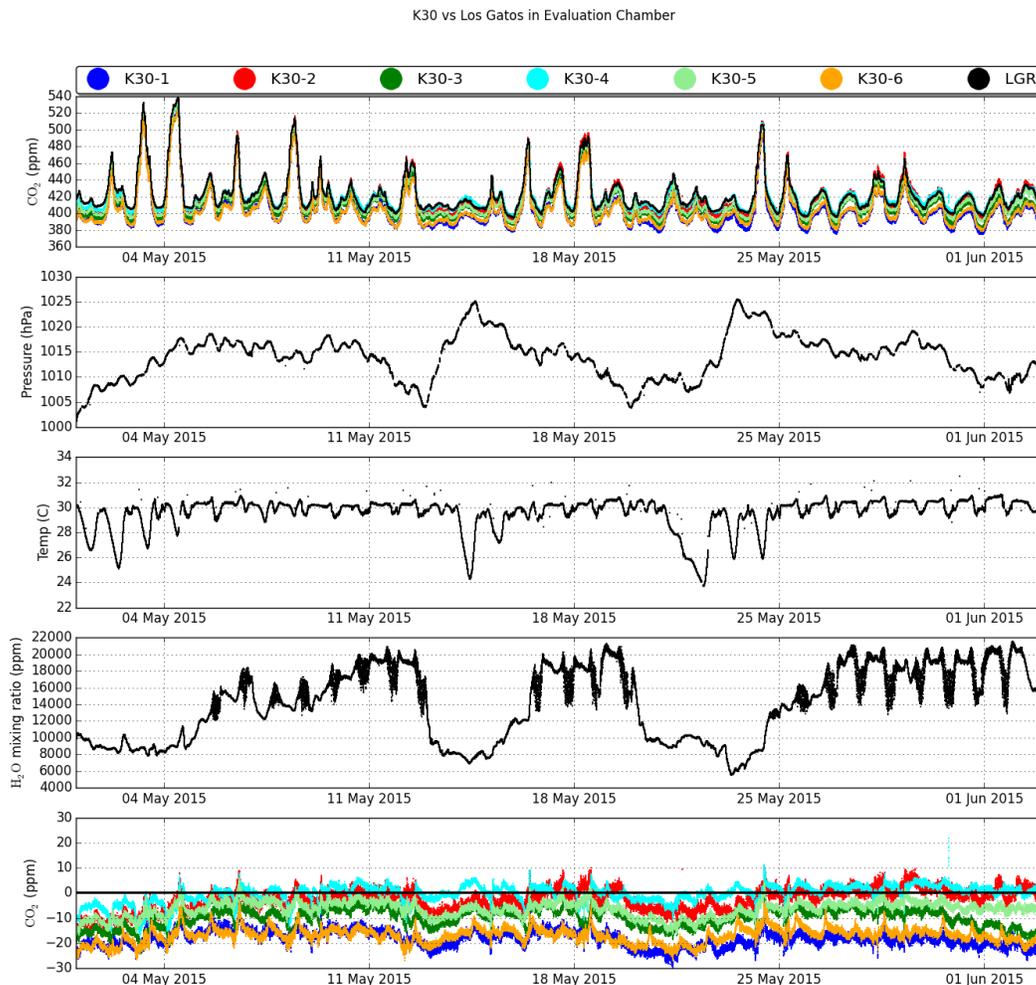
These preliminary datasets demonstrated that the K30 NDIR sensors have the potential to resolve changes in the ambient concentration of CO<sub>2</sub> at the 1 to 2 ppm level. However, some logistical issues should be addressed to more thoroughly assess their potential. First, it is difficult to seal these sensors in an airtight container due to their needs for both data collection and power. The issue that is most notable, as visible in Fig. 5, is that the blue K30 sensor plotted seems noisier than the red one. We found that some of the noise observed in certain K30s can be attributed to data transmission anomalies, e.g., spikes occurred whenever a nearby Wi-Fi adapter was active, but this was not present at all times or in each K30. It was later determined that these spikes were caused by a lack of sufficient power from the AC power supply, and by switching to a 2 amp 5 volt supply from the original 1 amp 5 volt one, these issues largely went away. Still, these problems were not noticeable in every sensor, which implies that some of these sensors perform better than others, due to minor variations in manufacturing, as well as minor damage that can arise during shipping. Therefore, there is a need for a process to quickly evaluate large sets of these K30 sensors, to determine which, if any, are unsuitable for our intended applications.

#### **4 Ambient evaluation chamber**

The need to quickly and effectively evaluate a relatively large number of sensors free from human influence led to the use of a rooftop observation room located just above the aforementioned laboratory. Because this rooftop room had limited access, and it was not part of the building's HVAC system, it served as an ambient evaluation chamber with minimal influence from human respiration. The room features a small, independent heating and cooling unit, but it was not operational during our evaluation period, rather the room was only slightly ventilated to allow ambient outdoor air to slowly diffuse into the room, with a small household box fan also in the room to ensure that the air is well mixed. Additionally, the room was not temperature controlled, because the HVAC unit was not in use. With the exception of daytime heating warming the room more than the outside air, the interior temperature took on a diurnal cycle similar to the outdoor temperature. This ventilation strategy was intentional so that the room then mimics the ambient CO<sub>2</sub> concentration of the surrounding atmosphere, and approximates the outdoor temperature and humidity, while protecting instruments from direct sunlight and inclement weather. This provides an advantage over controlled tests in a laboratory setting in that rather than just a multi-point calibration, comparing datasets over ambient concentrations and environmental conditions allows for a realistic evaluation of these instruments in more real world scenarios.

For a continuous, uninterrupted period of over four weeks from 1 May 2015 through 2 June 2015, twelve sensor packages as described in Sect. 2 were run alongside the same Los Gatos Fast Greenhouse Gas Analyzer used in winter 2014 in this room. For our reference dataset, we used the dry carbon dioxide (CO<sub>2 dry</sub>) output calculated by the LGR. This takes into account the

correction of the molar mixing ratio as if water vapor wasn't present. The raw CO<sub>2</sub> values were recorded from each K30, temperature and pressure were recorded from each BMP180 sensor, and water vapor mixing ratio was also recorded by the LGR. All of the observations were recorded every two seconds, and averaged into one-minute values. The complete time-series of data from the LGR and select K30s are shown in the top panel of Fig. 6 and a difference plot for each K30 time series versus the LGR in the bottom panel of Fig. 6.



**Figure 6.** Continuous time series data from the rooftop ambient evaluation chamber described in Sect. 4. Top panel: CO<sub>2</sub> mixing ratio observed by six K30 sensors as well as the Los Gatos Research Fast Greenhouse Gas Analyzer. Middle panels: observed atmospheric pressure, temperature, and water vapor mixing ratio, respectively. Bottom panel: difference of each K30 from the Los Gatos instrument.

Over this 32-day period, the LGR observed an ambient variation of CO<sub>2</sub> with an average value of nearly 423 parts per million, and a standard deviation of just over 23 ppm. The aforementioned previous CO<sub>2</sub> calibrations of the LGR during a month-long period (i.e., a multi-point calibration was conducted every 1 to 2 days) showed a remarkable stability of the analyzer (see Fig. 1).

There is distinct synoptic variation in the diurnal cycle observed, with the magnitude varying from as little as 10 to 20 ppm over 24 hours to as much as 100 to 200 ppm. Each of the K30s was successfully able to resolve the ambient variations in CO<sub>2</sub> over this evaluation period, although none of the K30s matched the LGR perfectly in both absolute concentration and relative change. However, without any correction or calibration, each K30 was well within the manufacturer's stated accuracy of  $\pm 30$  ppm  $\pm 3\%$  of the reading.

From the difference plot, there are some important things to note. First and foremost, each individual K30 sensor has a distinct zero offset. A few of the sensors are approximately the same as the LGR, but many can have an offset that is as much as 5% lower (20 ppm) than the LGR FGGA. The difference between each K30 and the LGR all have standard deviations between 3 to 5 ppm and root mean square errors (RMSE) between 3.6 to 18.5 ppm. This means that once accounting for the offset of each individual K30, the practical accuracy of the K30 CO<sub>2</sub> sensor can be within 1% of the observed concentration. Secondly, each K30 difference time series appears to feature two wave patterns, one with a period of around one week, and another with a period of approximately a day. Given that the cycles seem fairly consistent and are present in each K30, this suggests that the difference between the recorded values from the LGR and each K30 is not random, but instead that there are external factors that can be assessed for potential compensation in the K30 response.

## 5 Calibration and environmental correction

In Figure 6, the difference between the LGR and each K30 is shown alongside time series of environmental data from the evaluation chamber. Temperature and pressure are from one of the BMP180 sensors in the room and the water vapor mixing ratio is recorded by the LGR. Just like the difference plot, each of the environmental variables features two distinct wave patterns.

There is a diurnal cycle of each variable, as well as synoptic-scale variability attributed to weather systems that occurs on the order of one week. Because the observed differences and the environmental variables seem to be correlated on both short, and long time scales, we used statistical regression methods to correct the observed concentration of CO<sub>2</sub> from the K30 sensor to a value approximately that of the actual concentration. Generally, a multivariate linear regression is of the form shown in Eq. (1):

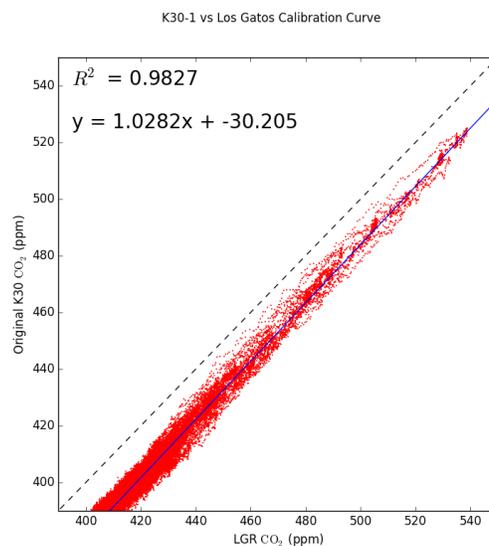
$$y = a_0x_0 + a_1x_1 + a_2x_2 + \cdots a_nx_n + b + \varepsilon_n \quad (1)$$

In this case, the K30 sensor measured value  $y$  is influenced by: the ‘true’ CO<sub>2</sub> value  $x_0$  (taken as the value from the high-accuracy LGR instrument), pressure, and other environmental variables  $x_1, x_2, \dots, x_n$ . A multivariate regression analysis can then be used to find the corresponding coefficients. In addition, in order to better identify the contribution from each individual factor, we also analyzed the data in a successive regression analysis, as described below.

### 5.1 Successive regression method

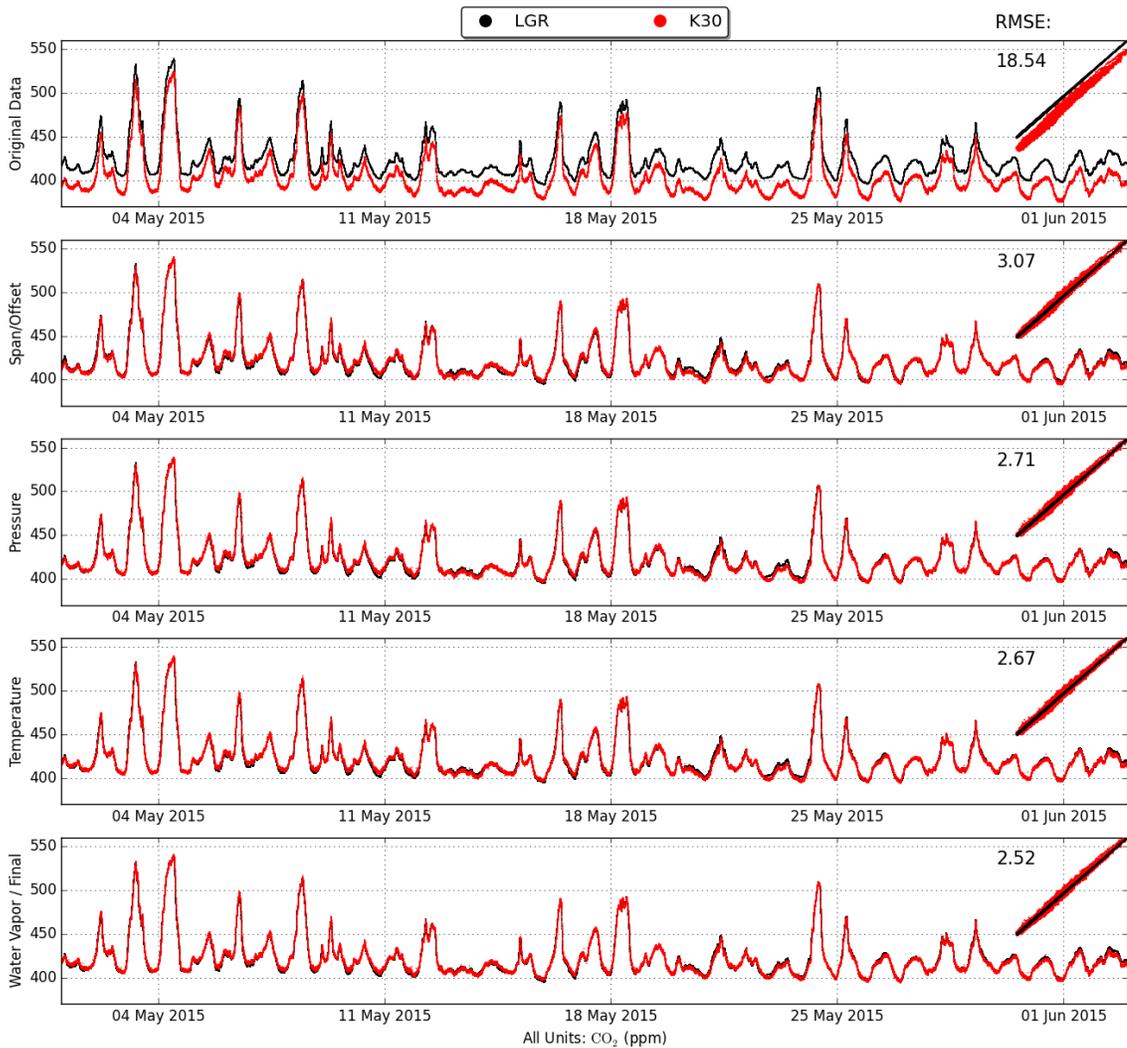
Applying Eq. 1 above for Los Gatos and K30 data only without atmospheric variables, we get the calibration curve shown in Fig. 7. For the first K30, the initial root mean square error

(RMSE) of the data was 18.54 ppm. After this initial span and offset regression, it dropped to 3.07 ppm. We then added the atmospheric variables one by one, using univariate regression method successively. After correcting for atmospheric pressure, the RMS difference between the K30 and LGR dropped to 2.71 ppm. Furthermore, including air temperature and water vapor mixing ratio dropped the standard deviation to 2.67 ppm and 2.52 ppm respectively. Therefore, using the successive regression method, the RMSE of the observed difference dropped from 18.54 ppm to 2.52 ppm. See Fig. 8 for the results and scatter plots for each step of the correction as well as Fig. 9 for a difference plot at each step for one K30. Another K30, that with the highest observed variability but relatively low offset, had the RMSE decrease from 4.96 ppm to 2.74 ppm. Similar results were observed for each K30 sensor evaluated.



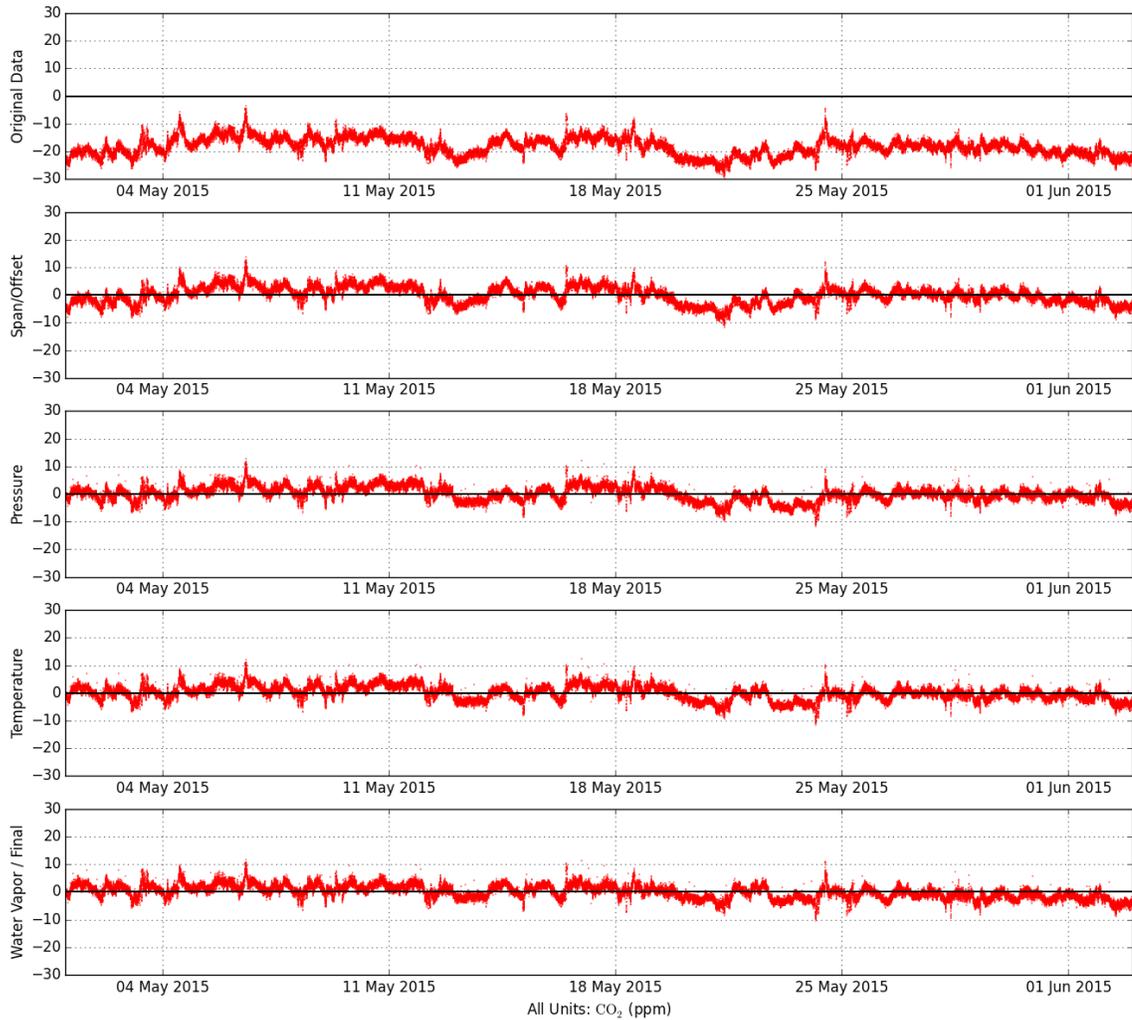
**Figure 7.** Calibration curve of K30-1 vs LGR before environmental correction.

K30-1 vs Los Gatos Successive Regression Results



**Figure 8.** A continuous time series as well as scatter plots for K30 #1 compared to the LGR during each step of the successive regression described in Sect. 5.1. From top to bottom: the original dataset, after correcting for span and offset, after correcting for pressure, after correcting for temperature, and finally, after correcting for water vapor. The root mean square error (RMSE) of the K30 data compared to the LGR at each step is annotated on the inset of scatter plot.

K30-1 - Los Gatos difference in Successive Regression

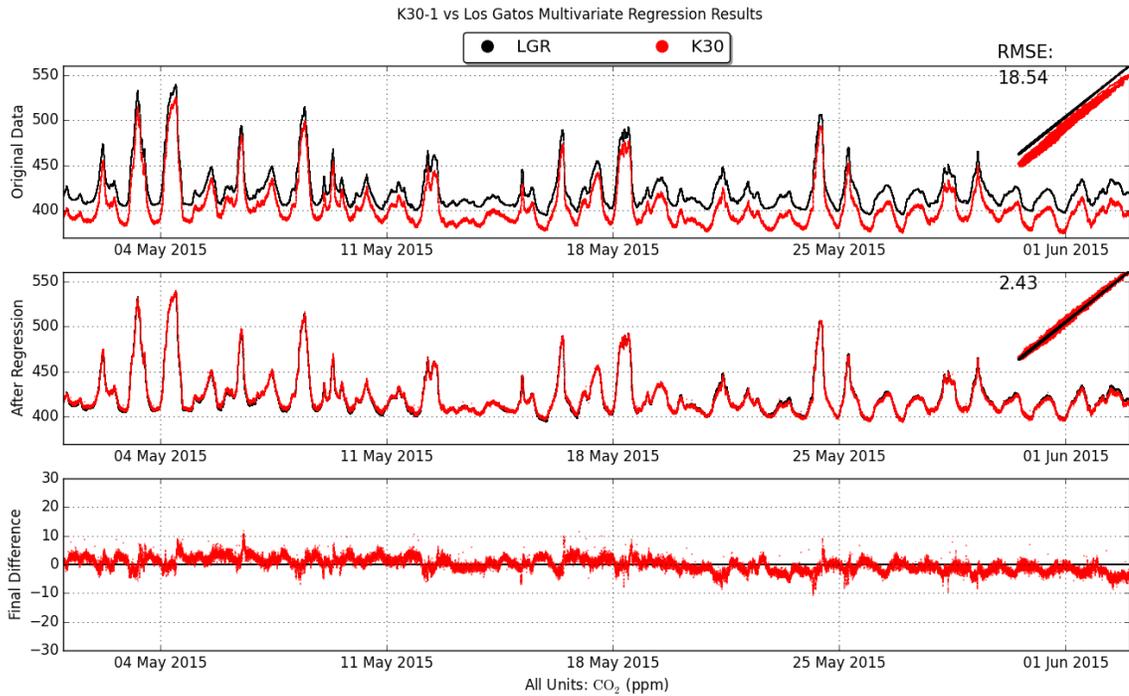


**Figure 9.** Difference plots for K30 #1 compared to the LGR during each step of the successive regression described in Sect. 5.1. From top to bottom: the original dataset, after correcting for span and offset, after correcting for pressure, after correcting for temperature, and finally, after correcting for water vapor.

## 5.2 Multivariate linear regression method

Alternatively, a multivariate linear regression statistical method can be used to calculate the regression coefficients for each K30 sensor. This results in five correction coefficients  $a_n$  and  $b_n$  where  $n$  represents each independent variable, the dry  $\text{CO}_2$  from the LGR, pressure  $P$ ,

temperature  $T$ , and water vapor mixing ratio  $q$ . Like the successive method above, these coefficients can be used similar to Eq. (1) along with the original K30 data,  $y$ , and the environmental variables to predict the true  $\text{CO}_2$  concentration observed.



**Figure 10.** A continuous time series as well as scatter plots for K30 #1 compared to the LGR for the multivariate regression described in Sect. 5.2. Top panel: the original data, middle panel: final time series after correction, and the bottom panel: difference plot between the corrected K30 dataset and the original LGR dataset. The root mean square error (RMSE) of the K30 data compared to the LGR before and after the regression is annotated to the upper left of the scatter plot.

Using the multivariate regression function provided by Python-SciPy-Stats (Jones et al., 2001), the same two K30's as described in Sect. 5.1 differences from the LGR were reduced to a RMSE of 2.43 ppm and 2.17 ppm respectively. Figure 10 shows the final results of the multivariate regression for the same K30 as in Fig. 8 and Fig. 9, as well as the difference between the corrected K30 dataset and the LGR. Like with the univariate method, similar results were observed from each K30 sensor evaluated.

### 5.3 Discussion

There are two things of note observed during this evaluation and analysis. First, both before and after the multivariate regressions, there are frequent large shifts in the sign of the difference between each K30 and the LGR. These changes occur at or around sunrise most days, and because of the rapid change in atmospheric CO<sub>2</sub> concentration, the ambient calibration chamber may not be well mixed during this time period. This, combined with the different response time of the K30s compared to the LGR, can lead to dramatic differences between what each K30 observes and what the LGR observes at the same timestamp. It is conceivable that if a more controlled environment were in place, the root mean square error of the total time period may be even lower than what was observed. Additionally, when using the successive regression method described in Sect. 5.1, the order of the independent environmental variables used in the regression was largely irrelevant to the end results. Rather than a particular variable having the most impact, the first regression always made the most significant reduction in error. This is most likely due to the two distinct variations observed in both the differences and the environmental variables, the slow synoptic-scale variation on the order of a few days, and the diurnal variation.

## **6 Conclusions**

The K30 is a small, low-cost NDIR CO<sub>2</sub> sensor designed for industrial OEM applications. Each of the sensors tested falls within the manufacturer's stated accuracy range of  $\pm 30$  ppm  $\pm 3\%$  of the reading when compared to a high-accuracy Los Gatos Research analyzer, but these ranges aren't particularly useful for scientific applications. However, once correcting for a zero-offset, and performing a regression analysis, the practical accuracy of these sensors is less than five parts per million, or approximately 1% of the observed value, with final root mean square errors of each K30 being 2-5 ppm. With errors in this range, these instruments could be used in a variety of scientific applications, including one where the cost allows for the spatial density of observations to better represent an area's CO<sub>2</sub> concentration than that of a single high-accuracy observation site.

## References

- Allan, D. W.: STATISTICS OF ATOMIC FREQUENCY STANDARDS, Proceedings of the Institute of Electrical and Electronics Engineers, 54, 221-&, 10.1109/proc.1966.4634, 1966.
- Bosch Sensortec: BMP180 Digital Pressure Sensor Datasheet, available at: <https://www.adafruit.com/datasheets/BST-BMP180-DS000-09.pdf> (last access: 29 December 2015), 2013.
- Briber, B., Hutyra, L., Dunn, A., Raciti, S., and Munger, J.: Variations in Atmospheric CO<sub>2</sub> Mixing Ratios across a Boston, MA Urban to Rural Gradient, *Land*, 2, 304, 2013.
- Eugster, W., and Kling, G. W.: Performance of a low-cost methane sensor for ambient concentration measurements in preliminary studies, *Atmospheric Measurement Techniques*, 5, 1925-1934, 10.5194/amt-5-1925-2012, 2012.
- Gas Sensing Solutions: COZIR Ultra Low Power Carbon Dioxide Sensor, available at: [http://www.gassensing.co.uk/media/1050/cozir\\_ambient\\_datasheet\\_gss.pdf](http://www.gassensing.co.uk/media/1050/cozir_ambient_datasheet_gss.pdf) (last access: 29 December 2015), 2014.
- General Electric: Telaire T6615 Sensor Dual Channel Module, available at: [http://www.avnet-abacus.eu/fileadmin/user\\_upload/Products\\_Menu/Amphenol/AmphenolAdvancedSensors\\_CO2\\_double\\_channel\\_module.pdf](http://www.avnet-abacus.eu/fileadmin/user_upload/Products_Menu/Amphenol/AmphenolAdvancedSensors_CO2_double_channel_module.pdf) (last access: 29 December 2015), 2011.
- Holstius, D. M., Pillarisetti, A., Smith, K. R., and Seto, E.: Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California, *Atmospheric Measurement Techniques*, 7, 1121-1131, 10.5194/amt-7-1121-2014, 2014.
- Hurst, S., Durant, A. J., and Jones, R. L.: A low cost, disposable instrument for vertical profile measurements of atmospheric CO<sub>2</sub>, Centre for Atmospheric Science, Department of Chemistry, University of Cambridge, 2011.
- Keeling, C. D., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M., and Meijer, H. A.: Atmospheric CO<sub>2</sub> and <sup>13</sup>CO<sub>2</sub> exchange with the terrestrial biosphere and oceans from 1978 to 2000: Observations and carbon cycle implications, *History of Atmospheric CO<sub>2</sub> and Its Effects on Plants, Animals, and Ecosystems*, 177, 83-113, 2005.
- Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO<sub>2</sub> emissions: Minimum site location requirements for the Los Angeles megacity, *Journal of Geophysical Research-Atmospheres*, 118, 1-8, 10.1002/jgrd.50135, 2013.
- Los Gatos Research: Fast Greenhouse Gas Analyzer (Enhanced Performance Model) Datasheet, available at: [http://www.lgrinc.com/documents/LGR\\_FGGA\\_Datasheet.pdf](http://www.lgrinc.com/documents/LGR_FGGA_Datasheet.pdf) (last access: 29 December 2015), 2013.
- Langridge, J. M., Ball, S. M., Shillings, A. J. L., and Jones, R. L.: A broadband absorption spectrometer using light emitting diodes for ultrasensitive, in situ trace gas detection, *Review of Scientific Instruments*, 79, 14, 10.1063/1.3046282, 2008.
- Piedrahita, R., Xiang, Y., Masson, N., Ortega, J., Collier, A., Jiang, Y., Li, K., Dick, R. P., Lv, Q., Hannigan, M., and Shang, L.: The next generation of low-cost personal air quality sensors for quantitative exposure monitoring, *Atmospheric Measurement Techniques*, 7, 3325-3336, 10.5194/amt-7-3325-2014, 2014.

- Raspberry Pi Foundation: Raspberry Pi Hardware Documentation, available at: <https://www.raspberrypi.org/documentation/hardware/raspberrypi/> (last access: 29 December 2015), 2015.
- SenseAir: CO<sub>2</sub> Engine K30 Specification, available at: [http://www.senseair.com/wp-content/uploads/2015/03/CO2-Engine-K30\\_PSP110-R7.pdf](http://www.senseair.com/wp-content/uploads/2015/03/CO2-Engine-K30_PSP110-R7.pdf) (last access: 29 December 2015), 2007
- Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., Patarasuk, R., and Razlivanov, I.: Toward quantification and source sector identification of fossil fuel CO<sub>2</sub> emissions from an urban area: Results from the INFLUX experiment, *Journal of Geophysical Research-Atmospheres*, 120, 292-312, 10.1002/2014jd022555, 2015.
- Wang, Y., Li, J. Y., Jing, H., Zhang, Q., Jiang, J. K., and Biswas, P.: Laboratory Evaluation and Calibration of Three Low- Cost Particle Sensors for Particulate Matter Measurement, *Aerosol Science and Technology*, 49, 1063-1077, 10.1080/02786826.2015.1100710, 2015.
- Yasuda, T., Yonemura, S., and Tani, A.: Comparison of the Characteristics of Small Commercial NDIR CO<sub>2</sub> Sensor Models and Development of a Portable CO<sub>2</sub> Measurement Device, *Sensors*, 12, 3641-3655, 10.3390/s120303641, 2012.
- Young, D. T., Chapman, L., Muller, C. L., Cai, X. M., and Grimmond, C. S. B.: A Low-Cost Wireless Temperature Sensor: Evaluation for Use in Environmental Monitoring Applications, *Journal of Atmospheric and Oceanic Technology*, 31, 938-944, 10.1175/jtech-d-13-00217.1, 2014.