Spring 1-1-2018

Enabling Community-Based Air Quality Science Through the Development of Sensor Systems, Resources, and Partnerships

Ashley Monika Collier-Oxandale
University of Colorado at Boulder, Ashley.Collier@Colorado.EDU

ENABLING COMMUNITY-BASED AIR QUALITY SCIENCE THROUGH THE DEVELOPMENT OF SENSOR SYSTEMS, RESOURCES, AND PARTNERSHIPS

By

Ashley Monika Collier-Oxandale

B.S., University of Colorado Boulder, 2013

M.S., University of Colorado Boulder, 2015

A thesis submitted to the Faculty of the Graduate School of the University of Colorado Boulder in partial fulfillment of the requirement for the of Doctor of Philosophy

Department of Environmental Engineering, 2018

This thesis entitled:

Enabling Community-based Air Quality Science through the Development of Sensor Systems, Resources, and Partnerships

Written by Ashley Monika Collier-Oxandale

has been approved for the Department of Environmental Engineering

_____

Michael P. Hannigan

_____

Jana B. Milford

_____

Date

The final copy of this thesis has been examined by the signatory, and I find that both the content and the form meet acceptable presentation standards of scholarly work in the above-mentioned discipline.

Collier-Oxandale, Ashley Monika (PhD, Environmental Engineering)

Enabling Community-based Air Quality Science through the Development of Sensor Systems, Resources, and Partnerships

Thesis directed by Professor Michael Hannigan

Abstract

Low-cost air quality sensor systems have the potential to provide entirely new information about our air quality given (1) the increase in temporal and spatial resolution that they facilitate, (2) their capacity to utilize many different sensor types in a single system, and (3) their accessibility that enables citizens to measure air quality for themselves. However, there are still many challenges associated with sensor use, including issues of sensor performance quantification and a need for best practices to guide the use of this technology, particularly for community-based research. This thesis addresses these challenges. This thesis includes the quantification of volatile organic compound using sensors to provide methane and non-methane hydrocarbon concentration estimates in complex environments. To support the development of best practices, multiple deployments allowed for the exploration of questions related to the influence the building-scale variability of pollutants on sensor system siting, choices in sensor data processing, and field calibration procedures. Additionally, education and outreach work utilizing sensors and involving partnerships with local communities are described with the focus on resources and lessons that could support future community-based air quality research.

This thesis also demonstrated the potential for sensor data. For example, sensor estimates of methane levels from a network deployed in rural Colorado revealed trends similar to those noted by other researchers using high-quality instrumentation and methods. In another example, sensor estimates of methane and total non-methane hydrocarbons levels, analyzed along with other sensor signals (i.e., from carbon monoxide and carbon dioxide sensors), helped to identify distinct pollutant sources on a fine temporal and spatial scale in a South Los Angeles neighborhood. The results and conclusions of this work support the continued development of this technology with the goal of collecting preliminary and supplementary information that may contribute to improved public and environmental health.

Acknowledgements

I am truly humbled by the support, guidance, and encouragement I have received throughout this work.

I would like to begin by offering sincere gratitude to my advisor, Professor Michael Hannigan, who found the perfect balance between developing my knowledge and skills and granting me the freedom and creativity that allowed my work to flourish. I would not be here without his guidance and I am lucky to be able to carry what I've learned from him into my career. I would also like to extend a generous thank you to my committee members: Jana Milford, Marina Vance, William Griswold, and Jill Johnston for sharing their time and expertise. Their insights shaped my thinking and helped me in the writing of this dissertation. I also feel so much appreciation for the team I have had the privilege of working with, including: Evan Coffey, Joanna Gordon Casey, Kyle Karber, Jacob Thorson, Ricardo Piedrahita, Kira Sadighi, Lucy Cheadle, and Nick Masson, among many others. Working in a collaborative environment with this team, full of curiosity and optimism, helped shape me as a scientist and brought joy to my work.

This work was enhanced by the support and assistance of many wonderful academic colleagues, including: Hannah Halliday, John Ortega, Daniel Knight, Jill Johnston, Bhavna Shamasunder, Christine Wiedinmyer, William Griswold, Massimiliano Menarini, Kevin Patrick, Michael Ostertag, and Sharad Vikram, as well as many other collaborators over the years. In addition to academic partners, I have so much gratitude for the community partners I have had the honor of working with, including: Nicole Wong, Sandy Navarro, George Ware, LaShonn Billingsley, Patti Iwasaki, Debbie Main, the numerous teachers and students I worked with in Delta County and Greeley County (in particular Ben Graves) as well as many other participants in our projects – working with these individuals helped to shape my perspective and values. I have also had the great fortune to work with those supporting community-based projects and outreach work, including: Raj Pandya, Natasha Udu-gama, Ben Kirshner, and Katya Hafich. These individuals have provided me with so much inspiration and direction over the years, I would like to extend a thank you for this. Notably this was in addition to all the support I have received from CU's Office of Outreach and Engagement throughout my graduate work. Furthermore, many aspects of this research would not have been possible without the help of regulatory partners, including: the Colorado Department of Public Health and the Environment (esp. Erik Mattson and Gordon Pierce), South Coast Air Quality Management District (esp. Andrea Polidori and his team), San Joaquin Valley Air District, San Diego Air Pollution Control District, and Delta County Public Health

(esp. Ken Nordstrom). As well as Ron Williams, Andrea Clements, and Eben Thoma with the US EPA. These partners facilitate important field work and help to sustain a dialogue around research into low-cost air quality sensors.

Last but certainly not least, I would like to thank all of my very dear friends and family for their unending support and encouragement. My husband, Joseph, who continually inspires me to new ideas and ways of thinking; I'm so thankful for the ways in which our conversations have nourished the growth of my work and for the richness he has brought to my life. My brothers and best friend (Eric, Dave, and Sheila) whose belief in me and offers of encouragement whenever it was needed helped to propel me forward. I would especially like to thank my father, John, who instilled in me the confidence to pursue engineering. I am so proud to be able to cite his graduate thesis in my thesis – a small acknowledgement of his incredible personal and professional accomplishments, which laid a foundation for my own successes. I'll conclude with a thank you to my mother, Monika, for teaching me grace and strength; I know she has been with me in spirit on this journey.

Funding sources that made this work a reality: National Science Foundation (under grant numbers CNS-1446912 and CBET-1240584), National Institutes of Environmental Health Sciences (under grant number: R21ES027695), CU Engage Graduate Fellowship Program, DISCOVER-AQ Project (NASA), and the University of Colorado Boulder's Office of Outreach and Engagement

Contents

List of Figures

6.12: Panel (a) includes the complete Validation 2 dataset from the Los Angeles deployment. The statistics for each are RMSE is 0.17 ppm and absolute percent relative error is 6.77 % for the unfiltered data (blue); RMSE is 0.15 ppm and absolute percent relative error is 6.19 % for the filtered data (green), with the reference data plotted in black. Panel (b) is a close-up of approximately 2 days illustrating an instance where the filtering helped to reduce an overprediction of methane concentrations.

6.13: Histogram of carbon monoxide data from the two deployments and scatter plot of carbon monoxide data vs. methane data also from the two deployments. Note that all data in these two plots are from reference instruments.

6.14: (a) Time series of methane and carbon monoxide data from the reference monitors and converted U-Pod sensor data. (b) Scatter plot of reference methane data (x axis) vs. U-Pod methane data (y axis) with the points colored by carbon monoxide vales. Together they further show the Figaro TGS 2600's cross sensitivity to carbon monoxide, illustrating how many over predictions correspond to instances when CO is at or above 0.5 ppm.

6.15: Time series of minute-resolution data from four Colorado sites and a map showing the location of the four U-Pod with respect to all active and inactive wells (COGCC, 2017).

6.16: Box plots of all data for each U-Pod, grouped into day-time and nighttime values. Note that whiskers are the 5th and 95th percentile values. These U-Pods are then further grouped by region.

6.17: Well sites (COGCC, 2015) are plotted along with the differences between the 90th percentile value for the main Platteville U-Pod (P1) and every other U-Pod during the day (a) and night (b). The color bar indicates the magnitude of the difference in units of ppm CH_4, and the same color bar scale is applied to both plots.

6.18: Figure 18. Mean methane value for each hour of the day, for each U-Pod color-coded by region.

7.1: The map (1a) illustrates the placement of the Platteville site with respect to the nearby cities Denver and Boulder, and oil and gas activity as indicated by active and inactive wells (34). The top right photo (1b) shows the inside of a Y-Pod (a newer version of the U-Pod), which depicts the sensors used and the general design of the system. There were some updates to the circuit board, but the sensors used were the same. The photo on the bottom right (1c) shows the placement of the two U-Pods on the NATIVE Trailer at the Platteville site.

7.2: The far-left panel (2a) depicts a time series including the benzene data from the PTR-MS, and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5^{th} and 95^{th} percentile respectively (in red).

7.3: The far-left panel (2a) depicts a time series including the summed aromatic data from the PTR-MS (benzene, toluene, C_8 and C_9 alkylbenzenes), and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5^{th} and 95^{th} percentile respectively (in red).

7.4: The far-left panel (2a) depicts a time series including the summed VOC data from the PTR-MS (acetaldehyde, acetone, formaldehyde, methanol, and the aromatic species), and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5^{th} and 95^{th} percentile respectively (in red).

7.5: The far-left panel (2a) depicts a time series including the methane data from the Los Gatos cavity ring-down instrument, and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5^{th} and 95^{th} percentile respectively (in red).

7.6: Cumulative distributions of relative sensor error for all fitted testing or validation data, the two dotted lines represent the fitted testing data for benzene and summed aromatic, respectively, with the low values removed from the reference data.

7.7: Complete residuals verses other, non-target VOCs. The complete residuals for the Model 2 results are in black. The blue and fuchsia points that represent two values held constant. For panel a, the constant benzene values selected were 0.75 and 1.6 ± 0.05 ppb. For panel b, the summed aromatics values selected were 19 and 38 ± 1 ppbC. For panel c, the summed VOC values selected were 47 and 67 ± 1 ppbC. For panel d, the methane values
selected were 2.5 and 3 ± 0.05 ppm.

7.8: Plots of error (RMSE) versus the coefficient of determination (R^2) for the target VOC or group versus a non-target, potentially confounding VOC or group. These values are calculated on the basis of a 1-hour moving window and the points are colored according to the average of the target pollutant or group for that hour. The dotted black line on each is the overall RMSE determined in Section 3.1.

7.9: The boxplots above illustrate the results of training on a randomly selected 15% of the data and testing on the remaining 85% 25 times, using each sensor individually, the two sensors as predictors, and the two sensors plus an interaction between them. Plots a and b depict the R^2 and RMSE respectively for the testing data. Note all of the data has been normalized and the whiskers are the 5^{th} and 95^{th} percentile respectively. The x-axis indicates the VOC or group being predicted.

7.10: Each panel displays two pollutants plotted together (from the reference data set). Panel a is the benzene versus the toluene while panel b is methane versus the benzene and panel c is carbon dioxide versus benzene. Points with certain sensor ratios are then colored according to the colors listed in each the legend. Other relevant information, for example a ratio of 2 for the toluene to benzene ratio or relationships from previous studies.

7.11: Histogram of complete sensor ratios for each hour using moving correlation window vs. the ratios for a select subset of data with high correlation between benzene and methane (> 0.85), a ratio of benzene:methane typically associated with oil and gas activity (< 1.0 ppb/ppm), and a significant change in methane (> 0.5 ppm).

8.1: Map of sampling sites and major sources of interest. Note that the Y-Pod locations have been approximated to the center of their respective blocks in order to protect participant identities. The first initial is indicative of

the site type: N – neighborhood site, R – near roadway site, E – near oil extraction site, c – ongoing co-location site.

8.2: Top left – Y-Pod interior, top right – two-Y-Pods deployed on a roof, bottom left – Y-Pods deployed at a co-location site near downtown LA, bottom right – Y-Pod deployed on a building above a roadway.

8.3: Single Sensor model for CH_4, the colors indicate individual Y-Pods, while the reference data is shown in black. The testing period is highlighted in yellow.

8.4: Multi Sensor model for CH_4, the colors indicate individual Y-Pods, while the reference data is shown in black. The testing period is highlighted in yellow.

8.5: Multi sensor model for TNMHCs, the colors indicate individual Y-Pods, while the reference data is shown in black. The testing period is highlighted in yellow.

8.6: Panel a shows additional validation data, concurrent to the field deployment, with both the single (in blue) and multi sensor (in green) models plotted with the reference data (in black). Panel b shows the single sensor fitted data plotted against the multi sensor data, the color indicates a normalized density of the data (yellow – high density, blue – low density). The black dotted line is the 1:1 line.

8.8: Paired data for the two sampling sites near the drill site, with co-located data plotted in purple and deployed data in green; panel a – single sensor for CH_4, panel b – multi sensor for CH_4, and panel c – multi sensor for NMHC

8.9: Hourly averages across deployment for all sensors (indicated in colors, with the exception of C1, which is indicated in grey), for the two methane plots, the averaged reference data, is indicated in black. Panel a – single sensor model for CH_4, panel b – multi sensor model for CH_4, panel c – multi sensor model for TNMHC

8.10: Roughly one day of hourly data from two MegaCities sites: CNP (near our C1 Y-Pod), and USC (near our R4 Y-Pod). The data was retrieved from https://megacities.jpl.nasa.gov/portal/data-access/. Note, the Y-Pod data was also hourly averaged, and the data plotted was from the multi sensor model.

8.11: Panel a – Original methane level estimates from three sites, Panel b – the same data with the baseline removed. All data was converted using the multi sensor CH_4 model.

8.12: Baseline removed data from four sites grouped by hour of the day (in local time), panels a-d include the single sensor CH_4 data, panels e-h include the multi sensor CH_4 data, and panels i-l include the multi sensor NMHC data. The whiskers on the box plots represent the 5th and 9th percentiles respectively.

8.13: Approximately one week of baseline removed data, with panel a including the single sensor CH_4 data, panel b including the multi sensor methane data, and panel c including the multi sensor NMCH data. The time stamp is local, and the yellow box highlights the event discussed in the text. To the right of each panel is a zoomed in version of the event highlighted in yellow.

8.14: Panel a includes baseline removed TNMHC, CO_2, and CO data from Site E2 as well as baseline removed TNMHC data from Site E1. All of the data is from 8/15/16 and the times listed are local times. Panel b includes a wind rose for only the period of enhancements in TNMHCs.

8.15: Panel a includes baseline removed TNMHC and CO_2 data from Site E2, with the wind direction data during the period of elevation highlighted in yellow shown in panel b. Panel c includes baseline removed TNMHC and CO_2 data from Site E1, with wind direction data from the period of elevations highlighted in yellow plotted in panel d.

8.16: Baseline removed data from the same four sites shown in Figure 12, grouped by hour of the day (again, in local time), panels a-d include CO_2 data, and panel e includes CO data from Site E2. The whiskers on the box plots represent the 5th and 9th percentiles respectively.

8.17: Scatterplots of baseline removed data, with the correlation coefficients for each set of data noted. The top row of CH_4 data is from the single sensor model, and the second row of CH_4 data is from the multi sensor model.

8.18: Complete CO_2 and CO data from Site E2, with a ref. ratio based on the CARB inventory included, illustrating the similarity between the observed data and the expected CO/CO_2 ratio for the LA area.

8.19: Baseline removed CO_2 and TNMHC data from Site E2, with a period highlighted in which an odor complaint was noted by nearby residents.

8.20: Baseline removed CO_2 and TNMHC data from Site E2 and Site E1, annotated with noise and odor complaints as well as observations of residents of heavy activity at the drill site.

8.21: The daily 50th and 95th percentile values plotted against the number of observations per day (i.e., summed activity reports, odor complaints, and/or noise complaints).

IV.1: Time series of Y-Pod data (shown in colors), and reference data (shown in black), for CO_2 (top) and O_3 (bottom).

IV.2: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining neighborhood-scale variability of CO_2.

IV.3: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining neighborhood-scale variability of O_3.

IV.4: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining building-scale variability of CO_2.

IV.5: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining building-scale variability of O_3.

IV.6: Scatter plot of differences between B1 and B5 (on the x-axis, for O_3, B1 – B5; and on the y-axis, for CO_2, B5 – B1), colored by the hour of the day. The dotted lines around the center axes indicate the respective RMSEs for CO_2 and O_3.

V.1. Scatter plots of P1 versus every other Y-Pod including both co-located data (in green) and deployed data (in blue).

V.2: Panel (a) includes time series of methane reference data and fitted sensor data. Panel (b) includes the methane data for the Validation 2 dataset binned by days. The yellow segments highlight periods when underestimations below atmospheric background were removed, which coincide with days where the dynamic range is less than the expected uncertainty (RMSE is 0.18 ppm).

V.1: Correlation plot for reference gases

V.2: Regression model residuals (from Section 3.1)

List of Tables

Chapter 1: Introduction

"The senses don't just *make sense* of life in bold or subtle acts of clarity, they tear reality apart into vibrant morsels and reassemble them into meaningful patterns."

~ Diana Ackerman, A History of the Senses

Sensors serve a purpose very similar to our natural senses, they make it possible to detect and extract very specific pieces of reality, which we can then inspect for meaningful patterns. These patterns can anchor our understanding and provide a basis for action. At this stage in our evolution, our natural senses, despite their sophistication, are not enough. Manmade sensors can supplement our senses or provide information in places we cannot go, for instance broadening the visible spectrum (Gålfalk et al., 2016) or placing them inside of a brain to begin to decode the mysteries of its functions (Collier & Mahoney, 1984). Recently, much work has gone into understanding and developing sensors for lower cost air quality monitoring. This development has resulted in tools that can provide us with entirely new information about our local air quality resulting from the ability to deploy these systems in networks and the ability to leverage signals from different sensor types. While these new tools will never replace our existing monitoring infrastructure, or the high-quality tools on which we base regulatory decisions – there is an enormous potential to supplement the information we already have. These new tools could enable quicker identification of air quality issues and more strategic use of high quality instrumentation helping us confront the causes of poor air across the globe.

According to the World Health Organization, in 2016 outdoor air pollution was estimated to have caused 4.2 million premature deaths in both cities and rural areas, and 91% of the world's population was estimated to be living in areas where air quality guidelines (defined by the WHO) were not being met (WHO, 2018). This statistic translates to 9 out of every 10 people breathing polluted air worldwide. While in some cases the cause of this pollution may be obvious, air quality monitoring is still a necessary piece of the puzzle as it can support the design of more robust solutions and be used to assess the effectiveness of different actions intended to reduce exposure. Furthermore, new technologies such as low-cost sensing systems will support the collection of data on smaller, more local scales, informing locally relevant action.

Currently sensor systems are well-suited for monitoring certain criteria pollutants. For example, there are several studies illustrating the robust performance of carbon monoxide and ozone sensors at ambient levels (Mead et al., 2013; Cross et al., 2017; Sadighi et al., 2018; Zimmerman et al., 2018). There has also been much research in

particulate matter sensors that utilize light-scattering detection principles (Johnson et al., 2018; Kelly et al., 2017). While these pollutant species are important for human health, the ability to use sensors for the detection of air toxics and important hydrocarbons, such as methane, would build our capacity to characterize air quality issues.

One study in rural Alaska demonstrated the capability of low-cost metal oxide VOC sensor to replicate diurnal trends at low, ambient concentrations (Eugster & Kling, 2012). Other studies in laboratory settings have illustrated promising performance from low-cost sensors for VOC detection; particularly when multiple sensors are used in an array or when techniques such as temperature-controlled operation are incorporated (Sauerwald et al., 2018; Schütze et al., 2017). These studies have demonstrated the ability of sensors to not only detect specific VOCs at ambient levels, but also to accomplish this amidst confounding gases (Leidinger et al., 2014). A few studies have also piloted the use of low-cost sensors in the field, illustrating their capacity to detect VOCs in more complex environments. One of these showed the detection of ambient benzene using an array and neural network calibrations (De Vito et al., 2008). Another pointed to the usefulness of PID sensors to detect short-term enhancements and how this could be leveraged by combing sensors with a tool that can provide speciated information (e.g., passive adsorption tubes (Thoma et al., 2016). However, there are no studies utilizing low-cost sensors for methane detection in complex environments, or studies investigating the ability of multiple VOC sensors to provide estimates of concentrations of individual and grouped or summed VOCs in the field.

The detection of hydrocarbons and VOCs are an important next step in the development of low-cost sensors. Methane is a potent greenhouse gas, and recently leaks along the production and distribution chain have become a concern as researchers are consistently finding that methane has been underestimated in the established emissions inventories (Miller et al., 2013; Marchese et al., 2015). Other VOCs are a concern as these can include hazardous air pollutants (HAPs) or air toxics, with known implications for human health. These are a concern as researchers have found that small scale variability exists among these pollutants based on proximity to sources. For example, in the RIOPA study researchers found that residential concentrations of certain air toxics were 1.5 – 4 times higher at homes less than 50 m from a source (Kwon et al., 2006). Also, concentrations of toxics have been detected above the EPA's benchmark for cancer and non-cancer risks in communities in recent years (Kwon et al., 2006; Wu et al., 2012). When considering potential emissions sources, a typical refinery may be operating according to regulations, however there may be another refinery that is an outlier and releasing a relatively large amount of emissions that affect nearby communities. Low-cost sensors could help to identify these outliers and hot spots more quickly. Instances like this are

of particular concern for communities disproportionately impacted by air pollution, communities that are often also typically minority and socioeconomically disadvantaged. Several studies over the years have demonstrated the reality of this disproportionate impact (Brown 1995; Souza et al., 2009; Marshall 2008). Additionally, when residents are concerned about a nearby potential source of pollution, they are eager to get involved and learn more (Brown 1992). Furthermore, the residents of these communities often have valuable experiential knowledge on their immediate environment and the sources that they live next door to.

AN EXAMPLE OF COMMUNITY-BASED AIR QUALITY RESEARCH

Community members in Northeast Denver were concerned about a number of local air quality issues. After a meeting was held discussing different approaches to air quality monitoring and their concerns, we organized a pilot project to screen for perchloroethylene and radon in a small sample of homes (TNH2H, 2015). Perchloroethylene was selected as residents had seen reports of PERC vapor intrusion near former dry-cleaning facilities affecting homes and childcare facilities in Denver. Radon was added by the suggestion of a community member. While we did not find evidence of high levels of perchloroethylene, we did find levels of radon above the US EPA's "action level" in 12 out of 15 homes (TNH2H, 2015). Although the more startling take-away was how many participants remarked that they had never been advised to test for radon, despite being long-term residents of the neighborhood, some up to 30 years. This is especially surprising because radon is a well-known issue in Colorado, with 1 out of every 2 homes in Denver County testing above the "action level" (CDPHE, 2018). Radon is also the leading cause of lung cancer in non-smokers in the US (US EPA, 2018).

This pilot project provides a clear example of a low-cost, accessible air quality monitoring method (off-the-shelf, short-term radon sampling kits), helping to identify an important and actionable issue in a community. This project was a partnership between Taking Neighborhood Health to Heart, an organization with experience engaging in community-based participatory research (CBPR), the Thriving Earth Exchange, an organization working to support community-based science, and researchers from our team at CU Boulder who received additional support from CU Engage. Using the results of this pilot project, the community-based organization, Taking Neighborhood Health to Heart (TNH2H), has continued to expand this project and conduct more radon sampling in the community as well as hold education and outreach events with the support of a US EPA Environmental Justice Small Grant (US EPA, 2017c). This pilot project serves as a great model for community-based participatory air quality research.

Utilizing a participatory method already established by TNH2H aided with recruitment of participants, ensured the success of crowd-funding efforts, and enabled us to implement sampling procedures that protected both participants and researchers (Main et al., 2012). Iterative data analysis and interpretation between partners helped to determine ideal next steps, including securing a source of financial support for mitigation in low-income homes through an existing emergency home repair program. Arguably the pilot project was conducted using modest funding (<$5000), but the support and participation of many partners led to its success and helped to facilitate the acquisition of funding from more conventional sources. It's possible that low-cost sensors systems could be used in a similar iterative fashion, first identifying issues and providing initial results that facilitate further investigations.

BENEFITS AND POTENTIAL OF LOW-COST SENSING SYSTEMS

Of course, as previously stated, sensor systems will never replace conventional air quality monitoring equipment, but there are gaps in the research that sensors may be well-suited to fill. For example, regulatory monitoring stations require not only costly instrumentation, but also personnel to carry out the appropriate Federal Reference Methods and ensure the collection of legally defensible data. A task that is especially important as this is the data used to determine whether or not regulations are being met. However, it would be cost prohibitive to deploy this type of equipment and these methods on neighborhood scales in every neighborhood with a concern. While there are reliable tools, maintained by research groups and utilized on a more as needed basis that can sometimes be accessed through partnerships, for example Picarro CRDS which can provide high quality methane data or a PTR-MS which can provide speciated high-time resolution hydrocarbon data, these tools are generally reserved for larger campaigns that pursue research questions of interest to the academic community. These tools would not typically be accessible to a community curious about the local effects of a particular emission source. Another option is passive sampling techniques, which are lower in cost and can provide high quality speciated information (Eisele et al., 2016; Johnston & Gibson, 2013). The trade-off here is that these tools cannot provide high-time resolution data and may not support the study of individual emission events.

Low-cost sensor systems can provide high temporal and spatial resolution data and they can be deployed in customizable networks. These networks are ideal for examining temporal and spatial variability. In one study, researchers utilized a baseline extraction technique to separate the regional and local trends in the data from a sensor network. These researchers were then able to compare the average rural vs urban baseline and learn about the transport

and impact of local sources on urban versus rural air. Other studies have confirmed the existence of variability in ozone concentrations on very small spatial scales, which could impact exposure levels and suggests that the nearest regulatory monitoring stations may not always be the best estimate an individual's exposure (Sadighi et al., 2018; Cheadle et al, 2017). Studies have also confirmed the ability of these networks to increase the granularity of air quality data, which has the potential to reveal new information and patterns (Shusterman et al., 2016; Schneider et al., 2017).

Linked to this idea of increased spatial and temporal resolution of air quality data are networks to support community-based investigations. Given the accessibility and relatively simple deployment and operation of these tools, they are also well suited to support community-based investigations (Shamasunder et al., 2018). The IVAN Air Monitoring Network in Imperial County utilizes a network of low-cost PM sensors to provide higher resolution data, supplementary to the existing regulatory network (English et al., 2016). This community, in a county with the highest rates of hospitalization for asthma for school-age children, utilizes this data to inform actions at the local schools – specifically decisions regarding whether or not students spend time outdoors (English et al, 2016). This network of sensors allows individual schools to make decisions based on more locally relevant information. In addition to the ability to provide higher-resolution data to communities, there is the potential to leverage community knowledge and expertise to add context to the data from these networks. As noted earlier, community residents are likely to know quite a bit about potential sources of air pollution in the community. This information in addition to observed odors or visible emissions can support the interpretation of data from sensor networks. However, this type of analysis would require new ways of collecting comprehensive observational data as well as ways to merge the quantitative sensor data and qualitative observational information.

Other promising sensor applications are fenceline monitoring or leak detection and personal exposure monitoring. Given their size, cost, and power requirements, low-cost sensors are ideal for fenceline monitoring and leak detection even in remote areas, for example on a well-pad in a rural area. Studies have demonstrated that sensors are capable of detecting large changes from baseline, particularly in situations where the surrounding environment has no other or relatively few sources (Mead et al., 2013; Thoma et al., 2016). A sensor could be used to continuously monitor for leaks at a single target source, and additional streams of information such as meteorological data can enhance confidence in the information gained from this type of sensor data. In terms of personal exposure monitoring, many studies have pointed out how using the nearest available regulatory monitoring to estimate exposure is inadequate (Wilson et al., 2005). Researchers in public health are excited about sensors not only because there is the

possibility of increasing the accuracy of exposure data, but also given the temporal resolution, researchers can begin to explore the impact of acute, high-level exposure events. A few studies have also examined sensor capabilities for this purpose (Jerrett et al., 2017; Piedrahita et al., 2014).

Sensors are also well-suited to support education and outreach (E&O). Again, given the cost, and relatively simple operation of these tools they work well in K-12 E&O classrooms. For example, low-cost sensing systems can help make complex and seemingly invisible concepts more concrete, like combustion chemistry. Students can burn different fuels types (e.g., by using simple items like a lighter and a coffee stir stick) illustrating the difference between complete and incomplete combustion. These differences are visible in the data by comparing the response from CO_2 and VOC/CO sensors. Access to sensor systems can also support student-led investigations, in addition to more structured activities. For example, the Air Quality InQuiry Program (AQ-IQ), which I helped to build along with partners from CU, the Delta School District, and the Delta County Health Dept, provides the support needed for students to design and conduct their own air quality studies. This program includes access to low-cost sensing systems, a project-based learning curriculum, and university mentors (cite). Through this program students are able to experience the complete research process from planning to interpreting their data and presenting their results, all the while using the same tools being used by academics. This program encourages students to research questions that are personally and locally relevant to their interests, for example they often investigate the indoor air quality in their school or emissions from products that they and their fellow students use, such as perfume. More accessible environmental monitoring tools would support more efforts such as this, and in addition to benefitting those directly participating in the E&O, the resources developed and lessons learned could translate to the use of sensors by the public more broadly.

While communities often wish to use sensors to investigate specific concerns, the projects also provide an opportunity for environmental education and a chance to improve scientific literacy. By participating in research projects utilizing sensors, individuals might learn more about the complexities of air quality monitoring and decision making as well as more about air quality in general. These types of projects could also have the added benefit of improving relationships between regulatory agencies and communities.

HOW DOES THE THESIS ENABLE SENSOR SYSTEMS TO REACH THESE VISIONS?

In addition to these benefits and potential applications for sensors, it is important to discuss the challenges and limitations associated with their use. A workshop organized and led by the MetaSense Project provides an overview of the current state of the sensor technology and the major issues requiring attention. Chapter 2 includes a summary of this workshop as well as a general review of low-cost sensing systems. This summary provides valuable background and context on the field of low-cost sensor research; it also highlights current needs and questions that my work presented here speaks to. Several of the issues noted during the workshop, which my work helps to address, include quantification of sensor performance, a lack of best practices, and the need for more case studies and examples of low-cost sensor system use. Chapters 3 and 4 provide information on my use of low-cost sensors for education and outreach work, including the development of a program around the use of sensors and assessments of that program. These two chapters also further highlight the benefits and potentials of sensor use.

Returning to the content of Chapter 2, workshop participants identified the establishment of best practices as an upcoming need. The US EPA has developed many valuable resources to help encourage the purposeful use of sensors in the Air Sensor Toolbox, which includes a helpful guide to planning a sensor study and several tools for analyzing and visualizing sensor data (US EPA, 2017a). Still, given the wide variety of sensor uses we are likely to see, there is value in examining these questions from different perspectives and in different contexts. For example, by studying the variability observed by multiple sensor systems placed around a single field site. The results of this type of study might contain lessons that could inform sensor system placement for future studies, which is the analysis shared in Chapter 5. Currently, the siting and placement of sensor systems is primarily based on what is safest and most convenient for the property owner or resident. However, if a sensor system is placed in a location where it cannot collect the appropriate data to address the research question, then it might not matter if the sensor data was carefully quantified and is of high data quality. Answering questions like this will further support the collection of purposeful and useful sensor data.

Another ongoing concern regarding the use of low-cost sensor systems has been related to quantification and ensuring reliable data quality. The main reason this concern persists is that most low-cost sensors are cross-sensitive to environmental factors (i.e., temperature and humidity) and sometimes other pollutants as well (Lewis, 2016). Previous studies have found that field calibration or field normalization, as opposed to laboratory calibrations, support the generation of calibration models that can mitigate some of the effects of these cross-sensitivities and provide

relatively robust estimates of the concentration of target pollutants (Piedrahita et al., 2014; Castell et al., 2017). Basically, the dynamic conditions experienced during a field deployment are difficult to simulate in the lab. Field calibrations rely on a co-location with high-quality reference instruments, typically before and after a field deployment. Then using a technique such as multiple linear regression, the co-located data is used to train and test a calibration model to predict concentrations of the target pollutant. This calibration model can then be used to convert raw sensor signals to useable concentration data. Some specific issues related to calibration during the workshop included the need to determine the extent to which field calibrations are transferable to new locations and standardizing the way in which uncertainty is calculated and reported.

My work quantifying sensor performance, presented throughout Chapters 6, 7, and 8, primarily focuses on the quantification of VOC sensors for methane as well as individual and grouped VOCs, in the context of field deployments. As previously mentioned, this is an understudied area and further research could open the door to broader use of these tools in the future. However, quantifying VOC sensors is in some ways more complicated than it would be for sensors designed to detect other pollutants because there are so many more potential confounders given the number and variety of VOCs potentially found in a city. A review article highlights the complexities of measuring VOCs, even with conventional methods, and provides an overview of worldwide profiles, further illustrating the potential for complex mixtures (Kumar & Viden, 2007). Consequently, in addition to assessing the predictive abilities of the calibration models generated using typical quantification techniques, I also consider sensor behavior across changing VOC compositions. Furthermore, the availability of datasets collected in different environments, such as urban Los Angeles and rural Colorado, may lead to results and lessons learned that are more broadly applicable.

A final observation shared by workshop participants was that more case studies and examples of sensor projects could be valuable to both researchers and community groups by helping them think about how to use sensors and plan studies. The work presented here in Chapters 6 and 8 includes several deployments, from different locations, and with different objectives. These vary from a deployment in rural Colorado, examining the spatial variability of estimated methane concentrations across an area containing various densities of oil and gas activity, to another in Los Angeles, examining methane and total non-methane hydrocarbon data from a sensor network in a high density residential area impacted by traffic and other local industries. Each of the studies provides the opportunity to explore what conclusions can be drawn from this data that are unique to this approach and technology, which will help to highlight how sensors can best support existing monitoring efforts. Therefore, these studies both provide adaptable

examples and they contribute to a bigger picture. Additionally, during these deployments and projects, resources were developed and lessons were learned that can hopefully be leveraged by future sensor projects helping to push the needle forward on what future projects are able to accomplish.

Preface to Chapter 2

The following chapter is intended to provide background and context for my dissertation work. The author list of this publication is not in order of the level of contribution, but rather the order in which contributions were added to the paper. I was responsible for writing Sections 5.3 (Opportunities for Community-Driven Science) and 7 (Workshop Conclusions). I also assisted with a review of and revisions to the whole summary paper. In addition to work on this publication, I led the planning and organization of the workshop with the help of the planning committee.

Chapter 2: Low-Cost Air Quality Monitoring Tools: From Research to Practice (A Workshop Summary)

Contributing Authors: Andrea L. Clements[1], William G. Griswold[2], Abhijit Rs[3], Jill E. Johnston[4], Megan M. Herting[4], Jacob Thorson[5], Ashley Collier-Oxandale[5], & Michael Hannigan[5]

[1]US Environmental Protection Agency
[2]University of California, San Diego
[3]Environmental Defense Fund
[4]University of Southern California
[5]University of Colorado Boulder

Publication Information: Sensors, 17(11), 2478. DOI: 10.3390/s17112478, 2017

ABSTRACT

In May 2017, a two-day workshop was held in Los Angeles (California, U.S.A.) to gather practitioners who work with low-cost sensors used to make air quality measurements. The community of practice included individuals from academia, industry, non-profit groups, community-based organizations, and regulatory agencies. The group gathered to share knowledge developed from a variety of pilot projects in hopes of advancing the collective knowledge about how best to use low-cost air quality sensors. Panel discussion topics included: (1) best practices for deployment and calibration of low-cost sensor systems, (2) data standardization efforts and database design, (3) advances in sensor calibration, data management, and data analysis and visualization, and (4) lessons learned from research/community partnerships to encourage purposeful use of sensors and create change/action. Panel discussions summarized knowledge advances and project successes while also highlighting the questions, unresolved issues, and technological limitations that still remain within the low-cost air quality sensor arena.

1. INTRODUCTION

In the United States, air quality has traditionally been measured according to a metric established by the United States Environmental Protection Agency (USEPA) using equipment that implement a federal reference method (FRM) or federal equivalent method (FEM). These devices cost tens of thousands of dollars and require significant infrastructure and trained personnel to operate. Within the last ten years, miniaturization and other technological advances have brought to market a number of low-cost (<$2500) sensors designed to measure atmospheric particles and gases. Although sensors cannot replace traditional FRM/FEM monitors, these sensors have created new opportunities for broadening access to ambient air quality monitoring for applications such as personal health and sub-

regional air quality assessment (Snyder et al., 2013; Hagler et al., 2014). Residents in Environmental Justice communities are particularly interested in using sensor technology to gather neighborhood-level data to illustrate the impact of specific emissions sources and magnitude of air quality issues affecting their communities.

Efforts relating to the application of this new generation of low-cost sensors have taken several forms. While much has been invested in the development of low-cost ambient air quality sensors, research to identify appropriate and purposeful use of sensors and networks and develop data analysis and visualization tools to process and interpret collected data is ongoing. In order to increase sensing accuracy, studies have sought to characterize how low-cost sensors respond under changing environmental conditions.

Accuracy and reliability have become unifying concerns. Sensor system development and implementation have been hampered by the complex nature of low-cost gas and particle sensor responses. Some sensors have exhibited vexing manufacturing variations and sensor sensitivity to environmental factors like temperature, humidity, and barometric pressure has proven difficult to model when the conditions range widely. The response of low-cost sensors has been shown to change as they age, in connection with how long they have been in operation. Issues with sensor drift have necessitated frequent recalibration, reducing their cost advantage. In addition, low-cost sensors have also proven sensitive to enclosure air exchange rates, complicating mobile deployments, which are a popular application of low-cost sensors because of their small size and low power requirements (Arfire et al., 2016). Sensors can be slow to respond to changes in pollutant levels, causing pollutant spikes encountered in mobile deployments to be underestimated. Finally, air quality sensors are often marketed for one pollutant, but exhibit cross-sensitivity to other pollutants. Ongoing research is investigating a number of sensors to identify cross-sensitivities and working towards resolution and quantification of individual gas species using advanced calibration techniques and comparison to FRM/FEM measurements (McKercher et al., 2017; Lewis et al., 2016; Piedrahita et al., 2104; Cross et al., 2017; Borrego et al., 2016; Jiao et al., 2016; Spinelle et al., 2105a; Spinelle et al., 2017a). Keeping abreast of this technology sector is difficult at best given that new sensors or versions are coming to market seemingly every day and our ability to evaluate the technology requires painstaking research.

All of these efforts have resulted in valuable lessons and the MetaSense research team (Griswold et al., 2015) convened this workshop with the leading practitioners in the field of low-cost air sensing to share insights and discuss open problems. The community of practice included academic researchers, industry professionals, non-profit groups, community-based organization, and regulatory agencies. The workshop was organized into four panels guided by the

following topics: (1) best practices for deployment and calibration of low-cost sensor systems, (2) data standardization efforts and database design, (3) advances in sensor calibration, data management, and data analysis and visualization, and (4) lessons learned from research/community partnerships to encourage purposeful use of sensors and create change/action. The following summarizes the workshop's discussions among panel experts and the community of practice and shares insights from two community-based organizations gathering air quality data in an effort to improve their communities.

2. PANEL 1, 'BEST PRACTICES' FOR DEPLOYMENT AND CALIBRATION OF LOW-COST SENSOR SYSTEMS

Significant effort has been devoted to exploring the use of low-cost air quality sensors. The low-cost and small size of these devices make them very attractive for increasing the spatial coverage of existing networks, deploying sensors in urban locations where small footprints are desirable, and mobile or semi-mobile sampling schemes. During Panel 1, panelists Andrea Polidori (South Coast Air Quality Management District, Diamond Bar, CA, USA), Ron Cohen (University of California, Berkeley, CA, USA), Jonathan Thornburg (Research Triangle Institute, Durham, NC, USA), and Angelo Bianchi (AQMesh, Stratford-upon-Avon, UK) discussed the exploratory research done to understand the appropriate use of sensors, evaluation efforts being conducted by a variety of institutions, current sensor performance expectations, the state of calibration research, and recent deployment efforts. The overall conclusion was that this field is in transition and more work and funding is needed to continue to realize the full potential of low-cost air quality sensors.

2.1. Categorization and Use of Sensor Systems

There appears to be a common misconception that low-cost air quality sensors are capable of measurements comparable to FRM/FEM measurements. Experience proves this is not the case with some sensors showing no correlation to FRM/FEM measurements, while others show reasonable correlations (R^2~0.7) (US EPA, 2017b; SCAQMD, 2015). The USEPA is not currently entertaining sensor applications for FEM consideration. In addition to poor correlations, sensors often have greater uncertainty. Thus, many sensors fall into an undefined space somewhere between qualitative educational and regulatory compliance measurements and future work may further define this space.

Although low-cost air quality sensors, in their current state, cannot be used for regulatory or compliance purposes, there are a number of appropriate and useful application for these low-cost tools (Hall et al., 2014; Williams et al., 2014). Sensors may not be able to report a sufficiently precise or accurate pollutant concentration to replicate FRM/FEM measurements but some correlate fair to reasonably well (R^2 = 0.4–0.8) (US EPA, 2017b; SCAQMD, 2015). Those that correlate can be used to supplement existing monitoring networks to increase spatial coverage and fill knowledge gaps. They can be used to measure smaller scale variations in spatial concentration or determine how a suspected source may be affecting a nearby community, frequent concerns of community-based organizations. Education and developing air quality awareness are natural applications of sensor technologies and provide a means by which citizens and students might learn about air quality issues, sources impacting air pollution, and variations in air quality in various environments such as work, home, and outdoors. Personal exposure monitoring is another emerging and exciting application for sensors, especially for individuals who are more sensitive to air pollution. Personal exposure monitoring may help an individual make decisions about the timing and location of daily activities like commuting and exercise, based on air quality data. Whether or not a particular sensor is being used appropriately is application dependent—the sensor model, the calibration/deployment procedures, and even data processing and interpretation should all be motivated by the research question.

2.2. Sensor Evaluation

Choosing an appropriate sensor is an important first step in any data collection effort. To date, three notable programs have been established to characterize the performance of low-cost air quality sensors and to make the results of such evaluations available to all potential users. In each case, evaluations are done objectively and the evaluating programs often outright purchase the sensors to ensure integrity. Sensors are often evaluated under real-world outdoor field conditions where sensors are placed alongside the traditional FRM/FEM equipment to which their data are compared. Laboratory testing is also employed for evaluations and involves exposing sensors to known pollutant concentrations within an environmental chamber. Evaluations may include variable temperature and relative humidity conditions as well as introduction of known or possible interfering pollutants. These results should be consulted in the search for an appropriate sensor.

The Air Quality Sensor Performance Evaluation Center (AQ-SPEC), operated by the South Coast Air Quality Management District (SCAQMD), was established in 2014 to evaluate the accuracy and usability of commercially available low-cost air quality sensors. At present, evaluations focus primarily on turn-key products that are ready for

immediate deployment and/or operation. All sensors evaluated in this program are operated outside in southern California field conditions, and reasonably performing sensors are also tested under controlled laboratory conditions with varied temperature and relative humidity conditions. Sensor evaluation reports and details about testing protocols are available on the AQ-SPEC website located at www.aqmd.gov/aq-spec (SCAQMD, 2015).

The US EPA Office of Research and Development (ORD) also conducts evaluations of low-cost air quality sensors and the evaluations results are one of the many resources made available through the USEPA Air Sensors Toolbox located at www.epa.gov/air-sensor-toolbox (US EPA, 2017a). The US EPA has undertaken a number of sensor evaluation efforts under outdoor field and controlled laboratory conditions. More field evaluations are undertaken than lab evaluations and most have been conducted at the Ambient Air Innovation Research Site (AIRS) test platform at the Research Triangle Park location in North Carolina. In addition to turn-key products, US EPA has also evaluated some component based sensors and have incorporated such sensors into a number of devices including the Village Green Stations, AirMappers, and several versions of the Citizen Science Air Monitors (CSAMs). These devices, and a small number of turn-key products, have been operated and evaluated over the course of several small to mid-sized field deployments connected to US EPA projects all over the country (US EPA, 2017a).

The Joint Research Center (JRC), as the European Commission's Science and Knowledge Service, has also conducted research evaluating low-cost air quality sensors via testing under controlled laboratory conditions in a state-of-the-art chamber and outdoor field deployments. Laboratory evaluations have focused on component based sensors, calibrations, and long-term experiments to give insight into long-term performance and drift. Field deployments have investigated normalization techniques too. Information about the chamber, testing protocols, sensor evaluations, and field deployments are mainly found through reports and in the scientific literature (Spinelle et al., 2015a; Spinelle et al., 2013; Spinelle et al., 2015b). Additionally, a number of academic researchers have published papers sharing lab evaluations (Wang et al., 2015) or field performance data on a limited selection of sensors (Austin et al., 2015; Duvall et al., 2016; Lin et al., 2015; Manikonda et al., 2016; Zikova et al., 2017).

Sensor performance evaluations have been extremely helpful for practitioners. However, under the current paradigm, the evaluating institutions pay for the sensors they evaluate, which is likely to be unsustainable in the long run with a rapidly changing marketplace and increasing costs as sensor systems get more complex. Debates about the path forward are ongoing and options include third party evaluation, sensor certification, or a program which would encourage manufactures to self-evaluate. Although these discussions are happening now, it is important to note that

such programs come with significant investments of time (e.g., developing test methods for each pollutant, getting manufactures on board) and money (e.g., start-up, program maintenance) and are likely to take many years to develop.

2.3. Current State of Sensor Performance

The sensor evaluation efforts described have provided insight into the current state of sensor performance and have elucidated areas where further research and development is needed. Evaluations show that currently available particulate matter (PM) sensors exhibit reasonable performance (select sensors approaching $0.7 < R^2 < 0.9$) (SCAQMD, 2015). Evaluations found that most PM sensors have minimal downtime, moderate inter-sensor variability, and show reasonable correlation with FRM/FEM measurements, although calibration and normalization is still needed. Many show temperature and humidity effects, especially at high humidity, and under-report at very high (>200 g/m3) concentrations (Crilley et al., 2017; Williams et al., 2014). Additionally, most sensors cannot detect very small particles (lower cutoffs between 0.3 and 1 m) and will miss ultrafine particles and smoke (Jovasevic-Stoianovic et al., 2015). Some work is ongoing to develop sensors capable of measuring particles in this small size range. Most of the current sensors detect particle counts rather than particle mass and must use an algorithm to report PM mass concentrations (Jovasevic-Stoianovic et al., 2015).

Evaluations show that gas-phase sensors exhibit acceptable data recovery but have more inter-sensor variability than PM sensors. When carbon monoxide (CO), nitrogen dioxide (NO_2), and ozone (O_3) are measured alone in a laboratory setting under controlled conditions without confounding gas species present, sensors exhibit reasonable to good correlation with FRM/FEM measurements ($0.8 < R^2 < 0.99$) (SCAQMD, 2015; Castell et al., 2017; US EPA, 2017b). Sensors that are cross-sensitive to multiple pollutants show low correlations with FRM/FEM measurements when operated in the field environment where a mixture of pollutants is present ($0.3 < R^2 < 0.9$) (SCAQMD, 2015; US EPA, 2017b). Temperature and relative humidity (RH) have a larger effect on gas phase sensors leading to decreased sensitivity in high RH conditions and degradation over time (Lewis et al., 2016; Masson et al., 2015a). Therefore, repeated field calibration of gas phase sensors is needed and is further discussed in the next section. To date, ambient concentrations of sulfur dioxide (SO_2), hydrogen sulfide (H_2S), methane (CH_4), and volatile organic compounds (VOCs) prove extremely difficult to quantify despite sensors that advertise sensitivity to these species.

2.4. Sensor Calibration

Sensor performance evaluations have indicated a need to calibrate sensor response if one wishes to compare one sensor's data to that of another sensor or to nearby regulatory monitoring data. Field normalization of sensor signals that have been collocated with FRM/FEM measurements is the most common method of calibrating sensor measurements. Linear regression is commonly used to normalize sensor signals to reference measurements, but there is no evidence that these correlations are transferrable to different locations. Environmental factors such as temperature, relative humidity, relative concentration of confounding pollutants, and particle sources and variation in particle size are all known to affect sensor response, so it is not surprising that these variables also change how sensor measurements compare with the reference.

Researchers are exploring different methods of calibrating sensors against reference measurements motivated by the known presence of complex nonlinear and cross-sensitive behavior of sensors. Field normalization techniques that attempt to address these complex behaviors include multi-linear regressions, non-linear multi-variate models, and machine learning (Cross et al., 2017; Spinelle et al., 2015a; Spinelle et al., 2017a; Masson et al., 2015a & 2015b; Esposito et al., 2016). These methods may make calibrations more transferrable between regions because they consider many of the factors known to influence sensor performance, though model extrapolation is a concern. Therefore, it will be important to calibrate over a wide range of environmental variables and pollutant concentrations. Future experiments may investigate if such calibrations can be performed in a laboratory setting (Cross et al., 2017). Methods for dealing with sensor aging, which can cause a range of issues from drift to sensor failure, are still largely underdeveloped.

Beyond field normalization to reference measurements, researchers have used their understanding of atmospheric chemistry to add another level of validation to the data produced by sensors. There is great potential for researchers to develop rules based on atmospheric chemistry/physics to 'check' sensor data and to share these resources with users. For instance, Ron Cohen shared that his group has been using VOC + NO_2 ozone formation chemistry to check some of their sensor data. Briefly, O_3 concentrations should fall to zero at night, if there is any NO present so nighttime sensor readings may point to a bias within the sensor measurements and monitoring changes in this minimum concentration may help detect sensor drift or more dramatic shifts in sensor performance.

2.5. Sensor Network Deployment

Numerous sensor deployments have been conducted in recent years ranging from residents investigating air quality in their homes or neighborhoods, to small networks looking at community-level concentrations, to large sensor networks covering cities or regions (Jiao et al., 2016; Lin et al., 2015; Shusterman et al., 2016; Moltchanov et al., 2015; Sun et al., 2015; Tsujita et al., 2005; Gao et al., 2105; Kaufman et al., 2017). Increasing network size leads to increasing complexity and exponentially increasing costs and effort for data analysis and visualization.

When deploying sensors for data collection purposes, there are a number of factors to consider. The EPA's Air Sensor Guidebook (Williams et al., 2014) may be a helpful resource to those designing a data collection effort using sensors. The following represent key considerations and 'best practice' recommendations.

- The research question needs to drive proper sensor selection. Consult evaluation reports during the sensor selection process to better understand how a sensor might be expected to perform given the environment and expected pollutant concentrations.
- Calibration is key to any successful deployment. Although researchers are still investigating a number of ways to calibrate sensors (Section 3.4), collocation of all sensors with nearby FRM/FEM is still an essential best practice. At a minimum, this should be done before any sensor deployment or field study. Repeating the procedure after a deployment will help quantify sensor drift and help bound uncertainty. Long-term deployments often rotate sensors through several short collocation periods to continually monitor for drift and sensor failure.
- Sensor failure and replacement is a concern especially for successful long-term deployment. Evaluation efforts have noted significant variation and failure of new sensors in the low-cost price range. Early deployments noted pre-mature failures and indications of short sensor lifetimes with declining performance within the 1st year of use. Users should purchase additional sensors to complete the deployment plan and continually monitor sensors for failure and declining performance.
- The research question and pollutant of interest should govern the size and siting of the sensor network. For instance, if the question involves how air pollutant concentrations vary in the outdoor and indoor environment, a small number of sensors may be needed and siting criteria would include considerations like weather, ventilation, sources, and obstructions. If however, one would like to reliably monitor concentrations

over a large area, sensor siting is still important but so is sensor redundancy, pollutant variation, and sensor density within the network.

Data collection and management is a not a trivial matter, especially as the size of a sensor network deployment grows and the data is collected more frequently. Panels 2 and 3 of the workshop were convened to discuss data issues and those discussions are detailed in the following Sections 3 and 4.

3. PANEL 2, DATA STANDARDIZATION EFFORTS & DATABASE DESIGN

The low-cost sensor revolution has been making air quality sensors affordable and available to large populations and community-based organizations. Users come from a variety of backgrounds and have varied objectives. The number of deployed sensors appears to be increasing over time. Currently, data from low-cost air quality sensors comes in a variety of formats sometimes without data labels, units, or metadata to easily understand and process the available information or to compare one dataset to another. Panelists Abhijit RS (Environmental Defense Fund, San Francisco, CA, USA), Andrea Clements (U.S. Environmental Protection Agency, Research Triangle Park, NC, USA) and Michael Hannigan (University of Colorado, Boulder, CO, USA) discussed the need for a harmonized approach to data management. The group discussed the value in developing and adopting data standards.

3.1. Data Standardization

A variety of low-cost air quality sensors and sensor systems are presently available. These sensor systems may measure one or more pollutants and/or environmental parameters, employing one of a variety of measurement techniques. Some sensor systems include onboard algorithms to transform raw data signals into pollutant concentrations. Each uses its own data structure to capture, store, and publish the data.

In order to efficiently store and process large volumes of data sourced from disparate sources, all the incoming data should be representable in a uniform and common structure and format. In practice today, this requires data transformation in order to integrate data from various sources. The idea of data standards plays a very important role in developing a large-scale data management system. Receiving data from sensors in standard data formats would save a lot of time and effort for everyone involved. If sensor system developers adopt data standards, in terms of both data formats and data quality (for example by reporting confidence intervals along with pollutant concentrations), deploying new sensors in the field could become easier by reducing the technical burden on the user and expanding the utility of the measurements they record.

Data format standardization in the air quality domain includes date and timestamp formats, standardized definitions of terms including pollutant names, units of measurement for pollutant concentrations and their interfering factors like meteorological parameters, and a minimum set of data elements to be recorded by the sensors and stored by the backend data system. It also includes data transfer protocols and file formats used for data exchange.

Sensor data currently exists in various formats—comma delimited (CSV) files, XML and JSON formats, database tables, PDF files, etc. Some of these files have headers indicating what is contained in each field or column and some don't. The date and timestamp in these files may or may not have a time zone designator; they may not take daylight saving changes (DST) into account; some may represent timestamps in UTC while others will report in local time. Sometimes, a date and timestamp is not reported at all. Additionally, the units of measurement vary among these datasets; some sensors report particulate matter (PM) concentration in mass (e.g., g/m3) and others in particle count. The data elements (fields or columns) contained in these datasets vary widely; some files have raw sensor signal or pollutant concentration measurements only while others include statistical summaries like mean and median alongside sensor measurements. Other issues connected to data quality include field duplication, data duplication, unexpected insertion of text character strings, data gaps, and irregular data reporting. Uniform procedures for addressing all of these challenges in every dataset would make it easier to integrate the data in order to perform analysis across data sets.

3.2. Air Quality Data Platform

Low-cost sensors can provide data with very high spatial and temporal resolution, which is not easily achieved with conventional instruments. Researchers and academics have been collecting air quality data for decades; in recent years, community organizations have been deploying sensor networks in their neighborhoods to monitor their local air quality and citizen scientists have been using sensors to learn about air quality in their immediate surroundings. However, most of this data is only available to the people who collected the data, and generally not available to a larger audience. These siloed data stores limit the extent to which data analytics can be performed on air quality data. Combining all these datasets and providing a framework amenable to sophisticated analysis would facilitate better understanding of air quality in many places and on many scales. This gained information could help influence behavioral changes that result in improved environmental protection and human health.

There is a need to develop a schema that facilitates air quality data aggregation and sharing. Such a schema could consist of a scalable cloud-based infrastructure, which could provide the capabilities for users to run their

computations and analyses instead of downloading data to their local systems for processing. A centralized system could catalyze development of software tools to analyze and visualize the data and make them available to all the users. Air quality researchers and sensor developers could look at wide varieties of pollutant concentration data in concert with factors like meteorology, land use, traffic, and emission sources that affect air quality. Community organizations and citizen scientists would be able to compare various neighborhoods and develop science driven policy recommendations founded on data. The data platform would be in a position to connect with other systems that host data relevant for air quality analyses like health informatics, real estate market, urban planning, emission inventories, and water quality; thereby, expanding the scope of use.

3.3. Bridging the Data Gap

The Environmental Defense Fund has convened the Air Sensor Workgroup (ASW), a broad-based group with participants from state and federal government, academic institutions, sensor developers, and other organizations and stakeholders interested in making air quality data open and Findable, Accessible, Interoperable, and Reusable (FAIR). The main objective of the ASW is to enable easy and efficient access to large volumes of air quality data for the common good. To achieve their vision, they developed Date and Timestamp Guidelines and are working on other relevant data standards. They are also developing a data platform to host and publish data collected from low- and medium-cost air quality sensors globally. The ASW does not have any commercial interests and the software and tools developed by them will be released as open source software and will be publicly available at no cost to the users. The ASW encourages users to leverage this data platform to make advances in auto-calibration of sensors and support scaling the sensor deployments in addition to other potential uses. More information about the ASW is available at www.edf.org/asw (EDF, 2018).

3.4. Future Needs and Directions

The air quality community needs to move away from qualifying the data as good or bad, and toward characterizing the exact qualities of the sensor data, including confidence in pollutant concentrations. Air quality measurement data should be supplemented with metadata. The data platform will need to be flexible with limited optionality to keep it simple for users. Some basic data quality validations could be performed by the data platform, but it will be up to the end users to determine whether the quality of data is good enough for their particular use.

The sensor calibration details are currently not published widely. This makes researchers wary of the reported measurements. That leads to additional testing and potential recalibration by advanced users. Additionally, air sensors may behave differently under lab conditions and field conditions, which may need to be taken into consideration while calibrating. Hence, providing more information about the out-of-the-box calibration will not only expedite the use of sensors but also create opportunities for improving the calibration methods and scalability of deployment.

One of the important questions is how an open-access data platform might impact local communities and environmental justice issues. Such a data platform may be used to develop products and services, and monetize them. The data by itself may not be monetized but the tools to process and visualize the data could be; the results of data analytics and the corresponding findings could find monetary value as well. While this may not financially benefit the data owners who contributed data to the data platform, a concern of some groups, it certainly helps to advance science and there are potential indirect benefits that data contributors might reap over time. An open-access data platform would allow researchers to perform analyses and then share the results with other platform users. Community groups could use those case studies to guide local action. Community groups may also be able to post their data and solicit assistance with analyses to develop actionable insights. Health scientists and other may be able to combine personal air quality exposures to health outcomes. Eventually, air quality data could be as widely available and interpretable as traffic or meteorological data. An open-access data platform might also lead to a variety of analyses and interpretations; some of them could seem contradictory. This may open up channels for further communication among the researchers and analysts, and might help in advancing science. Negative impacts have yet to be defined.

3.5. Summary

Sensor and sensor system developers and users conforming to data standards could facilitate the aggregation of data, making it possible to create a larger, richer dataset which could lead to the discovery of new insights. A common data platform could open up opportunities to integrate data from global sources leading to development of data products and applications that can help users understand air quality at the neighborhood scale. Given this vision, establishing data standards and complying with them is critical to harnessing value from the data measured by low-cost sensors.

4. PANEL 3, ADVANCES IN SENSOR CALIBRATION, DATA MANAGEMENT, AND DATA ANALYSIS AND VISUALIZATION

During Panel 3, panelists Michael Heimbinder (HabitatMap, Brooklyn, NY, USA), Sanjoy Dasgupta (University of California, San Diego, CA, USA), Nicholas Masson (Qsense Inc., Boulder, CO, USA), and Mark Potosnak (DePaul University, Chicago, IL, USA) posed several questions to guide the discussion regarding sensor calibration and data management, analysis, and visualization. The subsequent discussion focused on five key issues outlined and summarized in this Section. Despite outstanding data calibration, quality, and validation issues, participants agreed that there is great value in the data collected by sensors but that this data must be used wisely and with caution. Users were encouraged to collect supplemental data (e.g., metrological data, co-pollutants concentrations, traffic and other observational data) that might help in subsequent data interpretation efforts. Researchers were also encouraged to be open and honest in setting expectations and in explaining the appropriate use and current limitations of sensor technology. Repeatedly, community organizers mentioned the need for effective infographics and data visualization tools to help share data, interpret the results, and educate the public.

4.1. Data Quality

As discussed in Section 2, data from low-cost sensors are not equivalent to data from FRMs/FEMs, but rather than thinking of sensor data as "good data" if it compares well to FRMs/FEMs, it might be better to consider if data is "good enough" for the intended objective (Williams et al., 2014). For instance, to monitor spatial variation, the paramount consideration is that sensor measurements are comparable to one another. Sometimes, another factor (like how the body responds to a pollutant concentration) may have more uncertainty than the concentration measurements allowing for more flexibility in the sensor uncertainty. Thus, the necessary quality of the data should be considered in the study design process.

Quantification of uncertainty or confidence interval is essential for understanding and using sensor data. Generally, uncertainty is defined through collocation with FRM/FEM instruments, but statistical modeling may help determine the appropriate confidence intervals. The interval should fully capture the uncertainty in the data and the width of this interval can help determine the usefulness of the data. Researchers should be sure to consider whether the measurement uncertainty is driven by the sampling environment, systemic biases, or random error.

Network deployments may alter the data quality questions. Looking at the sensor data in aggregate may render smaller errors unimportant. Environmental factors can significantly influence sensor performance and are likely

to remain important. Information about traffic and expected sources may also be helpful in interpreting data. When considering data in aggregate, researchers can look for similar behavioral patterns among a number of sensors to verify changes and may be able to identify or confirm pollution sources. Although this approach may be helpful, care should also be taken not to exclude interesting data. Data that may seem to be outliers may actually be a signal deserving of future investigation.

4.2. Supplemental Data Collection

As previously mentioned, environmental factors such as temperature, relative humidity (RH), and the concentration of co-responsive pollutants are all known to affect sensor response. At a minimum, it is important that any data collection effort measure these essential variables. This realization has led to a rise in the development multi-sensor instruments (boxes, pods, systems, etc.). Many of the commercial instruments on the market today are attempting to measure all of these parameters and leverage the instruments to make as many measurements as possible. In many cases, the increased complexity of these instruments takes them from the low-cost sensor realm into a more expensive price range ($2000–$15,000), which also makes them more difficult for communities and citizen scientists to afford, especially if a large distributed sensor network deployment is needed to address research questions.

It is worth noting that temperature and RH measurements also have caveats. Many metal oxide and electrochemical sensors respond to temperature and RH, so measurements of these environmental variables in the air mass directly adjacent to the sensors (within the sensor enclosure if one is used) is very important. Some enclosures are not designed to dissipate heat and temperatures in their interior can differ greatly from the outdoor environment. However, the ambient temperature and RH also influence the chemistry that can affect levels of atmospheric pollutants. Thus, both measurements are important and care should be taken in designing sensor enclosures to minimize the difference between enclosure and ambient measurements.

As the community begins to consider aggregating sensor measurements to make them useful beyond the initial intended use, additional supplemental data may also prove important in interpreting the results. It is difficult to know what might be important when starting a small-scale study and the needs will vary depending on both the level and scope of analysis undertaken. For instance, information about the sensor (make, model, serial number, purchase data, time in service, etc.), position (GPS coordinates), results of collocation efforts, and the calibration equation used are essential. Environmental factors such as temperature, RH, pressure, dew point, wind speed, wind direction, and solar radiation may all assist in understanding sensor response and interpreting variations within the data. Urban or

near-roadway data interpretation may benefit from noise/sound data, traffic count, traffic pattern, and vehicle fleet information. Source inventories and source locations may be especially important in interpreting data near sources with episodic and transient behavior. Satellite data may help elucidate the influence of regional sources like wildfire or dust. Unfortunately, it is often impossible to collect all of this information during the course of a data collection effort, often due to cost, but some of this information may be available from other nearby sources (e.g., a weather station, local government website) or a previous study may have many of the same measurements that could give ballpark estimates. Often, researchers involved in a data collection effort will know the best sources for supplementary information and listing them in metadata for future reference would be helpful.

4.3. Working with Communities

It is important to understand that many communities are seeking assistance in further understanding their lived experience. They often look for scientific research partners to guide them in collecting and interpreting data. More discussion about these types of partnership and needs are discussed in Section 5.

Communities of all types and scales (neighborhoods to cities to states) are interested in collecting air quality data using sensors but may not be prepared to handle data calibration issues or the vast amount of data that comes with a large-scale deployment. Some participants attended this workshop just to learn which sensors would be most widely recommended and free of errors or issues. Several received the idea of a data repository with enthusiasm, partly because the burden of creating and hosting a database could be lifted.

4.4. Data Interpretation Needs

Researchers working with low-cost air quality sensors are generally aware of the quality and uncertainty associated with their sensor measurements. Many other users, especially more casual users, may need more assistance in understanding the limitations of the technology and interpretation of the data. Repeatedly, practitioners mentioned the need for effective infographics to help share data, interpret the results, and educate the public.

Experts in epidemiological research, including participants Rima Habre (University of Southern California, Los Angeles, CA, USA) and Michael Jerrett (University of California, Los Angeles, CA, USA), noted that low-cost air quality sensors are changing the type of exposure data available often pushing toward a goal of measuring personal exposures. Highly time-resolved data (seconds to minutes) creates opportunities for new research in deciphering the impact of acute exposures to various pollutants. Numerous researchers are exploring the development of apps and

websites aimed at helping people explore and interpret their personal exposures. Given the current state of sensor science, with relatively large measurement uncertainties, there is a concern from practitioners about encouraging citizens to change behavior based on sensor measurements. On the one hand, users observing spikes may be prompted to change their behavior resulting in reduced exposure. On the other hand, the reduced exposure may not result in an observable health outcome and users may be less likely to continue with their behavior changes as a result. Moreover, spikes in air quality data may result in users experiencing increased stress, possibly negating any other health benefit.

4.5. Sensors and Modeling

One of the motivations for measurements with low-cost sensors is to increase the spatial resolution of our atmospheric measurements to identify variation below the city or regional level, even down to the city block-level or below (Apte at el., 2017). Current modeling techniques struggle at this level due to the high dynamic variability of pollution sources, wind, obstructions, etc. (Berrocal et al., 2010). Because of the greater uncertainty associated with the low-cost sensor measurements, much work is still needed to determine if low-cost sensing can improve model performance and better describe personal exposure. There are some on-going efforts using machine learning to inform models with low-cost sensor measurements (Verma et al., 2011; Nikzad et al., 2012).

5. PANEL 4 AND A COMMUNITY PANEL, LESSONS LEARNED FROM RESEARCH/COMMUNITY PARTNERSHIPS

Communities are demanding a greater role in scientific research and decision-making that impacts their lives. Across the US and globally, residents continue to recognize that pollution sources impact their neighborhoods and exposure to pollutants may be causing health hazards for them based on where they live, work, and play. Further, communities are increasingly seeking tools to document these exposures and environmental health disparities. Currently, regulatory air monitoring systems generally do not assess neighborhood variability in air quality at a sufficiently refined spatial scale (Basu et al., 2004; Jerrett et al., 2005). The increase in the availability of low-cost air pollution sensors has increased the number of citizen scientists collecting and using air quality data to better characterize and understand their local environment. Education and involvement of communities in science and research is not only important for improving public health; it is also important for building awareness about the sources of air pollution, exposure pathways, and the association between contaminants and health endpoints (Wing, 2005). During Panel 4, panelist Jill Johnston (University of Southern California, Los Angeles, CA, USA), Nicole Wong

(Redeemer Community Partnership, Los Angeles, CA, USA), Vanessa Galaviz (State of California, OEHHA, Oakland, CA, USA), Ashley Collier (University of Colorado, Boulder, CO, USA), and Andrea Clements (U.S. Environmental Protection Agency, Research Triangle Park, NC, USA) outlined key aspects to forging a successful partnership with communities that are seeking information about ambient air quality and offered recommendations for how low-cost sensors can be deployed for research by citizen scientists. In addition, community leaders Sandy Navarro (People not Pozos, Los Angeles, CA, USA) and Luis Olmedo (Comite Civico Del Valle, Inc., Brawley, CA, USA) shared specific examples from their experiences as community leaders addressing air quality concerns (see Section 5.3 for more details about these community projects).

5.1. Building Scientific Literacy

Low-cost air quality sensors offer new opportunities to gather data about local air quality in an individual home, during a bike ride, or in various neighborhood parks simultaneously. Devices that measure real-time pollution and provide immediate feedback have the opportunity to serve as tools to build the capacity of residents to understand air pollution, spatial and temporal variability, and exposure patterns relevant to their community. Through this process, residents can learn about scientific methods, the ability to interpret data within a given context, and the potential links between air quality and health outcomes (Trumbull et al., 2000). Residents offer expertise to identify potential sources of air pollutants otherwise unknown to scientists or regulators as a result of their lived experiences and knowledge of their neighborhood. Collaboration and bidirectional dialogue is important to characterize the question, evaluate whether available low-cost sensors are appropriate for addressing that question, and design a method for collecting the data. Community members also contribute observational data or qualitative information to add context to recorded pollutant concentrations. It is key, however, that all parties understand both the advantages and limitations of low-cost sensing.

5.2. Leveraging Low-Cost Sensors

Low-cost sensor technology can be leveraged to advance the co-production of knowledge. During the initial phases of study design, a variety of expertise should be considered, such as that of community members, scientists, regulatory partners, and even representatives from potential sources of concern. Academic or regulatory partners can support communities to ensure appropriate sensor technologies are chosen considering: (1) the pollutant or source of interest, (2) spatial and temporal scale of interest, and (3) the "ease" of interpretability of the data. Collaborators should

discuss the design of defensible calibration techniques and collocation with regulatory monitors. Community members offer vital knowledge about important factors that may need to be considered when planning how to conduct observations of the problem (source, frequency, intensity, etc.) (Main et al., 2012), and possess vital community contacts to help with community engagement.

In order to build trust between community members, scientists, and regulatory bodies, panelists and attendees made several best-practice recommendations:

- Discuss, during the study design phase, responsibilities and expected outcomes with all key partners.
- Explain the capabilities of sensor measurements at the time of a partnership and educate all parties on the current challenges that remain for the field.
- Clarify the expectations of what sensor data can and cannot help elucidate, how the sensor data compares to "gold standard" FRM/FEM instruments, and what will happen with the data during and after the study. Such agreements should all be clearly outlined and accepted by partners.
- Outline the limitations of such data for use by regulatory agencies.
- Prepare residents for various potential outcomes based on their questions, such as negative or no results.
- Establish agreements regarding data sharing and ownership, communication of results, and publication during the study design phase.

Collaborative teams may find it valuable to include social scientists or to look to other disciplines for examples of useful formats for sharing data/results (Main et al., 2012), effective ways of communicating risk (Bickerstaff et al., 2004), or the principles of community-based participatory research (CBPR) established in public health research (Israel et al., 2012).

5.3. Opportunities for Community-Driven Science

Community-research partnerships can prompt action to prevent harmful exposures or improve local air quality. Innovators continue to advance low-cost sensor technology, but even with the existing limitations, sensor systems on the market now can still provide insight for communities aiming to gather data about ambient air exposures. For example, while exact concentration measurements may be fairly uncertain, relative difference within or between communities or before and after an event (e.g., engine changeover or reactivation of industrial source) may still be valuable depending on the questions and goals of a particular community. Similarly, sensors may be able to give

insight into spatial/temporal patterns as well as determine “hotspots” for future targeted studies with more sophisticated instrumentation.

During the workshop, two community leaders shared their thoughts on applications for and experiences with low-cost sensors. Sandy Navarro from People Not Pozos, a grassroots program that is part of Esperanza Community Housing based in South Los Angeles, described using sensors as tools to better identify local air exposures. Since 2010, local residents have complained of noxious odors and health symptoms (e.g., respiratory illness, fatigue, headaches, nausea, eye & throat irritation, dizziness, and spontaneous nosebleeds) (Sadd & Shamasunder, 2015). Many of those residents identified a nearby oil drilling site, situated across the street from one of Esperanza’s low-income housing buildings, as a source of odors and air pollution. After three years of official complaints and protests by this environmental justice community, investigators from the USEPA visited the site and discovered violations resulting in a shutdown of operations (Sahagun, 2014). People Not Pozos organized in response to this issue; including collaboration with researchers in order to collect environment and health data and training community residents to engage as researchers on the project. Nonetheless, community frustrations persist. The community has not received response to official complaints, sufficient data or access to collected data, nor easily understandable information in an accessible way (e.g., Spanish translations). Thus, an on-going project, in collaboration with another group of researchers, has deployed a small network of low-cost sensors to characterize neighborhood-scale air quality.

The subsequent discussion highlighted a common problem in public health and environmental justice: that need is often greater than the capacity. Particularly in a city like LA, given its size and density, public health officials are likely to be limited by available resources (including both equipment and time). A researcher from the local regulatory agency expressed precisely this sentiment, indicating that the agency is piloting new technologies to try and help expand their capacity to investigate community concerns and complaints. Among these new technologies are low-cost sensor systems, which may be able to serve as a sort of alarm. Additionally, he noted that communication is an issue and it’s likely the residents’ complaints and requests for data were probably not reaching the appropriate person. This example highlights an area where local regulatory agencies can assess the effectiveness of or improve their education and outreach efforts to the communities they serve.

In the Imperial Valley, Luis Olmedo with the nonprofit Comite Civico Del Valle shared another story describing the Identifying Violations Affecting Neighborhoods (IVAN) monitoring system. This system includes a platform for submitting and viewing environmental reports, as well as real-time data from a network of 40 air quality

monitors utilizing low-cost particulate matter sensors (Comite Civico del Valle, 2017a). On the website, users can receive air quality alerts that include recommendations for adjusting outdoor physical activity to reduce an individual's exposure. These recommendations are based on a scale that the IVAN team has developed that provides numeric and color indicators based on PM concentrations and potential health impacts (Comite Civico del Valle, 2017b). In the Imperial Valley, this system has been integrated into a school-based flag program that uses colored flags to indicate air quality and provides recommendations on outdoor activity. While recognizing this information is still limited, Luis asserted that this strategy enables individuals to make their own choices about their health and potential exposures. These monitors provide a picture of PM concentrations at a spatial scale and resolution previously unavailable allowing the community to take targeted action. Actions that are especially important given that Imperial County has the highest rates of asthma-related hospitalizations and emergency room visits among school-aged children of all counties in California (CDPH, 2015).

There remains a need to develop best practices for risk communication and visualization of air sensor data for residents. For example, in the case of real-time data, it is important to communicate the difference between short-term high exposures versus 24-h or weekly averages in pollutant levels. The use of real-time and personal monitoring with low-cost sensor provides an opportunity to better assess dose-response relationships to various health outcomes and more specifically study vulnerable and susceptible populations—such as asthmatics or those living in environmental justice neighborhoods. Ultimately, the results from sensor studies have the potential to help communities decide on actions they themselves wish to take to protect health (Minkler, 2010).

5.4. Summary

Overall, community-driven research using sensors is likely to benefit both community and scientists alike. In particular, the communities facing the greatest environmental exposure risks and health effects are demanding a greater role in researching, describing and prescribing solutions to address the local environmental hazards they face (Cole & Foster, 2001; Corburn, 2002). Coupled with technical expertise and air quality sensors, communities can play a central role in defining the problems, supplying local knowledge and interpreting the results in the context of the local reality. With communities' expertise, there are improvements in the relevance of research questions at the scientific level. Community research may also help to build trust and empower participants and community members, especially when the data is community owned and managed, giving them a 'seat at the table' with industry and

regulators. Moreover, using sensors in community science allows for real-world application of the research, allowing for people to make a difference and improve the health and lives of their community members.

6. STAKEHOLDER SMALL GROUP DISCUSSIONS

To conclude the workshop, attendees split into small groups to discuss what could be taken away from the workshop discussions. Attendees were given general guidance to focus on three core topics: existing resources to be shared and new resources that should be developed, important takeaways or best-practices that could be shared more widely, and important next steps for the field. A theme that emerged from these discussions was the need to improve communication between all stakeholders and how communication strategies could address the challenges highlighted in this paper. The variety of stakeholders and accelerating pace of research necessitate a variety of communication strategies to address the breadth of challenges in this field. In general, the group discussions focused on communications between and within two main groups of stakeholders: researchers and the participating community members.

6.1. Creating Dialogue in Community-Based Research

Communication between researchers using low-cost air quality sensors and engaged members of communities in which those sensors are deployed is both challenging and vitally important. A key challenge is creating a dialogue that brings all stakeholders to the table and values each member's knowledge and perspectives. Successfully establishing this dialogue will improve project relevance and data quality while identifying other areas of interest that might otherwise be overlooked.

Establishing realistic expectations between all parties at the onset of community-based research studies is paramount. Sensor limitations must be discussed and the measurements adequately contextualized. The community's concerns, objectives, and insights must be discussed. Some collaborations have found it helpful to establish Memoranda of Understanding (MoU) and/or Frequently Asked Question (FAQ) pages for their projects. All parties involved in the research should collaboratively develop such documents so that they are easy to understand (including language translation when necessary) and adequately capture the expectations and responsibilities of all project partners. Developing standard templates for and promoting the wider use of MoU in community-based research projects could help maintain positive relationships and create engaged communities that are more open to working with researchers.

An important expectation that should not be overlooked is the ownership and control of data collected during a community-based research project. Every community's expectations will be different and researchers should be mindful of the sense of ownership that community members may feel toward data that they were responsible for collecting. Researchers should share data in a manner that respects the wishes of the community in which it was collected.

Appropriately communicating data can be especially challenging given the developmental nature of these instruments. However, it is critical to develop data communication methods that allow community members to transparently access and understand data and uncertainty while providing adequate context. Enhancing access to the data will allow researchers and community partners to collaboratively draw insights from the data. These methods may come in a variety of forms and the communication of data to non-technical persons should also be a consideration when designing infographics and interpretation or visualization tools for community based research. By developing best practices for data visualization and communication, researchers can help communities to better quantify and communicate aspects of their lived reality.

6.2. Collaboration and Standards for Low-Cost Sensor Research

As research in the field of low-cost sensors accelerates, it will become increasingly important for researchers to harmonize terminology and data reporting formats. One effort that was widely discussed at the close of the workshop was the Data Platform initiative being led by the EDF (discussed in Section 3). This project attempts to create a database and schema that would allow users to openly share data collected using low-cost air quality sensors. A lengthy discussion focused on how standardizing a data reporting format based on sensor type could facilitate largescale comparison between research projects and allow researchers to test calibration models on a larger parameter space. The formatting guidelines could include details as simple as the date and time string format to the specific metadata that should be included.

Given interest in large scale comparisons between research projects, it is important that metadata include information about the sensor and its performance and measurement uncertainty. Most importantly, data quality elements such as precision, bias, detection limit, age, and calibration can be imbedded in the metadata. This information can give researchers more information about the measurement uncertainty and help them determine the usefulness of the data for a given application. It could also streamline collaborations between projects using a variety of different instruments and more generally improve communications between stakeholders.

Calibration presents another complicated challenge for communication and comparison. For example, when researchers compare different calibration techniques, it is important to be mindful of the statistical techniques used in these comparisons (e.g., RMSE, R-squared, correlation coefficient). It is likely that the best metric to compare and select a calibration model will depend on the application and goals of the project. The difficulty in fairly evaluating calibration methodologies, including physical methodologies and numerical methods was discussed at length. Physical methodologies, like collocation, have questionable applicability once sensors are moved into a new area to collect data. Once collocation calibration data has been collected, a variety of numerical calibration methods are currently utilized ranging from simple linear fits to complex machine learning and artificial neural networks (Cross et al., 2017; Spinelle et al., 2015a; Spinelle et al., 2017a; Masson et al., 2015a & 2015b; Esposito et al., 2016; Zimmerman et al., 2018).

From the workshop discussions, it was clear that additional research is needed to develop a comprehensive best practice for sensor calibrations. One idea for calibration method comparison could include sharing a collection of data from a variety of sensor technologies in a number of unique environments that include collocated reference data. These datasets could allow researchers to test novel methodologies and to compare their effectiveness against previous methods. This data could be released as sets of "training" datasets including reference data and "validation" sets without reference data to penalize over-tuning of calibration algorithms.

6.3. Going Forward

The low-cost air quality sensing workshop provided an excellent forum for researchers, regulators, manufacturers, and community advocates to discuss a sampling of the challenges and successes in this fast-growing field. Going forward, it will be important to continue to hold workshops like this and to include an even broader group of stakeholders in the discussions. This may include policy makers, electrical engineers, programmers, and others from both the United States and abroad. With such a fast-moving state of technology, it will remain important to collaborate with all parties to ensure that research projects are successful and relevant. Beyond workshops, there was also interest expressed in other methods of staying connected and sharing resources, for example via a listserv or wiki page. Toward that goal, Table 1 shares a list of existing resources that were compiled during the final discussion. This list is not comprehensive but may serve as a starting point. These are valuable resources for those involved in air quality and environmental justice and sharing others not listed here is equally important. Please note that inclusion or omission does not indicate an endorsement or lack thereof of these tools.

Table 1. Online tools that may be useful to low-cost air quality research stakeholders.

Category	Tool	Description	Website
Government	EPA's Air Sensor Toolbox	Resources to support communities and citizens selecting and using low-cost sensors	epa.gov/air-sensor-toolbox
	AirNow	Aggregated national data from regulatory grade monitoring stations and information on potential health impacts	airnow.gov
	EJSCREEN	Environmental justice screening and mapping tool	epa.gov/ejscreen
	RETIGO	EPA software for mapping collected air quality data	epa.gov/retigo
	C-FERST	Community-focused exposure and risk screening tool	epa.gov/c-ferst
	AQ-SPEC	Sensor evaluation program	aqmd.gov/aq-spec
	US EIA Energy Mapping System	Mapping tool to explore energy use and development across the United States	eia.gov/state/maps.php
	CalEnviroscreen	California specific environmental justice screening tool	oehha.ca.gov/calenviroscreen
Non-Profit and Academic	IVAN Monitoring System	Community-based monitoring system for environmental concerns	ivan-imperial.org/air
	FrackTracker	App for logging reports and observations around oil and gas activity	fractracker.org
	Institute for Health and Metrics Evaluation	Resource for exploring health data and statistics	healthdata.org
	EDF Air Sensor Workgroup Data Standards	Guidelines for sensor data formatting	edf.org/health/data-standards-date-and-timestamp-guidelines
	AQ-IQ Air Quality and Sensor Curriculum	K-12 curriculum supporting the use of low-cost sensors by students	teachengineering.org/curricularunits/view/cub_airquality_unit
	OpenAQ	Aggregated global data from regulatory grade monitoring stations	openaq.org
	SensorThingsAPI	Open source API for connecting IoT sensing devices, data, and applications over the web	opengeospatial.org/standards/sensorthings
Commercial	Envirosuite	Software for environmental management and planning	envirosuite.com
	Breezometer	Commercial API for displaying air quality indicators	breezometer.com

7. WORKSHOP CONCLUSIONS

This workshop provided the opportunity to reflect on the current state of low-cost air quality sensor research. The discussions made it clear that better communication within the field could help integrate the wide array of knowledge held by researchers, communities, and other stakeholders. There was consensus among the diverse group of attendees that, for the time-being, low-cost air-quality sensing was a complement, not a replacement for high-end sensing. Likewise, attendees agreed that to properly utilize low-cost sensing, it is critical to apply sensors in ways that complement the capabilities of the sensors. The group also expressed optimism that despite sensor limitations there are areas in which studies using low-cost sensors can make valuable contributions. In the next phase of low-cost air quality sensor research, a goal should be the bringing together of diverse sets of expertise to identify and tackle ongoing and emerging issues, especially via projects that lead to data driven actions and improved public health.

ACKNOWLEDGMENTS

The authors wish to acknowledge colleagues Massimiliano Menarini and Kevin Patrick from University of California San Diego, Tim Dye from Sonoma Technology, and Ronald Williams from U.S. Environmental Protection Agency who assisted with planning this workshop. Panelists Andrea Polidori, Ron Cohen, Jonathan Thornburg, Angelo Bianchi, Abhijit RS, Andrea Clements, Michael Hannigan, Michael Heimbinder, Sanjoy Dasgupta, Nicholas

Masson, Mark Potosnak, Sandy Navarro, Luis Olmedo, Jill Johnston, Nicole Wong, Vanessa Galaviz, and Ashley Collier-Oxandale were instrumental in setting the stage and leading the group discussions around the topics discussed here. This manuscript was improved through proof-reading and review by Joanna Casey (University of Colorado, Boulder), Vasu Kilaru (USEPA), Paul Solomon (USEPA), and our anonymous journal reviewers. The conduct of this workshop and the writing of this paper was supported in part by the National Science Foundation under grant numbers CNS-1446912 and CBET-1240584. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The views expressed in this article are those of the author(s) and do not necessarily reflect the views or policies of the U.S Environmental Protection Agency. Publication charges were paid by the U.S. Environmental Protection Agency.

Author Contributions: A.L.C., W.G.G., A.R.S., M.M.H., A.C.O. and M.H. assisted with planning this workshop. A.L.C., W.G.G., A.R.S., J.E.J., M.M.H., J.T., A.C.O. and M.H. all assisted with the development of this paper.

Preface to Chapters 3 & 4

The following two chapters provide information on my use of low-cost sensors in an education and outreach context. Both describe the development of an E&O program, assessments of this program, and plans to ensure the program's sustainability. These papers were intended to provide information for those interested in building a similar program. Additionally, the two chapters contain anecdotes and data speaking to the value of this program. There is some overlap in content between the two chapters – Chapter 4 was intended to be a continuation of and build off of Chapter 3. Thus Chapter 4 contains newer and more data and assessment information than Chapter 3.

Chapter 3: The North Fork Valley Project – A Project-based Learning Curriculum to Support the use of Next-generation Monitoring Technologies in Rural Communities

Contributing Authors: Ashley Collier[1], Daniel Knight[1], Katya Hafich[1], Michael Hannigan[1], Ben Graves[2], & Madeline Polmear[1]

[1]University of Colorado Boulder
[2]Delta County Public Schools

Publication Information: paper presented at the American Society of Engineering Education, Rocky Mountain Section Conference 2015 (April), Denver CO

ABSTRACT

Advances in sensor technology mean more research groups are developing, using, and improving monitoring systems that utilize low-cost sensors. However, these technologies also have the potential to make data collection more accessible, benefiting education and outreach programs. Through the North Fork Valley (NFV) Project, we are bringing next-generation air quality monitors into the classroom and developing supporting resources to help students engage in their own hands-on research. This project is a collaboration between the University of Colorado at Boulder, the Western Slope Conservation Center (Paonia, CO), and Colorado's Delta County schools. The objectives of the overall project include (1) collecting baseline air quality data using low-cost monitors (termed U-Pods) and (2) engaging local students in citizen science efforts.

The NFV is a rural Colorado community that faces potential impacts to air quality as a result of current and future energy development; however, rural communities typically lack historical air quality data, tools to monitor air quality, and connections to higher education and university resources. During year one of the project (Sept. 2013 – May 2014), we introduced the technology and began working with students across three high schools. In addition to the baseline monitoring, we also provided a set of portable monitors and helped students conduct their own air quality research projects. Impressed by the engagement and interest on the part of the students, we began building a project-based learning curriculum to support student driven air quality data collection. In year 2 (Sept. 2014 – May 2015) we were still working with three schools in the NFV, formalizing the curriculum, implementing it, and conducting assessment activities.

Assessment activities included student surveys, teacher surveys and questionnaires, and questionnaires for other stakeholders. We will also send the curriculum for review by an external team of teachers, researchers, and

others working in education/outreach. This paper discusses the curriculum and the results of our assessments, as well as unexpected project outcomes, such as the value of providing students with a window into academic-level scientific research and also with an opportunity to collect local data that allows them to engage at a new level with a global environmental issue – air quality.

1. INTRODUCTION

The North Fork Valley Air Monitoring Project began as a collaboration between the Western Slope Conservation Center (WSCC) in Paonia, CO and the University of Colorado, Boulder. The WSCC was interested in collecting baseline air quality data ahead of proposed oil and gas development in the region. There is a history of engagement with environmental issues in this community, for example the WSCC maintains over 10 years of volunteer-collected surface water quality data. However, the community has very little information about their air quality. The University of Colorado's Office of Outreach and Engagement connected the WSCC with a mechanical engineering research group working with low-cost air quality monitoring tools (an air quality lab under the direction of Michael Hannigan). The original objectives of the project included (1) piloting baseline air quality data collection using low-cost tools, and (2) engaging the community in citizen science. This project also provided one of the first opportunities for long-term deployment and use of the low-cost air quality monitors (U-Pods). Our research group hoped to use this opportunity to learn more about the performance of the U-Pods, as well as how people (citizen scientists and community groups) would like to use the technology. This project, now in year two, has grown into the foundation for a promising K-12 outreach program connecting high school students with university researchers.

This project is based in the North Fork Valley (NFV), which is located in Delta County on the Western Slope of Colorado. This is a rural and primarily agricultural community with several coal mines that have been in operation for over 30 years. Despite ample natural resources, students are at a disadvantage when it comes to accessing higher education. In 2010, American Community Survey Statistics revealed that 10%-20% of county residents have a college degree (US EPA, 2015), which may indicate that some students are growing up in homes where a college degree is not highly valued. Additionally, the county is more than a four-hour drive from Colorado's large state schools located on the Front Range, making K-12/university partnerships difficult to initiate and sustain. Together these factors may result in students having misconceptions about the world of academic research.

The recent collapse of the local coal industry (Markus, 2013) and the continual hemorrhage of jobs due to the mechanization of agriculture is causing serious disruptions in the local economy. During the 2013-2014 school year, 50.1% of the students were eligible for free or reduced lunch (CDE, 2015). Problems with the local economy can affect the entire school district; a poor economy leads to lowered enrollment, which in turn affects district funding. We feel that these factors make rural school districts, like Delta County, especially important places for K-12 outreach programs. The NFV project specifically exemplifies, how access to resources and support from university partners can result in positive STEM learning experiences for students and benefit all participants (the students, the teachers, and the researchers).

Year one of the project began with the installation of a network of stationary air quality monitors (U-Pods) and the next step was to encourage community engagement in the project. Several local high school teachers were interested in learning more and invited us into their classrooms. We brought more U-Pods into the classroom and introduced the students to the project. The teachers were interested in working together and we continued to partner together for the remainder of the school year. Thus began our K-12/university partnership. This partnership involved monthly visits from a CU graduate student maintaining the monitoring pilot project and supporting student research projects. Although our long-term vision involves community groups and citizens using U-Pods to collect useful air quality data, the technology itself is still developing, as well as the necessary supporting infrastructure necessary to use the technology (e.g., data organization and storage solutions, and data QA/QC mechanisms). We quickly realized how working with students could provide us with valuable information to help build the foundation of air quality citizen science while also providing a unique opportunity for students to learn more about academic research and participate in a hands-on STEM learning experience. Thus the project naturally grew into a mutually beneficial opportunity.

Following year one, the project was able to continue through the NSF-funded Sustainability Research Network AirWaterGas Project – a large-scale interdisciplinary research project including air and water quality researchers, public health experts, economists, and more all examining the benefits and challenges associated with oil and gas development. The NFV Project is continuing specifically under the Education and Outreach efforts of the AirWaterGas Project. Year two has provided us with the opportunity to consider how we can grow our successes into a sustainable program offered by the university. This paper describes the efforts of our second year of the project, primarily to build a project-based learning curriculum which will support the continued implementation and use of

the project. We describe our motivations, methods, thoughts on implementation and sustainability, and an overview of the assessment activities conducted.

2. THE TECHNOLOGY

Developments in sensor technology are making environmental monitoring tools more accessible, and ideal for STEM learning opportunities. Our monitor (the U-Pod, shown in Figure 1 below) is an open-source design and can be completely constructed for under $1000 per monitor (Mobile Sensing Technology, 2017). The sensors are electronic and data is collected continuously to an on-board mini-SD card. Another advantage, in addition to the cost, is that the U-Pod is capable of measuring multiple pollutants (i.e., carbon dioxide, carbon monoxide, volatile organic compounds, ozone, and nitrogen dioxide). This technology has the potential to increase the spatial resolution of data (e.g., deploying 20 monitors throughout a neighborhood could provide detailed information regarding the impact of emission sources on local air quality), and supplement our existing large-scale monitoring systems. Although the sensors used in the U-Pod are less accurate than current instruments used for regulatory purposes, if used in conjunction with higher quality instruments we may be able to learn a lot more about our air quality and ways of improving it.

Another exciting prospect is the use of this technology for citizen science. “Citizen science refers to the engagement of non-professionals in scientific investigations – asking questions, collecting data, or interpreting results” (Miller-Rushing et al., 2012). Citizen science can range from contributory (where participants collect data as part of a larger, researcher led project) to collaborative and co-created (where scientists and participants work together with both groups engaged in defining the question, collecting the data, and interpreting results) (Miller-Rushing et al., 2012). Our research group is interested in the long-term potential of this technology to be a part of collaborative and co-created citizen science. This vision, however, may be a few years off as researchers continue using and improving the technology. In the meantime, the U-Pods are ideal for educational applications. Using the monitors provides students with exposure to the fields of electrical engineering, computer science, and mechanical engineering in addition to the environmental engineering involved in air quality research. The U-Pods can be used by teachers to instruct and inspire interest in a variety of subjects. New technologies such as this one are extremely valuable for STEM education, not only because of their increasing affordability, but also because students can take an inside look at something they could build themselves – these tools are not the traditional “black box” instrument.

Figure 1: Monitor Deployed Outdoors (center), An Interior View of the Circuit Board and Sensors (upper left)

3. PROJECT-BASED LEARNING

We have chosen to utilize a project-based leaning (PBL) structure for out outreach program. “Project-based learning is a teaching method in which students gain knowledge and skills by working for an extended period of time to investigate and respond to a complex question, problem, or challenge” (Buck Institute for Education, 2018). It involves the application of knowledge, as well as the acquisition of new knowledge to investigate a driving question and it challenges students to manage their time and resources. In engineering education, projects seem to provide students with “a better understanding of the application of their knowledge in practice and the complexities of other issues involved in professional practice” (Mills & Treagust, 2003). Similar outcomes have been observed in K-12 education, “research clearly indicates that project-based learning is beneficial, with positive outcomes including increases in level of student engagement, heightened interest in content, more robust development of problem-solving strategies, and greater depth of learning and transfer of skills to new situations” (Holm, 2011). In light of these positive outcomes, we are using the model of PBL to drive the structure of our curriculum.

After an introduction to the capabilities of the U-Pods and the field of air quality research (through the curriculum discussed in the following section of the paper), students are tasked with designing their own investigation into air quality issues in the community. As students progress with designing their project, they stumble upon “need-to-knows;” questions about atmospheric chemistry and pollution that they need to answer in order to progress in their investigation. Students then collect and analyze their own data and present their research in a scientific-style poster. This project provides students with a research experience truer to life than the typical in-class experiments and

activities because in addition to the air quality research, students are responsible for defining and maintaining a schedule and working as part of a collaborative team.

A lack of resources in rural communities typically makes the implementation of projects like these challenging; however, if this can be overcome (e.g., through K-12/university partnerships) the benefits to student learning are impressive. During the first year, teachers reported (anecdotally) that engagement had increased and in particular from students who had not previously displayed a strong interest in their science classes. We believe part of the reason for this might be an increased sense of ownership over the project, possibly due to the authenticity of the questions students ask. For example, students investigate questions of their choosing in their own communities. They are being provided with the opportunity to take the air quality background knowledge they have gained and apply it to an investigation of their immediate environment.

4. CURRICULUM DEVELOPMENT

At the beginning of the project, we thought we might ask students to help us maintain our stationary monitoring network and collect baseline data. But there was greater engagement from the students (as reported by teachers) and interest from the teachers when we gave students U-Pods to use for other activities. We adapted our project to meet the needs of the community and we ended up with a more mutually beneficial project. For example, we continued to collect the baseline air quality data and were able to use that data in the classroom for teaching purposes. Alternatively, asking students to investigate their own research questions provided our group with more information on potential local air quality issues (e.g., agricultural practices that might affect air quality).

Some examples of how students were interested in using the U-Pods include measuring emissions from different vehicles, investigating the possible impact of train emissions by comparing two different locations in the NFV, and examining indoor air quality in different classrooms throughout their school. One group of students burned homemade biodiesel and regular diesel in the lab and compared these emissions; they were invited to compete at the state level in the Colorado Science Fair at Colorado State University (Ft. Collins, CO). In addition to competing, these students were able to visit the CSU campus, attend lectures, and were awarded an honorable mention in the competition. For these rural students, this experience had a significant impact on their desire to pursue higher education.

During year two of the project, we are building a curriculum that will support the continued implementation of this project. Five modules are under development. These modules are intended to provide enough background knowledge and enough of an introduction to the skills students will need to act as a springboard for their project phase. In other words, we want to empower students to ask complex and informed questions that will guide their research, but it is expected that they will seek out new knowledge and build on their skills in order to successfully complete projects (this is true to the objectives of PBL). The curriculum is not intended to be comprehensive, but is intended to support the teachers interesting in conducting challenging air quality research projects in their classrooms.

- Module 1: This module provides an introduction to air quality and the technology; the module is intended to communicate to students (1) why we care about air quality, (2) some basic information regarding air quality measurements (the different types of pollutants, what we can measure with the U-Pod, etc…), and (3) get students thinking about the connection between specific emission sources and the quality of our air.
- Module 2: This module focuses on data collection using the low-cost monitors and this is taught through an activity. Students begin by learning about the difference between complete and incomplete combustion, and how this is relevant to internal combustion engines in vehicles (mainly, the air to fuel ratio). Then students are then given a U-Pod to collect data on emissions by idling different vehicles, as a group we analyze and discuss the data, as well as our experimental design.
- Module 3: This module covers data analysis using excel; we have two prepared data sets (one with carbon dioxide data and one with ozone data), both include approximately 100 rows of data with time, pollutant concentration, and temperature. The students use this data to complete an activity that takes them through basic data analysis using Microsoft excel.
- Module 4: During this module students are given time to brainstorm project ideas, and discuss these ideas with their group members, their teacher, and graduate student mentors. They also write up an official project plan, using a provided template.
- Student Projects: At the point, students will conduct their projects, modules 1-4 are intended to provide the support and preparation necessary to empower students to do their own research.
- Module 5: This module is intended to help students process and finalize their projects. Students take a look at example posters, fill out a worksheet to encourage them to think critically about their project and the results, and the finish by using a template to create a scientific-style poster.

The series of modules or lesson then culminates in a poster session, open to other students, teachers, parents, and the community. The poster session challenges students to communicate their research and gives them an opportunity to show off what they have done. Each module includes complete instructions, presentations, worksheets, data sets, templates, links to more information, and suggested assessment as necessary. Additionally, the modules can be implemented independently, for example if a teacher is looking for a 1-2 class introduction to excel and data analysis. However, as a whole the modules provide students with the skills necessary to conduct their own research.

At this point, heavy involvement from university partners is required throughout the entire lesson, for example graduate student mentors lead or co-teach each module. We hope that the curriculum reaches a point where teachers are comfortable implementing it more independently, which will allow us take a more focused role and reach a greater number of students. Although we do see the relationships between high schools students and graduate students as being an integral piece of the project. We envision having a graduate students build a relationship with a class, by making as many in-person visits as is feasible, being available to assist and support teachers as necessary, and being available to advise students on project design, data collection and data analysis. However, we would like to build a strong enough foundation (including quality/comprehensive curriculum, and making the technology as simple to use as possible) that graduate students can support rather than lead, which we believe will also give teachers more independence and flexibility in implementation. This may also include teacher training sessions in the future, which would provide hands-on training for teachers prior to their use of the U-Pods in their classrooms. This model will allow us to reach a larger number of classes across Colorado.

5. CURRICULUM IMPLEMENTATION AND SUSTAINABILITY

We intend to make the curriculum available publicly and free of charge in August 2015. We also plan to organize a monitor rental program. Given that we are working with a developing technology (the U-Pod) – a rental program would likely be more successful than if teachers were to purchase U-Pods. A rental program will allow our research group to maintain the U-Pods and provide technical support as needed, while giving teachers access to the technology. Once again, the project continuing through the AirWaterGas Project has provided us with the luxury of resources and time to explore and build the sustainability of the project.

Working closely with teachers as we develop the curriculum and program, is helping us to design for simpler implementation. We are aware of the hurdles teachers face, especially when presented with curriculum promoting less

conventional methods of education (e.g., PBL), but we hope that through continuous teacher feedback we can facilitate the incorporation of this type of teaching. In addition to the curriculum, we hope to foster a community of teachers and learners who can work together. We would like to build (or utilize an existing tools) an online space where participants (teachers, students, and community members/groups using the U-Pods) can upload and share data, process and visualize data, and communicate with each other about their projects, and the tools. This is further down the road, but is certainly part of our long-term vision.

6. ASSESSMENT PLAN

At the end of the first year we collected a student opinion survey and interviewed participating teachers, however, this year we are more rigorously evaluating the project as it is implemented and developed using an assessment plan. The plan will evaluate both our curriculum and the outreach project as a whole. For example, in order to evaluate the lesson, each module will be sent for external review by a team of experts. This team includes a teaching expert, an air quality science expert, and a curriculum/outreach program expert. The review criteria itself is based on CLEAN Guidelines developed by CIRES to evaluate potential and proposed CLEAN curriculum (CIRES, 2015). The information from the external review team will be used to finalize the curriculum before it is made publicly available.

Summative assessment data will assist in evaluating our project's success and sustainability; this data will focus on student skills/attitudes, the project's impact, and opinions regarding the project (Knight & Sullivan, 2006). Summative assessment activities include student surveys, teacher surveys and interviews, and stakeholder interviews.

- Student Surveys
    - Four in total
    - Three accompanying Modules 1-3, administered following each module, copies of these three surveys are available in Appendices C, D and E
    - Surveys collect demographic information and include multiple choice questions that will assess students' attitudes toward STEM and higher education, as well as their confidence in new, specific skills developed during the modules
    - The final survey will be implemented after the completion of the entire lesson, and it will collect demographic information and attempt to assess the impact of participating in the project
- Teacher Surveys and Interviews

    - Administered after the completion of the entire lesson
    - This survey will assess their impressions of effectiveness of the project, and whether or not they felt they received sufficient support (both technical and regarding content)
    - It will include multiple choice survey questions and open ended interview questions
- Stakeholder Interviews
    - Administered after the completion of year two
    - Example of stakeholders – individuals associated with the WSCC and Delta County Public Health (another project partner)
    - Given the uniqueness of rural communities and the challenges the face, we would like to assess how effectively we are meeting the needs of the community through a series of open-ended questions

The summative assessment data will be used to evaluate our success in meeting the goals of the NFV Project specified in the assessment matrix, as well as the goals of the larger AirWaterGas Education and Outreach Project. This data will also be used to determine improvements that we can implement moving forward. For reference, Table 1 provides information on the scope of the NFV Project during years one and two.

Table 1: Summary Statistics for Engagement

	Year 1	Year 2
Number of Delta County Schools	3	3
Number of Teachers	3	3
Number of Students	40	46

Assessment data will help ensure we address Goals 1-3 from the AirWaterGas Education and Outreach assessment matrix (Appendix B) in future offerings of the program by helping us to determine how well we are meeting these goals currently. Preliminary data from year one student and teacher interviews indicates we are successfully reaching minority populations with our education and outreach activities, the minority population being an underserved rural Colorado community (addressing Goal 2). Of the three teachers who participated during year one, two teachers were interested in participating for a second year (the third teach expressed interest, but is no longer teaching for unrelated reasons). Additionally, we are bringing hands-on environmental and engineering educational opportunities to Delta County (a network stakeholder), also addressing Goal 1. Using both this preliminary information, and the data we are currently collecting, we can make improvements to future iterations of the program that will better ensure we meet our goals for this outreach program and it is contributing to the larger AirWaterGas program that has provided the funding allowing the project to continue.

7. ASSESSMENT RESULTS AND OUTCOMES

Teacher and students reactions to the first year of the project drove our desire to continue the project and explore ways to make it sustainable. Some observations shared by teachers during the year one interviews mentioned the following benefits: ownership of their projects by the students and increased engagement, a better understanding of the scientific process (especially of how science can be more of a non-linear and challenging process than is typically presented in high school), and students strengthened skills likely to help increase their success in higher education (e.g., problem-solving and analysis skills). Two out of the three teachers also noted, that their own participation was likely to ensure they would continue teaching in the district.

These outcomes are not unlike outcomes observed in other K-12 STEM outreach projects. Four outreach projects at the Colorado School of Mines were evaluated and researchers noted an increased interest in science, mathematics, and computer science, as well as an increased exposure to and interest in college (Moskal et al., 2007). These outreach projects taught the application of mathematics to science and engineering, but they also involved teaching through hands-on projects and also the use of graduate student mentors in K-12 classrooms, which is similar in structure to our program. It seems that although the content is somewhat different, the programs cited in this paper and ours are similar in approach. Preliminary data suggests we are seeing outcomes and benefits in our program that are similar to the outcomes and benefits of other K-12 STEM outreach programs, indicating that we are building a promising program.

In addition to the more or less anticipated outcomes, there have been several unexpected outcomes we are beginning to see as well. For example, Laursen et. al describe a 'leveling effect' to outreach activities, teachers observed that the new material had a way of challenging and engaging all students regardless of their typical level of performance or aptitude (2007). While air quality is typically taught in Environmental Science and AP Environmental Science classes (the types of classes we target for this project), the content required for this project tends to be more in-depth than students are used too. For example, students learn more about particular pollutants and air quality measurement techniques. In addition to that students are using a new and unfamiliar tool to make measurements (the U-Pod) and building skills with other 21st century tools (e.g., Microsoft excel, and tools on the Google platform like Documents and Fusion Tables) that facilitate online collaboration and project management. We have also heard multiple comments from teachers expressing their surprise at the engagement from typically lower performing students. It seems curricula like this may have the potential to research students that traditional methods do not.

Year two of the project will provide the opportunity to better confirm these outcomes. Although assessment is still underway, we have administered the student surveys accompanying Modules 1-3 and are able to share some preliminary results. The Project Intro Survey was distributed to the high school students following Module 1. The survey sought feedback on students' opinions and attitudes regarding air quality in general and in the context of the project. The results suggest that the presentation offered a valuable learning experience as students rated their understanding of the impact of air quality research and the amount they learned highly (4.43/5.00 and 4.56/5.00, respectively), as shown in Table 2. Open-ended questions were developed to gain qualitative feedback on the experience. The results revealed that students were very responsive to the hands-on nature of the activity. Using the U-Pod offered the opportunity for students to engage in the experiment and understand the application of the lecture material. One student commented," I love doing hands-on and that is how I learn" and another noted, "I liked how you could really see what was happening with the air" after using the U-Pod. The question regarding recommendations for improvement for the presentation also revealed strong satisfaction. Students commented, "It could not be made better, it was awesome" and "it was perfect, filled with really interested facts." One common suggestion was to shorten the length of the lecture to allow for more time for the interactive experiment. We intend to integrate this student feedback, by improving the modules as suggested and maintaining the successful aspects.

Table 2: Module 1 Student Survey Results

Project Intro Survey (n=55)	
Category	**Mean**
I am interested in learning more about air quality.	3.62
I want to know more about what an environmental engineer does.	3.38
I feel that engineers have made major accomplishments and advancements.	4.69
I would have a difficult time explaining this topic to another person.	3.09
I know where to go to find out more about this topic.	3.53
I understand the impact of air quality research on the community.	4.42

8. CONCLUSIONS

The NFV Project is the beginning of a promising K-12 outreach project. The project has evolved from a grassroots effort to work with rural Colorado students into a systematic project being designed for sustainability and longevity. By the end of this second year we will have a curriculum that supports the implementation of PBL in the

classroom, as well as student use of a new technology – the U-Pod. We believe our model of 'renting' the U-Pods will ensure teachers are provided with technological support and will help university researchers in building K-12 partnerships. We have been extremely fortunate in having expert educational partners who are willing to work closely with our group to weigh in on what works in the classroom and what does not. These partnerships will also be a key piece in the sustainability of the project going forward.

Our research group has learned a lot about building partnerships and working with rural communities and we hope to build on this work in order to reach other rural Colorado communities. Rural communities are underserved; they do not lack great teachers or talented students but are sometimes put at a disadvantage through a lack of resources partly caused by distance from large universities typically offering resources. We hope this project is able to provide a sustainable program that to makes more resources available to teachers and students, as well as providing a way for students to build personal connections with researchers. We are excited by the project's initial results thus and would like to see the kind of impact it can have in the long run.

Chapter 4: On the Development and Implementation of a Project-Based Learning Curriculum for Air Quality in K-12 Schools

Contributing Authors: Ashley Collier[1], Daniel Knight[1], Katya Hafich[1], Ben Graves[2], Michael Hannigan[1], & Madeline Polmear[1]
[1]University of Colorado Boulder
[2]Delta County Public Schools

Publication Information: In Proceedings - Frontiers in Education Conference, FIE (Vol. 2014). DOI: 10.1109/FIE.2015.7344246, 2015

ABSTRACT

The availability of low-cost sensors and environmental monitoring technologies is growing rapidly. While researchers are making use of these technologies and validating their results, there is also enormous potential for their use in education and outreach. Through the North Fork Valley (NFV) Project, we are bringing next-generation air quality monitors into the classroom and developing necessary resources to support student-driven research projects. This project is a collaboration between the University of Colorado at Boulder, community partners, and educational partners (Delta County School District 50, and St Vrain Valley School District) and is funded by the NSF-funded AirWaterGas (AWG) Sustainability Research Network Education and Outreach (E&O) efforts. During year two, we are formalizing the curriculum, implementing it in multiple school districts, and conducting assessment activities. This paper will discuss the curriculum, the results of our assessments thus far, and our vision for the project. The work completed during year two will help ensure we are building a sustainable project. Additionally, we intend to make the curriculum public with hopes that other research groups working with low-cost air quality sensors can utilize our materials and possibly our model for their own education and outreach.

1. INTRODUCTION

For the past two years our team has implemented a project that brings together innovative research and K-12 education. The North Fork Valley (NFV) Air Monitoring Project takes a low-cost, next-generation air quality monitoring instrument, still undergoing the research and development process, and explores its place in education and outreach. Partnerships like these provide a unique opportunity beneficial to both researchers and K-12 students. Researchers are able to learn about their work from a different perspective; in our case how the public would like to use our technology, and how well it can be adapted for this purpose. Students in these K-12/university partnerships are able to work in a hands-on way with a technology that is not a black-box and learn about university-level research

and the possibility of pursuing a degree or career in STEM. This program provides students with an introduction to disciplines like environmental and mechanical engineering, as well as the problem-solving that goes into investigating complex questions. Teaching students about air quality is not our final goal, but rather we hope to use air quality research as the vehicle by which students can learn more about the skills specific to STEM fields, both technical and professional. This paper will describe the project's evolution, the technology that makes this project possible, and our efforts to build a sustainable outreach program.

1.1 The Communities

Our initial school district partner was Delta County School District 50 (DCSD) located on the Western Slope of Colorado in the North Fork Valley. During the second year we have expanded to the St Vrain Valley School District (SVVSD) on Colorado's Front Range. Delta County is a rural and primarily agricultural community with several coal mines that have been in operation for over 30 years. However, the recent collapse of the local coal industry (Markus 2013) and the continual loss of jobs due to mechanization in agriculture is putting stress on the local economy. During the 2013-2014 school year, 50.1% of the students were eligible for free or reduced lunch (CDE, 2015). In terms of the perceived value of higher education, 10%-20% of residents have a college degree (US EPA, 2015). DCSD is also more than a four-hour drive from Colorado's large state schools located on the Front Range, which makes K-12/university partnerships difficult to initiate and sustain.

By contrast, SVVSD is a mixture of rural and suburban. This part of Colorado has experienced significant growth over the past few years due to the expansion of the oil and gas industry (Dunn, 2014). During the 2013-2014 school year, 37.4% of the students in SVVSD were eligible for free or reduced lunch (CDE, 2015). Additionally, the area is well served with university outreach programs as the schools in this district are less than an hour drive from the Denver Metro area and several major Colorado universities. The districts also differ demographically, with DCSD made up of a 10-20% minority population and SVVSD made up of 20-40% minority population (US EPA, 2015). Working with these two districts during year two of the project is helping our team to build a program that meets the needs of a variety of groups, especially underserved rural communities. We hope to build a curriculum both accessible to and suitable for the diverse geography and socio-economic demographics of Colorado.

1.2 NFV Project, Year 1

The North Fork Valley Air Monitoring Project began as a collaboration between the Western Slope Conservation Center (WSCC) in Paonia, CO, the University of Colorado, Boulder's Office of Outreach and Engagement, and the Hannigan Air Quality Research Lab (Mechanical Engineering, CU Boulder). Delta County lacks detailed historic air quality data, and the WSCC was interested in understanding current air quality in the area in light of proposed increased oil and gas development in the NFV. The original objectives of the project were to pilot low-cost tools for baseline air quality data collection and engage the community in citizen science.

During year one we engaged local high school students in assisting with the project, but found that they were more interested in asking their own questions than helping with the collection of baseline data. The project naturally grew into two parts: (1) the long-term, continuous collection of air quality data using a stationary network of monitors, and (2) student-driven air quality research projects. After seeing the benefits of this experience to students, particularly in a rural and underserved community (Collier et al., 2015a), we decided to continue for a second year in an attempt to build a sustainable outreach program.

1.3. NFV Project, Year 2

During year two, the project is continuing through the NSF funded AirWaterGas (AWG) Sustainability Research Network, specifically with the help of the AWG Education and Outreach (E&O) team (Knight et al., 2015). AWG is a large-scale interdisciplinary research project including teams studying air quality, water quality, water treatment, water quantity, natural gas infrastructure, social-economic systems, health effects, practices and policies, and E&O. Together all of these teams are examining the benefits, risks, and challenges associated with unconventional oil and gas development. Year two activities for the NFV Project included continued data collection in the NFV, curriculum development, implementation in multiple school districts, and overall project/curriculum assessment. Although the title of the program has not changed, we have expanded to an additional school district in an area of Colorado outside of the North Fork Valley.

Working with school districts affected by current and proposed oil and gas development helps us build bridges to facilitate the sharing of results from the larger AWG Project. For example, these communities may have specific questions regarding risks to local air/water quality, and because we have built trust between the community and our University, they may feel comfortable coming to us for information. Using this program to build relationships has the potential to help us get up-to-date, relevant research to these communities to inform decision making on the

local level. Finally, air quality serves as a vehicle to initiate conversations on complex issues surrounding energy development and use.

1.4 THE TECHNOLOGY: U-PODS

Developments in sensor technology are making environmental monitoring tools more accessible, and ideal for supporting STEM learning opportunities. Researchers at the University of Maine demonstrated through the GK-12 Sensors! Program that using sensors in secondary schools is beneficial to students, teachers, and graduate teaching fellows (Arsenault et al., 2005). For several years, our research group has been developing and using an air quality monitor that utilizes low-cost sensors for our own research. Our monitor (the U-Pod, shown in Fig. 1) is an open-source design and can be constructed for under $1000 per monitor (Mobile Sensing Technology, 2017). The U-Pod is capable of measuring multiple pollutants (e.g., carbon dioxide, carbon monoxide, volatile organic compounds, ozone, and nitrogen dioxide) continuously; data is then recorded to an on-board memory card. Although the sensors used in the U-Pod are less accurate than the more costly instruments used for regulatory purposes, if we leverage both the low-cost and high quality instruments, there is the potential to collect far more detailed information regarding our air quality than currently exists.

New technologies such as this one are extremely valuable for STEM education not only because of their increasing affordability, but also because students can take an inside look at something they could build themselves. These tools are not the traditional "black box" instrument because they utilize a relatively simple design and 'off-the-shelf' sensors. Students can open up the U-Pods and understand how they function, making the technology is ideal for hands-on learning. Using the monitors also provides students with exposure to the various fields of study that went into making it: electrical engineering, computer science, and mechanical engineering in addition to the environmental engineering involved in air quality research. Citizen science is another area where technologies like the U-Pod could have a large impact. Citizen science "refers to the engagement of non-professionals in scientific investigations – asking questions, collecting data, or interpreting results" (Miller-Rushing et al., 2012). Citizen science can range from contributory (where participants collect data as part of a larger, researcher led project) to collaborative and co-created (where scientists and participants work together with both groups engaged in defining the question, collecting the data, and interpreting results) (Miller-Rushing et al., 2012). Our research group sees using the U-Pods for educational applications as a way to move toward this vision of using low-cost technologies to support collaborative and co-created citizen science.

Figure 1: Image of the U-Pod, interior (top left) and exterior

2. A PROJECT-BASED LEARNING CURRICULUM

2.1 Project-based Learning

In addition to hands-on learning, we have found that the U-Pods are an ideal way to support project-based learning (PBL) driven by student-generated research questions. Project -based learning is a method of teaching in which students work for an extended period of time on a complex problem, question or challenge, and thereby gain knowledge and skills (Buck Institute for Education, 2018). In engineering education, projects better equip students to apply their knowledge in practice, as well as providing a better understanding of the complexities in involved in professional practice (Perrenet et al., 2000). Similar outcomes have been observed in K -12 education where research indicates that PBL leads to an increased level of student engagement, higher interest in content, and better problem-solving skills, and an increased depth of learning and ability to transfer skills to new situations (Holm, 2011).

During year one of the project, teachers anecdotally reported similar outcomes including increased engagement and greater depth of learning. For this reason, we chose to build a program and curriculum supporting the PBL teaching method.

PBL challenges students to apply their existing knowledge in new ways and acquire new knowledge to solve novel real-world problems. As students progress with their project, they stumble upon “need-to-knows;” in this case, questions about atmospheric chemistry and pollution that they need to answer in order to progress in their investigation. Additionally, since projects are conducted over a long period of time, students must manage working with a team, adhering to a schedule, and monitoring progress. We currently cover the bulk of the content (discussed in the following section) during the first semester and leaving the second semester for projects. The curriculum

introduces students to the U-Pods, and helps them build useful skills with all of the pieces supporting the PBL model (Collier et al., 2015a).

2.2 Curriculum Development

During the first year, it was interesting to observe the synergism between the continuous data collected through the stationary network and the student-driven projects. For example, we were able to use the baseline data to support student learning about air quality concepts in the classroom. Alternatively, student-led investigations provided our group with information on potential local air quality issues (e.g., agricultural practices like ditch burning). In addition to local knowledge, students use of U-Pods clarified what skills and background knowledge users need. This also led to the idea of developing a curriculum to support U-Pod use.

During year two, the curriculum used in the classroom is being reviewed and polished for public distribution. Five modules are currently under development. These modules provide students with background knowledge and an introduction to the skills necessary for successful projects including data analysis and how to present scientific results. The curriculum will empower students to ask complex and informed questions about local air quality issues. Each module (summarized in Table 1) includes guidance for teachers, presentations, worksheets, data sets, templates, links to more information, and suggested assessment as necessary. Together the five modules cover skills useful for this project and beyond it (e.g., working with large data sets in Excel, making a poster in PowerPoint, etc…). The series of modules culminates in a 'Science Symposium' or poster session, open to other students, teachers, parents, and the community. In line with the goals of PBL, the poster session challenges students to interpret and communicate their research to an authentic audience.

3. CURRICULUM IMPLEMENTATION

Lack of resources, knowledge, and skills is often as a barrier to technology integration in K-12 settings (Hew & Brush et al., 2007). To address these barriers, we are creating a U-Pod 'check-out program'. Given that the U-Pod is a developing technology – a rental program will likely be more successful than selling the U-Pods because this type of program will allow the researchers to maintain the technology. This type of program will provide teachers with access to resources while facilitating technical support and providing our group with a way to maintain and strengthen relationships with districts using the U-Pods. We feel this is a promising model, and plan to make our curriculum publicly available in Fall 2015 in the hopes that other groups working with low-cost sensors will replicate this model.

Table 1: Air Quality Curriculum Overview

	Module Description	Skills
1	Introduction to air quality research and the U-Pod (lecture and demo)	Background knowledge and vocabulary for air quality and engineering and concepts
2	Data collection (activity)	Using the U-Pod Critical thinking regarding sources of error and results
3	Data analysis (activity)	Basic statistics and data visualization information Using Excel for calculations and graphing Interpreting results using background knowledge
4	The Project (planning, data collection, data processing and analysis)	Planning and carrying out a research project (scheduling, logistics, designing data collection, etc…) Processing raw instrument data Analyzing data in Excel
5	Interpreting data and presenting results	Data interpretation Communicating results Making a scientific poster or presentation

4. PROJECT ASSESSMENT

4.1. Overview

Informal and anecdotal assessment data were collected at the end of year one. Assessment for the second year is more rigorous; formal assessment matrices have been developed for both the NFV Project and the AWG E&O team (Collier et al., 2015a). The NFV Project goals include developing/finalizing curriculum and conducting formal assessment of both the curriculum and the program. The goals for AWG E&O are broader in scope, for example one of the goals listed in the AWG E&O Assessment matrix is to 'Integrate AWG SRN Research, Education & Outreach Activities'. We are successfully reaching an underserved rural Colorado community (Delta County students) by facilitating the use of the U-Pod for E&O in addition to the research it is being used for internally. Year two data will allow us to better determine which goals we are addressing and where we can improve.

Summative assessment data will assist in evaluating the project upon the completion of year two; this data will focus on student skills/attitudes, the project's impact, and opinions regarding the project. Our methods are typical of those used to assess similar K-12 outreach projects (Knight & Sullivan, 2006). We are also engaging in formative assessment throughout the year, as we work with teachers one-on-one to ensure we are meeting the needs of their class and implementation is running smoothly. Summative assessment activities and our progress are presented in Table 2. Student surveys (4 total) include attitude questions and confidence questions intended to assess whether students are

developing the skills intended with each module. Multiple student surveys with repeated questions will allow for pre vs post analysis to provide insight into the impact of the program over the course of the year. Surveys and interviews from teachers will ask whether they would continue participating in this program, what benefits to their students they observe, if they received enough support, and their overall reflections. Input from other stakeholders will allow us to consider whether or not we are meeting the needs of the community. The external review team assessing the curriculum includes an air quality expert and a teaching expert. The criteria this team is using was derived from guidelines developed by Cooperative Institute for Research in Environmental Sciences (CIRES) to assess proposed curriculum for Climate Literacy and Energy Awareness Network (CLEAN) (CIRES, 2015). In all surveys where we have requested that participants rate an opinion or experience, a score of 4.00/5.00 (80%) is our target. Together all of this data will be used to not only evaluate year two, but also inform future implementations of the project, the overall structure of the program, and revisions to the curriculum before it is made publicly available.

4.2 Preliminary Results

Table 3 provides an indication of the scope of the project during years one and two. Following the completion of year one, all three of the teachers were interviewed and a simple student survey was distributed. The teachers shared the following observations regarding their students:

- Increased engagement in science classes
- Student ownership of their projects
- A better understanding of the scientific process (particularly the iterative nature of data collection and analysis that is not typically demonstrated in high school science classes)
- Students strengthened skills likely to help increase their success in higher education (e.g., problem-solving and analysis skills)

Teachers also mentioned that students benefited from working with a graduate student mentor, access to new classroom resources, and hands-on use of technology (both the U-Pod and tools such as Microsoft Excel, and the Google Drive platform to facilitate collaborative work). Finally, all three teachers expressed interest in continued participation in this program.

Table 2: Assessment Overview

Stakeholder	Type of Assessment	Progress
Students	Surveys (4 total, all students)	1st & 2nd complete (results available) 3rd & 4th incomplete (mid-may)
Teachers	Survey (1, all teachers) Interview (1, 2 teachers)	Both incomplete (mid-may)
Community (e.g., partner with Public Health Dept.)	Interview (1, 2 stakeholders)	To be given (mid-may)
Curriculum	Survey for external review team (1 per module)	1st & 2nd complete 3rd, 4th, & 5th in progress (may)

Table 3: Project Scope

	Year 1	Year 2
Number of Districts	1	2
Number of Teachers/Schools	3	5
Number of Students	40	81

Complete assessment data from year two will help us to evaluate whether or not outcomes and benefits we observed during year one are occurring repeatedly. In the meantime, we can provide a discussion of preliminary results. The Module 1 and 2 Survey results are available in Tables 4 and 5 respectively. The Module 1 Survey was distributed to the high school students following our first classroom visit. This survey sought feedback on students' attitudes regarding engineering and air quality and their reactions to the module. The results suggest that the presentation offered a valuable learning experience as students from District 1 (Delta County) rated their understanding of the impact of air quality research as a 4.60/5.00 and the amount they learned highly as 4.50/5.00, both meeting the goal of a minimum score of 4.00/5.00. District 2 (SVVSD) shared similar results with a rating of 4.31/5.00 for their understanding of the impact of air quality research and 4.60/5.00 for the amount they learned. Mean scores below 4.00 for both school districts are reflected for questions 1, 2, and 5, which indicates that these are areas where we may be able to improve students' attitudes over the course of the project. Responses to similar questions in the final student survey will help us understand if the project may have positively impacted students' attitudes toward air quality, attitudes toward environmental engineering, or their knowledge of sources for information on air quality and other STEM related questions. Lower scores were also observed for question 4, but this question is a reflection of students' comprehension of the material, and a lower score indicates comprehension. Additional open-ended questions in the Module 1 survey revealed that students responded most positively to the hands-on demonstration of the technology.

Table 4: Module 1 Survey Results

Question	Mean		P-value
	District 1[a] *(n = 20)*	*District 2*[b] *(n = 35)*	*Two Sample T-Test*
1. I am interested in learning more about air quality.	3.50	3.69	0.399
2. I want to know more about what an environmental engineer does.	3.35	3.40	0.868
3. I feel that engineers have made major accomplishments and advancements.	4.85	4.60	0.126
4. I would have a difficult time explaining these topics to another person.	3.10	3.09	0.965
5. I know where to go to find out more about these topics.	3.45	3.57	0.732
6. I understand the impact of air quality research on the community.	4.60	4.31	0.190
7. I learned a lot in this presentation.	4.50	4.60	0.570

Table 5: Module 2 Survey Results

Question	Mean		P-value
	District 1[a] *(n= 15)*	*District 2*[b] *(n = 30)*	*Two Sample T-Test*
1. After doing this activity, I feel more confident that I could use this technology to collect air quality data	4.27	4.30	0.887
2. I learned something new about the process of data collection and experimental design in a complex environment.	4.33	4.43	0.743
3. This activity led me to think about scientific investigations I might conduct, either for this air quality project or in another context.	3.67	4.27	0.047

[a]District 1 is Delta County School District 50, [b]District 2 is St. Vrain Valley School District

The Module 2 Survey (Table 5) was distributed to students after they participated in an activity to measure car emissions during which they were able to practice using the U-Pods. The data indicates that the desired mean score of 4.00/5.00 or higher mean was met for every question and for both districts (with the exception of one score for question 3 and District 1). The qualitative feedback indicated that students enjoyed the hands-on aspect and the opportunity to do the experiment themselves. One student commented, “I liked how we got involved in the activity” and another noted the change in teaching style was valuable and that “it was nice to get out of the classroom and learn a different way.

5. DISCUSSION

Engineering outreach programs “introduce students to the joys and frustrations of engineering” (Sullivan et al., 1999); this type of experience can have a multitude of benefits. Some students may be inspired to pursue a degree in engineering or STEM, while others (who may go on to work in entirely different fields) gain an increased understanding of STEM and its role in our society enhancing their multidisciplinary education. We believe that using the U-Pods for E&O has the potential to provide this type of experience. Comparing our preliminary data to results of

the evaluation of four outreach projects at the Colorado School of Mines indicates that we are building a promising outreach program (Moskal et al., 2007). Among the observed benefits to middle school students (the target audience) Mines researchers listed:

- Increased use of technology
- Increased classroom resources
- Availability of college role models
- Increased interest in science mathematics, and computer science
- Increased exposure to and interest in college
- Successful participation in academic competitions, including several awards (Moskal et al., 2007)

Many of the benefits listed above match or are similar to benefits our teachers observed during year one (mentioned in the Preliminary Results section).

Year two is providing more quantitative information regarding our project's impact and thus far the indication is that the modules are serving their intended purpose. Students responses resulted in mean scores of over 4.00/5.00 for questions 6 and 7 (Table 4), these questions cover the primary purpose of Module 1 – introducing students to air quality research and communicating why it is important. Lower scores (below 4.00/5.00) for interest/opinion questions regarding engineering and air quality are not uncommon and have been found in other studies where high school students were surveyed (Zarske et al., 2007). It is possible that 80% is an unrealistic expected score for questions such as these, or that a lack of exposure to engineering in traditional curricula results in lower scores. We plan to compare these initial scores to the scores of similar questions in the final survey; however, this may be an area where more research is necessary.

Module 2 is intended to prepare students for data collection. As indicated by the scores, students came away from the activity feeling more comfortable with the technology, and thinking about study design. The one exception is the low score from District 1 for question 3 (Table 5) regarding whether the activity led them to consider project ideas. This may be an indication that the module was implemented differently, and it would be worthwhile to add suggested discussion questions into the curriculum to address this need. Additionally, this mean score is significantly different from District 2. All other mean scores were not significantly different (p-values > .05) between districts, which indicates that students in both locations are reacting to the curriculum and project similarly. We will continue

to observe the differences and similarities between the school districts throughout the remaining surveys, as it will inform our program design, and specifically designing to meet the needs of a variety of school types.

6. LIMITATIONS AND FUTURE PLANS

The primary limitations of our assessment data are the lack of a control group and small sample size. These are limitations we will address in future iterations of the program. We already have plans for implementation in new districts during year three, which will provide a larger population. We would also like to refine our assessment tools and implement established tools developed by external groups, as well as examining how we can incorporate more established outreach frameworks into the program. Year three of the project will include the above, as well as implementation of the final version of curriculum, continued assessment, and building the U-Pod 'check-out' program. In the long-term, we are interested in working with school districts to track the long-term progress of students and their degree choices in college.

In addition to planning for future iterations of the project, we are considering program viability. One key to ensuring our program is sustainable is to refine our ideal role as a university partner. Although the relationships built between high schools students and the graduate student mentors are an integral piece of the project, currently a graduate student mentor leads or co-teaches each module. Ideally, we would like graduate student mentors to support rather than lead, which in turn gives teachers more independence and flexibility in implementation and promotes program sustainability. To facilitate this transition, we will add teacher training sessions providing hands-on experience prior to their using U-Pods in the classroom. This model will result in a more resilient program and allow our team to take a facilitator/resource role and reach a greater number of students and classes across Colorado.

We are also aware that sustainability beyond the AWG grant period is another challenge to consider. Karp and Gale report in their paper on long-term program sustainability that a means for continuing beyond the initial grant was forming partnerships with established organizations (Karp & Gale et al., 2009). These partners were then able to assist the program in continuing after it had been built through the initial grant. This may be a model that would work for our project as well, and we will continue to explore our options.

We see this program as having the potential to benefit participants in a lasting way. This program provides students with a window into engineering, academic research, and even public and environmental health. For example, a group of students from one of the high schools in Delta County incorporated a U-Pod into their science fair project.

They used the U-Pod to compare by-products from the combustion of homemade biodiesel to the combustion of regular diesel fuel. This group was subsequently invited to the state science fair at Colorado State University (Ft. Collins, CO). At the state level they were awarded an honorable mention and scholarship money for college. Their teacher also reported that visiting CSU increased their interest in higher education. While in Ft Collins, the students were able to participate in a state-wide academic competition, attend classes, and tour the campus. With support from our team and access to resources like the U-Pod, the students were able to go from making biodiesel to assessing the impacts of using biodiesel – moving toward an understanding of the entire lifecycle of the fuel. This provided these students with experiences similar to those of environmental engineers, and also an idea of the supporting disciplines (e.g., mechanical engineering and computer science). Utilizing the program for a science fair project provide an example of a group taking the U-Pod further than the intended PBL curriculum, which is something we hope to see more of and support in the future. We hope that through opportunities such as this, increased access to university resources, or even simply by connecting them with researchers we can inspire students to consider the wide variety of options available post high school.

7. CONCLUSION

The NFV Project is the beginning of a promising K-12 outreach program. The project has evolved from a grassroots effort to work with rural Colorado students into a systematic project undergoing implementation in multiple school districts, funded by the National Science Foundation, and aiming for sustainability and longevity. We believe our unique model of a U-Pod 'check-out' program will ensure teachers are provided with access to resources and the necessary technological support and facilitate building partnerships that may lead to future collaborations and opportunities for our University. This program has the potential to build and sustain a community of learners using the U-Pods and inspire students to consider careers in science, engineering, and public and environmental health.

ACKNOWLEDGMENTS

We would like to acknowledge our funding source the National Science Foundation (NSF AWG-SRN Project, CBET: 1240584); the University of Colorado Boulder's Office of Outreach and Engagement, as well as the other members of the Hannigan Air Quality Lab (Joanna Gordon, Ricardo Piedrahita, Nicholas Masson, and Evan Coffey) and all of our community and educational partners throughout Delta County and St. Vrain Valley School District.

Preface to Chapters 5 & 6

The following two chapters passed peer review and were published prior to their inclusion in this dissertation (publication information is available following the titles). However, minor revisions have been made, to the versions of the articles contained in these chapters, based on the suggestions of my committee members.

Chapter 5: Comparing Building and Neighborhood-Scale Variability of CO_2 and O_3 to Inform Deployment Considerations for Low-Cost Sensor System Use

Contributing Authors: Ashley Collier-Oxandale[1], Evan Coffey[2], Jacob Thorson[2], Jill Johnston[3], & Michael Hannigan[2]

[1]University of Colorado Boulder, Environmental Engineering
[2]University of Colorado Boulder, Mechanical Engineering
[3]University of Southern California, Keck School of Medicine

Publication Information: Sensors, 18(5), 1349. DOI: 10.3390/s18051349, 2018

ABSTRACT

The increased use of low-cost air quality sensor systems, particularly by communities, calls for the further development of best-practices to ensure these systems collect usable data. One area identified as requiring more attention is that of deployment logistics, that is, how to select deployment sites and how to strategically place sensors at these sites. Given that sensors are often placed at homes and businesses, ideal placement is not always possible. Considerations such as convenience, access, aesthetics, and safety are also important. To explore this issue, we placed multiple sensor systems at an existing field site allowing us to examine both neighborhood-level and building-level variability during a concurrent period for CO_2 (a primary pollutant) and O_3 (a secondary pollutant). In line with previous studies, we found that local and transported emissions as well as thermal differences in sensor systems drive variability, particularly for high-time resolution data. While this level of variability is unlikely to affect data on larger averaging scales, this variability could impact analysis if the user is interested in high-time resolution or examining local sources. However, with thoughtful placement and thorough documentation, high-time resolution data at the neighborhood level has the potential to provide us with entirely new information on local air quality trends and emissions.

1. INTRODUCTION AND BACKGROUND

As research into and the use of low-cost air quality sensor systems continues to expand there is great potential for this technology to support community-level investigations. Furthermore, given the nature of these sensor systems, such investigations provide data with increased resolution on both temporal and spatial scales. Ideally, such sensor systems offer greater insight into personal exposure (Jerrett et al., 2017), small-scale variability (Sadighi et al., 2018),

and local emission sources or potential 'hot spots' (Mead et al., 2013). One of the barriers to widespread sensor use has been concerns over data quality and reliability. There is a growing body of research demonstrating the ability of sensors to quantify pollutants at levels relevant to ambient investigations (Piedrahita et al., 2014; Zimmerman et al., 2018; Cross et al., 2017; Kim et al., 2017). However, other issues have received less attention, for example, strategies for siting low-cost sensors. Sensor deployment and siting considerations are particularly important because while it is sometimes possible to re-analyze or re-quantify sensor data as new techniques become available, it is rarely possible to re-collect data as environmental conditions and emissions impacting a site are dynamic in nature. Careful consideration prior to and documentation of the sensor siting process could not only aid in data processing and interpretation, but also help to ensure the collection of useful and relevant data.

Previous studies have demonstrated that pollutant variability can exist on small spatial scales utilizing six-minute or hourly averaged data. Variability has been observed across a street or within a few hundred meters, especially in more complex urban environments (Vardoulakis et al., 2005; Croxford et al., 1998). Therefore, it is reasonable to expect that where the sensor is placed on a building could influence the data collected. While there are strict guidelines for siting a Federal Equivalent Method/Federal Reference Method (FEM/FRM) monitor, no such guidelines exist for low-cost sensor systems. Further complicating the matter, siting a sensor system at a home or business can be challenging as convenience, safety, and aesthetics are all factors in the decision rather than simply the most ideal placement for accurate measurements. As guidelines for sensor system deployment are established, it would be valuable to both examine existing literature and learn from field deployments. For instance, in addition to guidelines for siting FRM/FEM monitors for regional-scale monitoring, the US EPA offers some considerations for monitoring at the urban (4-50 km), neighborhood (0.5-4 km), middle (100 m - .5 km), and micro (< 100 m) scales (US EPA, 2017d). Given that many sensor studies occur in cities, modeling studies examining how urban landscapes and buildings influence air flow and pollutant transport could also assist with strategic sensor system placement at a site (Hang et al. 2012; Vardoulakis et al., 2003). Finally, studies on near-source measurements or small-scale variability could be valuable examples, some of these utilize combined monitoring/modeling approaches and many specifically investigate near-roadway emissions and trends (Clougherty et al., 2013; Vardoulakis et al., 2005). Similar to regulatory monitoring and conventional research, communities interested in using sensors must define the scale and objective of their monitoring, whether is to compare local emission sources, understand potential exposures at a neighborhood

level, or characterize air quality across their city. Here we used a short deployment to better understand the effects of certain sensor system placement choices, which could then be used to advise communities based on their objectives.

A recent study in New Zealand, also utilizing a field deployment of sensor systems, thoroughly examined specific aspects of this question by determining the intra-site variability of paired O_3 sensors and the impact of siting conditions/type on the overall dataset (Miskell et al., 2017). Miskell and colleagues found that most factors examined, such as placement at a site (i.e., on a wall, balcony, or roof), land coverage beneath the instrument, or land-use designation at the site, had little impact on the observed intra-site variability (2017). Two factors–exposure to direct sunlight likely causing temperature differentials between paired sensors and local emission events–resulted in the greatest intra-site variability (Miskell et al., 2017). The team concluded that networks of O_3 monitors set up by citizen scientists can supplement existing reference networks and provide new information, as limited variability was introduced due to siting choices and this variability was minimal over typical reporting scales (e.g., hourly or 8-hour averaged data) (Miskell et al., 2017).

This study by Miskell and colleagues provided a comprehensive example of how sensor systems can support existing monitoring networks for O_3 and the impacts of siting choices in this context (2017). However, it is possible that communities may wish to use sensors for the collection of high-time resolution data on smaller spatial scales rather than the larger averaging times and regional scales studied by Miskell and colleagues (2017). To explore the impact of siting choices in this alternate context we undertook a small case study during a larger deployment of sensor systems in Los Angeles, CA, USA. We added four additional sensor systems to one sampling site to observe the variability across several sensors on one building. We compared this building-level variability to the neighborhood-level variability. This analysis includes data from both metal oxide O_3 sensors and non-dispersive infrared CO_2 sensors–providing the opportunity to examine a primary and a secondary pollutant. The differing spatial scales (neighborhood vs. regional) and higher temporal resolution (utilizing primarily minute-median data) as well as the addition of CO_2 data offers a small, complementary dataset providing additional information to inform recommendations for siting practices.

Furthermore, while there currently exist several valuable resources contributing best practices and supporting community-based investigations using low-costs sensors, such as the US EPA's Air Sensor Guidebook (Williams et al., 2014), South Coast Air Quality Management District's Sensor Performance and Evaluation Center (SCAQMD, 2015), and the Environmental Defense Funds Air Sensor Work Group (EDF, 2018), additional case studies examining

the questions of best-practices in different contexts will support the development of recommendations appropriate for the variety of uses likely to emerge. This need for more standards to guide all aspects of sensor use from planning to deployment to data analysis and interpretation has been cited as critical by academic, community-based, and regulatory researchers (Clements et al., 2017).

2. MATERIALS & METHODS

2.1. Deployment Overview (Sensor Systems, Siting, and Timeline)

The sensor systems utilized for this study, called Y-Pods (Hannigan Lab at CU Boulder, Boulder, CO, USA), contain several gas-phase and environmental sensors. This analysis utilizes data from the SGX (Corcelles-Cormondreche, Switzerland, formerly e2v) metal oxide semiconductor O_3 sensors (model MiCS-2611) and ELT non-dispersive infrared CO_2 sensors (model S-300) as well as data from environmental sensors (i.e., temperature and relative humidity). These sensor systems, or similar ones (e.g., the U-Pod, predecessor to the Y-Pod) operating the same sensors, have been used in prior sensor quantification and spatial variability studies (Sadighi et al., 2018; Casey et al., 2018a; Cheadle et al., 2017; Masson et al., 2015b; Collier-Oxandale et al., 2018a). Figure 1 includes a photo of the interior of a Y-Pod and an example of two deployed Y-Pods. The Y-Pods, and all previous iterations, include a fan to drive active air flow resulting in multiple air exchanges per minute. The observations presented here would likely need to be re-evaluated for a system relying on passive flow. More information on signal processing and sensor performance quantification is available below in Section 2.2.

As previously mentioned, this study was integrated into a larger field deployment in Los Angeles allowing us to leverage one of the existing study sites and ongoing sensor calibration efforts. The study area is primarily high density residential with schools and some businesses nearby. In addition to local traffic and businesses (such as restaurants) other emission sources include two major highways to the North and East of the sampling area. The diagram in Figure 1 illustrates where the Y-Pods (B2, B3, B4, and B5) were added to the building site (main sensor system–B1). Note, the placements vary with respect to elevation and proximity to obstructions – it was the objective of the study to compare a variety of placements, including non-ideal locations. Two Y-Pods were placed on the front of the building on a fire escape, two and three stories off the ground, and 6–12” from the side of the building.

The fire escapes at the front and back of the building are both constructed of metal and allow for free airflow through and around the structures. The main Y-Pod was elevated on the roof, on top of a structure housing the stairs,

close to the front of the building, and with no obstructions on any sides. The fourth and fifth Y-Pods were placed at the back of the building on another fire escape, one at the roof-level and the other three stories off the ground, again 6–12” from the side of the building. The back of the building is obstructed by a narrow alley that does not allow through-traffic; the lack of access to representative air flow makes the placement of B5 the least “ideal”.

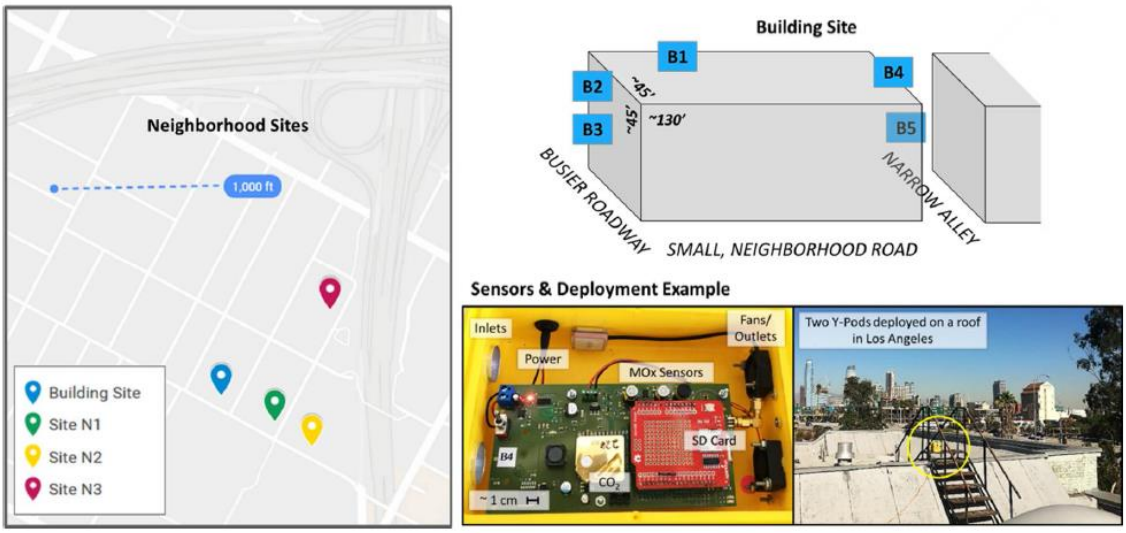


Figure 1: The map (left) indicates the sampling sites relevant to this paper, the diagram (top right) indicates where sensor systems were deployed at the Building Site, and the photos (bottom right) show the inside of a Y-Pod and a deployed Y-Pod.

Figure 1 also illustrates the location of several other neighborhood sites from which data was used in this analysis (N1, N2, and N3). These sensor systems were deployed on a relatively small scale with the furthest distance between any two neighborhoods sites being less than 1000 ft. It is important to note that the placement of N1, N2, and N3 at their respective sites also introduces some added variability as these placements differed site to site. The Y-Pod placement for N1 was most similar to B3 on a large second story balcony, on the side of a building open to the road. The Y-Pod placement for N2 was also most similar to B3–at the front of the building, on the street side, but set back by a small yard/driveway and lower in elevation (~10 ft off the ground). The Y-Pod placement for N3 was most similar to B1, placed on the roof of a multi-family residence.

This study relies on comparing co-located sensor data with spatially deployed sensor data, therefore we limited the data utilized to match the lengths of our co-located datasets meaning approximately three weeks of data were included in the analysis. Figure 2 shows the timeline of long-term sensor use, including time periods of co-location and periods of field deployment. The co-location of all sensor systems prior to the field deployment was used to understand neighborhood variability; this co-located time period is referred to as Week 0. The Week 0 co-location

occurred in a different part of Los Angeles at a regulatory monitoring site; this site is described in greater detail below in Section 2.2.2. For the first week of the building-scale variability study, the building Y-Pods (B2, B3, B4, and B5) were co-located with B1–this is referred to as Week 1. During this period the neighborhood Y-Pods (N1, N2, and N3) were already deployed to their field sites. Immediately following the first week of the field deployment the sensor systems were separated to their respective locations on the building and this is referred to as Week 2. The data from Week 2 was designated as the deployed dataset for both the neighborhood sites and the building sites.

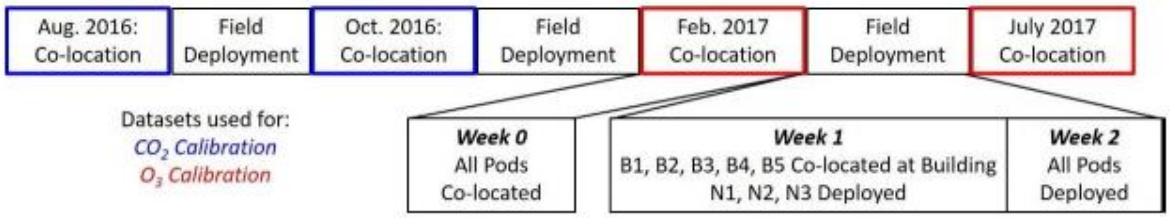


Figure 2. Timeline showing when co-location of sensors with reference instruments occurred and when deployments to field sites occurred.

2.2. Signal Processing and Sensor Quantification

Sensor signals were saved to a text file on a micro-SD card on the Y-Pod every 6–25 s, depending on the programming. As some of the metal oxide sensors used here require a warm-up period, the first half hour of data after a pod has been powered off for half an hour or more was removed. Minute medians were computed; using medians instead of averages removes any single extreme points likely the result of electronic noise. For both the CO_2 and O_3 sensors, voltage values were recorded to the SD card as ADC values. These voltages were used as is for the CO_2 sensor, but for the O_3 sensor they were converted to a normalized resistance prior to analysis (Sadighi et al., 2017; Piedrahita et al, 2014; Masson et al., 2015a). Note, all of the datasets for Weeks 0, 1, and 2 are complete with the exception of the O_3 data from Y-Pod N3, on which the O_3 sensor appears to have malfunctioned. Thus, this data has been excluded from the analysis.

Sensor signals were converted to concentrations using field calibration, which involves: (1) colocation with high-quality reference instruments; (2) the development of a calibration model using the air quality sensor signals, environmental sensor signals, and trusted reference data as well as a technique such as multiple linear regression; and (3) the evaluation of that model and its application to testing or validation data. Ideally; the sensors are co-located before and after the field deployment to better facilitate corrections for drift. It is common to incorporate environmental parameters into these calibration models as low-cost sensors are often cross-sensitive to temperature, humidity, and sometimes other pollutants (Cross et al., 2017). Note, each calibration model is specific to each Y-Pod,

utilizing only sensor signals from that board. This method of sensor performance quantification has been used by our research group as well as others (Sadighi et al., 2017; Spinelle et al., 2015a; Spinelle et al., 2017a) and with techniques such as linear regression, multiple linear regression, and machine learning (Zimmerman et al., 2018; Casey et al., 2018c). Details of the calibration employed here are presented below.

2.2.1. Quantification of CO_2 Sensor

For CO_2 sensor quantification, the Y-Pods were twice co-located with a LI-840A (Licor, Lincoln, NE, USA) placed at a regulatory monitoring site near downtown Los Angeles. The Licor LI-840A has an expected uncertainty of <1% of the reading as stated by the manufacturer, and the instrument is calibrated using a zero and two-point span calibration with gas standards. The Licor used in this study was calibrated prior to a deployment during the previous summer and was stored between these deployments. As a result of the time lag, we expect drift to have impacted the CO_2 reference data. However, as we are interested in sensor to sensor comparisons and the sensor data is baseline shifted (as described below), this drift is of minimal concern. These two co-locations with the Licor were 8 weeks apart and included 17 days total, 12 of which were used for calibration model training and 5 of which were used for model testing. In this instance more of the co-location data was designated for training in order to increase the robustness of the model and expand the environmental conditions for which the model was trained. The model used, Equation (1), included predictors for temperature (Temp), absolute humidity (AH), time (t), and the sensor signal or voltage (V) and solves for the CO_2 concentration (C):

$$C = (p_1 + p_3 * Temp + p_4 * AH + p_5 * t - V) * (-1/p_2) \quad (1)$$

Due to logistics and a lack of available reference data, both calibration co-locations occurred prior to the building-scale variability study (Figure 2). For this reason, further signal processing was necessary. Given that the CO2 calibration model is extrapolating in time, additional drift was expected. For this reason, the CO_2 data was converted using the calibration model and then this data was baseline corrected (to remove drift), and finally the 10^{th} percentile value from each Y-Pod was normalized to 400 ppm. We selected 400 ppm as it is the approximate atmospheric background concentration of CO_2 (NOAA, 2018). In light of the goals of this case study–comparing relative differences across co-located verses deployed sensors–this additional processing was deemed reasonable. Furthermore, the results illustrate the high correlation and agreement between co-located sensors post-processing as would be expected and is also present in the calibration data (Appendix IV).

2.2.2. Quantification of the O_3 Sensors

For O_3 sensor quantification, the Y-Pods were co-located with API/Teledyne 400 instruments (San Diego, CA, USA) at two different regulatory monitoring sites. The first site was in Los Angeles in a mixed-use area with some nearby housing and industry. The second site was outside of Los Angles in Shafter, a rural Californian community. These two co-locations occurred prior to and following the building-scale field deployment and therefore no additional signal processing was necessary. The model, Equation (2), used included predictors for temperature (Temp), absolute humidity (AH), time (t), the normalized sensor resistance (R/R0), as well as an interaction term between temperature and concentration, and solves for the O3 concentration (C). The interaction term is intended to address not only changes in baseline driven by temperature but changes in the magnitude of sensor response driven by temperature. This model has been demonstrated as well performing for this sensor in previous studies (Sadighi et al., 2018; Cheadle et al., 2017):

$$C = (p_1 + p_3 * T + p_4 * AH + p_5 * t - R/R0) * (-1/(p_2 + p_6 * T)) \quad (2)$$

3. RESULTS AND DISCUSSION

3.1. Field Calibration Results (Sensor System Uncertainty)

Table 1 below provides the performance statistics from the generation and validation of the calibration models. The complete statistics for individual Y-Pods as well as time series data are available in Appendix IV. For both CO_2 and O_3, there is relative consistency across the training and testing datasets. Additionally, the RMSE for the O_3 sensor was consistent with uncertainty typically cited for both this same sensor and other metal oxide O_3 sensors (Sadighi et al., 2018, Cheadle et al., 2017; Bart et al., 2014). A previous study using the CO_2 sensor in a portable sensor system found a RMSE ranging from approximately 9–16 ppm depending on the calibration model selected (Piedrahita et al., 2014).

Table 1: Performance Statistics as Compared to Reference Datasets (Averaged for all Y-Pods).

	Statistic	Training	Testing
CO_2	R^2	0.92 (0.03)	0.89 (0.06)
	RMSE (ppm)	8.33 (1.71)	10.09 (3.16)
	MB (ppm)	−0.02 (0.02)	3.89 (5.95)
O_3	R^2	0.97 (0.01)	0.94 (0.02)
	RMSE (ppb)	3.65 (0.42)	5.28 (0.86)
	MB (ppb)	−0.09 (0.02)	−2.30 (0.79)

R^2—coefficient of determination, RMSE—root mean squared error, MB—mean bias, with standard deviations in parentheses.

3.2. Neighborhood-Scale Variability

Comparing Week 0 (co-located) to Week 2 (deployed to field sites), there is increased variability in both the CO_2 and O_3 data. For CO_2, this variability is most extreme in the comparison between B1 and N3, which was also the site furthest away from B1 and closest to the highways. For this pair of sensors, the correlation decreases from 0.96 to 0.89 and the spread in the absolute differences as well as the median absolute difference increases, see Figure 3. This is not the case for the comparisons of B1 to N1/N2 where there is only a very small decrease in correlation. Examining the time series plots (available in Appendix IV) reveals differences in the variability seen in Week 0 versus Week 2. For Week 0 the variability seems primarily driven by offsets in which one Pod is biased low or high for a period, whereas for Week 2, the variability seems driven by differences in trends between the sites typically in the form of short-term enhancements. These enhancements present in the Week 2 data are likely sources or plumes impacting the sites unevenly.

For O_3, spatial variability across field sites was much more apparent. Although there was little change in the correlation coefficient, there was an increase in the spread in both the scatterplot and the boxplot (Figure 4). For Week 0, nearly all the absolute differences between B1 and N1/N2 were below the expected uncertainty (RMSE = 5.28 ppb). For Week 2, after the Y-Pods were spatially deployed the spread increased to well above the RMSE, see Figure 4. The time series plots (Appendix IV) confirmed that this increased variability was primarily driven by short-term dips in O_3 likely caused by localized destruction occurring in a NO_x plume. While it is possible that the differences in increased variability between the sensor types were in part due to CO_2 being a primary pollutant (thus less well-mixed) and O_3 a secondary pollutant (generally more well-mixed), it is also worth noting that the CO_2 sensor has a lower signal/noise than the O_3 sensor in this application. Another possible explanation for O_3 depletion could be loss to surfaces. This is important to consider as the Y-Pods were placed in close proximity to either the roof (which was a

sealed, finished roof) or the sides of the building (which were brick). Studies have illustrated how different surfaces drive different rates of ozone deposition (Simmons & Colbeck, 1990; Shen & Gao, 2018). Though loss of ozone to surfaces would most likely be expressed in the data as a gradual, continual decrease rather than quick depletions and recoveries.

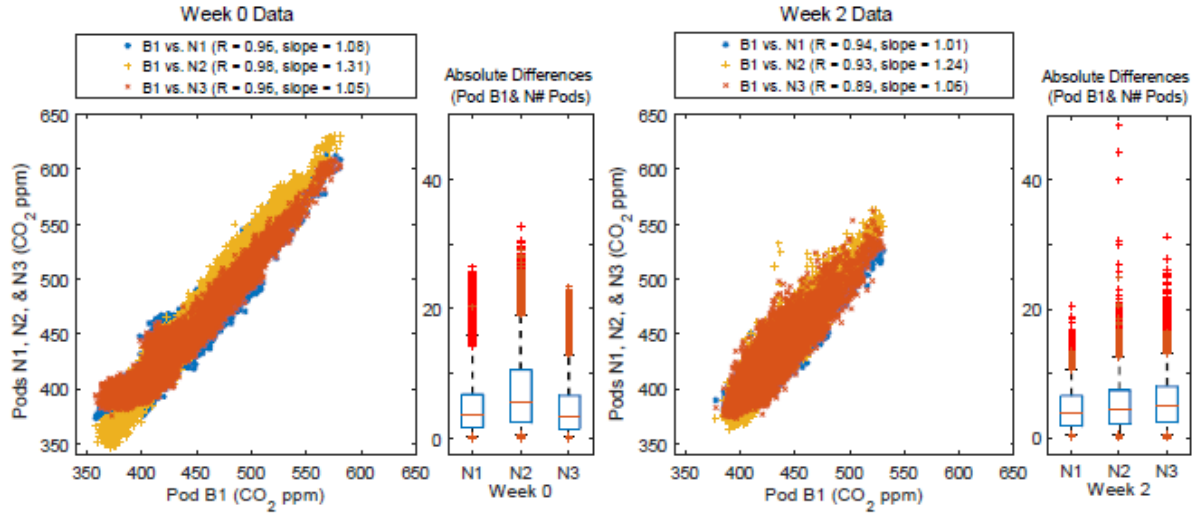


Figure 3: Scatter plots showing each neighborhood Y-Pod (N1, N2, and N3) vs. Y-Pod B1 for Weeks 0 and 2. The boxplots show the absolute differences between B1 and each of the neighborhood pods, with the whiskers at the 5th and 95th percentile respectively. Again, the data was shifted, placing the 10th percentile value at 400 ppm.

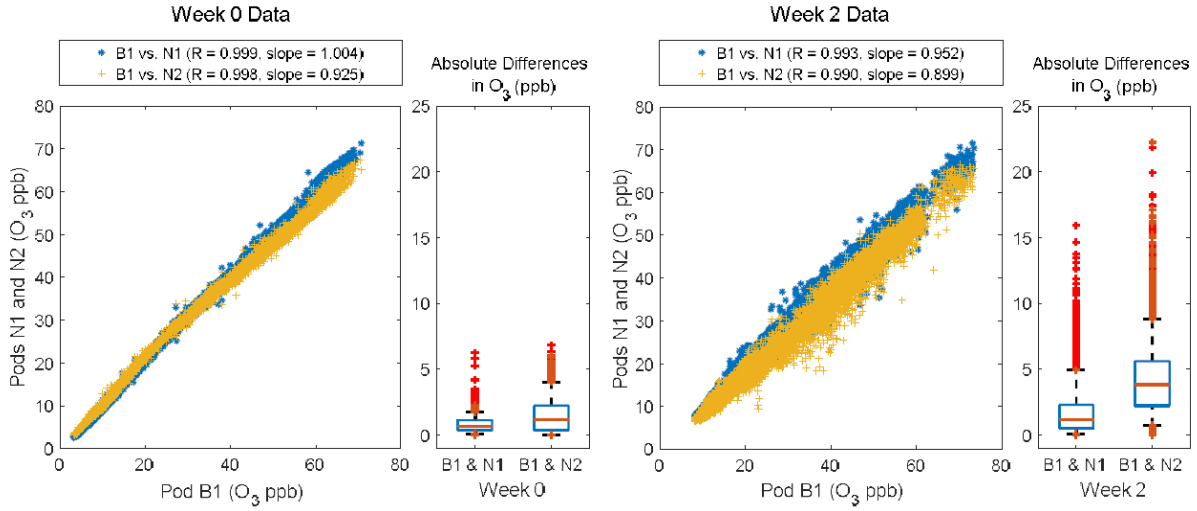


Figure 4: Scatter plots showing each neighborhood Y-Pod (N1, and N2) vs. Y-Pod B1 for Weeks 0 and 2. The boxplots show the absolute differences between B1 and each of the neighborhood pods, with the whiskers at the 5th and 95th percentile respectively.

3.3. Building-Scale Variability

Somewhat surprisingly spatial variability was also observed at the building-level for both sensor types when comparing Week 1 (co-located at the building) and Week 2 (deployed). For CO_2, there was a decrease in the

correlations on the same scale as occurred across some of the neighborhood sites (Figure 5). For O_3, again there are no significant changes to the statistics, but there is an increase in spread (Figure 6), similar to Figure 4. The time series (Appendix IV) showed the events driving these differences were short-term in nature and appeared to be driven by local emissions or transported plumes. This influence of nearby emissions events was observed by Miskell and colleagues as well (2017).

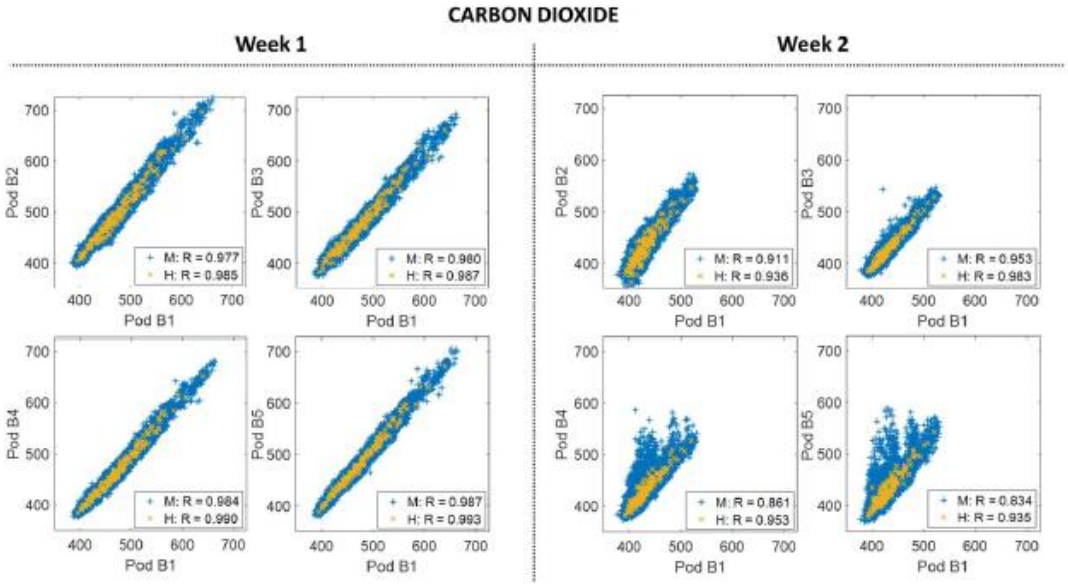


Figure 5. The scatter plots to the left show the correlation between Y-Pod B1 and each of the added building Y-Pods (B2, B3, B4, and B5) for both minute (M) and hourly (H) CO_2 data for Week 1 (co-located). The scatter plots to the right show the same correlations, again with minute and hourly data, but for Week 2.

Hourly-averaged data was added to both Figure 5 and 6 to determine whether this variability impacted data on more typical temporal reporting scales. Similar to Miskell and colleagues, the variability does not seem to impact the hourly O_3 data (2017). However, given the decreased correlation coefficients (particularly for sites B2 and B5), it appears there was some variability still present in the hourly-averaged CO_2 data.

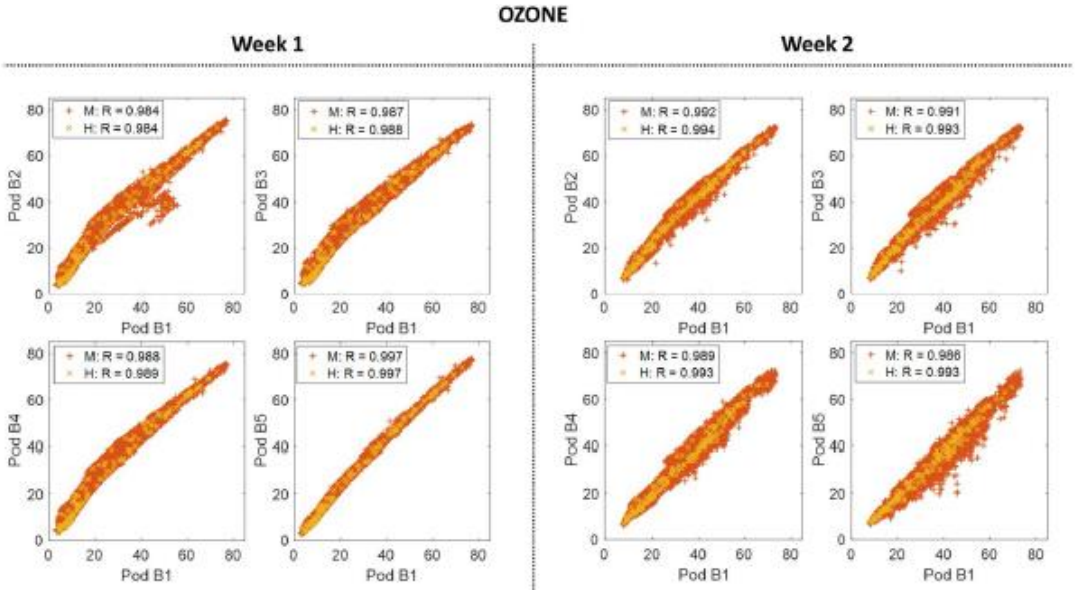


Figure 6. The scatter plots to the left show the correlation between Y-Pod B1 and each of the added building Y-Pods (B2, B3, B4, and B5) for both minute (M) and hourly (H) O_3 data for Week 1 (co-located). The scatter plots to the right show the same correlations, again with minute and hourly data, for Week 2 when they were spatially deployed around the building site.

For both pollutants, the most dramatic differences were between sites B1 and B4/B5, the two sites at the back of the building. Speaking with community partners from the project we determined that the building has both a natural gas hot water heater and natural gas dryers toward the back of the building where there are also pipes that appear to be venting these emissions. Sources on the building would seem to explain the large magnitude of the observed variability. By comparison, for the sites B2 and B3, which were on the front of the building above the road, there were occasional increasing spikes for CO_2 and decreasing spikes for O_3 that are smaller in magnitude. The range of responses observed in the sensors, along with this contextual information affirms that multiple pollutant sources were impacting the building in an uneven manner.

Providing further evidence for multiple sources, Figure 7 includes the absolute differences between Y-Pod B1 and B5 for CO_2 (in blue) and O_3 (in red). There are periods where the differences between CO_2 and O_3 were well-correlated indicating a shared source. Following this period were instances where the differences were primarily visible in one pollutant or the other. This lack of correlation likely indicates two separate sources, one with relatively more CO_2 and another with more NO. Furthermore, there were many instances where these differences between the two building sites were well above the RMSE values. In Figure IV.6 (Appendix IV) the spatial differences have been plotted in such a way as to highlight the temporal aspect of both the increases in CO_2 and decreases in O_3 at the B5 site. The correlation between differences in CO_2 and O_3 occur primarily in the evening hours, while the uncorrelated

periods result in enhancements during early morning and daytime hours. These temporal patterns also point to separate sources influencing sensor data.

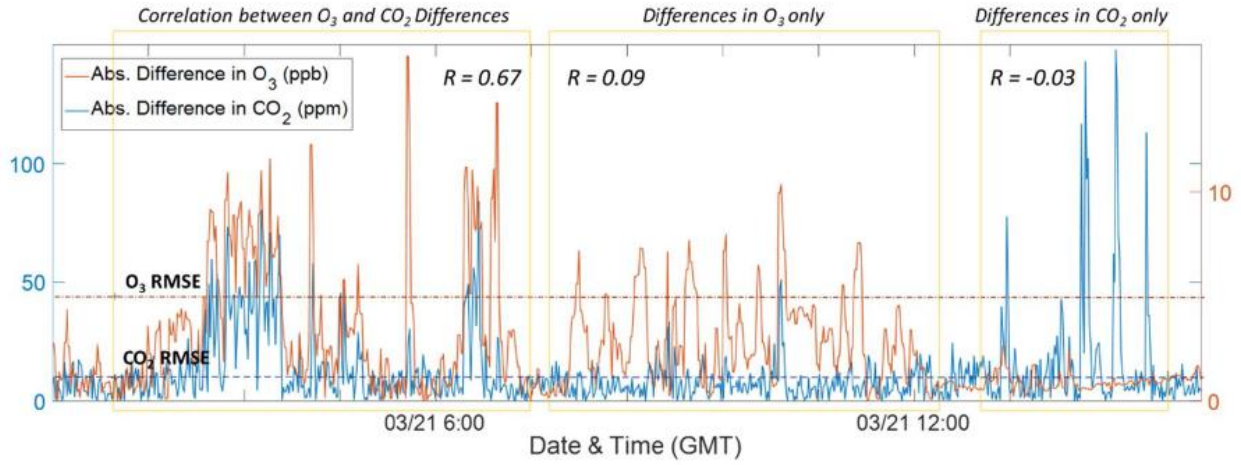


Figure 7: Time series of absolute differences between Y-Pod B1 and Y-Pod B5 for CO_2 (blue) and O_3 (red), the RMSE for both the CO_2 and O_3 sensors are indicated using dotted lines. The yellow boxes highlight periods where the differences in the two signals are well-correlated verses periods where the differences are occurring primarily in the CO_2 or O_3 signal. The correlation coefficient (R) has been added to contrast the different periods.

In addition to nearby emission events, Miskell and colleagues observed that direct sunlight causes thermal variations in the instruments causing variability (2017). We compared the internal temperatures in the Y-Pods to determine whether this could be a source of variability in our study as well. Figure 8 depicts the variability in light of temperature differences. Again, B1 was placed on a roof with no nearby obstructions meaning that it was exposed to more direct sun than B5, which was placed on a fire escape in an alley. In Figure 8, the internal temperature differences, between B1 and B5, less than three degrees Celsius were plotted separately from differences greater than three degrees Celsius. The line of best fit for the group with larger temperature differences (in yellow) illustrates a consistent bias in the data at low and high concentrations. This bias is visible in the time series as well, the B1 values are consistently greater than the B5 values when the temperature difference is above three degrees. Conversely, B1 and B5 are better matched in terms of long-term trends for smaller temperature differences. Although the calibration model does incorporate corrections for temperature effects, the model would be unable to account for the small differences driven by direct sunlight exposure as this would be difficult to control during co-location. The corrections incorporated into the calibration model are intended to deal with less acute temperature effects (e.g., diurnal patterns). One other possible explanation for this variability, mentioned previously, would be ozone loss to surfaces. It is possible that the different surfaces of the roof and walls of the building drive deposition at different rates. Although, Miskell and colleagues did

examine the impact of land coverage (i.e., grass, gravel, deck) below the instrument and did not find that this variable had a significant affect in intra-site variability (2017).

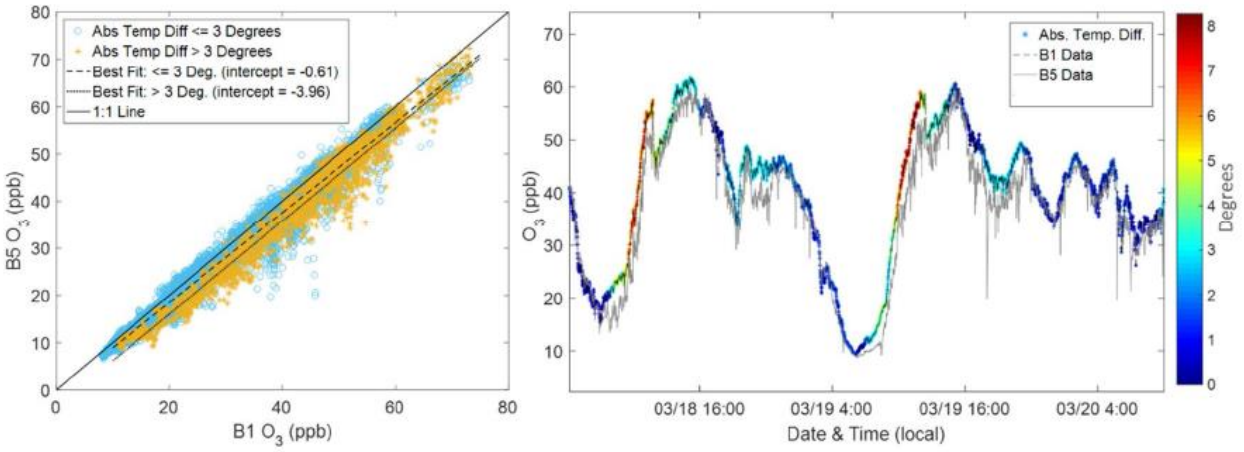


Figure 8: Two plots illustrating the effect of temperature differences between the pods. The scatter plot (left) depicts B1 vs. B5, separating points where the temperature difference between the two pods is less than and greater than three degrees Celsius. The time series (right), shows two days of data from B1 and B5 where the B1 data also has an overlay of temperature differences between the pods.

Siting choices and additional shading for the sensor systems could reduce this variability. Although some of the variability between building-sites can be attributed to thermal differences, it is important to recall that this variability is displayed as a bias rather than the larger spread associated with the variability driven by nearby emissions. Therefore, this variability would be unlikely to affect any conclusions about spatial differences due to sources in the same way the short-term enhancements would when examining high-time resolution data.

3.4. Impact of Siting Choices on Neighborhood Variability Analysis

In agreement with the findings of Miskell and colleagues, we have observed that local emissions or plumes can drive intra-site variability as well as temperature differences caused by exposure to direct sunlight (2017). Also, as with the previous study, this spatial variability does not impact O_3 concentrations on typical reporting scales (hourly or eight-hour averages for example). However, the same is not necessarily true for CO_2 suggesting it may be valuable to further investigate this aspect of variability for primary pollutants. The spatial variability observed here becomes especially important for communities interested in high-time resolution data, which may be used to assess exposure and/or understand the impact of local emission sources within a neighborhood. When high temporal and spatial resolution is of interest, incorrect placement could result in the inappropriate attribution of sensor responses or failing

to record emissions that are present. Figure 9 includes several days of data demonstrating the large magnitude of differences that can be observed across a single site.

To further explore the impact of the building-scale variations on the community-scale spatial differences, Figures 10 and 11 depict the average of the neighborhood sites with one building site selected and assumed to be representative for that location. The shading on the plot indicates the standard deviation for each mean. For the first case, in blue, Y-Pod B5 was selected as the building site Pod and for the second case, in red for minute median and green for hourly averaged data, B1 was selected. Similar to the previous comparison, there are minimal differences between the hourly O_3 datasets and only a few instances in the hourly CO_2 data where the mean of the B5 dataset differs beyond the standard deviation of the B1 dataset. However, examining the minute-median data for either pollutant, one might draw different conclusions regarding the neighborhood variability depending on which building site was selected. For example, one might anticipate more variability with B5 selected, or fewer local sources capable of scavenging O_3 with B1 selected. If examining the maximum daily CO_2 concentrations, the results for several days would differ. Regardless of which building site is selected, the diurnal trends are consistent potentially providing an indication of regional trends. Also, for the minute data, this difference between the datasets is more extreme for the CO_2 data possibly due to CO_2 being a primary pollutant and less well-mixed in the atmosphere.

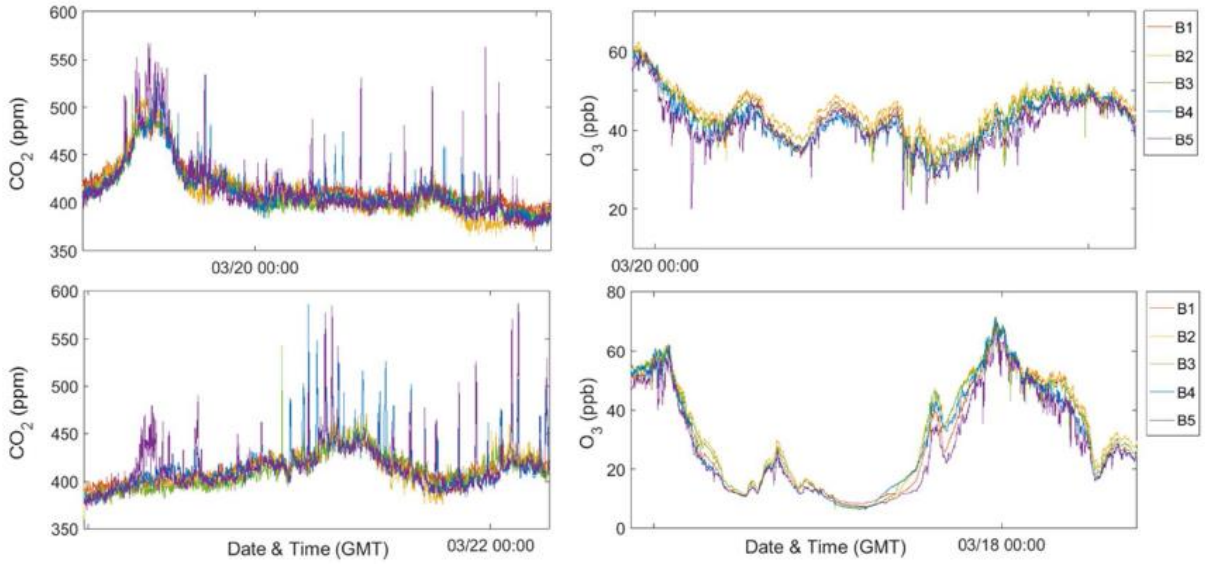


Figure 9: Time series of CO_2 data (top and bottom left) and O3 data (top and bottom right) each showing approximately one-day of data from the building sites during the Week 2 period.

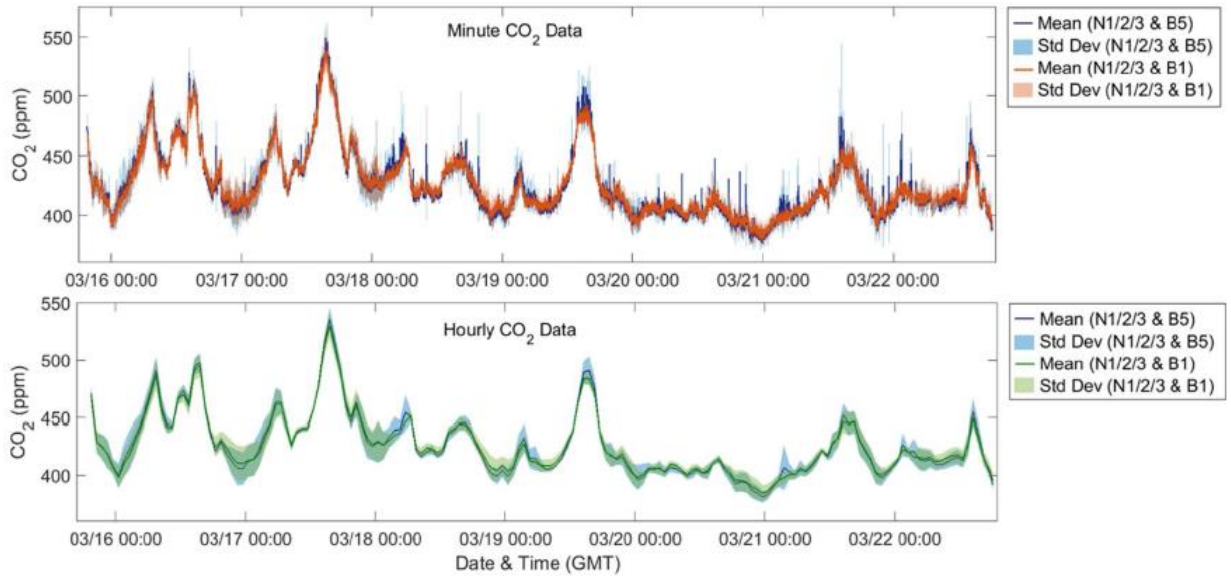


Figure 10: Time series of CO_2 data (top: minute-median, bottom: hourly-averaged), showing the mean and standard deviation of different sets of Y-Pods.

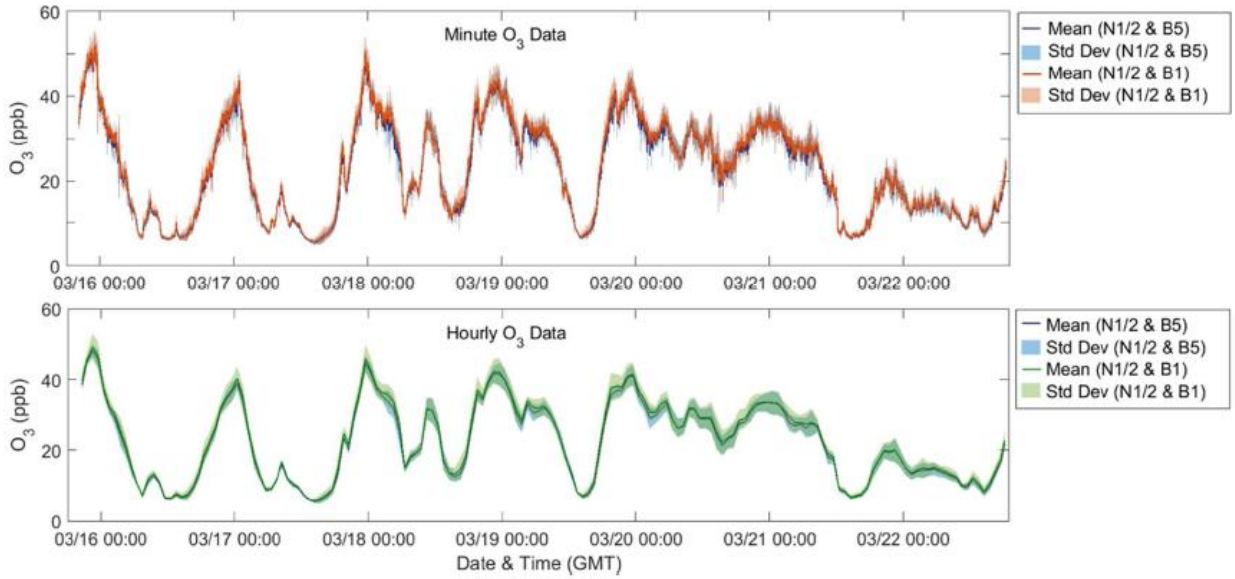


Figure 11: Time series of O_3 data (top: minute-median, bottom: hourly-averaged), showing the mean and standard deviation of different sets of Y-Pods.

3.5. Generalizability of Building-Scale Spatial Variability & Potential Recommendations

There are a few aspects of this study that limit generalizability: we used short periods of data, we only examined the variability around one building in Los Angeles (variability might look different around a different structure or in a different city), and the two sensors types we used rely on different operating principles. Given these limitations, there are still recommendations based on this analysis that can be made. As the following recommendations are intended for individuals or groups interested in conducting sensor studies, more general "best

practice" recommendations have been included as well. While some of these are more general, specifically the first and fourth ones, the results of the study nonetheless affirm their value. Furthermore, these suggestions complement the US EPA's existing recommendations for planning a study and siting sensors (Williams et al., 2014); though it would also be beneficial to supplement these with EPA recommendations for siting for regulatory and research studies (US EPA, 2017d). These recommendations are tailored to low-cost sensor studies involving high-time resolution data on a neighborhood or source-scale:

- Compare Sensors: Co-locating sensors in the field will support a better understanding of inter-sensor variability prior to their deployment, which will aid in attributing new variability introduced by the deployment of sensor systems to separate sites. These relative comparisons can also be valuable if there are problems with the calibration.
- Placement and Distribution: To study a particular emission source, place sensors upwind and downwind of the site of interest, at varying distances. Some of the sensors should have a line of sight to the emission source. Consider factors such as typical wind directions and potential obstructions, which may impact the transport of emissions. These placements should also minimize added variability when possible. For example, shading all sensor systems, placing them on the same sides of buildings, or placing them exclusively on rooftops could reduce the variability and biases that result from occasional direct sunlight.
- Supplementary Sensor Data: Consider using multiple systems or sensor types. The ability of sensors to capture variability on small spatial scales could be leveraged to aid in source identification by placing multiple sensor systems at a site with the objective of capturing local emissions with some systems and targeting exclusively regional trends with other systems. Leveraging data from multiple sensor types could also shed light on sources and emissions by studying the correlations or temporal patterns of data from sensors intended to measure different target pollutants.
- Document Deployment: Document your deployment in writing and with photos (take photos of the sensor systems from different angles and photos from the sensors of what they "see"). Learning about nearby activities could provide contextual information that can aid in data interpretation and reduce the misinterpretation of sensor data.

4. CONCLUSIONS

This deployment demonstrated how the variability in CO_2 and O_3, measured using low-cost sensors, across a single sampling site can be comparable to the variability across several sites in a neighborhood. However, this spatial variability occurs primarily in high-time resolution (<1 h) data as it seems to be driven by nearby emission plumes and occasional thermal differences. As Miskell and colleagues reported these differences do not persist at typical reporting scales (Miskell et al., 2017), but if a researcher or community is interested in high-temporal resolution data then this variability could become significant. This variability might also be more important to consider for studies taking place on smaller spatial scales, such as the neighborhood scale at which this study takes place, rather than larger regional scales.

While minute-level data is not currently utilized for regulatory purposes, this level of data can provide powerful preliminary and supplementary information when it comes to understanding the activities and experiences in a community and at local scales. Furthermore, the presence of building-level variability does not exclude sensors from being used in air quality investigations, but rather affirms their ability to detect these differences in trends. Through attention to siting and thorough planning/documentation, there is the potential to make more accessible the collection of data that could for example, inform detailed investigations into the impact of a single source on a neighborhood, track the transport of emissions through an area, or clarify the acute effects of brief, high-concentration exposures. These potential applications suggest that this new type of data, made possible by sensors, could eventually support improved public health.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (under grant numbers CNS-1446912 and CBET-1240584) and in part by the National Institutes of Environmental Health Sciences (under grant number: R21ES027695). We would like to thank project partners: Esperanza Community Housing, Sandy Navarro, partners at the University of Southern California and Occidental College, and all community member participants. We would also like to thank our regulatory partners for assistance with sensor co-locations, including access to reference monitoring stations and associated data: South Coast Air Quality Management District, San Joaquin Valley Air Pollution Control District, and California Air Resources Board (note regarding reference air quality data: this data has not passed through the normal review process, and is therefore not QA'd and is unofficial data).

Chapter 6: Assessing a low-cost methane sensor quantification system for use in complex rural and urban environments

Contributing Authors: Ashley Collier-Oxandale[1], Joanna Gordon Casey[2], Ricardo Piedrahita[3], John Ortega[4], Hannah Halliday[5], Jill Johnston[6], and Michael P. Hannigan[2]

[1]University of Colorado Boulder, Environmental Engineering
[2]University of Colorado Boulder, Mechanical Engineering
[3]Berkeley Air Monitoring Group
[4]National Center for Atmospheric Research
[5]NASA Langley Research Center
[6]University of Southern California, Keck School of Medicine

Publication Information: Atmospheric Measurement Techniques, 11, 3569–3594. DOI: 10.5194/amt-2017-421, 2018

ABSTRACT

Low-cost sensors have the potential to facilitate the exploration of air quality issues on new temporal and spatial scales. Here we evaluate a low-cost sensor quantification system for methane through its use in two different deployments. The first was a 1-month deployment along the Colorado Front Range and included sites near active oil and gas operations in the Denver-Julesburg basin. The second deployment was in an urban Los Angeles neighborhood, subject to complex mixtures of air pollution sources including oil operations. Given its role as a potent greenhouse gas, new low-cost methods for detecting and monitoring methane may aid in protecting human and environmental health. In this paper, we assess a number of linear calibration models used to convert raw sensor signals into ppm concentration values. We also examine different choices that can be made during calibration and data processing and explore cross sensitivities that impact this sensor type. The results illustrate the accuracy of the Figaro TGS 2600 sensor when methane is quantified from raw signals using the techniques described, with an uncertainty of 0.38 ppm observed in the Colorado deployment and 0.22 ppm observed in the Los Angeles Deployment. These results also highlight the need for further research into the cross-sensitivities that affect these sensors. Although, considering the limitations, these tools seem to be able to reveal air quality trends and events on small spatial and temporal scales – speaking to their potential to provide preliminary data that can inform more targeted measurements or supplement existing monitoring networks.

1 INTRODUCTION

1.1 Background and motivation

Given both the direct impacts on climate change and indirect impacts on human health, it is important to study increased atmospheric methane on varied temporal and spatial scales. Methane is an important greenhouse gas with 28 times the global warming potential of CO_2 over a 100-year lifetime (IPCC, 2015); moreover, the majority of methane emissions result from human activity (US EPA, 2017e). Researchers using ice core samples to measure historic methane levels found relatively stable atmospheric concentrations of approximately 0.695 ppm from 1000 AD until the Industrial Revolution (Etheridge et al., 1988), after which methane concentrations have grown to a present-day global average of 1.851 ppm (NOAA, 2017). This increased atmospheric methane not only intensifies climate change but also contributes to higher ground level ozone – a public health risk (Fiore, 2008). Multiple modeling studies have revealed the benefits of reducing methane emissions, which include decreased premature mortality from respiratory illness caused by ozone (West et al., 2006; Fang et al., 2013). A better understanding of emissions and sources could help in the effort to reduce atmospheric methane.

In 2015, production, storage, processing, and distribution of natural gas and petroleum were responsible for approximately one-third of methane emissions in the US (US EPA, 2017e). While all of the leaks along this chain, from production to distribution, contribute to climate change, vented and fugitive emissions of methane that occur at oil and gas production sites may raise concerns for nearby communities due to potential co-emission of hazardous BTEX (benzene, toluene, ethylbenzene, and xylene) compounds (Adgate et al., 2014; Helmig et al., 2014; Moore et al., 2014). Recent studies also suggest that methane emissions from the oil and gas sector are underestimated in current inventories (Miller et al., 2013; Wilcox et al., 2014; Zavala-araiza et al., 2015; Petron et al., 2014; Subramanian et al., 2015). Miller and colleagues found that methane emissions in US EPA inventories may be underestimated by a factor of 1.5 (Miller et al., 2013). It has been suggested that these discrepancies between measured methane and source-based inventory estimates may be explained by "super-emitters" – a small percentage of sites or equipment that contribute a large portion of the emissions (Wilcox et al., 2014; Petron et al., 2014). For example, a study in the Barnett Shale region found that at any given time, 2 % of facilities accounted for half of methane emissions and that these sites vary spatiotemporally (Zavala-araiza et al., 2015). As described in a recent review, smart-sensing systems designed to detect leaks and alert operators at the well pad level may aid in identifying these events as they occur

(Allen, 2014), speaking to the need for tools that can feasibly achieve useful spatial and temporal resolution for monitoring at the local or facility level.

1.2 Low-cost sensors for air quality monitoring

1.2.1 A place for sensors

Typically monitoring methods and technologies are driven by the research question of interest and available resources. For example, the National Oceanic and Atmospheric Administration (NOAA) has maintained a global monitoring network for methane for upwards of 30 years to study long-term atmospheric trends, seasonal cycles, and its global distribution (NOAA, 2017). Monitoring networks can also be built on smaller scales to study methane fluxes at the regional or city level; the Megacities Carbon Project is currently undertaking this work in Southern California, as is the INFLUX project in Indianapolis (Wong et al., 2015; Davis et al., 2017). Remote sensing provides a global picture and, given the spatial coverage, these data can highlight hotspots at the regional scale (Kort et al., 2014). However, interferences and satellite trajectories prevent truly continuous data collection for any single location. Aircraft campaigns and mobile monitoring using vehicles equipped with gas analyzers both allow for horizontally and vertically resolved spatial coverage at the neighborhood or facility level. Additionally, these methods facilitate the collection of high-quality data with precise instrumentation (Yacovitch et al., 2015; Karion et al., 2013). However, aircraft data typically represent a "moment in time" and changing meteorological conditions often limit the ability to repeat data collection. Ground-based mobile monitoring may be repeated more easily, but the data collected are often periodic in nature and intended for targeted studies. Currently the scientific and regulatory communities are limited in their capability to collect data continuously at the neighborhood or facility level. While it would be possible to site the same high-quality instruments utilized in global and regional monitoring networks at a local scale, this approach would be costly given the expense of the equipment, the siting requirements, and the expertise needed for operation.

Low-cost air quality sensing systems are potentially well suited to fill this role by providing continuous measurements in high-density networks at a local scale. Given their versatility and capacity to provide data of high spatial and temporal resolution, these systems could augment regulatory monitoring systems, aid in compliance monitoring (e.g., leak detection), or enable the public to formulate local strategies to reduce their exposure (Snyder et al., 2013). These systems are relatively easy to deploy and operate in nearly any type of location due to their size, low power requirements, and automated electronic data collection. These characteristics also make them more accessible for community-engaged research applications than conventional methods (Shamasunder et al., 2017). For example,

these systems could support a community collecting preliminary data, in partnership with researchers or local regulatory agencies, that could be evaluated for "hotspots" or correlated with community members' experiences (e.g., odors or health symptoms) – providing more information to support better understandings of complex air quality issues.

1.2.2 Previous sensor research

Several studies have demonstrated the ability of low-cost sensors to measure pollutants of interest at ambient levels. For example, CO, NO, and NO_2 have been measured in an urban sensor network with additional analysis demonstrating the ability to differentiate local emissions from regional trends (Mead et al., 2013; Heimann et al., 2015). In another example, researchers demonstrated the feasibility of collecting personal CO, NO_2, O_3, and CO_2 exposure data with uncertainty estimations using a portable, wearable system (Piedrahita et al., 2014). Several studies have also made use of sensors to study the spatial variability of O_3 on various scales (Sadighi et al., 2018; Cheadle et al., 2017; Moltchanov et al., 2015). Connected to this effort on sensor applications, there has been much work evaluating the performance of individual sensors (Masson et al., 2015a &2015b; Spinelle et al., 2015a & 2017a; Lewis at al., 2016) and demonstrating the performance of different calibration approaches (Zimmerman et al., 2018; Kim et al., 2018; Cross et al., 2017).

While many projects utilize sensors capable of detecting criteria pollutants, advances in the development of metal-oxide semiconductor (MOx) sensors have led to sensors capable of detecting methane in settings closer to ambient environmental conditions (Quaranta et al., 1999; Biaggi-Labiosa, 2012). Eugster and Kling (2012) demonstrated the ability of the Figaro TGS 2600 sensor to resolve diurnal methane fluctuations in a remote area of Alaska. A similar sensor, the Figaro TGS 2611-E00, was found to have an accuracy of 1.7 ppm in a laboratory setting for minute-averaged data, suggesting its suitability for detecting substantial methane leaks (Van den Bossche et al., 2017). These and similar metal-oxide volatile organic compound (VOC) sensors have also been utilized in other applications such as odor detection at landfills and electronic noses (Penza et al., 2015; Zhang et al., 2008).

This paper describes a methodology for collecting and quantifying data using Figaro TGS 2600 MOx sensors to examine ambient trends and methane enhancements on small spatial and temporal scales. Data from two field deployments are used to discuss the different considerations for calibrating and deploying these sensors. The first dataset was collected in Colorado during the FRAPPE/DISCOVER-AQ monitoring campaigns in the summer of 2014 (Pfister et al., 2017). This deployment primarily measured rural and semirural areas along the Front Range north of

Denver; important sources of methane in the area include oil and gas development and agriculture/ranching. The second dataset was collected in California near downtown Los Angeles in the late summer– early fall of 2016 as part of a community-based research project. This deployment was in a mainly urban area with high-density housing near two major roadways and urban oil extraction. With this work, we build on the previous study by Eugster and Kling (2012) by demonstrating the use of these sensors in more complex environments where they are likely subject to a greater number and variety of local and regional influences. We (1) demonstrate methods for sensor calibration and validation of the Figaro TGS 2600 MOx sensors using field co-locations, (2) examine different options and issues that arise in the calibration process, and (3) explore the potential for the data from these sensors to offer unique information. This paper is intended to explore ways of adapting this system to fit the needs and logistical constraints of different investigations in order to provide useful and relevant methane estimations.

2 METHODS

2.1 Instrumentation – low-cost sensor systems

In both deployments, embedded sensor systems termed U-Pods and Y-Pods (subsequent iterations of an open-source platform) were used for data acquisition (Mobile Sensing Technology, 2017). The main differences between the two versions were in the circuit board design and the programming, which was altered to improve reliability. Each U-Pod and Y-Pod (pod) was outfitted with multiple gas-phase and environmental sensors, listed in Table 1. The two Figaro VOC sensors were originally developed for monitoring in industrial applications where much higher pollutant concentrations are expected compared to ambient environmental monitoring. The following analysis will primarily utilize signals from one of these VOC sensors – the Figaro TGS 2600 MOx sensor. This is the same sensor used by Eugster and Kling in Alaska (2012), deployed here in environments characterized by complex mixtures including methane emissions and associated confounding gas species.

Table 1: U-Pod and Y-Pod sensor lists.

Sensor type	U-Pod	Y-Pod
Temperature and RH	RHT03 (also known as DHT22)	Sensirion SHT2
Temperature and pressure	47 Bosh BMP085	Bosh BMP180
Carbon dioxide	ELT S-100 NDIR	ELT S-300 NDIR
Ozone	SGX Corporation MiCS-2611	SGX Corporation MiCS-2611
VOC sensor 1	Figaro TGS 2600 MOx	Figaro TGS 2600 MOx
VOC sensor 2	Figaro TGS 2602 MOx	Figaro TGS 2602 MOx
Additional optional Sensors	Alphasense B4 series (CO, NO, NO_2, O_3, SO_2), Baseline Mocon PID	Alphasense B4 series (CO, NO, NO_2, O_3, SO_2), Baseline Mocon PID

These embedded sensor systems are housed in small weather-proof plastic cases (approximately 20 cm 25 cm 10 cm) with fans to pull ambient air through the enclosure and across the sensor surfaces resulting in multiple air exchanges occurring each minute. The systems in these weatherproof cases can be placed outdoors for long periods of time. They are powered using 12 V AC/DC adapters plugged into wall power, but can use car batteries and/or solar power in remote locations. All data are logged to an onboard micro-SD card. As configured, these pods draw roughly 11 W. These systems have been used in several other indoor and outdoor air quality studies (Casey et al., 2018a & 2018b; Sadighi et al., 2018; Cheadle et al., 2017). Figure 1 includes a labeled photo of a Y-Pod interior and a photo of two Y-Pods deployed.

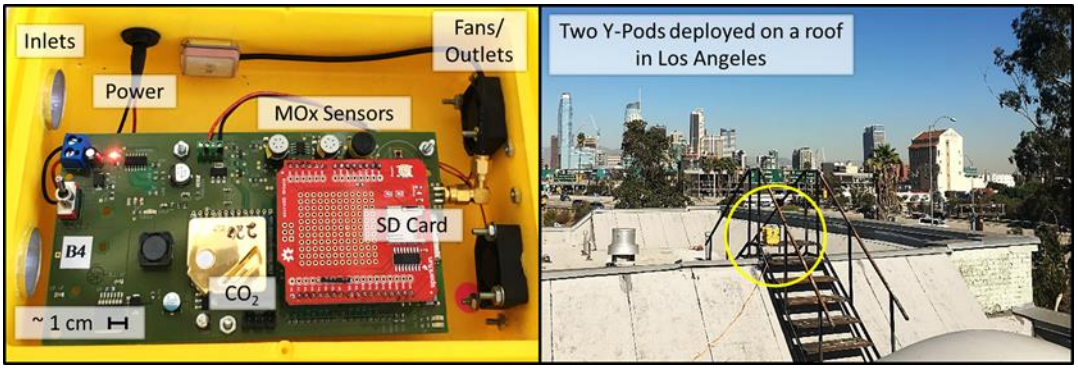


Figure 1: A labeled photo of a Y-Pod interior (a) and a photo of two Y-Pods deployed at a field site (b).

2.2 Deployment overview

As sensor systems are a developing technology, guidelines specifically for sensor deployment have yet to be established. In this section we describe our process for siting sensor systems. Once a site is selected, pod placement is chosen based on feasibility, access to air flow, and avoiding potential obstructions as much as possible to obtain samples that are representative of the area. As selecting sampling locations and setting up pods is typically a joint effort with community partners, different sites often require different approaches. For example, pods are typically placed on the roof of multi-story buildings, while we may place the pod on the edge of a first story roof or a fence of

a single-family home. Additional considerations include access to power, whether the instrument is obstructing a walkway or driveway, and safety of the residents. In both deployments discussed in this paper, site selection was guided by the research goals and access to representative air flow while also considering the preferences of the owner, tenants, or site manager.

In Colorado, the pods were used during the 2014 FRAPPE/DISCOVER-AQ campaign with the aim of characterizing small-scale spatial variability of pollutants. This deployment lasted roughly 1 month. The deployment of the pods was organized around a main site for the FRAPPE/DISCOVER-AQ campaign, the Boulder Atmospheric Observatory (BAO) Tower in Erie, Colorado. Fourteen pods were placed in an approximately 10 x 10 km grid, primarily to the northwest of the BAO Tower. The remaining four pods were placed to the southwest and northeast of the grid to provide regional comparisons, with measurements taken at the Golden National Renewable Energy Laboratory (NREL), Frederick, and Platteville. All Colorado sites are shown in Fig. 2.

Shown in Fig. 2 are the boundaries of the Wattenberg Gas Field and all active and inactive wells. In general, oil and gas activity increases in density moving from the southwestern side of the deployment region to the northeastern side, with the Erie sites on the edge of the gas field. Note that the Golden site has no nearby oil and gas activity, while the Platteville site is surrounded by a high density of wells. The pods were sited in rural and suburban areas primarily at homes, schools, or in open spaces with two monitors sited at a water reclamation facility. Of the 18 monitors, data from 15 were included in the following analysis. Three monitors were excluded because of extended power failure, temperature or humidity sensor failure, or MOx VOC sensor malfunction. Some of the remaining 15 monitors experienced occasional power loss, but all necessary sensors operated continuously.

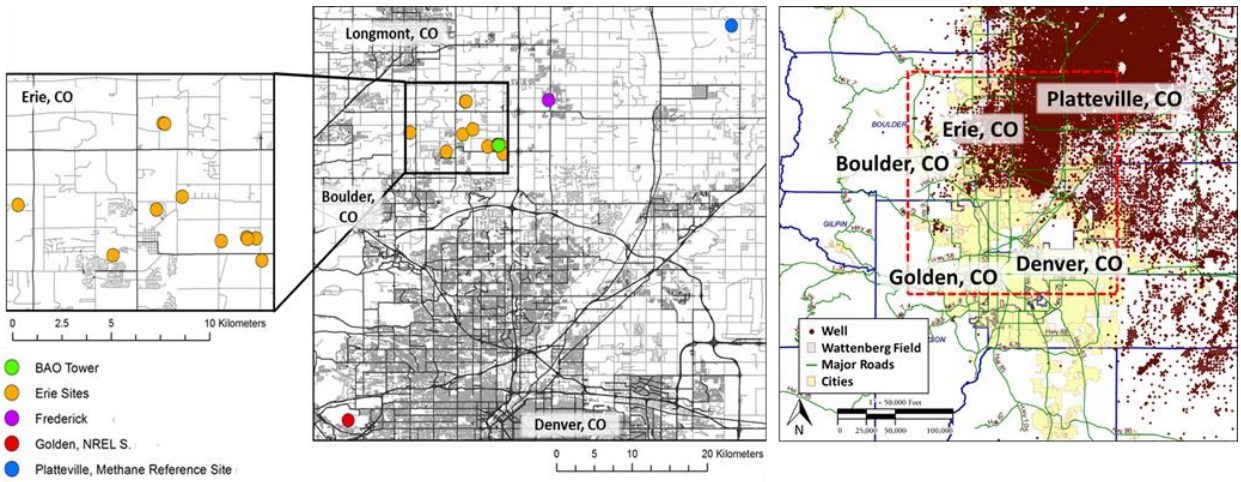


Figure 2: Maps of Colorado deployment area with every site displayed in panel (b) and the Erie sites displayed in panel (a). Panel (c) indicates active and inactive wells in the Wattenberg Gas Field along with major urban areas and counties; data are courtesy of the Colorado Oil and Gas Conservation Commissions (COGCC, 2017).

In Los Angeles, we partnered with two community-based organizations, Redeemer Community Partnership and Esperanza Community Housing, and deployed Y-Pods through-out a neighborhood south of downtown Los Angeles. This deployment lasted approximately 8 weeks. The community was specifically interested in deploying a monitoring network around an active oil extraction site. In this case, sites were selected at varying distances away from the drilling operation as well as varying distances from freeways, another potential source of pollutants (Fig. 3). Thirteen of the sites were within an approximately 5 x 5 km grid, and two additional sites were located further to the northwest and northeast. These two additional sites were utilized because they allowed for continuous co-location with reference instruments for validation purposes. The deployment area in Los Angeles was primarily urban and suburban with high-density residential areas, some commercial and industrial land use, and much higher-density traffic than the Colorado deployment area.

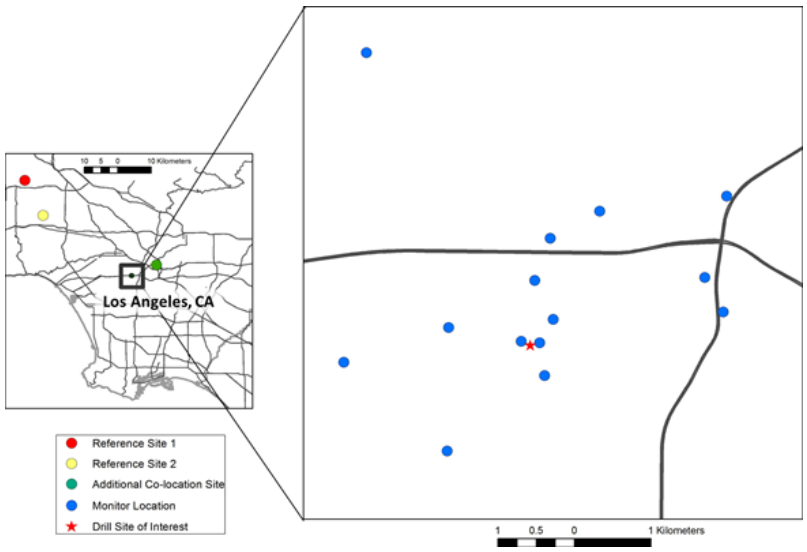


Figure 3: Map of Los Angeles deployment sites, showing (a) the deployment area and all sites where co-locations with reference monitors occurred and (b) the distribution of monitoring sites in relation to major roadways and the drill site of interest (note the monitor locations have been approximated to the center of their respective blocks to protect participant identities).

Evaluating the performance of the Figaro TGS 2600 MOx sensor in the context of these two deployments provides the opportunity to better understand its strengths and limitations. In Colorado, the sensor network covered a larger area and we examined methane trends with respect to regional differences in potential sources (the inclusion of the sites further to the northeast and southwest enlarge the sampling area to a more regional scale). In Los Angeles, the sensor network covered a smaller area to examine local methane trends and to attempt to distinguish emissions from point sources. Another important distinction between the two locations is the nature of the oil and gas activity. In Colorado, the deployment was in the SW portion of the Denver-Julesburg Basin, which produces a mix of natural gas, condensate liquids, and crude oil (US EIA, 2016a). This area also includes the Wattenberg Field, which ranked in the top 10 for both oil- and gas-producing fields in 2013 (US EIA, 2015). In Los Angeles, oil and gas activity refers primarily to crude oil production. California is the fourth top-producing state for crude oil (US EIA 2016b), and Los Angeles County is home to more than 5000 active oil wells (Sadd & Shamasunder, 2015). In both cases we expect methane to be emitted or co-emitted with other VOCs; we attempt to better understand local sources, methane may serve as a valuable indicator of emissions from these types of sites. The ratio of methane relative to other combustion products such as CO and CO_2 will likely be higher from sites related to oil and gas activity than from other local sources such as traffic (Nam et al., 2004; Popa et al., 2014; Peischl et al., 2013). While the two deployment locations

offer contrasting sampling environments, both locations offer complexity in terms of number and types of sources, geography, and typical atmospheric trends.

2.3 Sensor signal processing

The operating principle of MOx semiconductor sensors is based on a reducing gas changing the resistance of a semiconductor material in a simple resistance circuit (Sun et al., 2012). In clean air, the flow of current across the sensor surface is limited by donor electrons in the tin dioxide that are attracted to oxygen adsorbed to the sensor's surface. The flow of current increases when the target gas (e.g., methane) is present, thus reducing the amount of oxygen adsorbed to the sensor's surface (Figaro USA, Inc.). In other words, the resistance across the sensor decreases with increasing methane. In both the Y-Pods and U-Pods, the sensor voltage is continuously recorded to the SD card. Using Eq. (1), provided by the sensor manufacturer, we calculate the sensor resistance (R_s) at various concentrations (Figaro, 2005a). In this equation V_c is the circuit voltage, R_L is the load resistance, and V_{out} is the logged voltage. R_0 represents the resistance in clean air and the ratio of R_s/R_0 is typically used in the analysis of MOx sensor data (Eugster and Kling, 2012; Piedrahita et al., 2014). Gas sensor signals, temperature, humidity, and pressure are recorded to the SD card approximately every 6–25 seconds (depending on a pod's programming). This frequent data acquisition allows for the use of minute-median data in calibration and analysis. Unless otherwise stated, this is the time resolution used in our analysis and shown in this paper.

$$R_s = \frac{V_c * R_L}{V_{out}} - R_L \quad (1)$$

2.4 Sensor calibration, validation, and analysis

Field normalizations were used to generate calibration models for the sensors. Field normalization provides one approach to correcting for the cross sensitivities that low-cost sensors tend to exhibit with respect to temperature, humidity, and other trace gases (Spinelle et al., 2015a & 2017a; Sadighi et al., 2018; Masson 2015a, b; Wang et al., 2010). This method is implemented by co-locating low-cost sensor systems with high-quality reference instruments (typically regulatory-grade monitors) for a given period and then generating a calibration model using an approach such as linear regression. These calibration models predict the methane concentration (in ppm) based on the sensor signal (R_s/R_0) and other predictors. An advantage of calibrating sensors in the field as opposed to in a laboratory setting is that the models will be trained for the pollutant levels of interest and across the same dynamic temperature

and humidity values that a sensor will likely experience during field deployment. In a study involving personal air quality monitors, Piedrahita et al. (2014) successfully calibrated sensors and provided sensor-specific uncertainty estimates using this method.

In Colorado, we co-located U-Pods with a Los Gatos cavity ring-down spectrometer operated by the Penn State NATIVE Trailer team at the Platteville Atmospheric Observatory in Platteville, CO. In Los Angeles, we co-located Y-Pods with reference instruments at two different sites. The pre-deployment co-location was with a Baseline Mocon Series 900 Methane/Non-methane Hydrocarbon Analyzer located in a primarily residential suburban area of Los Angeles. The post-deployment calibration was with a Picarro cavity ring-down spectrometer located in a suburban/urban area with a mix of residential, retail, and industrial land use. Reference instruments at both Los Angeles sites were operated by the South Coast Air Quality Management District. The timelines in Figs. 4 and 5 illustrate when pods were co-located vs. deployed in the field and which data were used for the generation of calibration models (i.e., training data) versus the validation of those models (i.e., testing data). Note that for the Colorado deployment, both before and after the field deployment, the monitors were co-located in batches due to logistical constraints. Arrows indicate the movement of batches of monitors, and the "not in use" row clarifies whether pods were deployed. In addition, during the Colorado deployment, a single calibration model (a universal model) was developed based on the data from the "main" U-Pod, described in greater detail below. For the Los Angeles deployment, calibration models specific to each Y-Pod (sensor-specific models) were used.

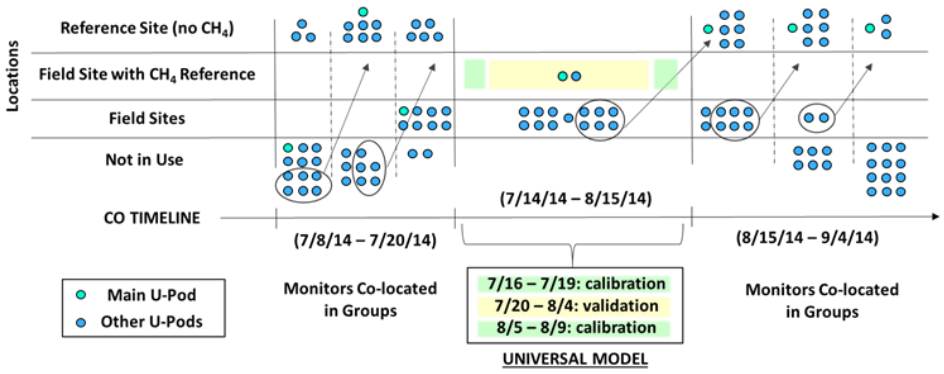


Figure 4: Timeline for Colorado, indicating when monitors were co-located together in batches before and after the field deployment and illustrating how two U-Pods were sited with a reference instrument during the field deployment (these data were used for calibration generation or training data versus model validation or testing data).

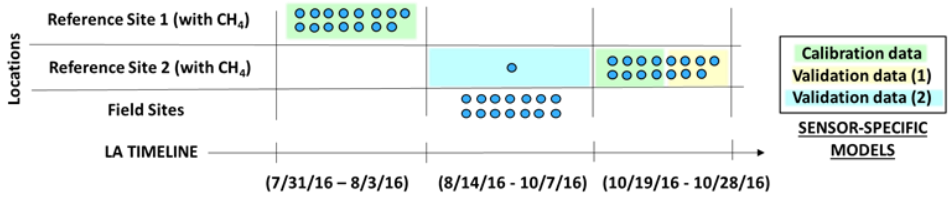


Figure 5: Timeline for LA indicating when Y-Pods were co-located together with reference instruments before and after the field deployment, as well as which data were used for calibration generation (training data) versus model validation (testing data).

The setup of Y-Pods for these co-locations was governed by limitations at the site. In Colorado, Y-Pods were mounted to the railing of the NATIVE Trailer (approximately 1.5 m above the trailer roof), which housed the reference instruments. The inlets to the reference instruments were approximately 2.5 m above the roof of the trailer and roughly 2 m away from the Y-Pods. For the first co-location in Los Angeles, the reference instrument was housed in a trailer in an open field. As we were not able to place the Y-Pods on the roof of the trailer, they were placed 0.75–1.5 m off the ground on the side of the trailer where the inlet was mounted. In this case, the Y-Pods were roughly 6 m below and 3 m to the side of the inlet. For the second co-location in Los Angeles, the reference instruments were housed inside of a building. In this case the Y-Pods were mounted to a railing roughly 1–2 m off the roof. However, the Y-Pods were also approximately 10 m away from and 1–2 m below the inlet, as this location was secure and out of the way of ongoing operations at the reference site. We would expect the variability between co-location setups to be most important for short-term spikes in CH_4 that do not pass over the Y-Pod and inlet evenly. As discussed in Section 3.1, our co-location site in Colorado experienced the most short-term CH_4 spikes, whereas the changes in CH_4 concentration at the two LA sites were more gradual in nature.

For both deployments, 4 days of data at the beginning and 4 days at the end of the co-location with reference instruments were used to generate the calibration models. Specifically, for Colorado, 4 days at the beginning and end of the field deployment were used for generating the quantification model. In Los Angeles, 4 days from both the pre- and post-co-location were used for model generation. The remaining data from co-locations were then used for model validation (approximately 18 days for Colorado and 4 days for LA). Table 1 lists the calibration models that were compared. Several models (the simpler ones) selected are commonly used in sensor calibration, while the more complicated models were selected based on predictors that aided in correcting for cross sensitivity and resulted in more normal residuals. The models are listed in order of their complexity, beginning with the addition of environmental parameters, then interactions, and then transformations. Regression analysis provides sensor-specific

coefficients for predictor variables. The models are then inverted so that gas concentration is expressed as a function of sensor signals and can then be used to predict pollutant concentrations using new data collected in the field. This inverted model approach is typical for field normalization (Piedrahita et al., 2014; Spinelle et al., 2015a & 2017a). Evaluation of model performance was based on the coefficient of determination (R^2) and the root mean squared error (RMSE), as well as an analysis of the residuals in relation to relevant environmental and air quality parameters. Validation data provide the opportunity to evaluate the consistency of each model's performance based on the same metrics and the addition of mean bias.

Table 2: Calibration Models

Model no.	Description	Model equation
1	Mdl1	$R_s/R_0 = p_1 + p_2(C)$
2	Mdl3	$R_s/R_0 = p_1 + p_2(C) + p_3(T) + p_4(H)$
3	Mdl4	$R_s/R_0 = p_1 + p_2(C) + p_3(T) + p_4(H) + p_5(T_i)$
4	Mdl4_1Int	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(H) + p_5(T_i)$
5	Mdl4_2Int	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(H) + p_5(T_i) + p_7(T \times H)$
6	Mdl4_1Int_Tr	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(\ln(H)) + p_5(T_i)$
7	Mdl4_2Int_Tr	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4\left(H^{-1}\right) + p_5(T_i) + p_7(T \times H^{-1})$
8	Mdl5_1Int	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(H) + p_5(T_i) + p_7(T_d)$
9	Mdl5_2Int	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(H) + p_5(T_i) + p_7(T \times H) + p_8(T_d)$
10	Mdl5_1Int_Tr	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(\ln(H)) + p_5(T_i) + p_7(T_d)$
11	Mdl5_2Int_Tr	$R_s/R_0 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4\left(H^{-1}\right) + p_5(T_i) + p_7\left(T \times H^{-1}\right) + p_8(T_d)$

Predictors (lower case p with subscripts): C is pollutant concentration (ppm methane), T is temperature, H is absolute humidity, T_i is continuous time, and T_d is categorical time of day; Mdl indicates the number of predictors; Int indicates the number of interactions; Tr indicates use of transformations. The predictor p_1 indicates an empirical constant.

Given the structure of each deployment and availability of co-located data, two different approaches to developing and applying calibration models were used: a universal calibration model vs. sensor-specific models. For the Colorado deployment, a universal calibration model was developed using the data from one sensor and this model was applied to all the sensors. As shown in Fig. 4, two U-Pods were co-located with the reference instrument throughout the field deployment. The data from one of these pods was used along with the following process: (1) generate a universal calibration model using data from the main U-Pod co-located with the reference instrument; (2) normalize all of the other U-Pods' raw sensor signals to the main U-Pod using data from when they were co-located together before and after the field deployment; and (3) apply the universal calibration model to the normalized sensor data from each pod. The second U-Pod co-located with the reference instrument allows for validation of this method.

For the Los Angeles deployment, sensor-specific calibration models unique to each Y-Pod were used. As shown in Fig. 5, the Y-Pods were all co-located together with reference instruments before and after the field deployment providing the opportunity to generate and use sensor-specific models. Additionally, one Y-Pod in Los

Angeles was deployed with a reference instrument throughout the field deployment pro-viding an additional set of validation data (referred to as Validation 2). These data offer the opportunity to calibrate the Los Angeles data using both sensor-specific calibration models and a universal calibration model – a direct comparison demonstrating the relative performance of these two methods. This offers an informative comparison as there may be instances where only one method is possible given logistics, such as access to reference instruments. Another advantage to this universal calibration model approach would be that the calibration model is not extrapolating in time as the training data would cover the complete field deployment period.

3 RESULTS AND DISCUSSION

3.1 Differences in reference data and environmental conditions that impact calibration

Different sampling environments necessitate the use of different strategies to produce the strongest calibration for each dataset. Reasons for this may be differences in local sources or meteorological trends. Figure 6a and b illustrate the difference in temperature and humidity values observed during calibration versus validation periods for both locations. In Colorado, the temperature and absolute humidity observed during the validation period are generally well represented by the data collected during the calibration period, although there are some high and low humidity values at certain temperatures that fall outside of the calibration parameter space. Conversely, in Los Angeles, the full range of temperature and humidity values observed during the validation period are captured in the calibration period. However, the Los Ange-les data have many temperature–humidity combinations that are unique to the validation period.

Other sensor limitations must be considered as well, for example relatively slow sensor response. A low-cost sensor with an operating principle relying on chemical reactions may not have time to fully detect a passing plume (Arfire et al., 2016), whereas this is not an issue for high-quality reference instruments that rely instead on optical properties. The manufacturer of the Picarro cavity ring-down spectrometer, for example, cites a gas response time under 3 seconds (Picarro, Inc., 2015), while Baseline Mocon cites a response time of less than 5 seconds for the Series 900 Methane/Non-methane Hydrocarbon Analyzer (Mocon, Inc., 2017). Given these quick response times and the high flow rates used for sampling by the reference instruments we would not expect a lag on the part of the reference instrument. The sensor failing to reach steady state when exposed to a short and high concentration plume, as a result of slow sensor response, would be more of a concern for calibration. This limitation may result in sensor data that is

fundamentally different from reference data, further complicating calibration model generation. One option for addressing this issue, explored below, is filtering very short-duration reference data features prior to model generation. Demonstrating the need for this step, Fig. 6d and e each show 3 days of data from the reference monitors in which the diurnal patterns are similar, but the Colorado data also include short-term enhancements or "spikes" in methane possibly from the oil and natural gas extraction activity in the study region. The histogram in Fig. 6c depicts the changes in methane values for each dataset from minute to minute, further highlighting instances in the Colorado data where methane levels change by 0.5 or even 1 ppm over the course of a minute. These differences in the environmental parameter spaces emphasize the need to customize quantification methods to each dataset.

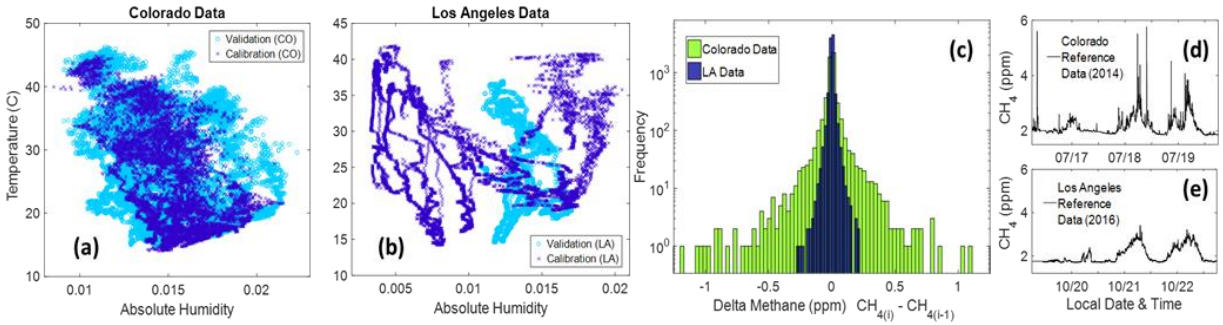


Figure 6: Panels (a) and (b) are the temperature and humidity values observed during calibration and validation periods in Colorado and Los Angeles, respectively. Panel (c) is a histogram of the changes in methane from minute to minute (for the reference data) and panels (d) and (e) are 3 days of minute-resolution data from the reference instruments in Colorado and Los Angeles, respectively.

3.2 Comparing calibration models

Table 3 contains the resulting statistics for each of the models described in Table 2 for one Colorado U-Pod and three Los Angeles Y-Pods. Three Y-Pods were selected randomly to facilitate analysis of the universal model method and provide an initial indication of calibration model consistency across different sensors. This table lists the R^2 and the RMSE as well as the mean bias for the validation data. In all cases, these are the statistics for the fitted sensor data (converted into ppm CH_4) versus the reference methane data. Note that the second value in the Colorado data is the result when a filter is used to remove short-term spikes from the reference and sensor data. Filtering the Los Angles reference data did not change any of the statistics for that dataset and therefore was not performed. This filter removes spikes that are greater than twice the past hour's standard deviation and last less than 5 min in duration.

Table 3: Calibration model generation and validation results

			Calibration data		Validation data		
No.	Model	POD	R^2	RMSE (ppm)	R^2	RMSE (ppm)	Mean bias
1	Mdl1	CO; Filt.	0.463; 0.469	0.391; 0.369	0.456; 0.459	0.398; 0.378	0.014; 0.007
		LA 1	0.244	0.561	0.550	0.438	−0.381
		LA 2	0.257	0.541	0.556	0.444	−0.388
		LA 3	0.229	0.582	0.552	0.434	−0.379
2	Mdl3	CO; Filt.	0.498; 0.507	0.367; 0.346	0.329; 0.331	0.576; 0.550	−0.035; −0.040
		LA 1	0.514	0.310	0.402	0.240	0.024
		LA 2	0.522	0.304	0.434	0.248	−0.079
		LA 3	0.552	0.286	0.481	0.225	0.003
3	Mdl4	CO; Filt.	0.574; 0.590	0.314; 0.292	0.392; 0.401	0.483; 0.453	−0.038; −0.035
		LA 1	0.500	0.319	0.380	0.252	0.070
		LA 2	0.479	0.331	0.368	0.249	0.045
		LA 3	0.526	0.301	0.438	0.250	0.096
4	Mdl4_1Int	CO; Filt.	0.596; 0.625	0.300; 0.271	0.423; 0.449	0.437; 0.383	−0.024; −0.011
		LA 1	0.747	0.184	0.518	0.224	0.047
		LA 2	0.765	0.176	0.558	0.212	0.026
		LA 3	0.752	0.181	0.527	0.225	0.069
5	Mdl4_2Int	CO; Filt.	0.588; 0.618	0.305; 0.275	0.432; 0.462	0.441; 0.384	−0.019; −0.011
		LA 1	0.753	0.181	0.496	0.229	0.053
		LA 2	0.776	0.171	0.536	0.218	0.035
		LA 3	0.752	0.181	0.527	0.226	0.067
6	Mdl4_1Int_Tr	CO; Filt.	0.593; 0.622	0.302; 0.273	0.425; 0.451	0.440; 0.385	−0.034; −0.019
		LA 1	0.737	0.189	0.512	0.341	0.264
		LA 2	0.761	0.177	0.457	0.389	0.316
		LA 3	0.745	0.184	0.525	0.380	0.316
7	Mdl4_2Int_Tr	CO; Filt.	0.588; 0.616	0.305; 0.275	0.440; 0.465	0.430; 0.379	−0.041; −0.027
		LA1	0.784	0.167	0.627	0.198	0.061
		LA2	0.813	0.151	0.667	0.202	0.096
		LA3	0.776	0.169	0.655	0.217	0.120
8	Mdl5_1Int	CO; Filt.	0.597; 0.625	0.300; 0.271	0.426; 0.451	0.435; 0.382	−0.021; −0.009
		LA 1	0.809	0.154	0.630	0.206	0.076
		LA 2	0.812	0.153	0.635	0.202	0.070
		LA 3	0.805	0.156	0.626	0.213	0.095
9	Mdl5_2Int	CO; Filt.	0.588; 0.618	0.305; 0.274	0.438; 0.466	0.437; 0.382	−0.014; −0.008
		LA1	0.827	0.146	0.571	0.225	0.092
		LA2	0.833	0.142	0.580	0.222	0.091
		LA3	0.819	0.149	0.579	0.229	0.108
10	Md5_1Int_Tr	CO; Filt.	0.594; 0.622	0.302; 0.272	0.428; 0.453	0.437; 0.384	−0.032; −0.018
		LA1	0.800	0.158	0.709	0.282	0.225
		LA2	0.807	0.154	0.630	0.341	0.285
		LA3	0.795	0.160	0.678	0.332	0.281
11	Mdl5_2Int_Tr	CO; Filt.	0.588; 0.617	0.305; 0.275	0.445; 0.467	0.426; 0.377	−0.038; −0.026
		LA1	0.820	0.149	0.756	0.160	0.040
		LA2	0.831	0.143	0.734	0.180	0.083
		LA3	0.804	0.156	0.731	0.190	0.104

Figures 7a and b provide a graphical representation of the same statistics from Table 3 and emphasize the differences in results between the two datasets. For the Colorado data, the greatest improvement in fit was observed when time was added as a predictor, but then the results level off with no major improvements as the models increase in complexity. However, the results consistently returned a higher R^2 and lower RMSE when short-term methane

spikes were filtered prior to model generation. In the Los Angeles dataset, there was continual improvement as models increased in complexity, with the most complex model producing a high R^2 and low RMSE as well as the most consistency across both the calibration and validations datasets.

Figures 8 and 9 provide plots of the "best-fitting" calibration models for each dataset based on regression statistics (particularly for the validation dataset), consistency across calibration and validation data results, and an analysis of the residuals. For the Colorado data, the selected model was the simplest well-fitting model, the fourth model, while for the Los Angeles data the selected model was the most complex model tested, the 11th model. With regards to both datasets, the selected model produced the most normal residuals, which also did not exhibit major trends in relation to the predictors, and they resulted in the closest one-to-one relationship between the fitted sensor data and the reference data. The time series plots also display the performance of the calibration model on the validation dataset. Note, the statistics for Mdl1 for the Colorado data appear to indicate that that model is best fitting, however, there were major issue with the residuals from this fit.

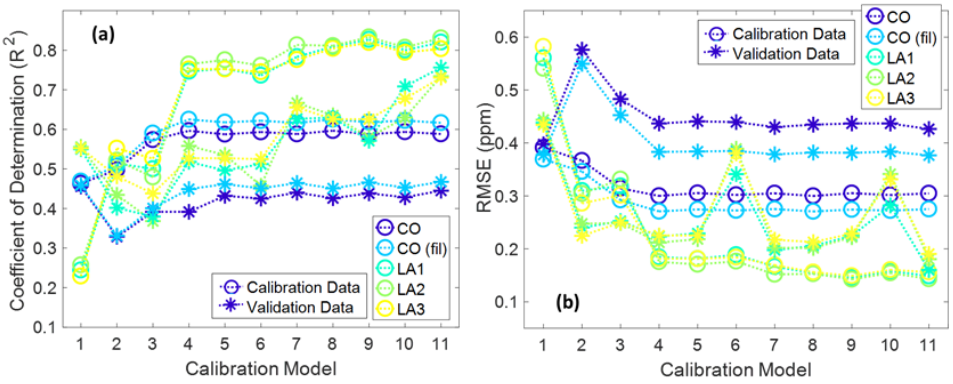


Figure 7: Plotted R^2 (a) and RMSE (b) for all models; circle markers indicate results from calibration generation (using the training data) and asterisk markers indicate results from the application of the models to the validation data (or testing data).

As demonstrated by these two datasets, calibration models are not "one size fits all". While the deployments in Los Angles and Colorado occurred at roughly the same time of year, the best-fitting calibration models and regression results proved to be quite different. This speaks to the need to consider the environmental and pollutant parameter space both when planning a deployment and when processing data. For example, more complex temperature and humidity behavior may require more complex corrections. Additionally, if there is little overlap between conditions observed during calibration, validation, and field deployment then the resulting calibrations will be less dependable. Likely, there are factors beyond environmental parameter space driving differences between sensor and

reference data. In that vein, it is important to explore the operational differences between the reference instrument and the sensors, including distance apart and proximity to significant sources. Here, we compensated for those operational differences by filtering "spikes" from the reference data; another modification could be to use a different averaging time such as hourly instead of minute data. This analysis demonstrates the importance of exploring different models, transformations of variables, and treatments of the data to find the model that provides the strongest methane estimates.

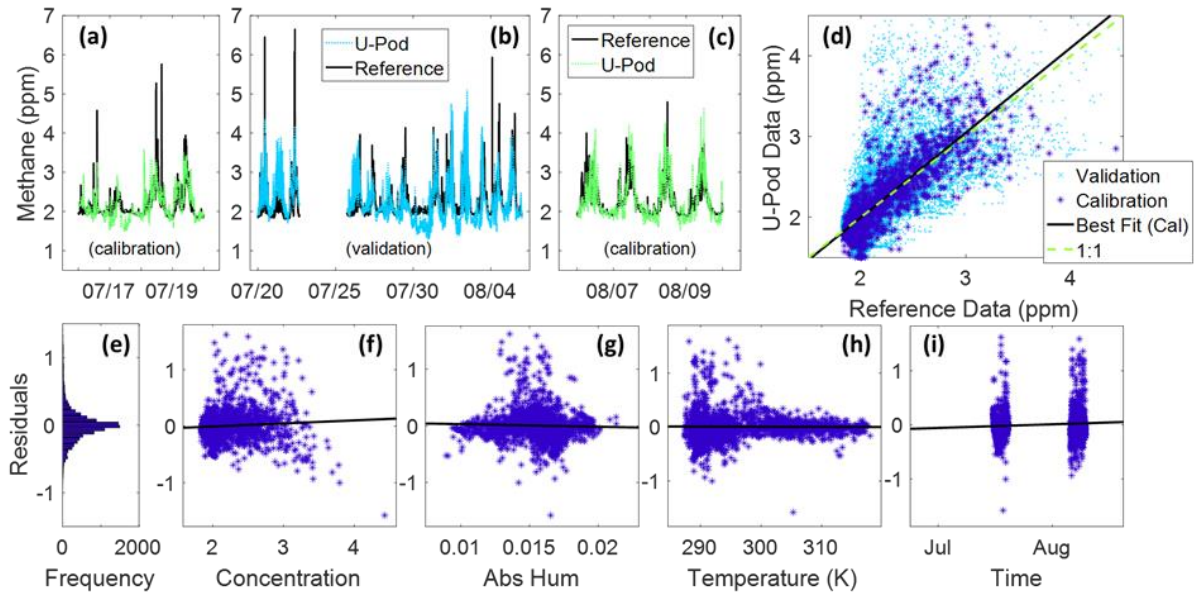


Figure 8: "Best-fitting" model (MDL4_1INT) for the Colorado data with residual analysis (for validation data, RMSE is 0.383 ppm and mean absolute percent error is 12.13 %). Panels (a–c) are time series of the reference data and converted sensor data. Panel (d) is a scatter plot of the same data. Panels (e–i) are the residuals from the calibration generation.

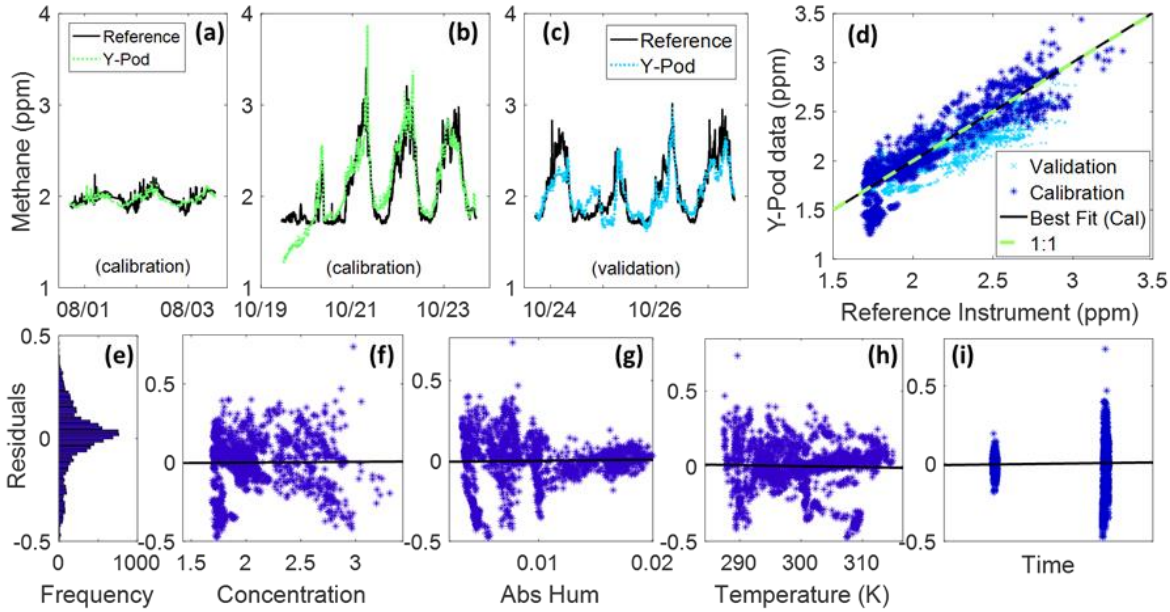


Figure 9: "Best-fitting" model (Mdl5_2Int_Tr) for the Los Angeles data with residual analysis (LA1 shown) (for validation data, RMSE is 0.160 ppm and mean absolute percent error is 5.75 %). Panels (a–c) are time series of the reference data and converted sensor data. Panel (d) is a scatter plot of the same data. Panels (e–i) are the residuals from the calibration generation.

One feature of the models that applied to both datasets was a correction for sensor drift over time, emphasizing the importance of collecting data that either bookend or span the duration of the field deployment. Even though the final models selected here differed, both included a correction for sensor drift over time and a pre-only or post-only calibration would not have allowed for this correction. To examine the consistency of this drift between sensors, we compared the linear drift from the three LA Y-Pods by examining data converted to concentrations using Mdl3, which does not include time as a predictor. The results were drift values of 0.009, 0.015, and 0.011 ppm week 1 for LA1, LA2, and LA3, respectively. While these numbers are similar to those reported by Eugster and Kling (2012), the total drift implied is less than or equal to our expected uncertainty for each Y-Pod making the estimates unreliable. Given the differences in the deployments and their lengths, we have a starting idea of drift and its consistency sensor to sensor, but a better understanding of drift as well as the effective lifetime of sensors will be important for future use of this and other MOx sensors.

3.3 Sensor-specific vs. universal calibration models

The additional validation data (Validation 2) collected during the field deployment in Los Angeles facilitates a comparison of the sensor-specific versus the universal calibration model approach. Several calibration models were

generated using this additional co-located data (Fig. 5), including the two models selected in the previous section as "best-fitting".

These models were then applied to normalized data from the other two Los Angeles Y-Pods included in the previous section. The raw sensor data from the Figaro TGS 2600 sensors were normalized using a simple linear regression (the R^2 values for these regressions were 0.989 and 0.999, respectively). Similar to the results from Section 3.2, the same model (MDL5_2INT_TR, model 11) emerges as the strongest for this particular dataset given that the validation statistics include the highest R^2 and lowest RMSE. An important note is that overall the results using this method are not as strong as the results seen using the sensor-specific models in the previous section. One reason for this may be that we are attempting to fit roughly 6 weeks, rather than 4 days, for the calibration model generation, meaning that the model is attempting to cover a larger environmental parameter space. This might also explain why the results for the "best-fitting" model are better for the validation period, which is much shorter. In any case, this calibration model approach provides useful information regarding methane levels (e.g., diurnal trends), as is demonstrated by Fig. 10, and this method can be used to con-vert the normalized signals from other sensors to a ppm value when logistics limit the potential for co-locating all sensors, whether due to time constraints or the limited availability of power and/or space at a co-location site. As the logistics of the Colorado deployment did not allow for sensor-specific calibrations, the universal calibration model approach is used below in Section 3.5 to convert the Colorado field data from all the U-Pods.

Table 4: Calibration and validation results for the universal calibration method (*normalized Y-Pod data).

		Calibration data		Validation data		
		R^2	RMSE	R^2	RMSE	Mean bias
MDL1 (1)	LA1	0.154	0.493	0.209	0.856	−0.396
	LA2*			0.223	0.835	−0.382
	LA3*			0.201	0.836	−0.377
MDL3 (2)	LA1	0.500	0.214	0.404	0.565	−0.249
	LA2*			0.438	0.581	−0.292
	LA3*			0.419	0.548	−0.256
MDL4 (3)	LA1	0.496	0.215	0.397	0.577	−0.261
	LA2*			0.434	0.589	−0.304
	LA3*			0.412	0.560	−0.268
MDL4_1INT (4)	LA1	0.463	0.230	0.372	0.622	−0.283
	LA2*			0.411	0.624	−0.324
	LA3*			0.387	0.606	−0.288
MDL5_2INT_TR (11)	LA1	0.477	0.223	0.486	0.334	−0.098
	LA2*			0.532	0.348	−0.139
	LA3*			0.529	0.336	−0.129

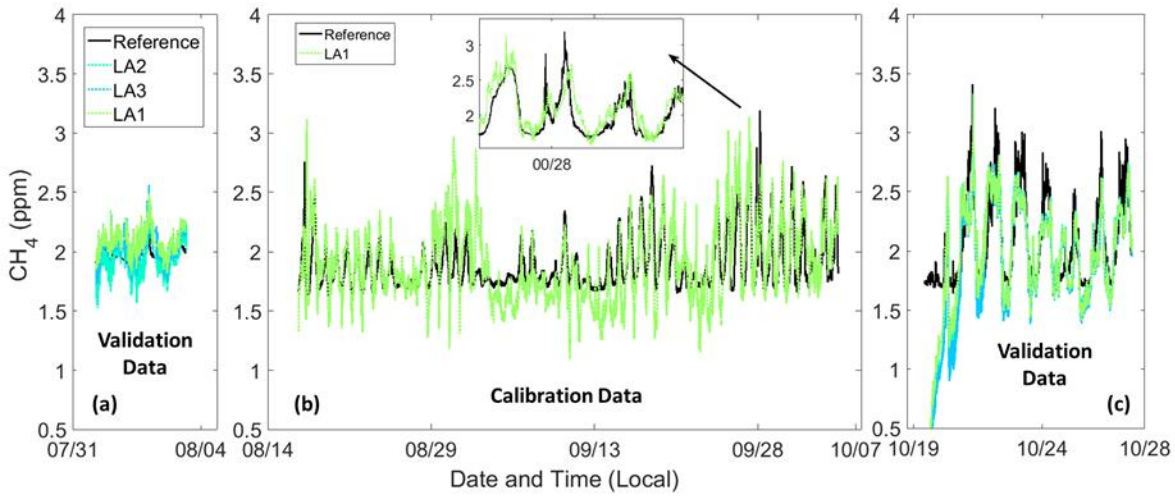


Figure 10: Time series of the universal calibration model (for MDL5_2INT_TR) generated using the Validation 2 period (b) and the application of the model to the pre- and post-deployment co-location data (a and c) for Y-Pods LA1 and the normalized data from LA2 and LA3.

3.4 Further sensor quantification considerations

Comprehensive best practices to guide the use of low-cost air quality sensors have not been established. A recent workshop for low-cost sensors outlined some of the concerns shared throughout the research community including deployment logistics, data formatting and sharing, and communication of uncertainty (Clements et al., 2017). With our datasets, we investigated three issues related to the development of best practices: the length of a co-location for a field normalization, additional dataset-specific filtering based on environmental parameters, and cross sensitivities to non-methane pollutants.

3.4.1 Length of co-location

Bootstrapping methods were applied to determine the variability and effectiveness of different co-location lengths for the Colorado data. A starting point in the complete dataset was randomly selected and consecutive data of varying lengths (0.5, 3, 7, or 14 days) were used to generate a calibration model. This model was then applied to the entire dataset for validation. For comparison purposes three different models were tested with 20 iterations for each model. The resulting statistics are plotted in Fig. 11 along with error bars for 1 standard deviation.

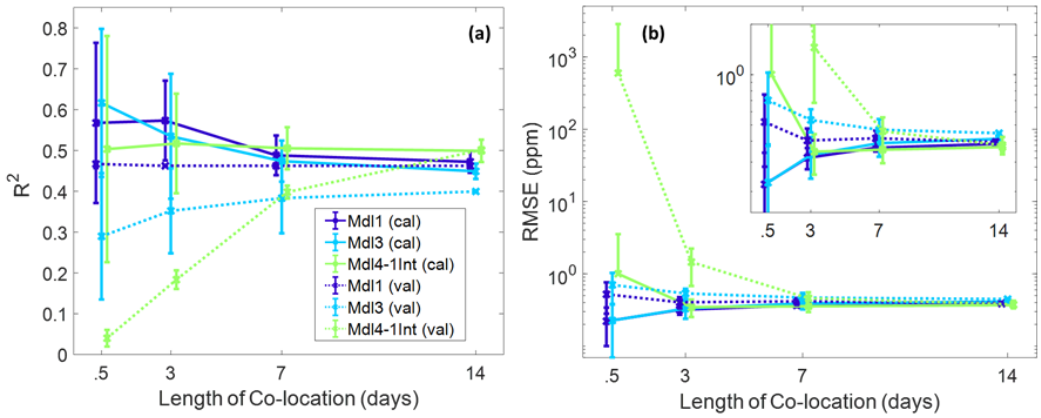


Figure 11: R^2 (a) and RMSE (b) for the calibration model generation (based on a given length of time) and the application of those models to the complete Colorado data. Note that calibration data, or training data, are selected using a random stating point in the complete Colorado dataset and the appropriate amount of consecutive data. The colors indicate the model, and solid lines indicate calibration results while dashed lines indicating validation results.

The simplest model (Mdl1), using only sensor signal and no environmental predictors, seems to perform consistently well for all lengths of time; however, the residuals reveal strong trends with temperature and humidity, indicating that these variables are not being corrected for. Given the analysis of the residuals, this model may provide useful information, but its implementation is also likely to be misleading. For example, this model may be useful in applications that do not require detailed analysis or decision making based on the data, such as education and outreach in a K-12 classroom where sensors are used for labs or student projects (Collier et al., 2015a). Taking into account residuals, Mdl3 provides some correction for temperature and humidity effects without overfitting on shorter co-location lengths. Mdl4_1Int, which includes time as a predictor, is the best performing model for co-location periods of 2 weeks. Given that time was a useful predictor in Section 3.2, the fact that the data are spanning 2 weeks is probably more important than having 2 full weeks of co-located data. This means that the co-location data must be long enough or span a long enough duration relative to the complete dataset in order to provide a time correction that does not lead to overfitting and poor performance on validation data. While greater complexity can provide a better calibration model, a sufficient amount of data must be used to avoid overfitting. Simply stated, the model selection should be appropriate to the data's characteristics and intended purpose.

3.4.2 The impact of model extrapolation

The additional co-located data from Los Angeles (Validation 2) facilitate a more in-depth exploration of the outlier residuals and approaches that could improve the predictive power of the calibration model. For example,

dataset-specific filters were applied to remove values where extrapolation is likely occurring in field data. Extrapolation in this case would be instances where one or more predictors are outside of the range of values used to train the calibration model. Table 5 provides the statistics that result from applying the calibration model with and without this added filtering. The unfiltered dataset statistics are the same results explored in Section. 3.2. All other statistics in the table were calculated after values not observed during calibration were removed. In the first filtered grouping in Table 5, instances where individual temperature or humidity values (primarily extreme values) not observed were removed. In the second grouping, all data combinations not observed during calibration were removed, meaning all instances where exact combinations of temperature, humidity, etc. were not observed. The final filtering option, shown in the fourth section, applies knowledge of atmospheric composition to assist with filtering. In this instance, the atmospheric baseline of methane was used to filter out low concentration values; the baseline was determined by the minimum value observed in the CH_4 reference data.

The final filtering approach should only be applied to sensor data selectively. Removing improbable values from sensor data that fall below zero or a known baseline may be a useful or even necessary strategy in certain situations. In dealing with air quality data, there are examples of additional processing being used to reduce negative values (Hagler et al., 2011) and examples of guidelines to remove negative values below a given threshold (US EPA, 2016). For work with sensor data, if the focus of the analysis is to understand enhancements over background captured by sensors, then removing improbably low values can elucidate these results. If preliminary data are being shared with the public, then flagging and removing improbable values can reduce confusion. Given the challenges in sensor quantification, this second example in particular warrants consideration by those using sensors in partnership with communities. However, it is also likely that these underestimations contain valuable information about sensor behavior and sensitivity; removing these values will also introduce a negative bias to the data. Accordingly, when using this type of processing, researchers will need to be clear about why this approach is useful and valid. For this dataset, every instance where underestimations are removed by the filter coincides with days having a dynamic range of methane less than the expected uncertainty for the sensor, which indicates that these underestimations may be connected to a limit of detection issue. Figure V.1 (Appendix V) demonstrates this association.

Table 5: Additional filtering to improve calibration model performance on field data (specifically the Validation 2 dataset). AH is absolute humidity, WS is wind speed, WD is wind direction, and AP is atmospheric principles.

	Data	R^2	RMSE	Mean bias	n
Unfiltered datasets	Calibration	0.820	0.149	−0.002	9822
	Validation 1	0.756	0.160	0.040	5461
	Validation 2	0.527	0.166	−0.032	71 411
Using only values represented in the calibration data	Temperature	0.529	0.167	0.031	70 340
	AH	0.529	0.166	0.031	71 551
	Resistance	0.548	0.160	0.028	70 851
	WS, WD	0.526	0.167	0.031	69 578
Using only paired values seen in the calibration data	Temperature, AH, resistance	0.552	0.160	0.026	69 473
	Temperature, AH	0.564	0.157	0.030	51 233
	Temperature, AH, resistance	0.574	0.160	0.063	20 059
	Temperature, resistance, WS, WD	0.715	0.188	0.091	7645
Adding AP	CH_4 baseline applied	0.552	0.160	0.047	71 736
Selected filtering	Temperature, AH, resistance, AP	0.581	0.153	0.041	69 473

Additional filtering at nearly each stage yields some improvement in statistics, with the removal of the complete data combinations not seen during calibration resulting in the largest improvements (when we look at the R^2 values), but this method also removes a significant portion of the data. The combination of applied knowledge of atmospheric composition and the removal of extreme individual values not observed during calibration yields improvements while maintaining a substantial amount of the data. This result, labeled “selected filtering”, suggests that this more conservative version of filtering may be sufficient. Not only did this filtering result in a RMSE that is lower than the RMSE for first validation dataset (0.1525 and 0.1601 ppm, respectively), but also these improvements are visible in a plot of the data. Figure 12a provides an overview of the complete dataset and highlights where some of the under-predictions are corrected for before 7 and 17 September, likely driven by the filter utilizing atmospheric principles. Figure 12b shows a close-up of a couple of days illustrating a reduction in overprediction, driven by the filter for either temperature, humidity, or resistance values.

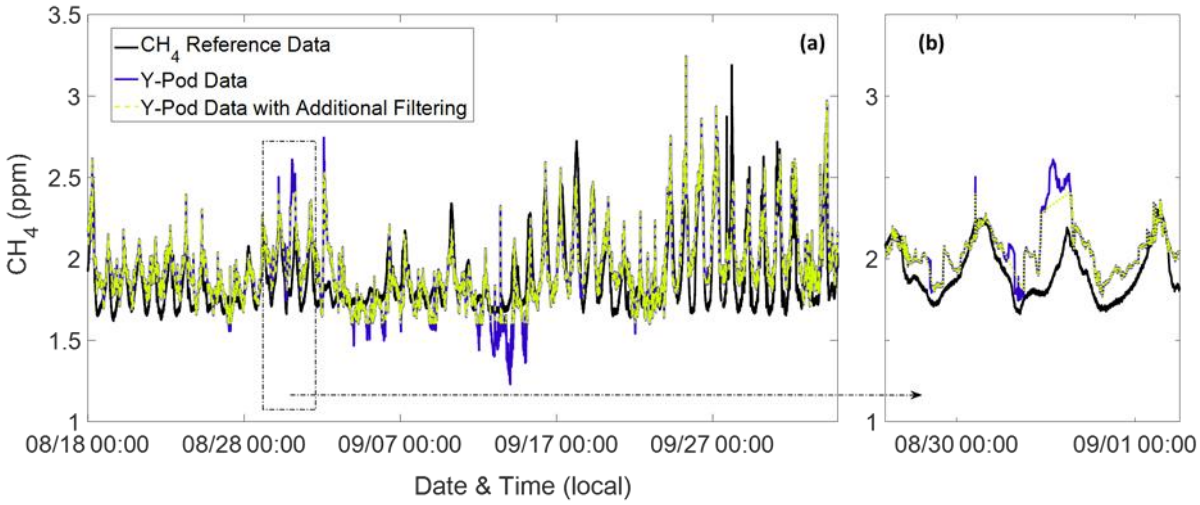


Figure 12: Panel (a) includes the complete Validation 2 dataset from the Los Angeles deployment. The statistics for each are RMSE is 0.17 ppm and absolute percent relative error is 6.77 % for the unfiltered data (blue); RMSE is 0.15 ppm and absolute percent relative error is 6.19 % for the filtered data (green), with the reference data plotted in black. Panel (b) is a close-up of approximately 2 days illustrating an instance where the filtering helped to reduce an overprediction of methane concentrations.

As was demonstrated with the Validation 2 dataset, we expect that applying the same filtering to each deployed sensor's data should result in more reliable field data from all of the sites. Thus in addition to filtering data prior to calibration by removing short-duration enhancements in the reference data, filtering converted sensor data can improve the reliability of calibrated data. While the bounds for this type of filtering should be dataset-specific, this step could easily be an automated addition to low-cost sensor quantification procedures.

3.4.3 Sensor cross sensitivities

Another common concern for low-cost sensors is cross sensitivities to other gases, in addition to known cross sensitivities to environmental factors. As discussed by Eugster and Kling (2012), the Figaro TGS 2600 sensor is reported to be sensitive to carbon monoxide as well as a few other hydrocarbons (Figaro, 2005a). This is not surprising, as each of these species can act as a reducing gas at the sensor surface and therefore also reduce the resistance to electron flow. While Eugster and Kling (2012) did not examine CO specifically given the absence of potential sources in their deployment area, they did perform an analysis of variance examining the effects of CO_2 and found no significant impacts. We applied the same analysis techniques to minute-resolution data to examine the impacts of other gases, specifically CO, O_3, and a few VOCs. Given the information provided by the sensor manufacturer, we expected a cross sensitivity to CO, but not to O_3; this analysis provided an opportunity to check these assumptions. Table 6 includes the resulting explained variance from each ANOVA, all of which included environmental parameters

and time along with the following differences: set 1 is CH_4 only, set 2 is CH_4 and CO, set 3 is CH_4 and O_3, set 4 is CO only, and set 5 is the combined (CH_4 + CO) predictor.

The overall results varied between the deployments. For example, absolute humidity explained a high percentage of the variance in Los Angeles, while the temperature and humidity both played a role in the Colorado data. It's worth noting that in some instances the percent of variability explained by the environmental parameters is equal to or greater than the variability explained by the pollutants. This trend has been seen in other studies as well (Eugster and Kling), and suggests that it might be worth considering housing for sensors that controls temperature and humidity reducing the amount of correction to the sensor signal that needs to occur.

A commonality was that the sensor exhibits a cross sensitivity to CO, but not to O_3. In both cases, the inclusion of O_3 resulted in a higher percentage of variance being attributed to the residuals, and the variance explained by the O_3 concentrations was 0.3 and 2.7 % for Los Angeles and Colorado, respectively. In contrast, the inclusion of CO in the ANOVA for the Colorado data resulted in a decrease of the variance explained by CH_4 from 29.2 % to a still significant 21.8 %, while 15.0 % was explained by the new CO predictor. Notably, this set of parameters also resulted in the lowest portion of the variance being left to the residuals, suggesting that it provided the strongest set of explanatory parameters among these five sets. The inclusion of CO in the ANOVA for the Los Angeles data yielded somewhat different results with the explained variance dropping drastically for CH_4 and being quite low for CO as well, at 2.6 and 4.2 %, respectively. This result is likely explained by the temporal correlation between the two gases obscuring the importance of each individually. The CO concentrations in Los Angeles were higher than those observed in Colorado and well correlated with the CH_4 data as demonstrated in Fig. 13. Further supporting this conclusion, parameter set 5 included a combined "CH_4 + CO" term and resulted in a higher portion of the variance explained through this term at 19.8 % versus CH_4 alone (16.8 %) or CO alone (18.1 %). This set also resulted in the lowest portion of variance left to the residuals. The lack of correlation between the Colorado CO and CH_4 allows us to examine the impacts of the CO cross sensitivity more closely. Figure 14 shows a portion of the Colorado data with both reference and U-Pod methane plotted along with carbon monoxide data from a reference monitor. In Fig. 14a spikes in CO correspond with overpredictions of methane (most notably on 8 August) and the scatter plot in Fig. 14b highlights how overpredictions seem to coincide with elevated CO concentrations.

Table 6: Explained variance from ANOVA analyses on Figaro TGS 2600 resistance values (Rs/R0) for different parameter sets.

	Los Angeles					Colorado			
Source of variation	Set 1	Set 2	Set 3	Set 4	Set 5	Set 1	Set 2	Set 3	Set 4
Temperature	0.3 %	0.1 %	0.0 %	0.2 %	0.08 %	12.2 %	9.6 %	1.5 %	27.9 %
Absolute humidity	61.5 %	72.5 %	63.7 %	61.5 %	62.4 %	6.8 %	10.2 %	10.6 %	6.4 %
Time	0.0 %	0.0 %	0.1 %	0.0 %	0.0 %	8.3 %	8.9 %	10.8 %	3.7 %
CH_4	16.8 %	2.6 %	14.2 %	–	–	29.2 %	21.8 %	20.5 %	–
CO	–	4.2 %	–	18.1 %	–	–	15.0 %	–	19.2 %
O_3	–	–	0.3 %	–	–	–	–	2.7 %	–
$CH_4 + CO$	–	–	–	–	19.8 %	–	–	–	–
Residuals	21.4 %	20.7 %	21.8 %	20.3 %	17.8 %	43.5 %	34.5 %	53.9 %	42.8 %
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %

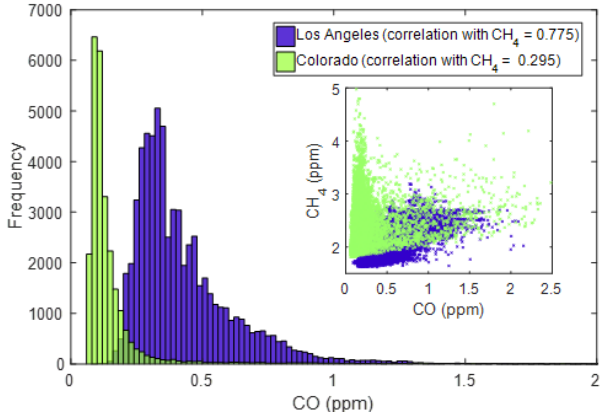


Figure 13: Histogram of carbon monoxide data from the two deployments and scatter plot of carbon monoxide data vs. methane data also from the two deployments. Note that all data in these two plots are from reference instruments.

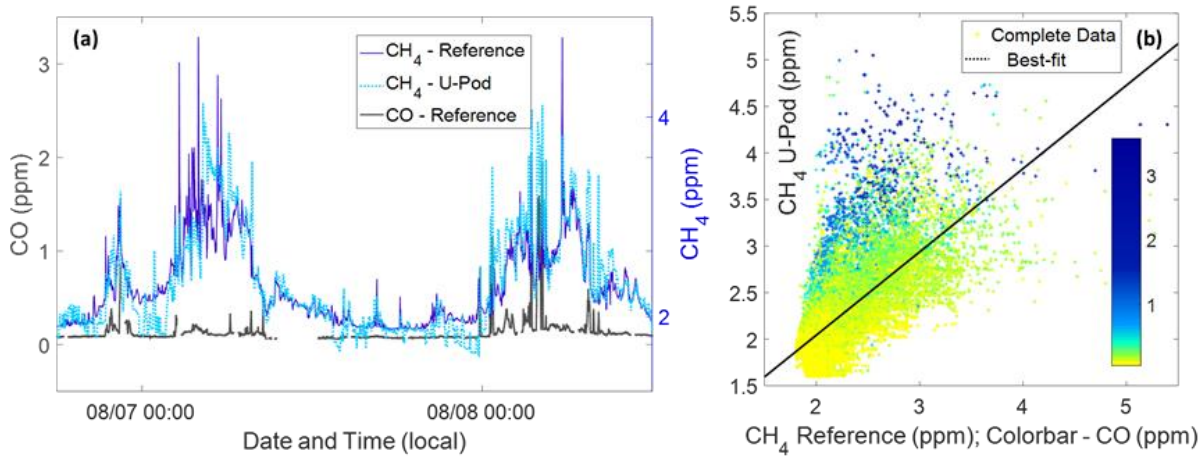


Figure 14: (a) Time series of methane and carbon monoxide data from the reference monitors and converted U-Pod sensor data. (b) Scatter plot of reference methane data (x axis) vs. U-Pod methane data (y axis) with the points colored by carbon monoxide vales. Together they further show the Figaro TGS 2600's cross sensitivity to carbon monoxide, illustrating how many over predictions correspond to instances when CO is at or above 0.5 ppm.

In addition to this observed cross sensitivity to CO, we expect that other hydrocarbons may affect the sensor response as well. This would be an important consideration for measurements made in areas with oil and gas activity where the pollutant mixtures may be complex. At the PAO site there was also a proton-transfer-reaction quadrupole mass spectrometry (PTR-MS) providing speciated VOC measurements (Halliday et al, 2016). Future work will provide a more in-depth analysis of VOC sensitivity and selectivity for the two MOx sensors we are using (Chapter 7); however, we have included here a preliminary look at this cross sensitivity to other hydrocarbons. Table 7 provides the results of another sensitivity analysis in which the explanatory power of a few speciated VOCs is examined. For simplicity, one VOC from different well-correlated groups was selected (e.g., benzene was selected out of the aromatic species). This analysis illustrates that VOCs (particularly acetaldehyde and benzene) do help to more fully explain the variance in the sensor signal, but they do not displace methane. This is most apparent for parameter sets 5 and 6, in which we see the variance explained by residuals increase slightly and the variance explained by temperature increase quite a bit as this factor compensates for the missing methane. When methane is added back in for parameter set 7, along with all three VOCs and CO, the variance explained by the residuals is at its lowest and the variance explained by methane is at 10.1 %, higher than the percentages for the individual hydrocarbons. Thus, the Figaro TGS 2600 sensor seems to be cross sensitive to carbon monoxide and some hydrocarbons, effects that should be considered or mitigated in future uses of this sensor to estimate methane.

Table 7: Explained variance of the Figaro TGS 2600 resistance values (R_s/R_0) for parameter sets including different VOCs.

Source of variation	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7
Temperature	9.6 %	17.4 %	10.1 %	10.9 %	33.7 %	28.9 %	16.3 %
Absolute humidity	10.2 %	11.7 %	14.2 %	7.5 %	6.2 %	9.0 %	12.6 %
Time	8.9 %	4.5 %	9.2 %	4.4 %	3.0 %	2.8 %	5.8 %
CH_4	21.8 %	12.3 %	14.2 %	18.5 %	–	–	10.1 %
CO	15.0 %	13.6 %	14.4 %	19.7 %	–	9.8 %	11.4 %
Acetaldehyde	–	7.5 %	–	–	13.4 %	8.9 %	6.5 %
Benzene	–	–	4.6 %	–	8.4 %	6.4 %	4.0 %
Methanol	–	–	–	0.9 %	0.4 %	0.1 %	0.3 %
Residuals	34.5 %	33.0 %	33.4 %	38.2 %	34.9 %	34.1 %	33.2 %
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %

Despite sampling in more complex environments than previous deployments of this sensor (Eugster and Kling, 2012), we are still seeing a sizable proportion of the sensor data's variance explained by ambient methane concentrations. Although these cross sensitivities need to be addressed to discern which signals are driven by methane

versus other pollutants, the Figaro TGS 2600 sensors are reacting in part to changes in ambient methane again providing useful methane estimates for applications where methane concentrations are needed with resolutions on the order of 0.2–0.4 ppm (the uncertainty determined in this analysis). Given the variety of low-cost sensors available, using the Figaro TGS 2600 sensors in a sensor array could provide additional signals at each deployment site facilitating more reliable data. Including multiple sensor signals in a neural network calibration approach may also improve the accuracy of the calibrated data (Zimmerman et al., 2018; De Vito et al., 2008; Huyberechts et al., 1997). Future analysis of the data collected in Los Angeles and continued use of this sensor in areas with complex mixtures will require carbon monoxide and non-methane hydrocarbon impacts be considered.

3.5 Ability to assess spatial variability in the northern front range of Colorado

The universal calibration approach along with the "best-fitting" calibration model (Fig. 8) was used to convert the field data from the sensors deployed in Colorado. Following the same procedure outlined in Section 2.4 and examined in Section 3.3, the raw voltage values from each Figaro TGS 2600 sensor, from the post-calibration period, were normalized to the sensor signals in U-Pod P1 (the main U-Pod) using sensor-specific simple linear fits. The calibration model was then applied to these normalized sensor data along with the temperature and humidity data from each U-Pod. An additional step was taken to detrend each set of converted sensor data by removing the best-fit linear trend from the whole dataset. It was necessary in this instance because the time correction incorporated in the calibration model appeared to be over- or under-correcting for different sensors. The choice to continue using this model was based on both the performance of the model observed in Section 3.2 and the fact that time appears to be a useful predictor for the Colorado data given the cross-sensitivity analysis in Section 3.4.3. Data from the pre-calibration period when six sensors were co-located with U-Pod P1 were used to verify that the application of this detrend function was appropriately correcting for the under-or over- correction of the model. One possible explanation for this difference in drift between the sensors is that 5 sensors were new and while the other 10 (including the one in U-Pod P1) had been previously deployed. This difference in drifts was not observed in the Los Angeles data (Section 3.3), which utilized all new sensors at the start of the deployment and were operated for the same amount of time throughout the deployment. The final step in preparing these data was to filter out data where the temperature and humidity values were outside of those ranges observed during calibration and to remove data where concentration values were lower than an expected minimum (atmospheric background) was observed, similar to the analysis in Section 3.4.2. In this case a conservative 1.6 ppm was used, roughly half of our RMSE below background methane

levels. For this analysis, filtering out implausibly low values highlights the differences in methane enhancements between the field sites. The largest amount of data removed from any U-Pod dataset as a result of this filtering was approximately 6 %.

Table 8: Statistics for Colorado data converted using the universal model method.

	Post-co-location		Deployed to field sites	
	R	RMSE (ppm)	*R*	RMSE (ppm)
Co-located pair; P2 to P1	0.961	0.210	0.904	0.241
Mean of all U-Pods to P1	0.866	0.180	0.025	0.556

Table 8 presents statistics illustrating the correlation coefficients and RMSE for converted sensor data during co-location versus field deployment for both the 2 U-Pods continuously paired and the mean of all 15 U-Pods. Here correlation coefficients were used as fitted sensor data was being compared to fitted sensor data rather than reference data. The result is high correlation when pods are co-located and low correlation when U-Pods are deployed to their field sites, high-lighting that (1) there is consistency in the data provided through the universal calibration model and (2) we are seeing quite a bit of variability across the field sites. Additionally, the RMSE for co-located U-Pods is less than the error we expect given the RMSE of 0.3832 ppm for the validation dataset. The details of each individual sensor versus P1 are available in Fig. V.1 and Table V.1, in Appendix V. These details demonstrate the extent of inter-sensor variability for co-located sensors and the increase in variability for deployed sensors. It's worth noting here that at least some of this variability may be due to variability in environmental parameters, though given the magnitude of the variability it seems likely that at least a portion of it is due to differences in pollutant levels and trends. While there is some variability among correlation coefficients, for nearly all sensors the periods of enhanced methane fall along the 1:1 line and most offsets occur at lower methane concentrations. Additionally, all RMSEs for co-located sensor fall below our expected uncertainty, while the RMSEs for deployed sensors is larger than this uncertainty (with the exception of the P1–P2 pair, which was co-located during the deployment).

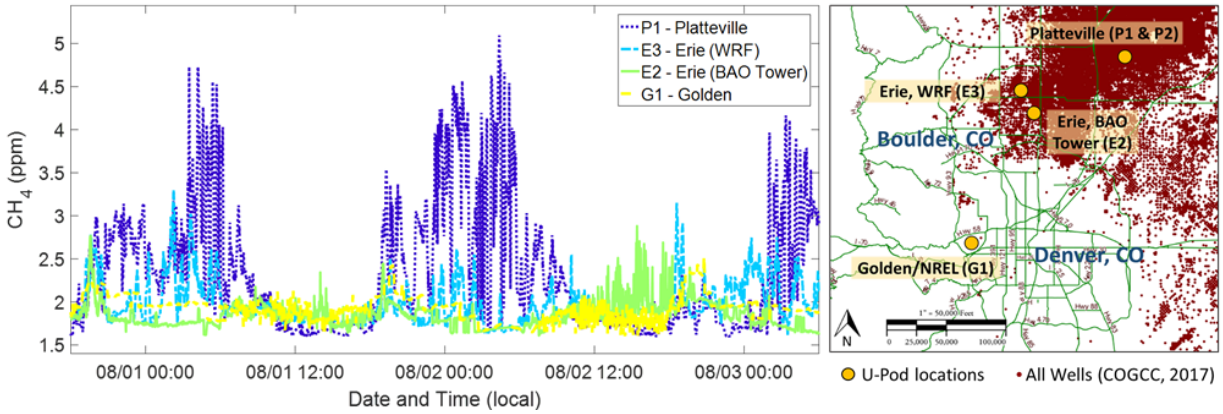


Figure 15: Time series of minute-resolution data from four Colorado sites and a map showing the location of the four U-Pod with respect to all active and inactive wells (COGCC, 2017).

This process provided minute-resolution methane estimates from 15 field sites, allowing for analysis of spatial data over different temporal scales. For example, Fig. 15 includes roughly 2 days of data from four different deployment sites: P1 was our primary U-Pod located at the Platteville site, U-Pod E2 was located at the Boulder Atmospheric Observatory, U-Pod E3 was located at a water reclamation facility, and the U-Pod G1 was located at the Golden site. Even from this small-time frame of data, we can see major differences between the sites. For example, there was a clear diurnal trend with elevated methane each night at the Platteville site. The high time resolution also allows us to observe short-term daytime increases at different sites, which were more sporadic and likely due to local emissions as there is typically more atmospheric mixing in the daytime (Bamberger et al., 2014). In contrast, the Golden site (U-Pod G1) exhibited relatively little variability in methane with differences between this site and the others well above our RMSE, suggesting significant differences in methane concentrations between these sites. Figure 15b provides a reminder of where the oil and gas wells of the Wattenberg Field are in relation to these U-Pods. These high-resolution data (minute-median) allow for the study of individual emission events and possibly their correlation to nearby activity or regional trends.

Figure 16 shows the day and night methane concentrations for each site throughout the deployment, grouped by region. This figure also highlights the ambient background for methane the U-Pod RMSE (0.3832 ppm) for this dataset on either side to illustrate that the enhancements above back-ground were well beyond our expected error. The sensor in Golden (U-Pod G1) exhibited little variability across both daytime and nighttime values, whereas all the sites in Erie, Frederick, and Platteville exhibited larger ranges and larger nighttime increases in methane likely contributed to by local or regional sources. At the majority of sites, over 50 % of the data fell within the RMSE of

typical background levels of methane. Though the middle 50 % of the nighttime data appears slightly shifted upward for U-Pods E11 and E3 in Erie (the two pods located near the water reclamation facility). This trend was even more pronounced at the sites in Frederick and Platteville. Recall the well density show in Figs. 2 and 15, illustrating no oil and gas activity around Golden, whereas we see higher-density activity in Erie and Frederick, with the highest density of activity around Platteville suggesting that one possible source driving this elevated methane is emissions from oil and gas activity. We observe this trend at night when atmospheric mixing is more limited and the planetary boundary layer is lower.

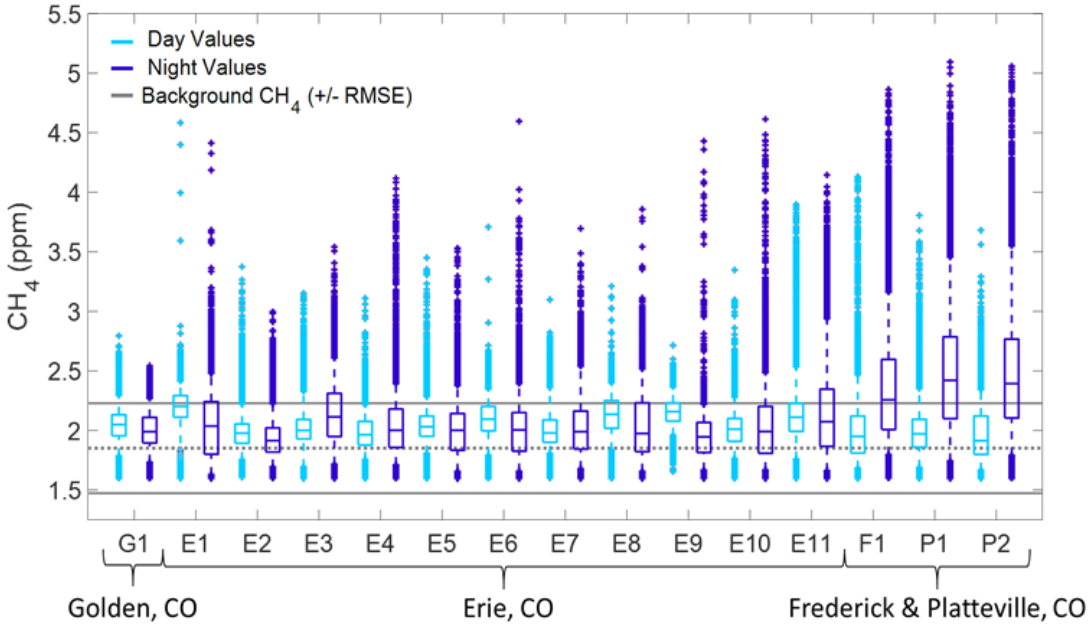


Figure 16: Box plots of all data for each U-Pod, grouped into daytime and nighttime values. Note that whiskers are the 5th and 95th percentile values. These U-Pods are then further grouped by region.

Figure 17 further illustrates this point by showing the difference in 90th percentile values between the main U-Pod (P1) and all other U-Pods during the day (left) versus at night (right). The daytime differences are small, within 0.2 ppm for all sites, possibly indicating effective daytime mixing. However, at night there is a clear gradient across the sites with little difference between the pods in Platteville and in-creasing differences as we move to the edge of the gas field and outside of it, with an approximately 1 ppm difference for the Golden site and the site furthest west in Erie. Throughout the Erie field sites, the two U-Pods furthest north show the smallest difference with the Platteville pods after the Frederick pod, which was located much further into the gas field. These two pods were also the ones located at the water reclamation facility and therefore subject to an additional local methane source. Interestingly,

however, the U-Pod furthest west in the Erie area was the only U-Pod in that grid located on the west side of the county line. This placed the pod in Boulder County during a time when a moratorium beginning in 2012 was in effect (Boulder County, 2016). This moratorium severely limited new oil and gas development in the county. Although we cannot conclusively say the observed difference in 90th percentile nighttime values is the result of differing methane trends on either side of the county line, it is an indication of a question possibly worth revisiting using other data collected during the FRAPPE/DISCOVER-AQ campaigns. More importantly, this example demonstrates how low-cost sensors can offer preliminary or supplementary data to help inform and guide future work.

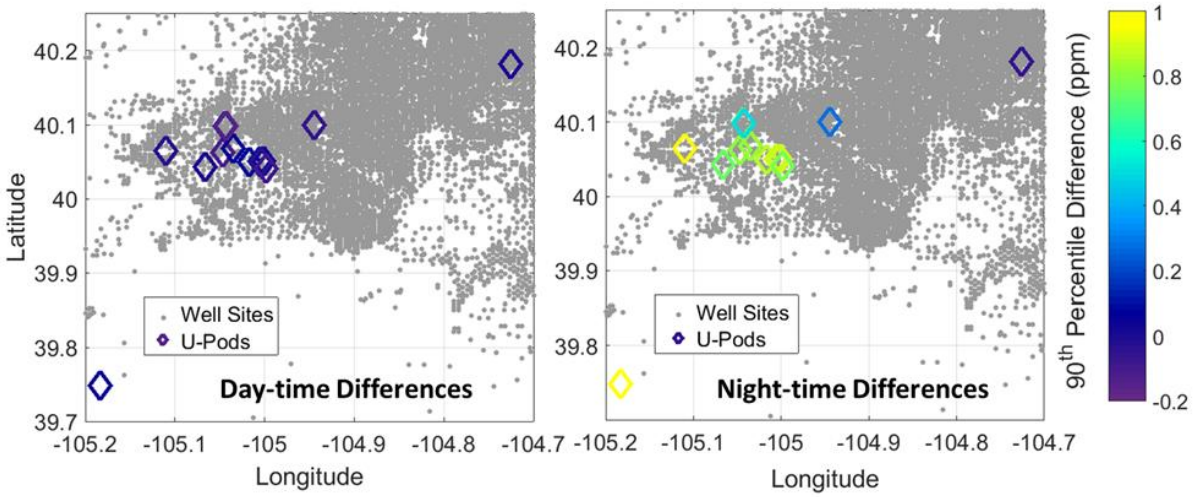


Figure 17: Well sites (COGCC, 2015) are plotted along with the differences between the 90th percentile value for the main Platteville U-Pod (P1) and every other U-Pod during the day (a) and night (b). The color bar indicates the magnitude of the difference in units of ppm CH_4.

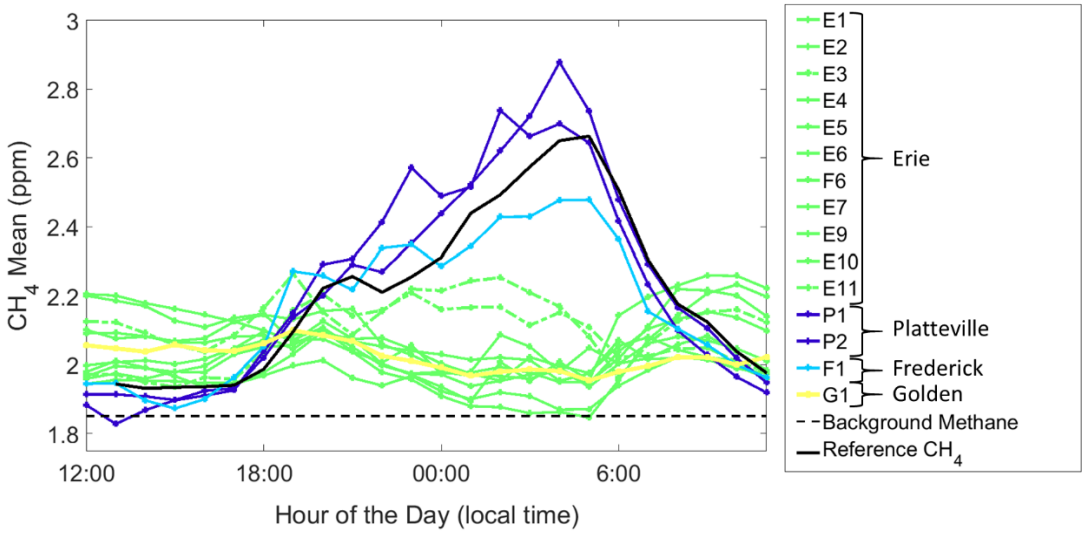


Figure 18: Figure 18. Mean methane value for each hour of the day, for each U-Pod, grouped and color-coded by region.

Figure 18 provides another overview of the field data. In this plot, each hour of the day is averaged for each pod using all available data – providing an indication of the diurnal patterns at each site. Again, we are seeing the nighttime increase in methane occurring at the Platteville site and to a lesser extent an increase at the Frederick site. These increases continue to be well above background and the estimated measurement error, which supports the conclusion that nighttime methane pooling was occurring in this location – a conclusion which is further supported by the observations of other researchers. Another study also conducted during the FRAPPE/DISCOVER-AQ campaign found elevated levels of benzene at the Platteville Atmospheric Observatory, occurring primarily at night. These elevated benzene concentrations were attributed to local oil and gas activity, as opposed to another source such as traffic, and the movement of the planetary boundary layer (Halliday et al., 2016). Figure 18 indicates that something similar may have been occurring with methane at the same site, likely driven by one or more sources and the fluctuations of the planetary boundary layer. Another study using data from 2013 found the mean level of light alkanes in Platteville elevated 5– 6 times above levels in Erie and 9–15 times above levels in downtown Denver (Thompson et al., 2014). This trend of elevated alkanes in Platteville and lower levels in Erie also agrees with the gradients apparent in Figs. 17 and 18 as we see the highest elevations in Platteville, moderate elevations in Frederick, and lower levels across our Erie sites. Overall, this confirms the ability of these low-cost sensors to provide unique information, in this case information regarding regional methane trends that is supported by studies that used more conventional monitoring instruments and sampling methods.

4 CONCLUSIONS

A common response to the question, "How good is low-cost sensor data?" is "it depends". It depends on what question you are trying to answer, what data you intend to collect, how you would like to use the data, and what supporting measurements are available. As demonstrated by the quantification system applied to the two deployments examined in this paper, the use of low-cost sensors, certainly in the short-term, is likely to be heavily application-dependent and sensors should be calibrated and quantified to meet the needs of a given research question and in response to the conditions of a particular deployment. As low-cost sensor systems become easier to deploy and data processing becomes more automated, these systems have tremendous potential. Their low-cost and portable nature allows for quick deployment across varied spatial scales, especially small, localized scales. Sensor data can already highlight potential "hotspots", which could lead to better allocation of resources or the detection of potential air quality

issues sooner. When used in this context, the sensor system described herein can provide a useful estimate of methane concentrations that may serve as preliminary or supplementary data. In Los Angeles, we were able to provide a methane prediction despite the complexity of sources and this methane signal has the potential to provide some insight into what is happening at the neighborhood level, although special attention will need to be paid to likely confounders and cross sensitivities. In Colorado, we were able to generate a dataset that can be examined on various temporal and spatial scales as well as data able to characterize regional trends that concur with the observations of other researchers. While more research into cross sensitivities and other deployment issues is certainly necessary, this sensor system currently provides a potentially powerful tool for understanding methane in communities near sources. Furthermore, this is a tool that is complementary to conventional monitoring methods.

ACKNOWLEDGEMENTS

Funding provided through the MetaSense Project (NSF Grant CNS-1446912), the AirWaterGas Project (NSF-SRN CBET: 1240584), and the DISCOVER-AQ Project (NASA). Publication of this chapter was also funded in part by the University of Colorado Boulder Libraries Open Access Fund. Thank you to all project partners during the DISCOVER-AQ/FRAPPE campaigns (NCAR, NOAA, CDPHE, US EPA), all project partners in Los Angeles at the University of Southern California Keck School of Medicine, Redeemer Community Partnership (Nicole Wong), Sandy Navarro, William Flores, Esperanza Community Housing, and Occidental College. Additional thanks to all research and regulatory partners who assisted with site access and reference data: the NATIVE trailer team, Colorado Department of Public Health and the Environment, and South Coast Air Quality Management District. Thanks to Christine Wiedinmyer (CIRES) for the deployment maps included in this paper (Section 2.2). Thanks to all monitor site hosts in Colorado and Los Angeles, and current and former members of the Hannigan Research Lab, especially Nicholas Masson and Drew Meyers for their work on the U-Pod/Y-Pod hardware and software, and Evan Coffey and Kira Sadighi for their assistance with the deployment in LA.

Chapter 7: Understanding the ability of low-cost MOx sensors to quantify ambient VOCs

Contributing Authors: Ashley M. Collier-Oxandale[1], Jacob Thorson[2], Hannah Halliday[3], & Michael Hannigan[2]

[1]University of Colorado Boulder, Environmental Engineering
[2]University of Colorado Boulder, Mechanical Engineering
[3]NASA Langley Research Center

Publication Information: manuscript in preparation

ABSTRACT

Volatile organic compounds (VOCs) present a unique challenge in air quality research given their importance to human and environmental health, and their complexity to monitor resulting from the number of possible sources and mixtures. New technologies, such as low-cost air quality sensors have the potential to support existing air quality measurement methods by providing high time and spatial resolution data. This higher resolution data could provide greater insight into specific events, sources, and local variability. Furthermore, given the potential for differences in selectivities for sensors, leveraging multiple sensors in an array format may even be able to provide insight into which VOCs or types of VOCs are present. During the FRAPPE/DISCOVER-AQ monitoring campaigns, our team was able to co-locate two sensor systems, using metal oxide (MOx) VOC sensors, with a proton-transfer-reaction mass spectrometer (PTR-MS) providing speciated VOC data. This dataset provided the opportunity to explore the ability of sensors to estimate specific VOCs and groups of VOCs in real-world conditions, e.g., dynamic temperature and humidity. Moreover, we were able to explore the impact of changing VOC compositions on sensor performance as well as the difference in selectivities of sensors in order to consider how this could be utilized. From this analysis, it seems that systems using multiple VOC sensors are able to provide VOC estimates at ambient levels for specific VOCs or groups of VOCs, it also seems that this performance is fairly robust to changing VOC mixtures, and it was confirmed that there are consistent and useful differences in selectivities between the two MOx sensors studied. While this study was fairly limited in scope, the results suggest that there is the potential for low-cost VOC sensors to support highly resolved, ambient hydrocarbon measurements. The availability of this technology could enhance research and monitoring for public health and communities impacted by air toxics, which in turn could support a better understanding of exposure and actions to reduce harmful exposure.

1 INTRODUCTION

1.1 Background

Volatile organic compounds (VOCs) are ubiquitous in daily life; from the naturally occurring scents of flowers blooming in the spring to VOCs resulting from human activity, such as BTEX emissions from vehicles, compounds emitted when cooking, and even fragrances in cleaning supplies and personal care products (McDonald et al., 2018). In addition to their ubiquitous nature, VOCs are wide ranging in terms of potential risks to our health. Many VOCs that pose a danger to human health are classified as Hazardous Air Pollutants (HAPs) by the US EPA (Woodruff et al., 1998). For example, two of the more toxic and prevalent compounds from the HAPs list are benzene and formaldehyde, both of which pose a variety of risks from acute toxic effects to long-term carcinogenic risks depending on the level of exposure (Suh et al., 200). The inhalation of benzene on shorter time scales can result in neurologic symptoms, such as dizziness, drowsiness, headaches, and unconsciousness (Suh et al., 2000). Formaldehyde has been cited as a concern for indoor air quality as it is a respiratory and sensory irritant (Rumchev et al., 2007). One study have found that children exposed to a median level of 20 $\mu g/m^3$ or more of benzene in their homes were eight times more likely to have asthma than children living in homes with lower levels of benzene; a similar link was found between formaldehyde and asthma with researchers finding a 3% increase in risk of having asthma for every 10 $\mu g/m^3$ increase in formaldehyde exposure (Rumchev et al., 2007). Additionally, researchers have found that both compounds rank among the highest in terms of cancer risk when personal exposure across microenvironments and different exposure pathways are taken into consideration (Loh et al., 2007).

Beyond these two compounds there are many other VOCs that warrant concern. The RIOPA Study measured VOC concentrations outside of homes to examine the impact of proximity to nearby emission sources on exposure. This study found certain VOC levels elevated 1.5-4 times above ambient levels for homes less than 50 meters from a source (Kwon et al., 2006). Among the VOCs studied during RIOPA were benzene and perchloroethylene, another known carcinogen. Researchers also found that living within 25 m of a source of either of these two compounds increased the lifetime upper-bound cancer risk by 50-200% as compared to living more than 250 m from a source (Kwon et al., 2006) – highlighting the importance of studying VOCs at increased spatial resolutions. A more recent study found that EPA determined cancer risks as well as non-cancer neurological and respiratory risk benchmarks were exceeded in two environmental justice communities for several compounds, including benzene (Wu et al., 2012). Beyond the health effects, VOCs are often the cause of odor complaints, such as those tied to industrial activity. Even

if there are no confirmed health effects from a pollutant, exposure to odors can cause quality of life issues and have hidden societal costs such as stress-related physical disorders (Beloff et al., 2000).

Given the prevalence of hazardous VOCs, more measurements and data could help inform actions to reduce VOCs and the public's exposure. However, given the spatial and temporal variability for VOC sources and complex mixtures of VOCs that occur in the environment, new approaches may be needed to supplement existing methods. Currently there are a variety of methods to quantify ambient VOCs, including real-time instruments, whole air sampling techniques, and passive methods capable of providing accurate, speciated measurements (Krol et al., 2010; Kumar & Viden, 2007). However, relying on a single high-quality instrument may miss important spatial patterns, and using a distributed method such as sorption tubes that provide time-averaged values may miss important temporal patterns. Next generation monitoring technologies, such as low-cost air quality sensors, used in combination with conventional techniques are an approach that may be able to help address these needs. Low-cost sensing systems often cost orders of magnitude less than conventional instruments on a per-unit basis and are simpler to deploy and operate making them particularly well-suited to provide preliminary or supplementary data for community-based projects or projects in partnership with environmental justice communities where resources may be limited (Shamasunder et al., 2018). Deployments of these sensing systems have already demonstrated the capacity to provide information on pollutant variability at small scales (Cheadle et al., 2017; Sadighi et al., 2018; Collier-Oxandale et al., 2018b), to differentiate regional trends from local emissions (Heimann et al., 2015), and to support personal exposure monitoring (Piedrahita et al., 2014; Jerrett et al., 2017). However, sensor performance quantification is an ongoing challenge for this technology. While some studies have demonstrated success quantifying sensors for gas-phase or criteria pollutants (Zimmerman et al., 2018; Casey et al., 2018c; Cross et al., 2017), this task may be more complicated for VOC sensors given that ambient VOCs exist in complex and dynamic mixtures.

1.2 Previous VOC Sensor Research

One of the reasons quantification is a challenge for low-cost sensors is their cross-sensitivity to environmental factors, such as temperature and humidity, and also to confounding gases (Lewis et al., 2016). Addressing this issue of cross-sensitivity is more complicated for VOC sensors compared to single chemical sensors as there are many more confounding species to consider and calibration models will need to be trained to target specific VOCs or groups of VOCs (Lewis et al, 2016). Several reviews provide an overview of the different sensors and systems available for measuring VOCs, including information on strengths and limitations, and performance evaluations based on

laboratory tests and examples found in the literature (Spinelle et al., 2017b; Spinelle et al., 2017c, Szulcynski & Gebicki, 2017; Williams & Kaufamn, 2015). These reviews also discuss the cross-sensitivity and selectivity issue; when considering the potential for sensors to make ambient measurements of benzene, the reviews noted that most sensors lack the selectivity and sensitivity for these measurements when sensors were tested individually (Spinelle et al., 2017b; Spinelle et al., 2017c). Researchers have also confirmed via laboratory tests that the limit of detection for most electrochemical and MOx sensors is too high for ambient measurements, and while photo ionization detector (PID) sensors are capable of lower detection limits with linear responses, these suffer from cross-sensitivities caused by interfering compounds. Similarly, laboratory evaluations conducted by the US EPA seemed to indicate there is the potential for these systems to record environmentally relevant levels of VOCs, however, they also found that only two of the five systems tested seemed capable of detecting VOCs below 25 ppb. While much of the current literature seems to suggest that the sensors currently available either lack low detection limits or the selectivity to pick out the compounds of concern, there are examples of deployments and laboratory studies presenting some innovative techniques for using these sensors and analyzing the data that could yield the useful information.

For example, DeVito and colleagues (2008) applied a neural net calibration to an array consisting of five different MOx sensors and provided a relatively stable benzene prediction with less than 2% error, for a period of 6 months. Furthermore, this study was conducted at a stationary monitoring site near a road, where the sensor system was subject to ambient temperature and humidity variations as well as varied concentrations of other VOCs (De Vito et al., 2008). Speaking to the potential for improving sensor quantification, the results of this study meet the Data Quality Objectives (DQO) outlined by the European Air Quality Directive for indicative benzene measurements, which call for a relative error less than 30% (Spinelle et al., 2017b). An earlier study by Wolfrum and colleagues (2006) demonstrated the use of MOx sensors (the Figaro TGS 2602) in arrays to differentiate and quantify three different VOCs (toluene, acetone, and isopropanol) in a laboratory setting. In addition to detecting these compounds at sub-ppm levels (approximately 0.1 – 1 ppm), analysis confirmed the sensor array's ability to predict individual pollutants in the presence of a confounding VOCs further speaking to the potential for VOC sensors. Another study by Eugster and Kling (2012), using a similar MOx sensor (the Figaro TGS 2600), demonstrated the detection of ambient methane in a remote area of Alaska throughout dynamic environmental conditions.

Other techniques being piloted to improve the capabilities of low-cost sensors include temperature-controlled operation (TCO) and/or utilizing a pre-concentrator (Schutze et al., 2017). TCO makes use of the fact that different

gases react with the surface of the sensor at different temperatures as well as the dynamic response after cleaning the surface via heating to elevated temperatures (Schutze et al., 2017). In one such study, researchers found that by applying TCO to a sensor system with MOx sensors, they were able to achieve an accuracy of ±0.2 – 2 ppb depending on the target gas concentration (benzene) and level of confounding gas(es) (Saurwald et al., 2018). Another study, also using TCO and multiple MOx sensors, demonstrated the ability to differentiate VOCs (benzene, formaldehyde, and naphthalene) at the ppb level, even in the presence of a confounding gas (ethanol) at much higher concentrations than the target gases (Leidinger et al., 2014). This differentiation was achieved using Linear Discriminant Analysis and was able to correctly classify the gas 95-99% of the time for concentrations of 4.7, 100, and 20 ppb for benzene, formaldehyde, and naphthalene respectively (Leidinger et al., 2014). In addition to TCO, another technique under consideration is the addition of an open pre-concentrator where a target gas could accumulate on an adsorbing material and then be thermally desorbed for analysis (Schutze et al., 2017; Leidinger et al., 2016). A system such as this would facilitate lower detection limits for MOx sensor systems. In addition to different techniques for collecting and processing sensor data, another option is to combine low-cost sensors with other measurement techniques. For example, a study in Philadelphia involved combining low-cost sensors (in this case PID sensors) and passive adsorption tubes (Thoma et al., 2016). The passive adsorption tubes provided speciated, quantified VOC data, while the sensor data along with meteorological information provided valuable information regarding pollutant trends and emission sources (Thoma et al., 2016). While there are many challenges associated with the use of VOC sensors, the potential this technology has to complement current monitoring efforts necessitates the exploration of these innovative solutions.

During the FRAPPE/DISCOVER-AQ campaign in Colorado, our team placed two low-cost sensor systems at the Platteville Atmospheric Observatory (PAO), co-located with a proton-transfer-reaction mass spectrometer (PTR-MS) that provided speciated VOC data. Each sensor system included two different MOx VOC sensors in addition to other gas-phase and environmental sensors. It is this combination of the availability of speciated VOC data and the dynamic environmental conditions of a field deployment that make this dataset and the subsequent analysis unique. There are numerous studies exploring the performance of these types of sensors in the lab when exposed to different VOCs and even complex VOC mixtures. There are several field studies examining the deployment of these sensors, however, these studies tend to involve a single VOC reference instrument (e.g., benzene) or target VOC. This dataset will help to further inform best practices and procedures for using these sensors thanks to the added complexity

of our reference data. In this paper we explore the quantification of these sensors for individual and grouped VOCs, we also examine the different selectivities of the two sensors to better understand how these differences can be leveraged, and finally we try to understand how consistent sensor performance might be across different atmospheric compositions. These results build off previous work that involved quantifying one of the MOx sensors for ambient levels of methane, allowing us to explore the advantages of multi-sensor systems (Collier-Oxandale et al., 2018).

2. METHODS

2.1 Deployment Overview

In the summer of 2014, during the FRAPPE and DISCOVER-AQ campaigns (Pfister et al., 2017), our team deployed a network of low-cost sensors systems in an attempt to quantify the small-scale spatial variability of pollutants. However, these measurement campaigns also provided the valuable opportunities to co-locate our systems with high quality, reliable reference instruments providing the opportunity to improve sensor quantification and validation. One such co-location was at the Platteville Atmospheric Observatory (PAO), where two sensor systems, termed U-Pods, were co-located with the NATIVE Trailer maintained by researchers from Penn State (Halliday et al., 2016). The two U-Pods were co-located on the roof of the NATIVE trailer for approximately one month from mid-July to mid-August, Figure 1c includes a photo of the two U-Pods. This site offered a unique dataset given the potential for different types of VOCs. The PAO is in a rural area to the northeast of populated urban areas and surrounded by nearby oil and gas activity; there was the potential for typical traffic and urban emissions as well as emissions from oil and gas activity, and possibly even from local agriculture. Figure 1a illustrates the site's placement in relation to nearby cities and active/inactive oil and gas wells.

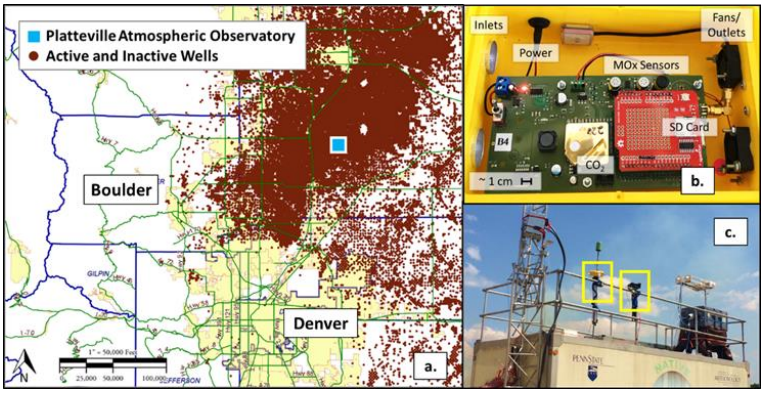


Figure 1: The map (1a) illustrates the placement of the Platteville site with respect to the nearby cities Denver and Boulder, and oil and gas activity as indicated by active and inactive wells (34). The top right photo (1b) shows the inside of a Y-Pod (a newer version of the U-Pod), which depicts the sensors used and the general design of the system. There were

some updates to the circuit board, but the sensors used were the same. The photo on the bottom right (1c) shows the placement of the two U-Pods on the NATIVE Trailer at the Platteville site.

This deployment is described in greater detail in our previous paper (Collier-Oxandale et al., 2018a). While VOC quantification was not the original intent of the deployment, it was something we were able to explore from the unique dataset provided through this co-location. Due to this analysis not being anticipated, techniques such as TCO were not incorporated; however, we did examine the use of a multi-sensor system in the context of relatively simple deployment and sensor performance quantification procedures. Thus, this work provides an opportunity to learn about VOC sensor potential at a fundamental level, under typical field conditions.

2.3 Reference Measurements

As previously mentioned, there were many reference instruments in operation at the PAO site, including a proton-transfer-reaction mass spectrometer (PTR-MS) that provided high-time resolution data for the VOC species listed in Table 1 (Halliday et al, 2016; Gouw & Warneke, 2007). In addition to the speciated VOC data, the Penn State NATIVE Trailer was outfitted to measure the other pollutants listed in Table 1 as well. Halliday and colleagues provide a detailed description of the operation of the PTR-MS during this campaign as well as more information on the other measurements occurring in the NATIVE Trailer (Halliday et al, 2016). All of the reference data was retrieved from the Discover-AQ data repository (NASA, 2015).

Table 1: Reference data utilized in this analysis and the instruments used to collect that data

Reference Pollutant	Instrumentation	Reference Pollutant	Instrumentation
Acetaldehyde	PTR-MS	Methane (CH_4)	Los Gatos CRDS
Acetone-Propanol	PTR-MS	Carbon Monoxide (CO)	Thermo CO Analyzer
Benzene	PTR-MS	Carbon Dioxide (CO_2)	Licor 7000
C_8 Alkylbenzenes	PTR-MS	Nitric Oxide (NO)	Thermo NO_y Analyzer
C_9 Alkylbenzenes	PTR-MS	Nitrogen Dioxide (NO_2)	Environment SA Analyzer
Formaldehyde	PTR-MS	Ozone (O_3)	Thermo O_3 Analyzer
Methanol	PTR-MS	Hydrogen Sulfide (H_2S)	PTR-MS
Toluene	PTR-MS		

2.2 MOx Sensors and the U-Pod Platform

MOx sensors are composed of a metal-oxide surface (often tin dioxide), a sensing chip to measure changes in conductivity, and a heater. The general mechanism is that oxygen molecules adsorb to the metal oxide surface, trapping electrons. When the sensor comes in to contact with the target reducing gas these oxygen molecules react

and are removed allowing the electrons to flow and increasing the conductivity across the surface (Wang et al., 2010). While the principle is simple, complications arise during quantification of pollutant concentrations, as the reactions occurring on the sensor surface are impacted by changes in temperature and humidity (Wang et al., 2010, Sun et al., 2012) as well as the fact that these sensors are cross-sensitive to gases other than the target gas (Spinelle et al., 2017c). Adding to this complexity, the nanostructure of the metal-oxide surface itself can also influence sensitivity and selectivity (Sun et al., 2012; Shen et al., 2018). These sensors were developed for and are typically used in scenarios where high pollutant concentrations would be expected, such as in an industrial setting or inside a vehicle engine.

Two VOC sensors are incorporated into our sensing platform, the TGS 2600 and the TGS 2602 (Figaro, Inc.). They are advertised for the detection of "air contaminants"; the manufacturer specifies a few different contaminants to which they are sensitive. These include methane, carbon monoxide, iso-butane, ethanol, and hydrogen for the TGS 2600 (Figaro, 2005a). Hydrogen, ammonia, ethanol, hydrogen sulphide, and toluene sensitives are specified for the TGS 2602 (Figaro, 2005b). Granted several of these contaminants are not VOCs, however, as we are utilizing these sensors for the detection of VOCs we will continue to identify them as VOC sensors. Additionally, both datasheets list a typical detection range of approximately 1-30 ppm (Figaro, 2005a & 2005b). However, as indicated by previous studies, sub-ppm levels of detection appear possible for both the TGS 2600 (Eugster & Kling, 2012) and the TGS 2602 (Wolfrum et al., 2006). The other environmental and gas-phase sensors used in the U-Pod are listed in Table 2 for reference.

Table 2: Complete list of sensors used in the U-Pod

Sensor Type	U-Pod
Temperature & Relative Humidity	RHT03 (aka DHT22)
Temperature & Pressure	47 Bosh BMP085
Carbon dioxide	ELT S-100 NDIR
Ozone	SGX Corporation MiCS-2611
VOC Sensor 1	Figaro TGS 2600 MOx
VOC Sensor 2	Figaro TGS 2602 MOx
Additional Optional Sensors	Alphasense B4 series (CO, NO, NO_2, O_3, SO_2), Baseline Mocon PID

The U-Pod is an embedded sensor system based on an open-sources design developed and assembled by our lab (Mobile Sensing Technology, 2017). These systems are housed in a small weather proof case, approximately 20cm x 25cm x 10cm, and use fans to pull air over the sensors and facilitate active flow. The U-Pods draw roughly 11 Watts of power and were powered by 12V AC/DC power adapters for this deployment, however they are capable of being powered by car batteries and/or solar power if remote deployment is necessary. The data is logged to an onboard

micro-SD card at a rate of one data point every 6-25 seconds, depending on how the system is programmed. Figure 1b includes a diagram of the interior or a Y-Pod, a newer version of the technology utilizing the same sensors. U-Pods and newer versions of the system (the Y-Pods) have been used in several indoor and outdoor air quality studies that included sensor quantification and an examination of spatial variability or air quality trends (Cheadle et al., 2017; Sadighi et al., 2018; Collier-Oxandale et al., 2018b).

2.4 Data Processing and Analysis Rationale

The variable voltage values associated with the changing conductivity of the sensors are recorded to a micro-SD card and this voltage is then converted to a normalized resistance value (R_s/R_0), which is typically the form used for sensor data analysis (Eugster & Kling, 2012; Piedrahita et al., 2014). The resistance is first calculated using Eq. 1, provided by the sensor manufacturer. In this equation R_s is the changing resistance in the sensor driven by the concentration of the target gas, while V_c is the circuit voltage, R_l is the load resistance, and V_{out} is the logged voltage. The R_0 value is typically the sensor resistance in clean air, and this value is used to normalize the resistance values. When calibrating in the field R_0 is identified as the maximum resistance value for the training period, or when the air is cleanest. For the following analysis this normalized term, R_s/R_0, is used. During processing, minute-medians are calculated from the sub-minute raw data, and warm-up data (the first half hour of operation after a U-Pod has been powered off for any period over half an hour) is removed.

$$R_s = \frac{V_c * R_L}{V_{out}} - R_L \qquad (1)$$

To facilitate analysis, the minute-median data was matched to the reference data using the nearest minute. This matched sensor (Table 2) and reference data (Table 1) was then five-minute averaged in blocks, in order to reduce the potential for any lags resulting from issues with time matching the data, particularly between short term spikes. Additionally, if three or more minutes were missing from either the sensor dataset or the reference dataset, then the whole five-minute average was excluded from the analysis. Occasional gaps in the reference datasets last from a few minutes to a few hours and vary by instrument; missing data was typically due to calibration events. One of the U-Pods, identified as P1 for this study, experienced a power failure resulting in approximately three days of data loss; the remaining data is complete. The second U-Pod, identified as P2 in this analysis, experienced intermittent power failure resulting in approximately 12 days of data lost in total out of the 22 days of deployment. However, the sensors remained fully functional throughout the deployment, despite power failures. While this analysis primarily utilizes

data from U-Pod P1, the data from P2 still provides an opportunity to validate our observations drawn from the P1 analysis.

To assess a VOC sensor's capabilities for use in the field, we applied typical quantification and analysis techniques. Given the cross-sensitives previously mentioned, researchers have found 'field calibration' or 'field normalization' to be a promising method to mitigate cross-sensitivities and calibrate for a target pollutant. Field calibrations are implemented by co-locating low-cost sensor systems with high-quality reference instruments (typically regulatory-grade, Federal Reference Method/Federal Equivalence Method monitors), often before and after a field deployment, and then generating a calibration model using an approach such as multiple linear regression or machine learning (Sadighi et al., 2018; Zimmerman et al., 2018; Cross et al., 2017). This technique allows predictive calibration models to be built for the conditions which sensors will experience in the field, such as diurnal environmental trends and background pollutants. While laboratory studies are valuable for understanding sensor capabilities and limitations in a controlled environment, researchers have continually observed that field as opposed to lab calibrations provide better pollutant estimations (Piedrahita et al., 2014; Castell et al., 2017). Therefore, the co-located data from the PAO site was used to conduct a typical field calibration by selecting a portion of the data from the beginning and the end of the deployment to build calibration models. The remaining data was used as testing data.

The models selected utilize multiple linear regression (MLR). While more complex machine learning techniques have proved very successful (Zimmerman et al., 2018; Casey et al., 2018c; De Vito et al., 2008), we wanted to start with a more fundamental and simple quantification method. Models were trained to predict benzene, summed aromatic species, summed total VOC species, and methane. For summed signals, the ppbC values for each compound were calculated as the number of carbons in the compound multiplied by the volumetric concentration of the same compound; these ppbC values were then summed (Chen et al., 2014). The summed aromatics signal was calculated as the sum of the ppbC concentration values for benzene, C_8 alkylbenzenes, C_9 alkylbenzenes, and toluene. This signal was calculated by summing the ppbC values for all of the available species measured by the PTR-MS: acetaldehyde, acetone, benzene, C_8 alkylbenzenes, C_9 alkylbenzenes, formaldehyde, methanol, and toluene.

Table 3 lists the multiple linear regression models utilized. Model 1 is a simple model including the two MOx VOC sensors, an interaction between the two sensors, and environmental predictors (e.g., temperature and humidity), and time to address drift. While Model 1 is the same for each target pollutant, Model 2 is different for each target group. For each Model 2, predictors were added to improve the resulting statistics and residuals for a given target

pollutant or pollutant group. In addition to simulating a field calibration, we also utilized bootstrapping and analysis of variance to get a more fundamental understanding of the sensors' selectivities and the consistency of their behavior.

Table 3: Multiple linear regression models used

Model Identifier	*Model*
Model 1: for all	$C = p_1 + p_2 * VOC_1 + p_3 * VOC_2 + p_4 * Temp. + p_5 * Abs.\ Hum. + p_6 * Time$
Model 2: for benzene	$C = p_1 + p_2 * VOC_1 + p_3 * VOC_2 + p_4 * Temp. + p_5 * Abs.\ Hum. + p_6 * Time + p_7 * VOC_2(V) + p_8 * (Temp. * VOC_2)$
Model 2: for aromatics	$C = p_1 + p_2 * VOC_1 + p_3 * VOC_2 + p_4 * Temp. + p_5 * Abs.\ Hum. + p_6 * Time + p_7 * VOC_2(V) + p_8 * (Temp. * VOC_2)$
Model 2: for VOCs	$C = p_1 + p_2 * VOC_1 + p_3 * VOC_2 + p_4 * Temp. + p_5 * Abs.\ Hum. + p_6 * Time + p_7 * (\ln(Temp.) * Abs.\ Hum.)$
Model 2: for methane	$C = p_1 + p_2 * VOC_1 + p_3 * VOC_2 + p_4 * Temp. + p_5 * Abs.\ Hum. + p_6 * Time + p_7 * CO2$

Model predictors: VOC1 – Figaro 2600 R/R0, VOC2 – Figaro 2602 R/R0, Temp – temperature (degrees C), Abs. Hum. – absolute humidity, Time – continuous time (to address drift), VOC2(V) – Figaro 2602 voltage signal, CO_2 – carbon dioxide concentration (in this case from the reference data); C is the concentration value being solved for, either an individual or group of VOCs

Clarifying how target VOCs and groups of VOCs were selected: benzene and summed aromatic species (including C_8 and C_9 alkylbenzenes) were selected for health reasons, as discussed in Section 1. While benzene health risks are the most well understood, the other common aromatic species (e.g., the BTEX compounds: benzene, ethylbenzene, toluene, and xylene) also present similar concerns for human health (Adgate et al., 2014). The summed total VOC signal was selected to provide some insight into the sensors' capacity to predict total non-methane hydrocarbons (TNMHCs). This signal was calculated by summing the ppbC values for all of the available species measured by the PTR-MS: acetaldehyde, acetone, benzene, C_8 alkylbenzenes, C_9 alkylbenzenes, formaldehyde, methanol, and toluene. This type of measurement may be useful in an area concerned with a broad array of air toxics or when used in combination with a method of VOC speciation.

3. RESULTS & DISCUSSION

3.1 Field Calibration Performance

In the following sections, we show the results of the two MLR models for predicting each of the target VOCs or groups of VOCs. In each case, a timeseries is included to illustrate qualitatively the models' ability to predict trends and VOC concentrations, while regression statistics note the performance of the model across training and testing data and any changes from Model 1 to Model 2. The training periods have been highlighted in yellow and are the same training and testing periods used in the previous methane quantification work (Collier-Oxandale et al., 2018a). Also

included are scatterplots to highlight improvements in testing data from Model 1 to Model 2, and boxplots of the residuals (observed – predicted) to show where the majority fall (despite the wide variance apparent in the scatterplot). Additional residual plots are available in the appendix (Figures VI.2). There are also plots of raw sensor data versus some of the reference data, providing an idea of trends prior to fitting, available in the appendix (Figures VI.11).

3.1.1 Predicting Benzene and Summed Aromatic Species

Figures 2 and 3 present the results for benzene and the summed aromatic level quantification. Overall, both models capture the diurnal trends and short term elevated concentrations, although the models under predict the highest concentration events. Model 2 performs better than Model 1, with an R^2's of 0.67 and 0.64 for benzene and summed aromatics respectively. In both cases, Model 2 pulls some of the more extreme values closer to the 1:1 line. For the aromatics, Model 2 also results in closer fitting values at low concentrations. Furthermore, the RMSEs for the Model 2 testing data, 0.52 ppb and 11.25 ppbC for benzene and summed aromatics respectively, seem low enough to enable useful predictions at ambient levels given that they are below the full dynamic range observed in this dataset. The underprediction of the benzene and summed aromatic peaks is most likely due to a limitation associated with sensor response time. The response time for the PTR-MS has a 1 second per species integration time during the 1-minute measurement cycle (Gouw & Warneke, 2007), however the MOx sensors respond more slowly as they are relying on a chemical reaction. This means that a sensor may not be able to reach steady state in the time it takes for a plume to pass.

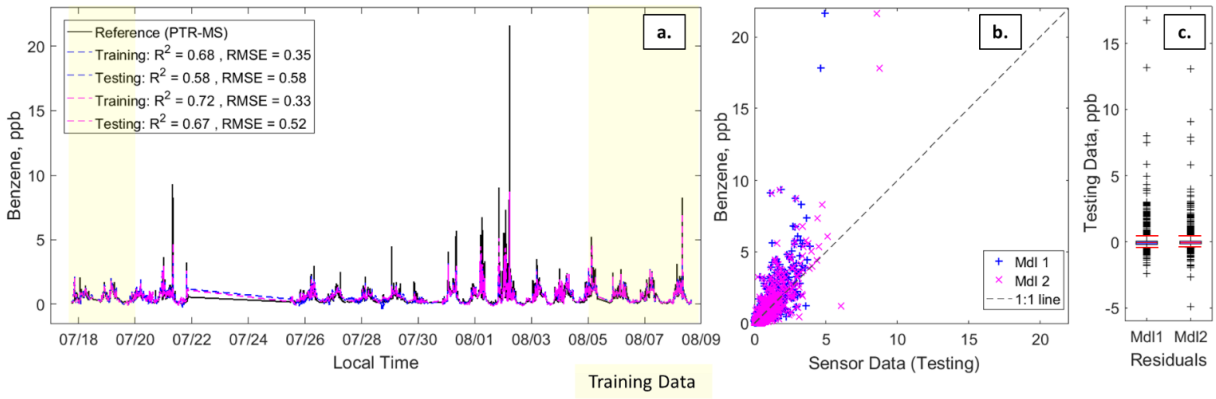


Figure 2: The far-left panel (2a) depicts a time series including the benzene data from the PTR-MS, and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5th and 95th percentile respectively (in red).

Despite this limitation, these two figures seem to suggest that these sensors can provide real-time estimates on aromatics and possibly even BTEX level estimates, a measure that could be especially valuable for exposure and health studies. Given that ethylbenzene and xylenes are C_8 alkylbenzenes, the relatively strong performance of the summed aromatics prediction supports the idea that these sensors are suited for a more targeted BTEX concentration estimates. Additionally, it is possible that more advanced analytical techniques, such as neural networks, could better preforming models, as show in the work of De Vito and colleagues with MOx sensors and benzene (2008).

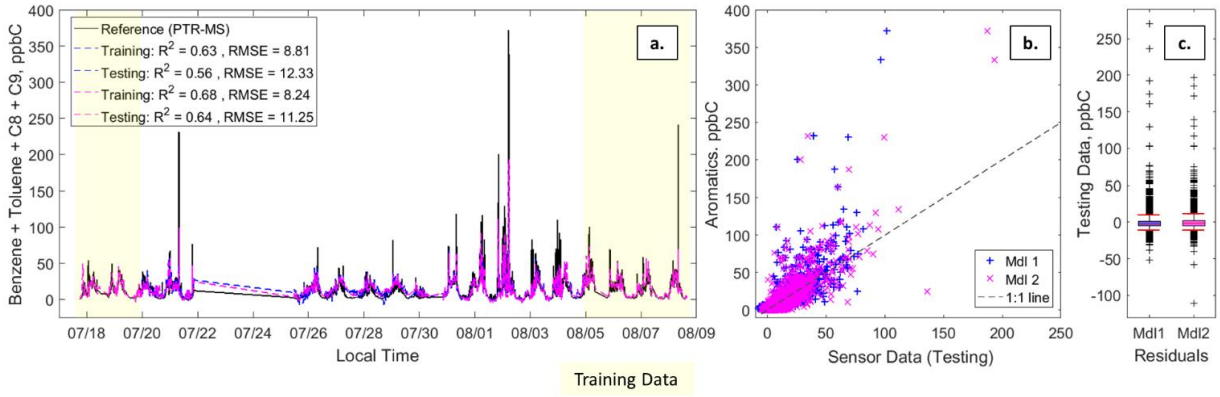


Figure 3: The far-left panel (2a) depicts a time series including the summed aromatic data from the PTR-MS (benzene, toluene, C_8 and C_9 alkylbenzenes), and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5th and 95th percentile respectively (in red).

3.1.2 Predicting Summed VOCs

Figure 4 illustrates the performance of both models for the sum of all VOC compounds available from the PTR-MS. Again, there are improvements with the more specialized Model 2, which corrects for some of the over-predictions. The resulting RMSEs of 13.38 and 12.78 ppbC and reasonably high R^2 values of 0.59 and 0.62 for Models 1 and 2 respectively, suggest this performance is suitable for certain ambient studies as the uncertainty is again well below the observed dynamic range. This analysis also demonstrates the stability of these models, as the signal being predicted is a sum whose precise composition is varying in time. For instance, the PTR-MS signals from the aromatic species are well-correlated with each other (Figure VI.1), but the aromatics are not well-correlated with any of the OVOCs, meaning the relative amounts of these compounds vary, potentially making the task of signal prediction more challenging. These results may also support the assertion that these sensors could be suited to make TNMHC measurements, a measurement sometimes made by regulatory agencies in areas where VOCs are a concern and real-time data is desired. Though speciated VOC measurements, made using canisters for example, are typically more

common. In terms of limitations, the models again under-predict the short-term peaks and this summed VOC estimate is noisier than for the previous two target VOCs (benzene and summed aromatics).

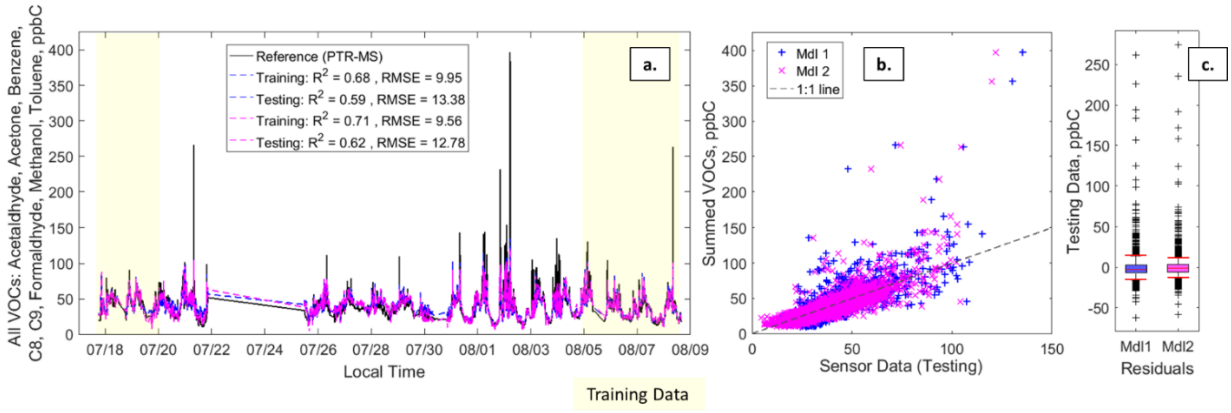


Figure 4: The far-left panel (2a) depicts a time series including the summed VOC data from the PTR-MS (acetaldehyde, acetone, formaldehyde, methanol, and the aromatic species), and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5th and 95th percentile respectively (in red).

3.1.3 Predicting Methane

Figure 5 depicts the performance of the models for methane quantification. As with the previous VOCs and groups of VOCs, the methane calibration models are able to predict periods of elevated methane and indicate some of the shorter-term methane plumes. There are also the same limitations as noted in the previous section, mainly the under-prediction of peaks. Expanding on our previous methane quantification work, which utilized only the Figaro TGS 2600 sensor, including the second VOC sensor improves our ability to predict methane levels. The R^2 and RMSE for the testing data in this previous work was 0.50 and 0.383 ppm respectively (Halliday et al., 2016). Simply adding an additional sensor to the model results in an R^2 and RMSE of 0.58 and 0.24 ppm for the Model 1 testing data. Model 2 is even better performing.

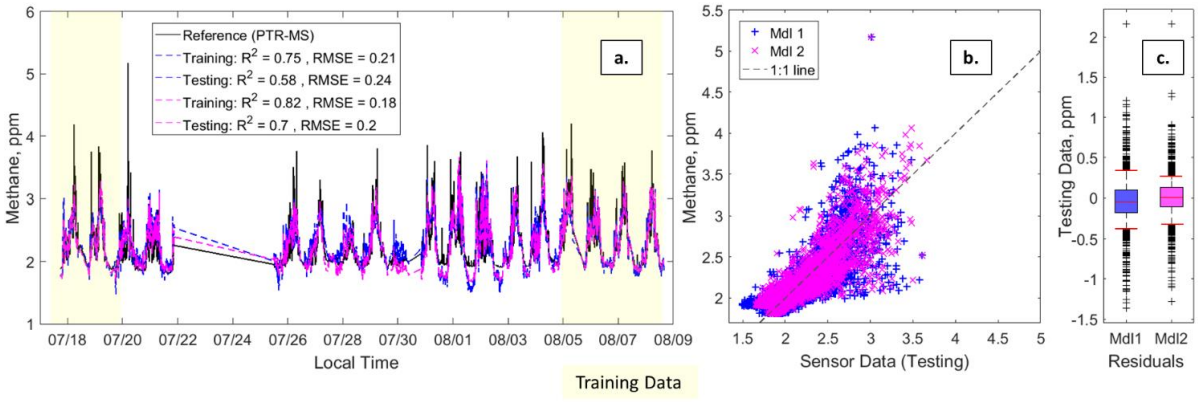


Figure 5: The far-left panel (2a) depicts a time series including the methane data from the Los Gatos cavity ring-down instrument, and the fitted sensor data from Model 1 (in blue) and Model 2 (in fuchsia). The text box includes the following statistics: coefficient of determination (R^2) and them the root mean squared error (RMSE) in that order for each testing and training dataset. The middle panel (2b) depicts a scatterplot of the testing data for Model 1 (in blue) and Model 2 (in fuchsia), the 1:1 line has also been added. The furthest right panel (2c) depicts boxplots of the residuals with the whiskers at the 5th and 95th percentile respectively (in red).

While the inclusion of a second VOC sensor does help to better target and predict methane, there is still room for improvement as is evidenced by the curvature in Figure 5b. As previously noted, it is possible that non-linear models or the use of a more advanced machine learning technique could facilitate further improvements. However, the overall results indicate that the use of multiple gas-phase sensors does help mitigate the cross-sensitivities noted in the previous paper and improve the performance of the calibration model (Collier-Oxandale et al., 2018a).

3.1.4 Model Accuracy, Specialization, and Robustness

The cumulative distributions of the relative error (in percent) for each set of fitted data (including the testing period only) is shown in Figure 6. Considering the relatively simple deployment and quantification procedures, this figure emphasizes the utility and potential for these MOx sensors. Applying the 30% relative error DQO for indicative benzene measurements required by the European Air Quality Directive to our measurements, upwards of 98% and 84% of the methane and summed VOC estimates meet this benchmark (Spinelle et al., 2017b). For benzene and the summed aromatics, this number is lower, 43% and 38% respectively. However, these larger relative errors seem to be primarily driven by fairly small differences in low observed concentrations and the associated predictions. For example, a 100% relative error resulting from an observed value of .5 ppb and a predicated value of 1 ppb. If low values are excluded from the datasets (instances where the reference data is below 0.5 ppb for benzene and below 10 ppbC for the summed aromatics, roughly the RMSE for each dataset), then the proportion of data meeting the

benchmark increases to 67% for benzene and 63% for summed aromatics. Excluding these low values might be reasonable for a study using sensors as the higher concentrations are most likely what would be of interest.

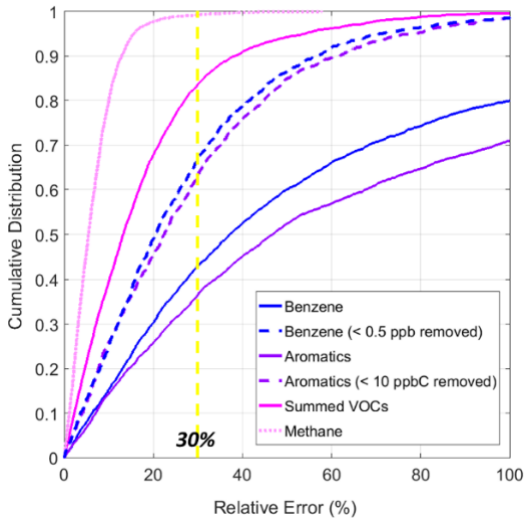


Figure 6: Cumulative distributions of relative sensor error for all fitted testing or validation data, the two dotted lines represent the fitted testing data for benzene and summed aromatic, respectively, with the low values removed from the reference data.

Given the limitations caused by cross-sensitivities, understanding how consistently sensor perform across changing compositions of VOCs is of high importance. While laboratory studies have illustrated the potential to identify specific pollutants in the presence of known confounders (Leidinger et al., 2014; Leidinger et al., 2017), the complex nature of field data requires a different approach. In Figure 7, the residuals from each Model 2 are plotted against all remaining VOCs. Figure 7a depicts the residuals for the benzene model versus the ppbC sum of all remaining non-methane VOCs. Figure 7b depicts the summed aromatic residuals versus the ppbC sum of all non-methane and non-aromatic species. Figure 7c depicts the summed VOC residuals versus methane, and in Figure 7d the methane residuals are plotted versus the ppbC sum of all non-methane VOCs. For each plot, two concentrations, indicated by colored markers, were selected to be held constant; these are roughly the 75th and 95th percentile values for the predicted datasets. The intention of this analysis was to examine the effects of varying levels of VOCs on model performance. More plots of residuals verses reference VOCs are available in the appendix (Figures VI.3).

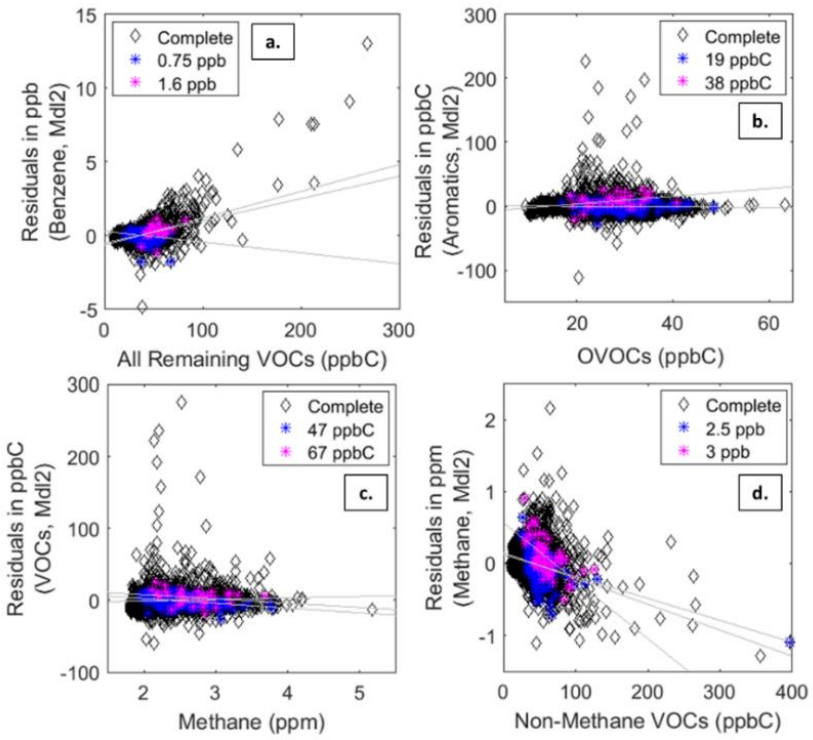


Figure 7: Complete residuals verses other, non-target VOCs. The complete residuals for the Model 2 results are in black. The blue and fuchsia points that represent two values held constant. For panel a, the constant benzene values selected were 0.75 and 1.6 ± 0.05 ppb. For panel b, the summed aromatics values selected were 19 and 38 ± 1 ppbC. For panel c, the summed VOC values selected were 47 and 67 ± 1 ppbC. For panel d, the methane values selected were 2.5 and 3 ± 0.05 ppm.

Figure 7b and 7c suggest that the predictions are robust across changing concentrations of other VOCs, as there are no positive or negative trends. In Figure 7c, the residuals seem to be largest for lower levels of methane and smaller for high levels of methane, which is possibly suggesting that summed VOCs are easier to estimate in the presence of higher concentrations of methane. However, this pattern is less apparent when we hold the estimated VOC concentrations constant. The negative and positive trends in Figures 7a and 7d emphasize that there is room for improvement in the models that predict a single compound. In Figure 7a, the positive trends once again highlight under-prediction of the highest benzene elevations. In Figure 7d, the negative trends suggest that over-predictions by the model may be driven by cross-sensitivities that are not adequately corrected for as these over-predictions correspond to higher summed NMVOC concentrations. However, Figures 7a and 7d plots still do not display any clear patterns, such as a well-defined linear relationship, again indicating some robustness amid changing VOC compositions.

In an effort to further understand model robustness and specialization, we explored model fit statistics during times when the target VOCs were well versus poorly correlated with a potentially confounding species. For this

analysis, we applied a 1-hour moving window to the data and calculated (1) the R^2 between the target VOC(s) and another/other VOC(s), (2) the RMSE for our Model 2 results for that hour, and (3) the average target VOC(s) concentration for that hour. These plots, Figure 8, help to confirm whether or not our models are being specialized to our target VOC(s). For example, if our models are predicting VOCs in a more general sense and not specializing for our targets, we would expect lower RMSE's to correspond to higher R^2's and higher RMSEs to correspond to lower R^2's, or periods when more differentiate was happening between the two signals. Conversely, if the models are being specialized to the target VOCs we would expect to see RMSEs that are more or less independent of correlations between different target VOC(s). In general, the results seem to point toward specialization.

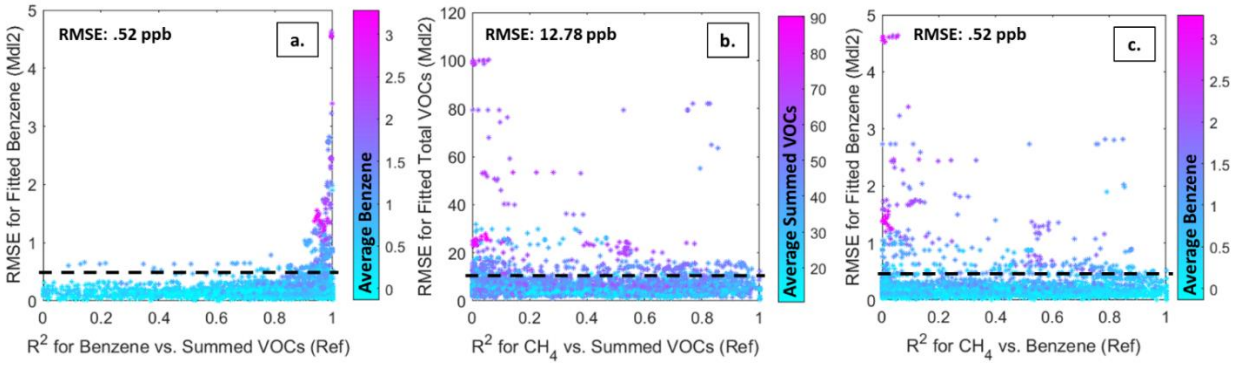


Figure 8: Plots of error (RMSE) versus the coefficient of determination (R^2) for the target VOC or group versus a non-target, potentially confounding VOC or group. These values are calculated on the basis of a 1-hour moving window and the points are colored according to the average of the target pollutant or group for that hour. The dotted black line on each is the overall RMSE determined in Section 3.1.

For our Model 2 for benzene, most of the RMSE's for each hour are below the overall RMSE for the model. For the points that are above the overall RMSE, many of these are for the high benzene concentrations – the underpredicted peaks. Furthermore, these higher benzene RSME values actually occur when we see high correlation between the summed VOCs and benzene, suggesting that the model is not simply fitting a more general VOC signal. Conversely, the high RMSE values do occur when there is low correlation between benzene and methane (Figure 8c), but there is still no clear trend of high RMSE for low correlation and low RMSE values for high correlation. For the summed VOCs, there is also not a clear relationship between correlation and RMSE values.

These results provide supporting evidence that the models are becoming specialized to the intended target VOC or group of VOCs. Table 4 provides further support for this point by illustrating lower correlation between fitted datasets for Model 2 versus Model 1 and when there is lower correlation between the reference datasets. The pollutant pair in Table 4 worthy of a closer look is the benzene and summed aromatics. In this case these two reference data

sets are very highly correlated, and it's not possible to confirm that the sensors are predicting benzene specifically, or a more general BTEX signature. The additional plots of residuals for the benzene Model 2, in Appendix VI, do indicate trends with regards to the benzene data and the other aromatic species. Therefore, while it seems that the models can be trained to predict specific or different groups of VOCs, for highly correlated species or groups (that the sensors are selective for) it may be more difficult to make this distinction. Sensor users should be careful to not over-assign meaning to signals that are more likely to be indicative of VOC types or groups rather than specific species.

Table 4: Coefficient of determination (R) between reference data pairs and fitted sensor data pairs.

Pollutant Pair	PT-RMS Data	Fitted Sensor Data (Model 1)	Fitted Sensor Data (Model 2)
Benzene vs. VOCs	0.906	0.936	0.877
Benzene vs. Methane	0.464	0.901	0.826
Benzene vs. Aromatics	0.957	0.989	0.987
Methane vs. VOCs	0.531	0.902	0.864

3.2 Sensor Selectivity & Consistency

The availability of speciated VOC data also allowed us to compare the selectivity of each sensor. Given that the manufacturer lists different compounds as target gases for each sensor, we expected a difference in the selectivity between the two VOC sensors. Furthermore, it's likely that differences in selectivity aided the specialization of the calibration models in Section 3.1.

We used analysis of variance to determine what the differences in selectivities might be and the results of this analysis are listed in Tables 5 and 6. The application of this technique is similar to previous studies where it was used to determine the effects of confounding species (Collier-Oxandale et al., 2018a; Eugster & Kling et al., 2012). These tables list the results of multiple runs in which different variables were included to determine their ability to explain the variance in the raw sensor signals. For instance, the first run includes only environmental parameters and time and for both sensors this set of predictors leaves the highest percentage of variability to residual error. Additional runs include all available reference signals (Run 2), or all available TNMHCs (Run 3). The final run in each case selects one VOC out of highly correlated groups (e.g., benzene for the aromatics and acetaldehyde for the aldehydes) and also results in the smallest portion of variance left to residual error, hence providing the strongest sets of explanatory variables. Comparing the two tables, the consistently important explanatory variables are highlighted in light grey. In Table 5, we see that acetaldehyde, benzene, methane and carbon monoxide are consistently important predictors for the Figaro 2600 signal, with methane and carbon monoxide as the most important. In Table 6, the percent

of variance in the Figaro 2602 sensor explained by benzene becomes more dominant, while carbon monoxide and the aldehydes are no longer important predictors. The lack of sensitivity to carbon monoxide for the Figaro 2602 sensor could be especially valuable in identifying the effects of this cross-sensitivity and correcting for it. It is important to note here that we do not expect the Figaro 2602 to be sensitive to pure methane (based on laboratory tests and manufacturer information), so it is probable that this response is driven by other light alkanes co-emitted and correlated with methane. Overall, this analysis confirms a difference in selectivities, which supports the idea that these two sensors can together be leveraged in VOC and source identification.

Table 5: Analysis of variance results for the Figaro 2600 sensor signal (R_s/R_0)

	Actd	Acet	Benz	C_8	C_9	Form	Meth	Tol	CH_4	CO	CO_2	O_3	NO_2	H_2S	Temp	AH	Time	Err
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	42.65	4.14	5.35	47.86
2	2.23	0.06	3.23	0.04	0.01	0.37	0.27	0.67	6.42	10.16	0.08	0.23	0.08	0.26	8.18	19.71	6.84	41.16
3	3.45	0.43	5.30	0.07	0.65	1.55	0.12	0.91	-	-	-	-	-	-	31.20	10.00	5.16	41.14
4	3.48	1.09	3.80	0.00	0.21	1.58	0.32	0.78	9.18	-	-	-	-	-	19.34	13.26	9.11	37.84
5	2.58	0.42	2.75	0.02	0.00	0.61	0.11	0.60	10.42	8.85	-	-	-	-	16.41	15.62	7.61	33.99
6	-	-	-	-	-	-	-	-	8.27	16.86	0.00	0.00	3.35	0.02	3.66	21.13	5.60	41.11
7	-	-	-	-	-	-	-	-	-	18.45	4.81	0.00	4.37	0.21	3.81	19.65	3.22	45.47
8	6.63	-	5.42	-	-	-	0.28	-	9.36	9.77	-	-	-	-	17.06	14.04	6.60	30.83

Table 6: Analysis of variance results for the Figaro 2602 sensor signal (R_s/R_0).

	Actd	Acet	Benz	C_8	C_9	Form	Meth	Tol	CH_4	CO	CO_2	O_3	NO_2	H_2S	Temp	AH	Time	Err
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	16.42	3.34	1.77	78.47
2	0.07	0.89	1.91	0.18	0.00	0.36	0.39	0.07	3.95	1.39	0.60	0.72	0.53	0.12	40.91	3.38	3.83	40.68
3	0.79	1.16	3.31	0.02	0.91	0.92	0.95	0.11	-	-	-	-	-	-	34.16	3.71	7.43	46.53
4	0.54	0.42	1.47	0.02	0.23	1.06	0.54	0.03	12.32	-	-	-	-	-	44.37	2.48	3.54	32.96
5	0.39	0.60	1.17	0.04	0.12	1.36	0.65	0.01	12.21	0.74	-	-	-	-	44.54	2.38	3.90	31.90
6	-	-	-	-	-	-	-	-	7.05	4.13	1.38	0.80	0.08	0.75	34.89	1.51	1.51	47.90
7	-	-	-	-	-	-	-	-	0.00	4.90	10.94	0.57	0.00	1.83	29.57	1.28	2.67	48.24
8	0.83	-	10.63	-	-	-	0.77	-	11.04	0.34	-	-	-	-	42.70	2.04	2.65	29.00

Additional ANOVA results are available in the appendix (Figures VI.4), including results of this analysis conducted on subsets of the data to test the robustness of our conclusions. The different subsets include the complete data, day vs. night data, and periods of elevated concentrations of specific compounds. Essentially these figures reinforce the conclusions drawn from the results above. Even across different subsets of data, the Figaro 2602 seems to be more responsive to aromatic species and the methane signal, while lacking sensitivity to carbon monoxide and the OVOCs. The Figaro 2600 is consistently responsive to methane, aromatic compounds, carbon monoxide, and to a

lesser extent the aldehyde species. For the subsets where we see low percents of variance explained by predictors, this may be due to relatively lower concentrations of the pollutants revealed to be important (Figure VI.5 illustrates the main differences in the subsets of data). As the conclusions drawn from the ANOVA results seem to be consistent across different runs and subsets of data, this supports the likelihood of consistency in sensor selectivities.

In addition to insights into selectivity these results also reiterate the importance of the cross-sensitivities to environmental parameters like temperature and humidity. As indicated in Tables 5 and 6 temperature and/or humidity often explain a greater percentage of variance in the sensor signal than the pollutants of interest. While the residuals from the regression analysis in Section 3.1 indicate that the models seem to be adequately correcting for temperature and humidity effects (Figures S2), the effects of these parameters are complex. We know that temperature can not only impact the rate of reactions occurring at the sensor surface, but also the rate of desorption [Schutze et al., 2017; Sun et al., 2012]. This behaviour means that temperature has the potential to impact the rates of response and recovery for the sensors as well as the magnitude of responses. Figures S6 provides an impression of these complexities. However, even though the pollutants explain a smaller percentage of the variance in sensor signal, performing the regression analysis from Section 3.1 with the sensor data excluded illustrates the value of the signals from the MOx sensors. Table 7 lists the results of the regression analysis for each Model 1 with all VOC sensor data excluded; this table also includes the original results in parentheses. Relying solely on environmental sensor data results in a higher RMSE and a lower R^2, particularly for the testing data. Plots in the appendix (Figures VI.7) further illustrate how excluding the sensor signal not only eliminates our ability to predict short-term enhancements in pollutant levels, but also reduces our ability to accurately track diurnal patterns.

Table 7: Model 1 regression statistics, excluding MOx sensor data (original results, including MOx sensor data)

	Training		**Testing**	
	R^2	**RMSE**	**R^2**	**RMSE**
Benzene (ppb)	0.22 *(0.68)*	0.55 *(0.35)*	0.08 *(0.58)*	0.83 *(0.58)*
Aromatics (ppbC)	0.24 *(0.63)*	12.6 *(8.81)*	0.09 *(0.56)*	17.72 *(12.33)*
Summed VOCs (ppbC)	0.26 *(0.68)*	15.13 *(9.95)*	0.07 *(0.59)*	20.72 *(13.38)*
Methane (ppm)	0.42 *(0.75)*	0.32 *(0.21)*	0.22 *(0.58)*	0.35 *(0.24)*

To better understand the consistency of the observed differences in sensor selectivities, we utilized bootstrapping for model training while also moving through different combinations of VOCs as predictors. Here each scenario was run 25 times with 15% of the dataset (in three-hour blocks) randomly selected for training and the remaining 85% of the data used for testing. Figure 9 shows the results for the testing data. The intended prediction

data sets were calculated by first summing the species in ppbC or ppmC and then normalizing the resulting sum to better make relative comparisons. These results continue to support the conclusions drawn thus far. The Figaro 2602 (noted as Fig_2) is better at predicting BTEX compounds than the Figaro 2600 (noted as Fig_1), however, the opposite is true when predicting methane and carbon monoxide. The poorest performance results from the predictions of summed, oxygenated VOCs. The model including both sensors and an interaction term nearly always provides the best results, indicating that power in leveraging the difference in selectivities between the two sensor types.

On a final note, regarding the consistency of these conclusions, this bootstrap analysis as well as the original regression analysis was repeated for the second U-Pod (P2) co-located at this site (the results are available in the appendix, Figures VI.8 and VI.9). In these results we see similar trends and behavior, but with poorer performance and greater variability. This poorer performance is likely due to the more fragmented nature of the data, as well as the possibly that the intermittent power failures affected the sensor signal enough to decrease the performance. Overall, the similarities in results suggest consistency within sensor types.

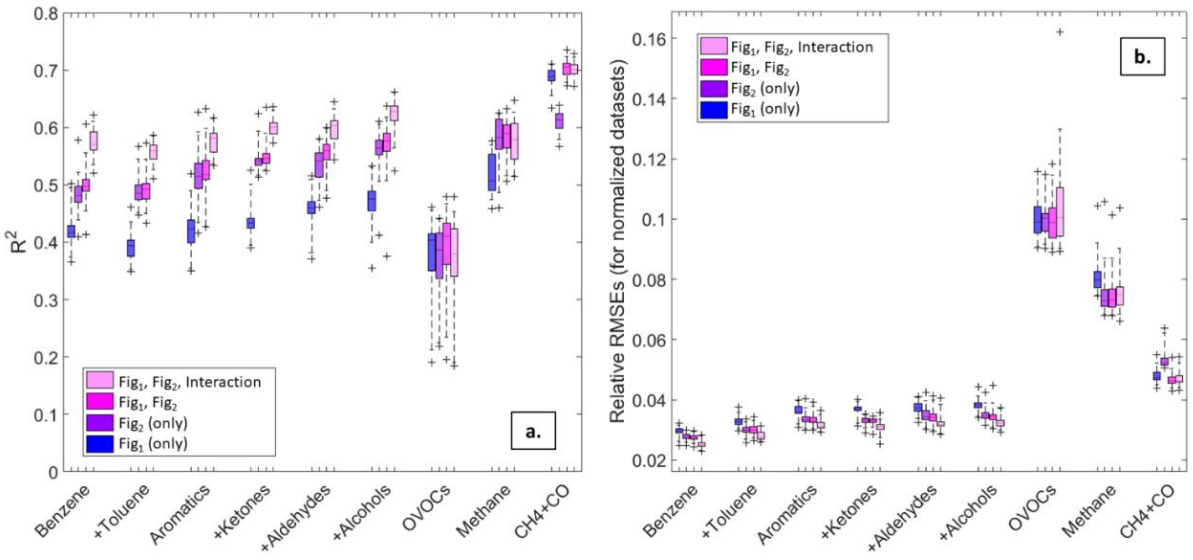


Figure 9: The boxplots above illustrate the results of training on a randomly selected 15% of the data and testing on the remaining 85% 25 times, using each sensor individually, the two sensors as predictors, and the two sensors plus an interaction between them. Plots a and b depict the R^2 and RMSE respectively for the testing data. Note all of the data has been normalized and the whiskers are the 5th and 95th percentile respectively.
The x-axis indicates the VOC or group being predicted.

3.3 Leveraging Sensor Differences to Learn About Potential Sources

Given the observed differences in selectivities, it is possible that sensor arrays may be able to provide useful information even in the absence of co-locations and field calibrations. For example, the ratios between different VOC sensors may be able to provide insight into VOC types or sources. This approach may be especially powerful if used

in conjunction with methods such as passive sorption tubes; sensors could indicate emission events and a rough idea of VOC composition and then sampling tubes could provide more quantitative speciation.

Here we compared sensor ratios to reference pollutant ratios and trends. For this analysis, the baseline was identified and removed from both the sensor and the reference data using a technique applied to sensor data by Heimann and colleagues (2015). The purpose of this baseline removal was to isolate short-term emissions and remove the larger regional/diurnal trends. We then calculated the R_s/R_0 ratio as the Figaro 2600/Figaro 2602 and removed ratio values deemed 'insignificant'; insignificant values were identified as ratios where the data from one or both sensors was below a given threshold. The threshold was 0.05 for the Figaro 2600 and 0.1 for the Figaro 2602. These thresholds were calculated as the average difference between paired values from the two co-located U-Pods P1 and P2. This was necessary as a low ratio could result from either the Figaro 2602 values being high or both the sensor values being low, in the latter case a low ratio may have been misleading. Finally, the different ratios were examined for correspondence to specific patterns in the reference data. Figure 10 below notes a few interested ratios and trends. A complete look at these ratios in relation to reference data is available in the Figures VI.10.

While the results of this analysis are limited to this single deployment, they suggest that this approach has potential. In Figure 10a, we see that a low ratio, indicated in fuchsia (meaning a greater response from the Figaro 2602 than from the Figaro 2600), corresponds to a higher toluene to benzene ratio. As Halliday and colleagues observed (2016), toluene to benzene ratios above 2.0 are more likely to be the result of traffic, while lower ratios are more likely to be indicative of oil and gas emissions. If this VOC sensor ratio were to consistently indicate high toluene to benzene ratios over other concentrations and VOC compositions, this could be a powerful tool for differentiating between traffic and oil and gas emissions. Further supporting this point, a larger ratio indicated in yellow (meaning a greater response from the Figaro 2600 than from the 2602) falls much lower than the toluene to benzene ratio of 2.0, with a ratio of approximately 0.89. When examining this same ratio with regards to the benzene to methane relationship (Figure 10b), it falls close to ratios observed in other studies in oil and gas areas. Thus, this analysis provides two examples supporting the idea that this ratio may be indicative of oil and gas emissions. One final observation, in Figures 10a and 10c, we see the ratio that falls closest to a 1:1 relationship for the sensors (in green) both falls below the toluene:benzene ratio of 2.0 and corresponds to many of the enhancements in benzene that seem to occur independent of enhancements in carbon dioxide. While there are no clear relationships between the benzene and carbon dioxide, this behaviour may illustrate an ability to indicate oil and gas emissions that are not the result of

combustion, but rather evaporative emissions. Though of course these observations would need to be demonstrated as consistent across different locations and with respect to differing background VOCs to ensure their reliability – the results here are limited to this single time and place.

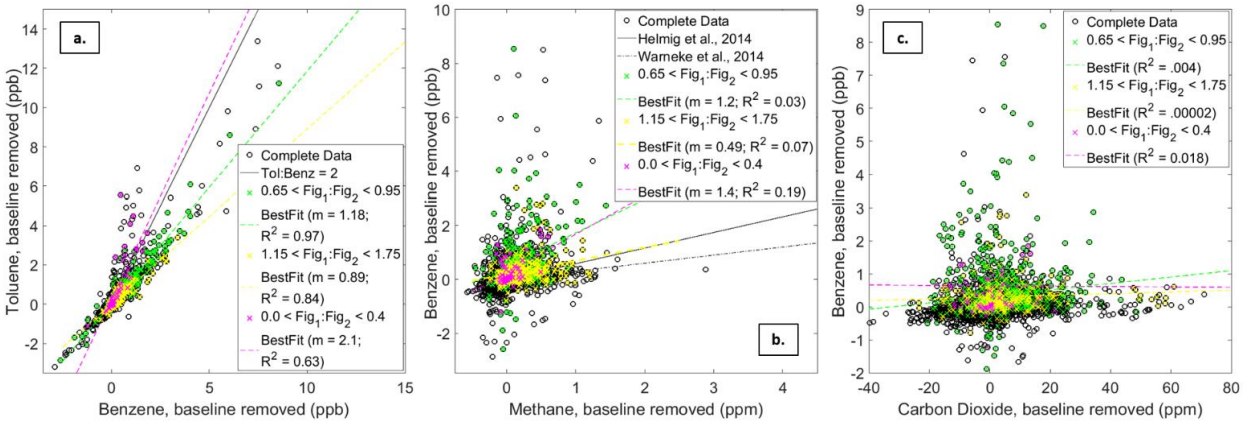


Figure 10: Each panel displays two pollutants plotted together (from the reference data set). Panel a is the benzene versus the toluene while panel b is methane versus the benzene and panel c is carbon dioxide versus benzene. Points with certain sensor ratios are then colored according to the colors listed in each the legend. Other relevant information, for example a ratio of 2 for the toluene to benzene ratio or relationships from previous studies.

Returning to Figure 10b, the relationship for the ratio in yellow is weak given the poor correlation. However, by utilizing a moving correlation to essentially extract periods where the relationship between the benzene and methane is stronger, indicating a likely shared source, we can better assess association of the yellow sensor ratio seems to benzene:methane ratios indicative of oil and gas areas. Using a moving correlation with an hour-long window, we calculated the correlation coefficients for each hour, the average ratios of benzene:methane and Fig_1:Fig_2, and the range of methane for that hour. Figure 11 depicts the distribution for the complete hourly averages of VOC sensor ratios, and a subset. In the selected subset, hourly averages correspond to an R higher than 0.85, a benzene:methane ratio < 1 ppb/ppm (the typical range observed in an oil and gas area), and a range in methane > 0.5 ppm. The shift in the distribution highlights that the emissions likely from oil and gas seem to be associated with the VOC sensor ratio of 1.15-1.75, which supports the idea that these sensor ratios may be able to differentiate oil and gas sources of hydrocarbons from traffic sources.

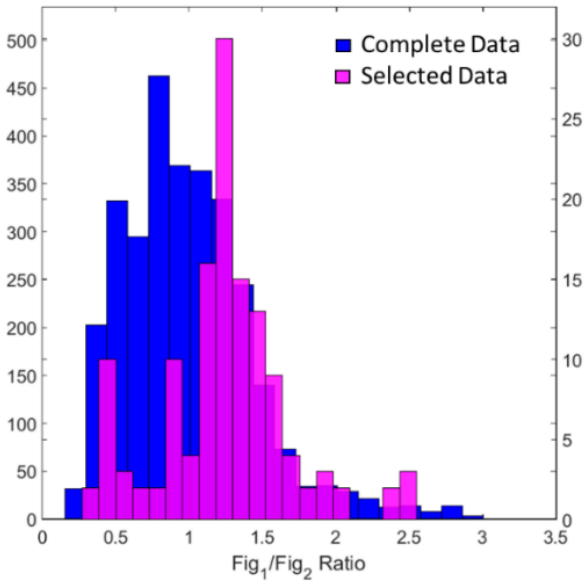


Figure 11: Histogram of complete sensor ratios for each hour using moving correlation window vs. the ratios for a select subset of data with high correlation between benzene and methane (> 0.85), a ratio of benzene:methane typically associated with oil and gas activity (< 1.0 ppb/ppm), and a significant change in methane (> 0.5 ppm).

4. CONCLUSION

While more field research is necessary, here we have provided an overview of MOx VOC sensor potential. Not only were calibration models capable of providing predictions relevant for ambient studies, but also these models appear to be specialized to the target pollutants and robust across changing compositions of other VOCs. Furthermore, this analysis confirmed a difference in selectivity between two MOx VOC sensors, a difference which can be leveraged in the development of calibration models, to identify and mitigate cross-sensitivities, and potentially in source classification. Cross-sensitivities to confounding species are currently a major concern, for low-cost sensors in general and in particular for VOC sensing. However, given the differences in selectivity it seems that multiple sensors could be used to strategically determine the gases most likely affecting a sensor. For example, a carbon monoxide sensor and the Figaro 2602 could help to confirm whether methane is the main driver of a response from the Figaro 2600. Furthermore, if the demonstrated association between sensor ratios and source types is shown to be consistent, multi-sensor devices could be a powerful tool for collecting preliminary or supplementary data in areas affected by numerous and complex sources – like environmental justice communities.

Low-cost field deployed MOx VOC sensors have the potential to provide information in support public health research, community-based environmental justice studies, or even supplement research by the regulatory or academic community. Given their cost and the relative ease of deployment, these tools can provide information at higher spatial

and temporal resolution than is currently available. Even considering the uncertainties and limitations discussed in this paper, these tools could guide exposure studies or provide a better idea of the impact of nearby sources on overburdened communities. Often in environmental justice communities, lacking resources, even cursory information on VOCs and local emissions could be valuable. These types of sensors could also supplement conventional monitoring approaches. For example, regulatory agencies sometimes utilize TNMHC measurements and MOx sensors may be able to supplement these instruments again by providing greater spatial resolution. Multi-sensor systems could also provide time-resolved information, adding to data collected using a speciated method such as VOC canisters or passive sorption tubes. Future research will hopefully explore these applications as well as further quantify the capacity and limitations of these sensors, however, the usefulness demonstrated here speaks to the potential MOx sensors have to provide new insights into the complex and dynamic VOC types and sources impacting our lives and communities.

Acknowledgments

Funding provided through the MetaSense Project (NSF Grant CNS-1446912), the AirWaterGas Project (NSF-SRN CBET: 1240584), and the DISCOVER-AQ Project (NASA). Thank you to all project partners during the DISCOVER-AQ/FRAPPE campaigns (NCAR, NOAA, CDPHE, US EPA). Additional thanks to all research and regulatory partners from the NATIVE trailer team, Thanks to all monitor site hosts. current and former members of the Hannigan Research Lab.

Chapter 8: Using Gas-Phase Air Quality Sensors to Disentangle Potential Sources at a Neighborhood Scale in Los Angeles

Contributing Authors: Ashley Collier-Oxandale[1], Nicole Wong[2], Jill Johnston[3], & Michael Hannigan[4]

[1]University of Colorado Boulder, Environmental Engineering
[2]Redeemer Community Partnership
[4]University of Southern California, Keck School of Medicine
[3]University of Colorado Boulder, Mechanical Engineering

Publication Information: manuscript in preparation

ABSTRACT

In the late summer of 2016, our team deployed a network of low-cost air quality sensing systems in partnership with community-based organizations, in a neighborhood in South Los Angeles. Residents of this community are concerned about possible emissions from local oil and gas activity, however in addition to these potential emissions, the neighborhood is also subject to a complex mixture of pollutants from other nearby sources including major highways. For this deployment, metal-oxide VOC sensors were quantified to provide methane and total non-methane hydrocarbon concentration estimates. This data along with other sensor signals, meteorological data, and community member observations was used to examine the composition and likely origin of emissions. We found that the sensor network displayed expected environmental trends and highlighted short-term elevations on CH_4 and/or TNMHCs which we were then able to investigate more closely. This study confirmed the existence of multiple sources impacting homes near the oil and gas site of interest, including a source that emits volatilized or vented hydrocarbons as opposed to hydrocarbons associated with combustion. We hope that this deployment can also serve as a model for how information from multiple sensor signals can be leveraged to better understand sources in complex areas, potentially supporting more community-based air quality research.

1. INTRODUCTION

Much of the research into low-cost air quality sensors has focused on quantifying sensors for criteria pollutants. However, moving beyond criteria pollutants and adding the capacity to estimate Hazardous Air Pollutants and greenhouse gases will expand our ability to identify sources and study important research questions related to health and exposure. Expanding the list of pollutants we are able to measure is also compatible with a shift to studying

air quality at a more localized level. The implementation of the Clean Air Act and subsequent work led to immense improvements in air quality, particularly at the regional scale. These improvements, on larger temporal and spatial scales, have allowed for a shift in focus to examining issues at a higher resolution. Studying air quality at this scale may necessitate examining emissions events and individual sources, making the capacity to measure a variety of pollutants valuable. Furthermore, these sources can carry risks to public and environmental health, for example, the potential for the emission of air toxics, including compounds such as benzene and formaldehyde. These compounds can originate from a number of sources outdoors and indoors and both rank high in terms of cancer risk when potential exposure pathways are taken into consideration (Loh et al., 2007). Alternatively, methane, a potent greenhouse gas, carries environmental concerns and recent studies have found methane emissions to be underestimated in current emissions inventories (Miller et al., 2013). Studies demonstrating the existence and impact of leaks along the entire production and distribution chain raise concerns (Marchese at al., 2015; Philips et al., 2013). For example, a study from 2015, in the Barnett Shale determined that roughly half of methane emissions were originating from 2% of oil and gas facilities, pointing to the need to identify individual sources having relatively larger impacts (Zavala-araiza at al., 2015).

As studies continue to focus in on these smaller scales, we are learning more about the small-scale spatial variability of pollutants. When considering this variability of pollutants, one important source is major roadways and the health of those living in close proximity to them. A review from 2007 noted the consistent observation of steep pollutant gradients near highways, ranging from 2 – 400 meters (Brugge et al., 2007). Within these gradients, the higher concentrations of certain pollutants can have an impact on public health. For example, one study cited in the review found that benzene and summed VOCs were twice as high outside of homes in areas with heavy traffic versus low traffic (Fischer et al., 2000). Sources other than roadways have also been shown to have an impact on local air quality and exposure as well. In the RIOPA study, researchers measured concentrations of a variety of VOCs, aromatic hydrocarbons (BTEX), methyl ter butyl ether (MTBE), and tetrachloroethylene (PCE), and examined the resulting concentrations in relation to proximity to sources (Kwon et la., 2006). The study found that residential ambient concentrations of VOCs were 1.5-4 times higher within 50 meters of a respective source (e.g., gas stations for BTEX, or dry cleaners for PCE) (Kwon et al., 2006), emphasizing the potential for small scale variability of pollutants near sources. Another study, utilizing passive sampling techniques to examine spatial variability around a refinery in Philadelphia, found significantly higher concentrations of BTEX and styrene on the fence line of the refinery as

opposed to further way (Mukerjee et al., 2016). In terms of scale, the data utilized was from sampling sites within a mile of the refinery (Mukerjee et al., 2016). New monitoring technologies are emerging, which will support more in-depth studied into small-scale spatial and temporal variability. One example is low-cost air quality sensors, which, while still a developing technology, can be deployed in networks to provide high spatial and temporal resolution data. A solution such as this could be especially valuable in places where the sources of interest exist amid a complex mixture of point and mobile sources.

1.1 Background

Los Angeles, CA has long been the subject of much research and action regarding air pollution due to the numerous sources, continued growth, and effects of the topography and meteorology on pollution accumulation and dispersion (Kunzli et al., 2003). One industry that communities are particularly concerned about in terms of potential impacts to local air quality is urban oil and gas activity. Los Angeles county holds more than 5000 active oil wells, with 850 of these located within the City of Los Angeles (Sadd & Shamasunder, 2015). Rapid development over the years has led to a "conflict in land usage" (Chilingar et al., 2005), which has resulted in the active wells and production sites in close proximity to high density residential areas and public services such as schools and hospitals (Shamasunder et al., 2018). In some instances, an oil well head may be within 60 ft of a residence (Shamasunder et al., 2018). Furthermore, residents living near some of these active facilities, on occasion, have been known to report adverse physical symptoms such as nose bleeds, headaches, and respiratory issues as well as other complaints such as disruptive noises or odors (Sahagun, 2013). Given the potential for the release of air toxics from these facilities, such as BTEX (benzene, toluene, ethylbenzene, and xylene) compounds (Adgate et al., 2014; Helmig et al., 2014; Moore et al., 2014), it is understandable that a facility in close proximity to a home might evoke concern. In addition to a potential for direct health effects, depending on the compounds emitted, the presence of odors could lead to decreased property values, or even stress-related physical disorders (Beloff et al., 2000).

Motivated by a concern for their health, residents of these communities often seek ways to collect more information or data to better understand their exposure and potentially facilitate action to reduce exposure (Brown, 1992). One example is a project where residents worked with academic and regulatory partners to deploy a network of lower cost air quality monitors that could provide more locally relevant information than the existing monitoring network, which in turn could be used to inform actions of local schools to reduce students' exposure (Wong et al., 2018; English et al., 2016). The increasing availability of more accessible environmental monitoring technologies,

along with greater engagement from academic and regulatory researchers could support more community-based air quality research projects. In turn, this type of research could lead to greater insights into human exposure and the effects of pollution in complex environments as well as locally relevant actions reduce exposure.

1.2 Low-cost Air Quality Sensors – An Overview

Next generation monitoring technologies, such as low-cost air quality sensors, have the potential to provide preliminary data to inform targeted studies, to supplement our existing monitoring networks with additional data, and to aid in the quicker detection of hotspots (Snyder et al., 2013). The cost of these systems (typically $500 -$5000 ea.) makes the deployment of networks of sensors more feasible, leading to datasets with much high spatial and temporal resolution. Furthermore, because of the lower costs and the relatively simple deployment and operation procedures, this technology is well-suited to support community-based investigations (Shamasunder et al., 2018). However, an ongoing challenge associated with the use of low-cost sensors is quantification and ensuring/assessing data quality.

These sensors exhibit cross-sensitivities to environmental parameters, like temperature and humidity, as well as confounding pollutants (Lewis et al., 2016). Significant research, both in the lab and the field, has gone into understanding and mitigating these cross-sensitivities in order to provide reliable calibrations for low-cost sensors (Masson et al., 2015a & 2015b). There are also many promising examples of sensors quantified for the detection of criteria pollutants (Mead et al., 2013; Piedrahita et al., 2014; Sadighi et al., 2018; Hagan et al., 2018) and hydrocarbons (Eugster & Kling, 2012; Collier-Oxandale et al., 2018a; De Vito et al., 2008; Leidinger et al., 2014; Collier-Oxandale et al., 2018c) at ambient levels. Research around sensor quantification also continues to explore the application of innovative techniques, such as machine learning or temperature-controlled operation, to improve the capabilities and quantification of low-cost sensors (Casey et al., 2018c; Cross et al., 2017; Zimmerman et al., 2018; Sauerwald et al., 2018; Schultze et al., 2017; Kim et al., 2017). Beyond quantification there are also example of sensors being used to study spatial variability of pollutants, to examine indoor air quality, and to support personal exposure monitoring (Sadighi et al., 2018; Cheadle et al., 2017; Casey et al., 2018a; Jerrett et al., 2017). As quantification continues to improve and best practices are established, it is possible that sensors could play a valuable role in addressing environmental inequities by helping to identify communities or areas overburdened by air pollution and potentially the likely sources of that pollution. A summary from a recent workshop provides a good overview of the current state of air quality sensor research (Clements et al., 2017).

In this paper, we utilize data from metal oxide (MOx) VOC sensors to provide quantitative estimates of methane (CH_4) and total non-methane hydrocarbon (TNMHC) concentrations (Collier-Oxandale et al., 2018a & 2018c). These estimates are then utilized along with other gas-phase sensor data and meteorological information to reveal air quality trends at the local level in a South Los Angeles neighborhood. These additional data streams are utilized to better understand the composition of emissions events and determine likely sources. To our knowledge, this research is the first to use sensors calibrated for CH_4 and NMHC sensors in a complex urban environment to study pollutant variability and individual emission events. Furthermore, the use of multiple sensors and analysis techniques shared in this paper could serve as a model for deploying sensors with communities to investigate local concerns around issues such as odors and industrial activity. This paper includes an evaluation of the sensor quantification efforts, an overview of the spatial and temporal variability seen in the sensor data, and analysis of individual emission events informed by supplementary data and in relation to local observations by community members.

2. METHODS

2.1 Study Overview

The low-cost sensing systems were deployed for a period of 8 weeks, from late summer through early fall, in a community in South Los Angeles. Figure 1 illustrates the sampling sites in relation to sources of interest; the sites are distributed on either side of and varying distances away from the major highways indicated by black lines and an active oil extraction site indicated by a red star. Fourteen of the sites were located within a roughly 5x5 km area. The two additional sites, were further away and were selected as they allowed for continuous co-location with reference instruments providing additional quantification validation data. The neighborhood in which the sensor systems were deployed is primarily high density residential with some commercial and industrial land use. While only the major highways are indicated, some other sites were located on fairly busy roads and vehicle emissions certainly would originate from these roads too. Other potential local sources of VOCs would include, most prominently, gas stations and dry cleaners. There are also other active and inactive extraction and processing sites in the area, though there are no other active drill sites near any sampling site other than the one noted in the map.

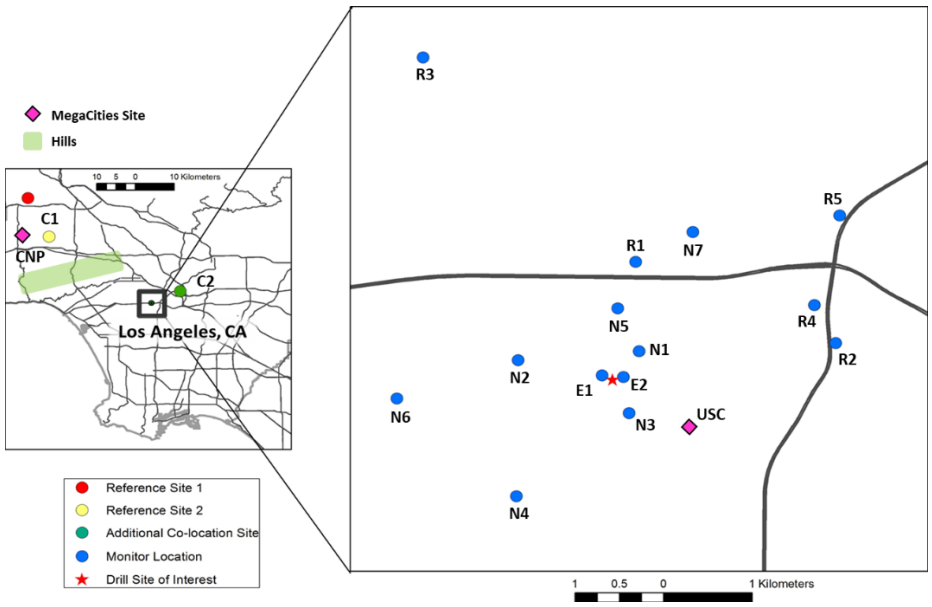


Figure 1: Map of sampling sites and major sources of interest. Note that the Y-Pod locations have been approximated to the center of their respective blocks in order to protect participant identities. The first initial is indicative of the site type: N – neighborhood site, R – near roadway site, E – near oil extraction site, c – ongoing co-location site.

2.2 Local Partnerships & the Communities

This project was conducted in a participatory manner, in partnership with two local community-based organizations. Redeemer Community Partnership, a community development corporation, has been organizing local residents around the drill site of interest and has been active in the community since 1992. Esperanza Community Housing has been active in the community since 1989 and played a large role in bringing awareness to violations at a formerly active drill site to the attention of regulators (Brown, 1995). Together we worked with community partners to plan the project, choose the sampling sites, and conduct the sampling. Partners also assisted with finding and hiring a local field technician to check the sensor systems periodically and collect the data. As part of the project a Memorandum of Understanding (MOU) was developed and signed by all partners to ensure a mutual understanding of the limitations of sensor technology and study objectives as well as to ensure ongoing communication, particularly around the dissemination of the results of the project.

This deployment spanned primarily two communities, West Adams and University Park. In both of these communities, problems such as poverty and housing insecurity contribute to an overburdening of residents. West Adams is made up of 87% residents of color, including 58% Latino and 20% African American. Furthermore 68% of residents live 200% below the poverty line (Shamasunder et al., 2018). University Park is predominantly Latino at 76% and here 72% of residents are living 200% below the poverty line (Shamasunder et al., 2018). Prior studies have

demonstrated the reality of environmental inequalities resulting in higher levels of exposure to air pollutants in minority and socioeconomically disadvantaged communities (Souza et al., 2009; Marshall, 2008; Wang et al., 2010). In some cases, these higher levels of exposure are observable via smell or sight, or can cause physical symptoms (though of course not all odors signify dangerous compounds or dangerous levels of these compounds). For that reason, we were also interested in the local knowledge and observations of community members. We wanted to consider whether or not observations correlated with pollutant enhancements and vice versa, and whether these observations could add valuable contextual information to trends in the sensor data. Following the deployment, we were provided with a log of observed activity at the drill site and notes of noise and odor complaints. This log provided another, more qualitative stream of data to fold into our analysis.

2.3 Low-Cost Sensing Systems

The low-cost sensing systems used for this study were designed and assembled in our lab at the University of Colorado Boulder. These systems include commercially available sensors of various types, including metal oxide semi-conductor, electrochemical, and non-dispersive infrared sensors for gas phase pollutants as well as sensors for environmental parameters. U-Pods and Y-Pods are iterations of the same design with some minor differences in the circuit board design and Arduino programming, with the design objective of increased reliability of performance for the Y-Pod. Table 1 lists all sensors that can be included in the system, while Figure 1 includes a diagram of the inside of a Y-Pod and several photographs of Pods deployed in the field and at a co-location site.

Table 1: Complete list of sensors used in the U-Pods and Y-Pods

Sensor Type	U-Pod	Y-Pod
Temp. & RH	RHT03 (aka DHT22)	Sensirion SHT2
Temp. & Pres.	47 Bosh BMP085	Bosh BMP180
Carbon dioxide	ELT S-100 NDIR	ELT S-300 NDIR
Ozone	SGX Corporation MiCS-2611	SGX Corporation MiCS-2611
VOC Sensor 1	Figaro TGS 2600 MOx	Figaro TGS 2600 MOx
VOC Sensor 2	Figaro TGS 2602 MOx	Figaro TGS 2602 MOx
Additional Optional Sensors	Alphasense B4 series (CO, NO, NO2, O3, SO2), Baseline Mocon PID, SEN-08942 RoHS weather meter (for wind speed and direction)	Alphasense B4 series (CO, NO, NO2, O3, SO2), Baseline Mocon PID

These systems are housed in a weather-proof case, approximately 20 cm x 25 cm x 10 cm. with electronics fans facilitating active flow across the sensor surfaces and resulting in multiple air exchanges per minute. The Pods are typically powered using 12 V AC/DC power adapters, although they can also be powered via a solar panel or car

battery. Data is logged to an onboard micro-SD card multiple times per minute. More information on these systems and their operation is available in other papers, as these systems operating these sensors have been utilized in a number of previous studies (Masson et al., 2015b; Sadighi et al., 2018; Collier-Oxandale et al., 2018a & 2018b; Casey et al., 2018a; Cheadle et al., 2017; Casey et al., 2018b; Castell et al., 2017).

For this deployment, these systems were placed at sites agreed upon by researchers and community partners. Placement of the systems at each site was partially limited by safety, convenience, and resident preference, although efforts were made to select ideal locations; locations elevated off the ground and with access to representative air flow. As observed in a previous study, intra-site variability is a possibility and is often driven by sources within close proximity to the site (Collier-Oxandale et al., 2018b). For this reason, the specific placement of each monitors was be taken into consideration during our analysis. We were also careful to ensure that the two Y-Pods adjacent to the drill site of interest had a line-of-sight to the potential source.

Figure 2: Top left – Y-Pod interior, top right – two-Y-Pods deployed on a roof, bottom left – Y-Pods deployed at a co-location site near downtown LA, bottom right – Y-Pod deployed on a building above a roadway.

In terms of data completeness, there were two periods of continuous data loss from two Y-Pods that do not appear to have effected data quality. However, one of the Y-Pods did experience repeated, sporadic power loss and this Y-Pod has been excluded from the analysis. Another Y-Pod displayed a problem with a VOC sensor that was only apparent during the calibration and resulted insignificantly poorer than the calibration results for all other Y-Pods. This Y-Pod was also excluded from the analysis, and as a result data from 14 Y-Pods was utilized in the analysis. While the analysis primarily utilizes data from the Y-Pods, which were new systems with new sensors, supplementary

U-Pods were sometimes co-located with Y-Pods to provide additional streams of data and this information will be included on an as-needed basis.

2.4 Sensor Signal Processing & Quantification

Metal oxide (MOx) semi-conductor sensors rely on a reducing gas to remove oxygen molecules adsorbed to the sensor surface, which then changes the conductivity by lowering the resistance (Wang et al., 2010). This variable resistance is then indicative of the concentration of the reducing gas or target gas in the atmosphere. While simple in principle, quantifying the concentration of the target gas is complicated by the fact that changes in ambient temperature and humidity affect these reactions occurring on the sensor surface (Wang et al., 2010; Sun et al., 2012), as well as the fact that other reducing gases may also affect the sensor surface (Spinelle et al., 2017c). These sensors were originally developed for scenarios with high concentrations of the target gas(es) – scenarios in which the effects of environmental variables and confounders would be small in comparison to the effects seen from the target gas. As such, intensive sensor quantification and robust calibration are a necessity when attempting to utilize these sensors for gases at ambient concentrations.

2.4.1 Field Calibration

While laboratory tests are useful for determining the capabilities of sensors, researchers have continually found that field co-locations allow for the generation of more accurate calibration models (Piedrahita et al., 2014; Castell et al., 2017). The reason for this is likely the large and dynamic range of environmental variables and presence of background pollutants, which cannot be adequately replicated in the lab. A field co-location particularly near the intended deployment site, provides data that trains the calibration model to detect the target pollutant amid environmental and background conditions similar to those that will be experienced during a field deployment. A field calibration or normalization involves the physical co-location of sensor systems with reliable reference instrumentation, typically prior to and following a field deployment. The concurrent data is then utilized along with a technique such as multiple linear regression or machine learning to generate and evaluate a calibration model, which can be used to convert field data to usable concentration values. For this study, support from local regulatory agencies allowed for both pre-deployment and post-deployment co-locations listed in Table 2. Note the final co-location site for TNMHCs was much further north, outside of Los Angeles, which while not ideal given the different background

conditions, was the best option for the post-deployment co-location. Furthermore, based on the results presented in Section 3.1, we believe this reference dataset still supports useful data quantification.

Table 2: Details of co-locations with reference instruments

Dates	Reference Instrument	Pollutants	Location
Pre: 8/1 – 8/4/2016	[a]Baseline Mocon Series 900 Analyzer	CH_4, TNMHCs	Los Angeles, CA
Post: 10/20 – 10/28/2016	[a]Picarro cavity ring-down spectrometer	CH_4	Los Angeles, CA
Post: 11/2 – 11/12/2016	[b]Synspec Alpha 115	TNMHCs	Shafter, CA

[a]operated by SCAQMD, [b]operated by SJVAD

2.4.2 Signal Processing and Calibration Models

As stated earlier, the raw sensor signals are recorded to a micro-SD card as voltages, which are then converted to a normalized resistance. This process is described in greater detail in previous works (Collier-Oxandale et al., 2018a & 2018c). This resistance is then used to generate the calibration model. In terms of other processing, the raw values were converted to minute-medians and the warm-up data was excluded. Warm-up data is defined as the first half hour of data after the Pod has been turned off for a half hour or longer. The processed sensor data was then used alongside the reference data to generate a calibration model. Table 3 lists the calibration models utilized in this study. Note, the calibration models rely on data from the Figaro TGS 2600 and Figaro TGS 2602, made by Figaro Inc., along with environmental data to convert sensor signals. The quantification, selectivities and cross-sensitivities of both sensors are described in greater detail in the previous papers works (Collier-Oxandale et al., 2018a & 2018c).

The most notable difference between the two models used for methane is the later incorporates data from multiple sensors while the former relies on only a single sensor. The former model was determined to be the best-fitting model in a previous work on methane quantification as it seemed to correct for complex temperature and humidity effects well works (Collier-Oxandale et al., 2018a); in this study we also utilized inverted models that necessarily result in a 1:1 relationship between pollutant concentration and sensor signal. In the subsequent study, we examined the usefulness of multi-sensor models, which required solving for pollutant concentration rather than a model for sensor signal works (Collier-Oxandale et al., 2018c). In this second study, the multi-sensor models helped correct for cross-sensitivities and provided better estimates of target species works (Collier-Oxandale et al., 2018c). Both were included in this analysis, allowing us to compare the two methane concentration estimates. This previous study also illustrated how similar models, utilizing the same sensors can result in calibration models specialized for

the target species or groups of species works (Collier-Oxandale et al., 2018c), which is why the same model form was used for CH_4 and TNMHCs in Table 3.

Table 3: Calibration Models

Model Name	Equation
Single Sensor CH4	$Fig1 = p_1 + p_2(C) + T(p_3 + p_6(C)) + p_4(H^{-1}) + p_5(T_i) + p_7(T * H^{-1}) + p_8(T_d)$
Multi Sensor CH4	$C = p_1 + p_2(Fig1) + p_3(Fig2) + p_4(Fig1 * Fig2) + p_5(T) + p_6(H) + p_7(T_i)$
Multi Sensor TNMHC	$C = p_1 + p_2(Fig1) + p_3(Fig2) + p_4(Fig1 * Fig2) + p_5(T) + p_6(H) + p_7(T_i)$

Predictors (lower case p with subscripts) are Fig1 – R_s/R_0 for the Figaro 2600 sensor, Fig2 – R_s/R_0 for the Figaro 2602 sensor, C – pollutant concentration, T – temperature, H – absolute humidity, T_i – continuous time. The predictor p_1 indicates an empirical constant.

2.4.3 Supporting Data & Additional Processing

Signals from other sensors were also utilized to better understand the composition of the emissions events eliciting a response, specifically from the electrochemical carbon monoxide sensors and non-dispersive infrared carbon dioxide sensors. These signals are quantified similarly to the VOC sensors, using field calibration. However, given that these sensors have been used in previous studies by our group and by others (Mead et al., 2013; Piedrahita et al., 2014; Casey et al., 2018b; Cross et al., 2017; Zimmerman et al., 2018; Jerrett et al., 2017; Collier-Oxandale et al., 2018b), and that the most important information from these signals is examining enhancements and relative amounts of CO_2 and CO, the work presented here focuses on the VOC sensor quantification. Details regarding the performance of CO and CO_2 sensors is available in Appendix VII.

One additional processing technique was incorporated into this study to highlight the short-term enhancements, likely associated with local emission events, baseline removal. We utilized a technique developed to extract the baseline from atmospheric data (Ruckstuhl et al., 2012). Heimann and colleagues (2015), have demonstrated the application of this technique to low-cost sensor data, specifically for electrochemical carbon monoxide sensors. Using this technique, these researchers illustrated the ability to separate regional trends (represented by the baseline) from local trends (apparent in the baseline-removed data). More details on the procedures and an example are available in Appendix VII. It's worth noting that while this is one technique for identifying and removing the baseline, many other techniques have been developed in the field of digital signal processing and particularly given the complex nature of sensors systems, a smarter processing technique might be worth considering (Smith, 1997).

3. RESULTS & DISCUSSION

3.1 Sensor Performance Quantification

3.1.1 Quantification Model Results

Table 4 summarizes the results of sensor quantification efforts for the two methane models and the non-methane hydrocarbon model. When compared to quantification results from previous studies, the results are very similar suggesting consistency in the behavior of these MOx sensors. For instance, the average R^2 and RMSE for the testing data for the three Y-Pods used in our previous study (Collier-Oxandale et al., 2018a) were 0.740 and 0.177 ppm respectively. When this quantification is expanded to include the other 12 Y-Pods available the new averages are nearly unchanged. Comparing the results for the two multi sensor models reveals similarities as well. For CH_4, the multi sensor model exhibits improvements over the single sensor model, which is the same outcome as was observed when the two models were applied to a different dataset works (Collier-Oxandale et al., 2018a & 2018c). For TNMHCs, the average R^2 is similar to those that resulted when sensors were calibrated to predict a summed VOC signal works (Collier-Oxandale et al., 2018c). The RMSE's for both training and testing datasets presented here are higher, however this is not surprising as the dynamic range of the TNMHCs (~300 ppb) is much higher than the range for the summed VOC signal (~30 ppb). The complete statistics for each individual Y-Pod are in the supplementary.

Table 4: Resulting statistics from calibration model generation and validation, averages across 15 Y-Pods

	Training			*Testing*		
	R^2	**RMSE**	**MB***	**R^2**	**RMSE**	**MB**
Single Sensor – CH_4 (ppm)	0.812	0.153	0.001	0.737	0.178	0.025
Multi Sensor – CH_4 (ppm)	0.880	0.111	0.000	0.802	0.157	0.074
Multi Sensor – NMHC (ppb)	0.598	31.15	0.011	0.458	46.35	7.62

*MB – mean bias

Figures 3, 4, and 5 depict the results for all Y-Pods and for each model for the collocations. From these plots, it is clear that there is little inter-sensor variability. These figures also help to illustrate the strengths and limitations of each model. For example, all the models seem to reproduce diurnal trends as well as some of the shorter-term peaks, but there is some underestimation of peaks, particularly for the multi sensor models.

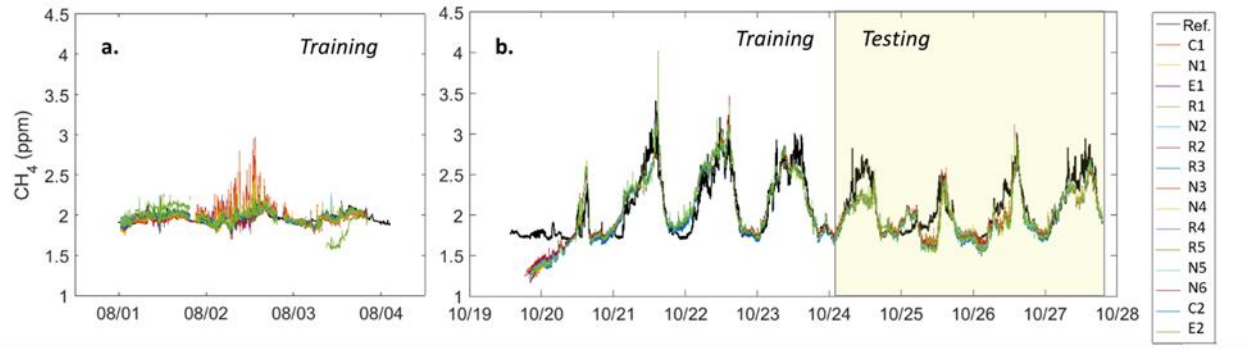


Figure 3: Single Sensor model for CH_4, the colors indicate individual Y-Pods, while the reference data is shown in black. The testing period is highlighted in yellow.

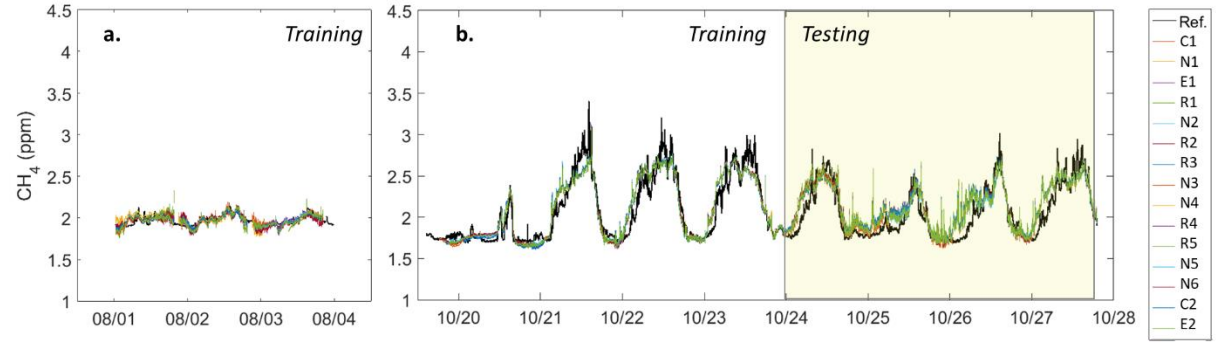


Figure 4: Multi Sensor model for CH_4, the colors indicate individual Y-Pods, while the reference data is shown in black. The testing period is highlighted in yellow.

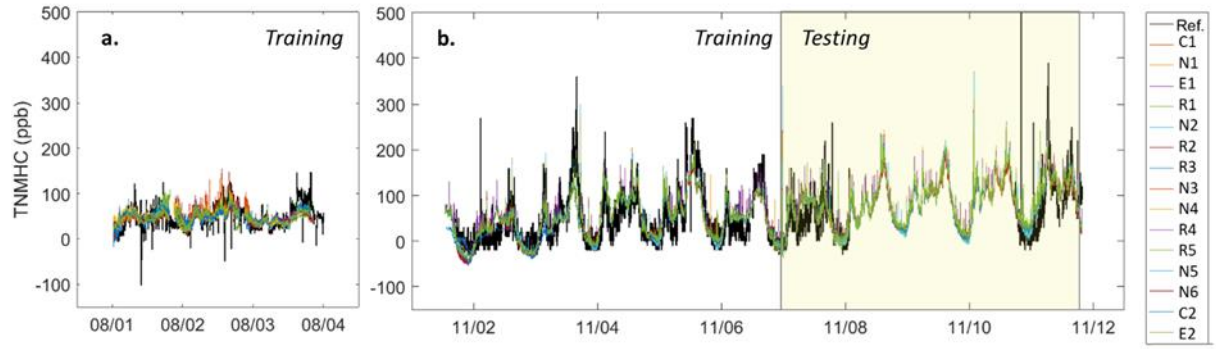


Figure 5: Multi sensor model for TNMHCs, the colors indicate individual Y-Pods, while the reference data is shown in black. The testing period is highlighted in yellow.

3.1.2 Comparing Methods

From the statistics presented above, it is clear that the multi sensor model provided improved methane level estimates; Figure 6 explores where these improvements occurred. Figure 6, panel a illustrates that the use of the multi sensor model reduced the underpredictions observed with the use of the single sensor model, providing a better estimate of the baseline. Additionally, the multi sensor model reduced some of the overpredictions, possibly caused

by confounding compounds. Figure 6, panel b shows that the paired fitted points fall primarily along the 1:1 line. This would suggest that the majority of paired points were fairly similar between the two model applications and most of the differences impact a small amount of the data.

Although the multi sensor model seems to produce more reliable methane level estimates, there may be unique information in the single sensor CH_4 model that makes it worthwhile to continue to consider this dataset throughout the analysis. One possibility that was discussed in our previous study works (Collier-Oxandale et al., 2018c), where we examined quantification for individual and grouped VOCs, was that models were being trained to quantify a more generalized VOC signal instead, or in addition to, their target compound. We found that even when using the same model form, the models for each compound or compound group were becoming specialized works (Collier-Oxandale et al., 2018c); however, correlations between VOCs made it challenging to conclude that confounding was not present in the quantification. For example, in this work, while the multi sensor model is an overall improvement, it seems that adding the second VOC sensor may be introducing some responses to VOCs other than methane; these appear to be visible as short-term spikes not present in the single sensor CH_4 data or the reference CH_4 data (Figure 6). As such, we thought it would be valuable to retain and examine the results from both methane models going forward. A more comprehensive evaluation of the single vs. the multi sensor model in relation to reference data would have been conducted had the post-deployment co-location offered both datasets concurrently as opposed to two separate co-locations for each species of interest. Though the datasets do overlap from the pre-deployment colocation, the methane and TNMHC reference data from this run both have a very small dynamic range.

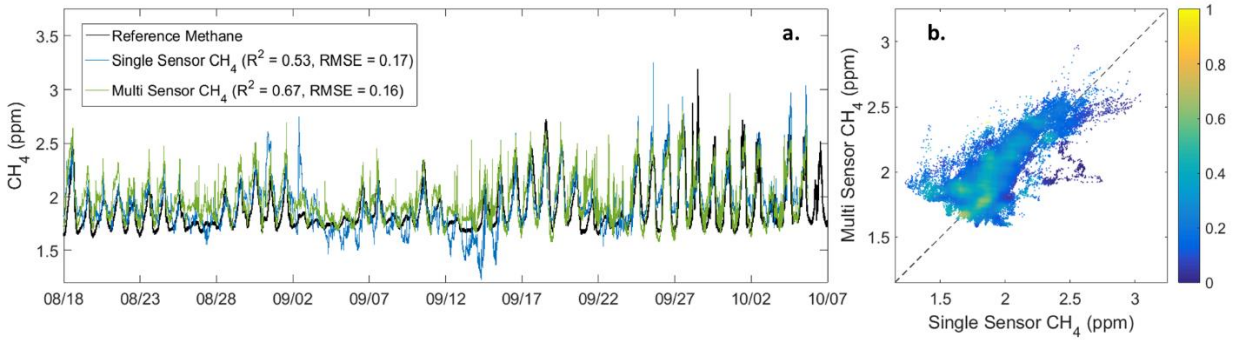


Figure 6: Panel a shows additional validation data, concurrent to the field deployment, with both the single (in blue) and multi sensor (in green) models plotted with the reference data (in black). Panel b shows the single sensor fitted data plotted against the multi sensor data, the color indicates a normalized density of the data (yellow – high density, blue – low density). The black dotted line is the 1:1 line.

3.2 Overview of Spatially Deployed Sensor Data

3.2.1 Variability Captured by Sensors

Comparing co-located verses deployed sensors, shown in Figure 8, reveals the variability introduced when sensors are moved to field sites, even when those sites are relatively close at approximately 140 m apart, which is the difference between the E1 and E2 sites, on either side of the drill site. Furthermore, the differences, and similarities, in the deployed data are significant with both correlated and uncorrelated elevations in methane and TNMHCs above the expected uncertainties (0.2 ppm and 50 ppb respectively). The high degree of correlation between co-located MOx sensors seems to be a reliable characteristic of these sensors and makes them well-suited to detect and study spatial variability (Collier-Oxandale et al., 2018b; Cheadle et al., 2017; Sadighi et al., 2018).

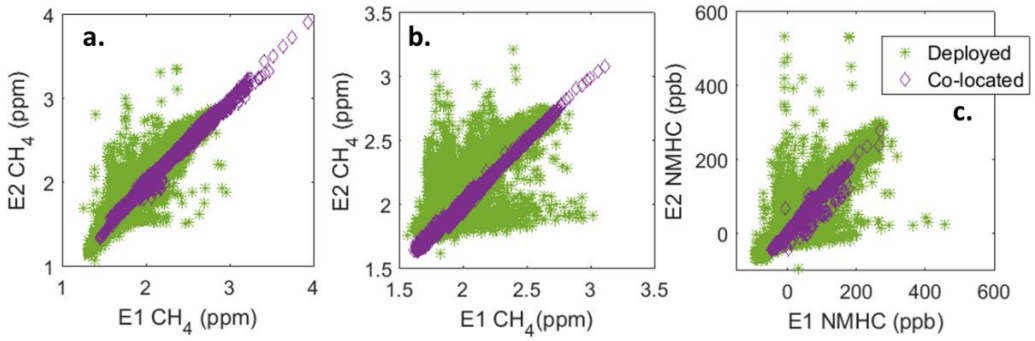


Figure 8: Paired data for the two sampling sites near the drill site, with co-located data plotted in purple and deployed data in green; panel a – single sensor for CH_4, panel b – multi sensor for CH_4, and panel c – multi sensor for NMHC

3.2.2 Spatial & Temporal Patterns

Across the Y-Pods there are similarities in the temporal trends. Averaging each hour for the whole deployment reveals that both the CH_4 and TNMHCs increase at night and decrease during the day (Figure 9). This fluctuation is driven by the planetary boundary layer that lowers at night, causing the accumulation of air pollutants that are emitted at the ground, and rises during the day, facilitating mixing, dispersion, and reduced concentrations. Figures 9a and 9b also include the reference CH_4 data collected throughout the deployment. While this reference data was collected north of our deployment sites, we would still expect to see similar diurnal trends across all sites. Additionally, one Y-Pod (C1) was co-located with the reference instrument throughout the deployment; the concentrations shown for C1 most closely matched the reference data. While there is no reference TNMHC signal from the deployment, it is evident that this data also follows a similar diurnal pattern. It's worth pointing out the

negative values present in Figure 9c, as well as the values below the known baseline for methane, particularly in Figure 9a. Though there is an uncertainty of approximately 0.17 ppm for the CH_4 models and 46 ppb for the TNMHCs, underestimations of this level are likely due to inadequacies in the models.

To provide further evidence that the sensors were replicating expected environmental trends, we compared data from sensors to data collected through the MegaCities Project (Wong et al., 2015), see Figure 10. This figure illustrates how features in the diurnal patterns can vary by region and that these differences are also reflected in the sensor data. For 10/5/16 – 10/6/16, we observed a distinct difference in the timing of when the CH_4 levels increase to their nightly peak across the two sites; these increases were almost six hours apart with the increase in the CH_4 at the CNP site and the C1 Site occurring first. The lag apparent in the USC data was also observed in our data from the R4 Site. The areas where these sites are located are separated by the hills that lie northwest of downtown LA making it reasonable that differences in topography and meteorology could drive different diurnal patterns on either side. It is also worth noting that differences between R4/USC and C1/CNP are not surprising given differences in siting conditions and locations as well as differences in the instruments.

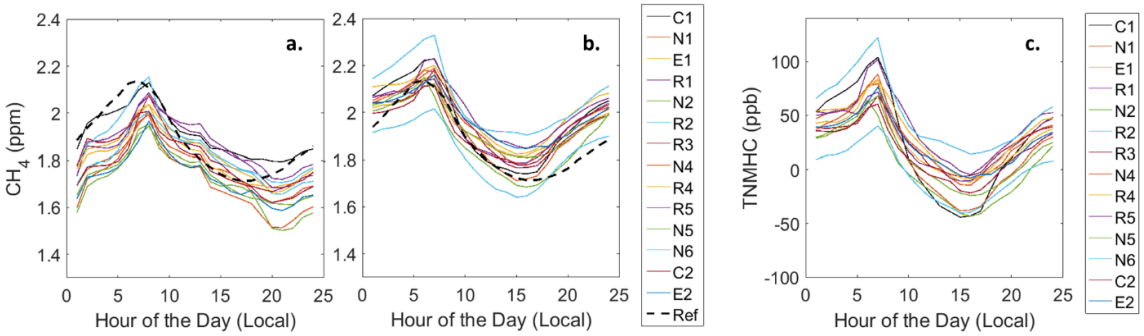


Figure 9: Hourly averages across deployment for all sensors (indicated in colors, with the exception of C1, which is indicated in grey), for the two methane plots, the averaged reference data, is indicated in black. Panel a – single sensor model for CH_4, panel b – multi sensor model for CH_4, panel c – multi sensor model for TNMHC

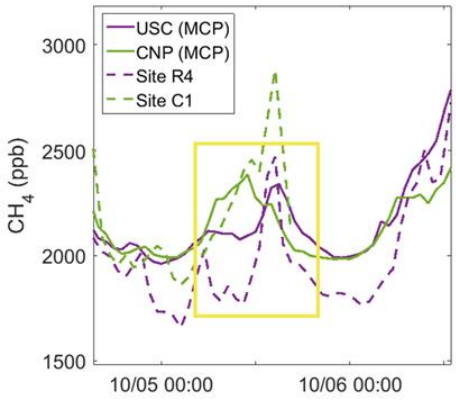


Figure 10: Roughly one day of hourly data from two MegaCities sites: CNP (near our C1 Y-Pod), and USC (near our R4 Y-Pod). The data was retrieved from https://megacities.jpl.nasa.gov/portal/data-access/. Note, the Y-Pod data was also hourly averaged, and the data plotted was from the multi sensor model.

Figure 11 provides an example of data from 3 sites, with the top panel depicting the converted sensor data and the bottom panel depicting the same data with the baseline removed. Notably, the short-term enhancements do not occur across all the sites at the same time, suggesting that there may be local sources impacting one sensor system and not the others. This figure also emphasizes how even though the short-term enhancements are larger than the expected error, it is easy for these features to get lost amidst the large diurnal swings. This figures also demonstrates why it can be difficult to identify differences between sites using passive sampling tools. While passive sampling methods can provide quantified, speciated hydrocarbon information, the levels of background pollutants that can accumulate during boundary layer fluctuations may overwhelm short-term events of interest. Sensors can prove to be valuable in this type of application as they allow for the detection of these events over very short time scales. The utility of combining these two sampling methods as has been demonstrated by other researchers (Thoma et al., 2016).

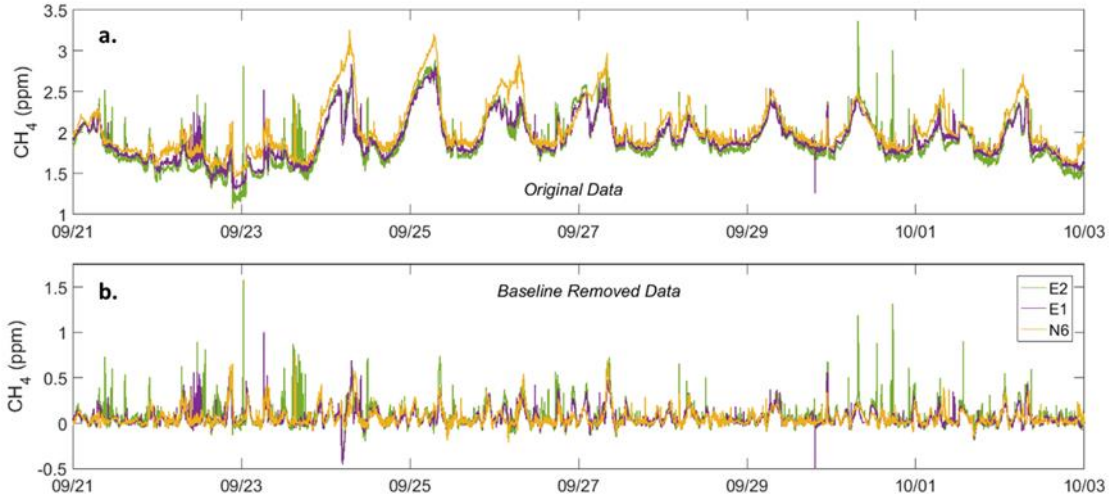


Figure 11: Panel a – Original methane level estimates from three sites, Panel b – the same data with the baseline removed. All data was converted using the multi sensor CH_4 model.

Looking more broadly across all the Y-Pods, there appears to be a trend of lower correlation among Pods that are further apart and higher correlation for Pods closer together. Table 6 illustrates this (1) by comparing all the pods in the study area to the Y-Pod C1 at the reference site (across regions), and (2) by comparing the Y-Pods in the study area to one of the neighborhood pods, N5, (within neighborhood), and (3) by comparing the two Y-Pods at the extraction site less than 150 meters apart (across block). To further emphasize this point, both the correlation coefficient for the complete dataset and the extracted baseline are compared. Generally, the R for the pairs of baseline data is higher and the correlations increase as distance between Y-Pod decrease. These are two reasonable results, as the complete datasets include the short-term events which are more likely to differ between sites than the diurnal/regional trends. We would also expect less variability in both datasets for Y-Pods close together as they will have more shared influences.

Table 6: Average correlation coefficient between Y-Pod baseline data across varying spatial scales

	Single Sensor CH_4		**Multi Sensor CH_4**		**Multi Sensor TNMHC**	
	Complete	Baseline	Complete	Baseline	Complete	Baseline
Across regions	0.6891	0.7164	0.7157	0.7602	0.5976	0.5895
Within neighborhood	0.8103	0.8498	0.7359	0.8084	0.7173	0.7580
Across block	0.9314	0.9867	0.8568	0.9527	0.8726	0.9741

3.3. Sensor Data in Relation to Potential Sources

In this section, the high-resolution spatial and temporal nature of the sensor data as well as additional signals are leveraged to better understand these short term-enhancements in methane and TNMHCs, including their probable composition and direction of origin.

3.3.1 Differences Across Site Categories

Figure 12 depicts the complete data from four sites, sorted by hour of the day. Site R3 is above a busy roadway, site N5 serves as a background site as it approximately 260 m from the nearest highway and 970 m from the drill site, essentially between these two potential sources. The final two sites, E2 and E1, are to the immediate east and west of the drill site. In terms of similarities, there are elevated levels of CH_4 and the TNMHC concentration estimates for all of the sites in the morning hours, from 6 – 8 am. This enhancement is likely a combination of morning rush hour traffic and the boundary layer having not fully lifted yet causing these emissions to accumulate (recall from Figure 9 that this occurs until around 10 am). As might be expected, there are also enhancements across the sites beginning around 6 pm and continuing through the evening hours, likely the result of evening traffic. However, in addition to these daily patterns, occurring most prominently at site E2 and occasionally at E1 are late morning and afternoon enhancements that are also above the expected error. These afternoon enhancements even pull the 95th percentile values higher for the E2 data, and lead to more variability across the sites for this time period. Additional plots with only the 95th percentile values are available in Appendix VII and clearly shown these trends.

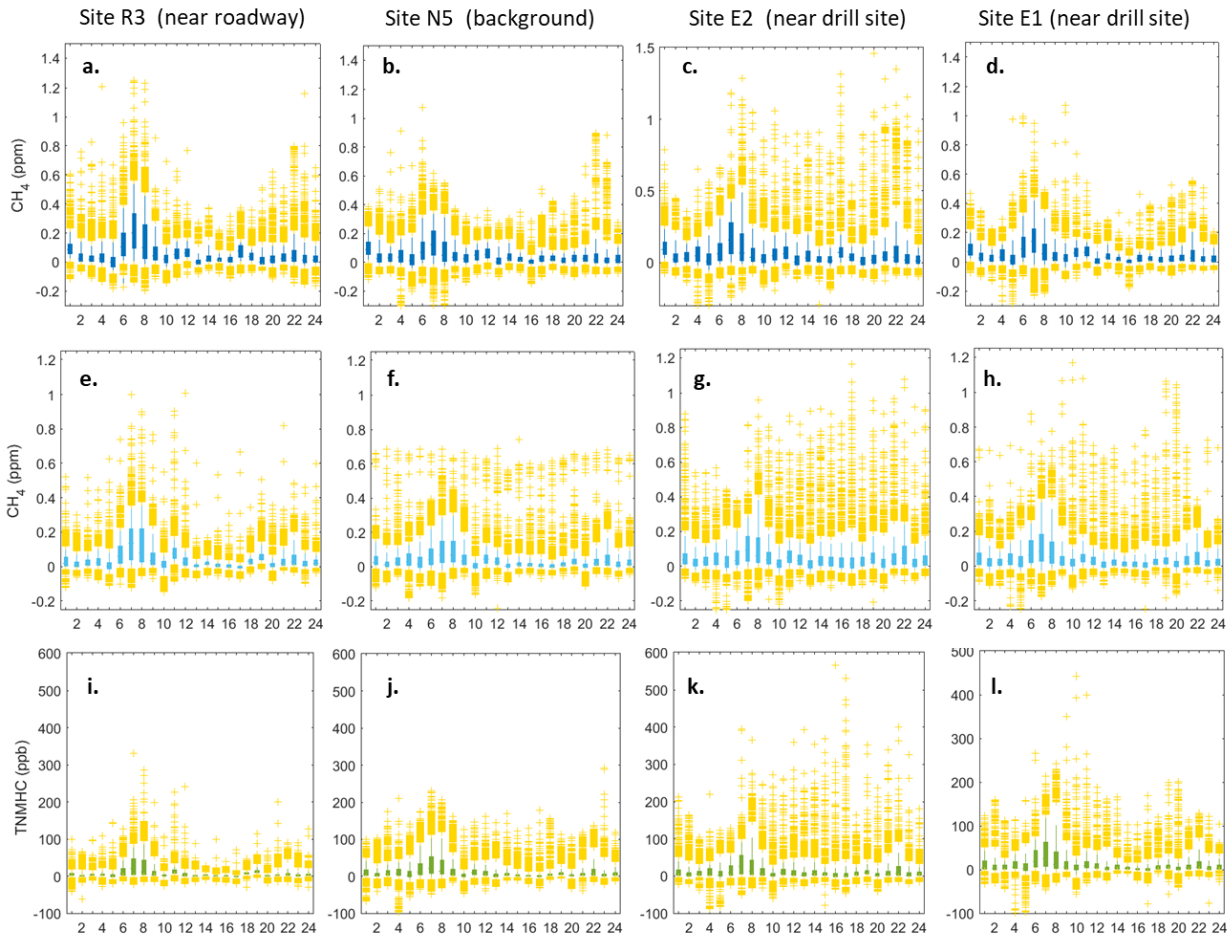


Figure 12: Baseline removed data from four sites grouped by hour of the day (in local time), panels a-d include the single sensor CH_4 data, panels e-h include the multi sensor CH_4 data, and panels i-l include the multi sensor NMHC data. The whiskers on the box plots represent the 5th and 95th percentiles respectively. The top and bottom 5th percent of the data are indicated by yellow markers across all plots.

The concentration enhancements do not occur across all of the sites included in Figure 12 and the timing of the enhancements can be spatially dependent. These changes in time and space were also observed by researchers during a study on methane spatial variability in relation to sources; that team concluded that measurements of ground level methane can provide information on sources up to 8.4 km away at night, but only 240 m away during the day (Bamberger et al., 2014). Those researchers emphasized the capacity for daytime mixing to disperse pollutants, supporting the idea that these afternoon enhancements must be the result of a local source. Those conclusions further support the idea that enhancements we observed in the early morning hours were likely from morning traffic sources nearby and further away.

Regarding the early morning enhancements observed specifically in the methane signal, while we would not expect large methane emissions from vehicles, it is possible that the sensor system was also being influenced by light alkanes or carbon monoxide. The Figaro 2600 has demonstrated a cross-sensitivity to carbon monoxide works (Collier-Oxandale et al., 2018a & 2018c). The same hour of day box plots, but for all of the deployment sites, are available in the supplementary for comparison. There were no locations with the magnitude and frequency of afternoon enhancements as those observed at Site E2.

In Figure 13, we focus on one of these afternoon-evening enhancements, as indicated by the yellow box. The baseline removed data from sites E1 and E2 are plotted together, and the inclusion of both methane models along with knowledge on the selectivity of the two sensors provides some insight into composition. This probable emissions event was estimated to be most drastic in the data from both multi sensor models, with changes in concentration beyond our expected error (approximately .2 ppm for CH_4 and 50 ppb for TNMHCs). The fact that the event was not predicted to the same extent by the single methane model, utilizing only the Figaro 2600, may indicate that it is primarily made up hydrocarbons that the second VOC sensor is more reactive to, as this sensor has different selectivities than the Figaro 2600. As determined in a previous study, the Figaro 2602 was found to be more well-suited for detecting larger aromatic hydrocarbons, and the Figaro 2600 was found to be more well-suited for detecting light alkanes like methane works (Collier-Oxandale et al., 2018c). As a counter example, following the event in the yellow box, there were repeated enhancements in all three panels, from roughly late afternoon on 8/17 through late afternoon on the 8/20, occurring at only Site E2. Since we observe these enhancements methane in panel a as well suggests that these enhancements may have been the result of a more diverse mix of hydrocarbons that included methane and/or lighter alkanes.

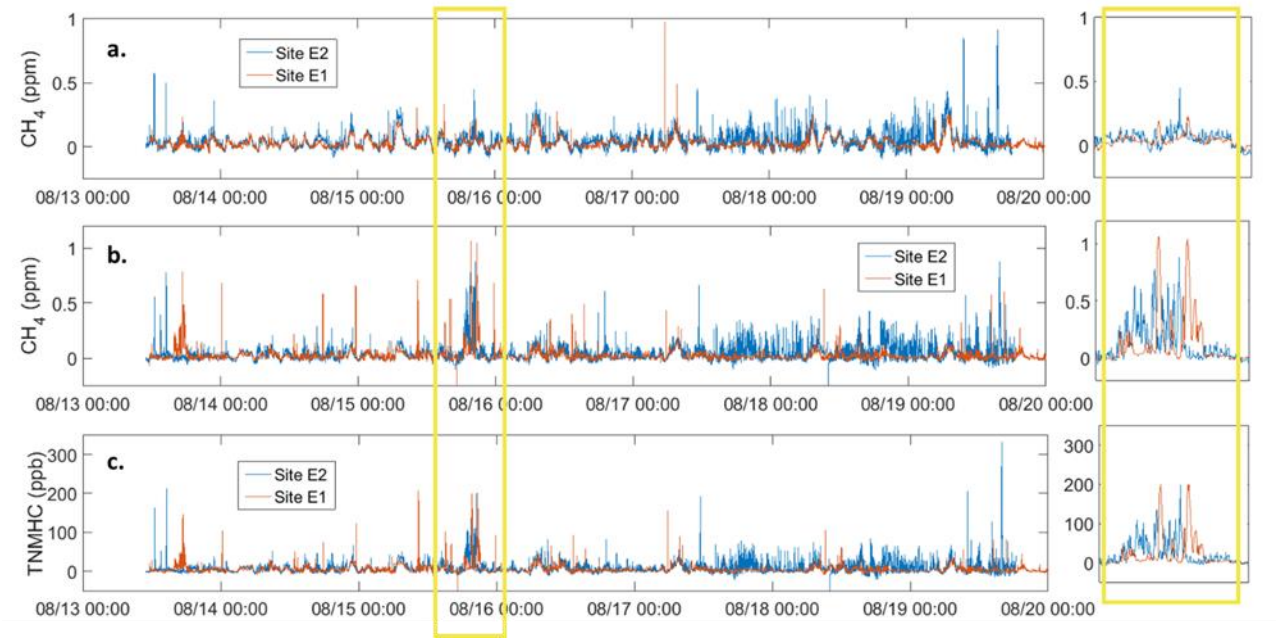


Figure 13: Approximately one week of baseline removed data, with panel a including the single sensor CH_4 data, panel b including the multi sensor methane data, and panel c including the multi sensor NMCH data. The time stamp is local, and the yellow box highlights the event discussed in the text. To the right of each panel is a zoomed in version of the event highlighted in yellow.

In addition to the likely presence of larger hydrocarbons, we see that the timing of the events was not perfectly correlated between the sites. The enhancements alternated between the two sites, possibly suggesting a shared source between the two sites and shifting wind directions. By incorporating additional streams of data, we can more fully understand this emission event.

3.3.2 Utilizing Additional Data Streams

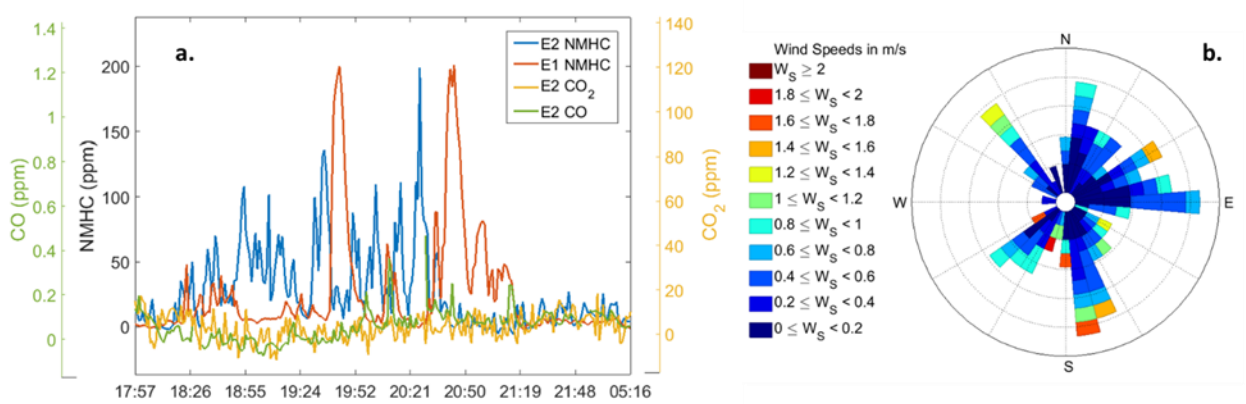


Figure 14: Panel a includes baseline removed TNMHC, CO_2, and CO data from Site E2 as well as baseline removed TNMHC data from Site E1. All of the data is from 8/15/16 and the times listed are local times. Panel b includes a wind rose for only the period of enhancements in TNMHCs.

Figure 14, panel a, shows the same event highlighted in Figure 13, with carbon dioxide and carbon monoxide data added. Figure 14, panel b, depicts the wind data recorded at Site E2 during only the emission event. The added data helps to further clarify the emission composition, something especially important also due to the possibility of confounders. Both carbon monoxide and carbon dioxide are combustion by-products, with carbon monoxide as indicative of incomplete combustion and carbon dioxide as indicative of complete combustion. In Figure 14, there does not appear to be a significant response from either CO or CO_2 that is also correlated with this emission event – suggesting the TNMHC response is likely the result of volatilized or vented hydrocarbons. As volatilized or evaporative hydrocarbons are common drill site emissions (Moore et al., 2014; Warneke et al., 2014), these results further support the idea that this emission event may be the result of activities at the drill site. The wind rose confirms the presence of active and shifting wind directions during this period, suggesting emission transport was possible. However. it's certainly possible that there were other sources of volatilized hydrocarbons in the area.

Examining the data from multiple sensors throughout the field deployment provides other periods where the TNMHC data was uncorrelated with combustion by-products giving insight into emissions composition. Figure 15, panel a, shows both a significant TNMHC enhancement independent of any increases in the CO_2 signal, and following that, an enhancement that is correlated with the CO_2 signal. Panel b in Figure 15 includes the wind direction data from the period of enhancement indicated in the yellow box, and confirms that there are examples of enhanced TNMHCs associated with winds coming from the west, indicated in the red circle. Figure 15, panel c provides another example,

in this case there are enhancements in TNMHCs independent of enhancements in CO_2 as well as enhancements in CO_2 uncorrelated with the TNMHC signal. In panel d, there are again enhancements in TNMHCs this time associated with wind from the east. Given that Site E2 is the to east of the drill site and Site E1 is to the west of the drill site, these examples provide further evidence that these enhancements may be originating from a source between the two sites.

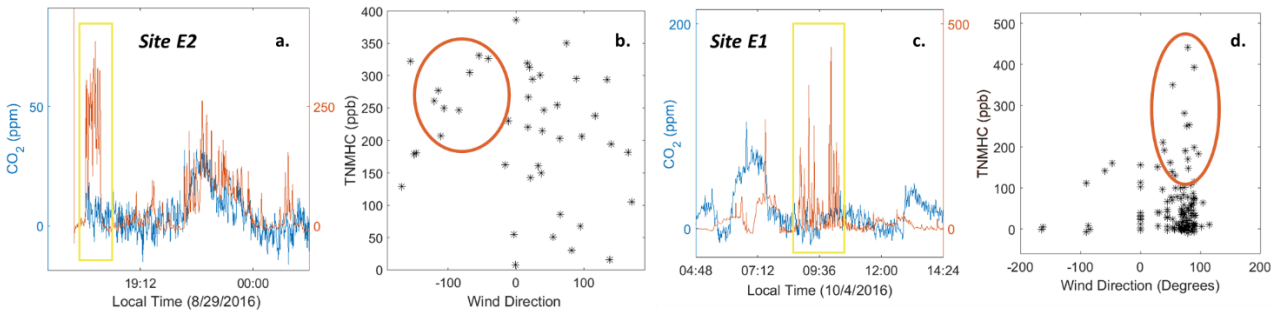


Figure 15: Panel a includes baseline removed TNMHC and CO_2 data from Site E2, with the wind direction data during the period of elevation highlighted in yellow shown in panel b. Panel c includes baseline removed TNMHC and CO_2 data from Site E1, with wind direction data from the period of elevations highlighted in yellow plotted in panel d.

In addition to examining individual events, a more quantitative idea of the differences between sites can be seen by counting the peaks, or enhancements at each one. Table 7 includes the number of peaks at each site, for the same timeframe, for a smaller and a larger peak height, and for the complete data as well as the afternoon only hours. Note, the baseline removed data was used to count peaks above the two thresholds and both peak thresholds are above our expected uncertainty. Site E2 consistently has some of the largest numbers of peaks. Conversely, Site E1 does not exhibit the highest quantities of peaks for whole dataset, but when we select for the afternoon hours, this site rises in the ranks. Similar to Figure 12, this table suggests that sites E1 and E2 are seeing relatively more enhancements in both CH_4 and TNMHCs in the afternoon hours, when the atmosphere is more well-mixed, and sources are therefore more likely to be local.

Table 7: Number of Peaks Above a Threshold

	CH_4 (Single Sensor, ppm)				CH_4 (Multi Sensor, ppm)				TNMHC (Multi Sensor, ppb)			
	All data		Afternoon Only		All Data		Afternoon Only		All Data		Afternoon Only	
	> .25	> .5	> .25	> .5	> .25	> .5	> .25	> .5	> 50	> 100	> 50	> 100
N1	442	76	11	0	338	53	15	3	731	192	57	6
E1	320	31	14	0	421	105	63	22	744	206	78	20
R1	428	27	2	0	304	6	7	0	666	177	45	7
N2	505	81	6	0	625	182	11	3	884	237	43	3
R2	509	39	5	0	280	8	2	0	644	167	33	0
R3	735	124	6	0	418	24	4	2	523	79	3	0
N4	528	84	10	0	457	91	32	12	837	257	47	11
R4	417	24	10	0	279	5	0	0	655	110	19	0
R5	484	57	13	1	524	51	59	7	1326	305	149	18
N5	361	22	16	1	357	86	57	26	872	233	109	11
N6	337	40	4	2	383	46	34	7	690	165	82	9
E2	712	120	82	31	546	73	121	25	918	221	155	46

3.3.3 Further Examining the Composition of Emission Events

When examining the trends related to composition across the entire deployment, the added CO and$_2$ data further explains the previously observed temporal patterns. For CO_2, enhancements across four selected sites occur primarily in the same early morning hours (6 – 8 am) with some occurring in the evening hours, and few enhancements in the afternoon (Figure 16, panels a-d). The elevated CO_2 in the early morning hours and evening hours, supports the idea that these enhancements are the result of traffic emissions, as CO_2 emissions would be expected from vehicles. The CO data from Site E2, reveals enhancements only during the early morning hours (Figure 16, panel e). The CO_2 and CO data further affirm that afternoon enhancements in hydrocarbons at Site E2 are primarily the result of volatilized or vented hydrocarbons.

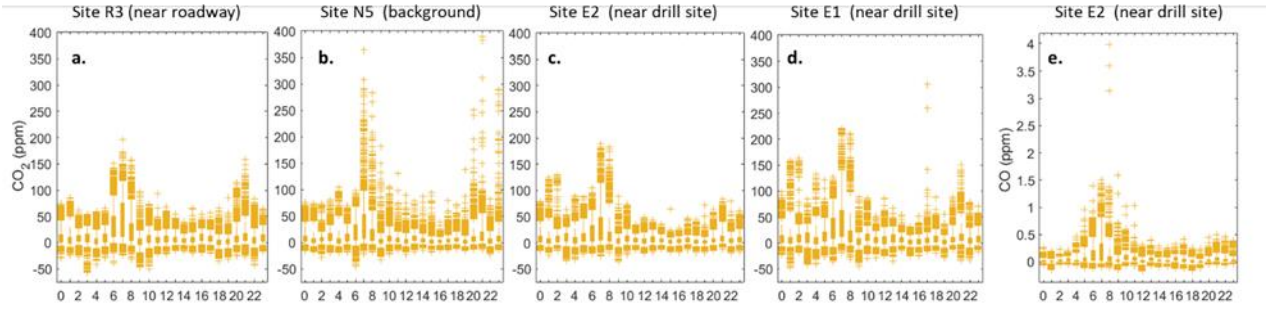


Figure 16: Baseline removed data from the same four sites shown in Figure 12, grouped by hour of the day (again, in local time), panels a-d include CO_2 data, and panel e includes CO data from Site E2. The whiskers on the box plots represent the 5th and 95th percentiles respectively.

The scatterplots of one atmospheric constituent versus another are shown in Figure 17 and they provide further insight into the potential sources. Some of the scatterplots exhibit distinct and separate trends likely the result of a combination of sources impacting the specific site. At first pass, there are at least two prominent sources, one emitting combustion emissions and one emitting volatilized or vented hydrocarbons. The enhancements resulting from volatilized hydrocarbons are reflected in the data from all quantification methods, suggesting a mix of hydrocarbons as opposed to a single species or group.

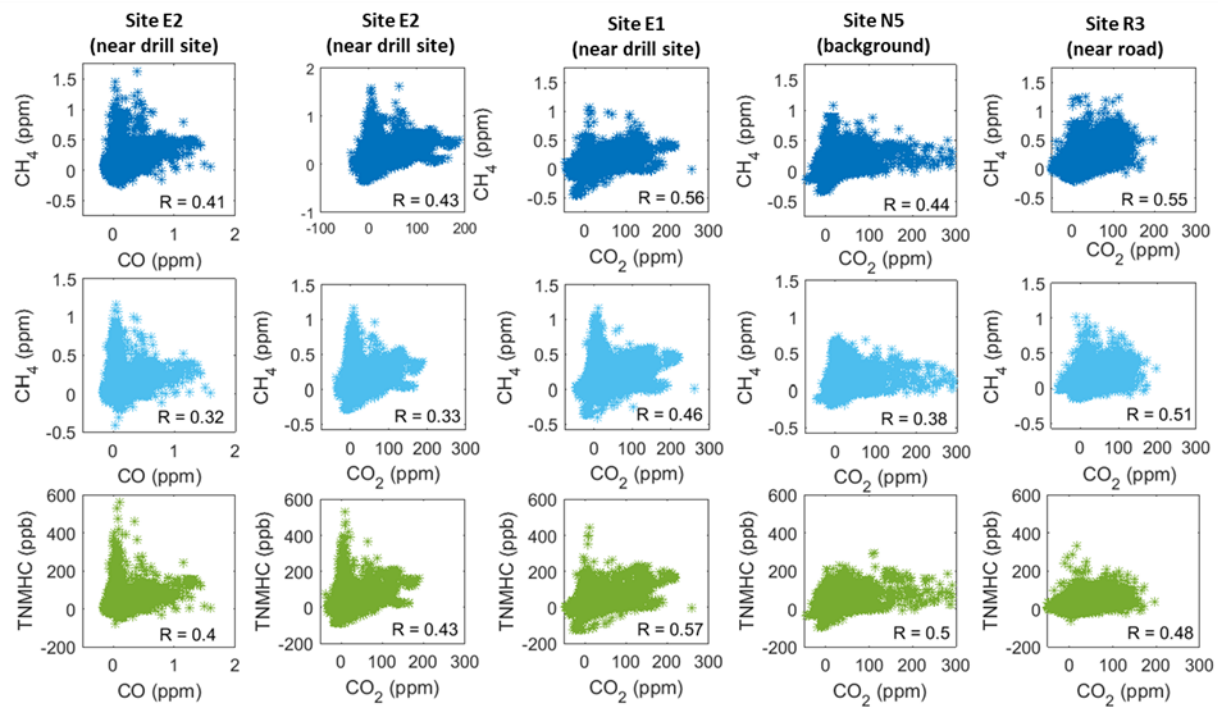


Figure 17: Scatterplots of baseline removed data, with the correlation coefficients for each set of data noted. The top row of CH_4 data is from the single sensor model, and the second row of CH_4 data is from the multi sensor model.

To explore the origins of the CO, we compared the complete CO and CO_2 data from Site E2, see Figure 18. In this plot there is a much higher correlation between the two atmospheric constituents, suggesting a single source for the combustion by-products observed at Site E2. A reference ratio has also been added to this plot, that represents the expected CO/CO_2 ratio based on the CARB emission inventory for the South California Air Basin; this ratio was calculated by researchers quantifying sources of methane over Los Angeles and this CO/CO_2 ratio was found to be consistent with aircraft measurements (Peischl et al., 2013). As this site is not on a major roadway it's reasonable that it would primarily experience well-mixed CO and CO_2 levels consistent with levels seen throughout the basin.

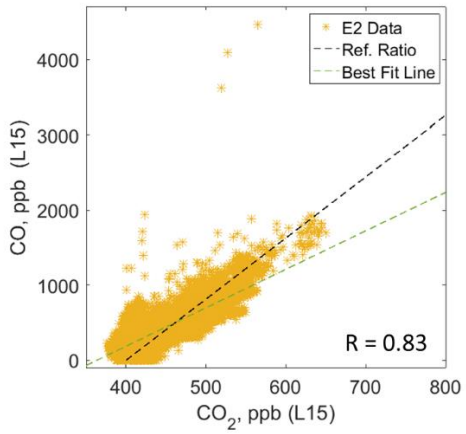


Figure 18: Complete CO_2 and CO data from Site E2, with a ref. ratio based on the CARB inventory included, illustrating the similarity between the observed data and the expected CO/CO_2 ratio for the LA area.

3.4 Joining Sensor Data and Local Knowledge

3.1.3 Additional Gas-Phase Sensor Data

As previously stated, the additional sensor data was quantified and incorporated on an as needed basis. Table 5 includes the summary statistics from the quantification of the CO_2 and CO data, with the number of Pods or sensors being averaged in the final column. These uncertainties presented here are similar to those of other studies using the same sensors. Studies quantifying the CO_2 sensor have found RMSEs of 10.1 ppm (Collier-Oxandale et al., 2018b) and standard errors of 9.4 – 16.8 ppm (Piedrahita et al., 2014). While studies utilizing the CO sensor have found RMSE's of 0.10 ppm (Casey et al., 2018a) and standard errors 0.28 – 0.44 ppm (Piedrahita et al., 2014). Plots and more details are available in the supplementary.

Table 5: Results of the generation and validation of models for the prediction of CO_2 and CO

	Training			Testing			
Pollutant	R^2	RMSE (ppm)	MB (ppm)	R^2	RMSE (ppm)	MB (ppm)	n (Y-Pods)
CO_2	0.930	9.41	0.126	0.807	14.79	-8.49	4
CO	0.841	0.11	0.002	0.603	0.09	-0.036	1

3.4.1 A Preliminary Attempt

As described earlier, in addition to sensor data, our community partners provided us with observational information from residents around the drill site. Our team was interested in whether or not this qualitative information could provide more context to the trends seen in the sensor data. And, yes, there are examples in the data of

observations occurring concurrent to periods when elevated levels of CH_4 and/or TNMHCs are also seen. For example, on 8/30 an odor complaint was noted around 3:00 pm, an afternoon during which enhancements in TNMHCs were also observed (Figure 19). In another example, odor complaints, noise complaints, and periods of heavy activity line up with enhancements in TNMHCs occurring at both Site E2 and E1.

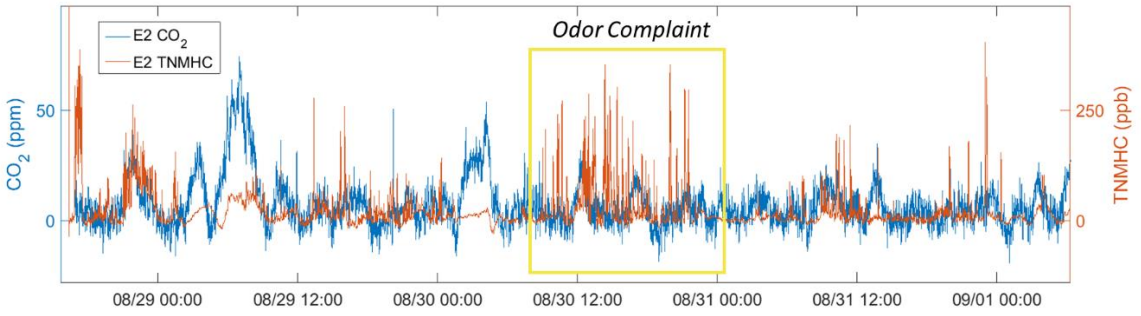


Figure 19: Baseline removed CO_2 and TNMHC data from Site E2, with a period highlighted in which an odor complaint was noted by nearby residents.

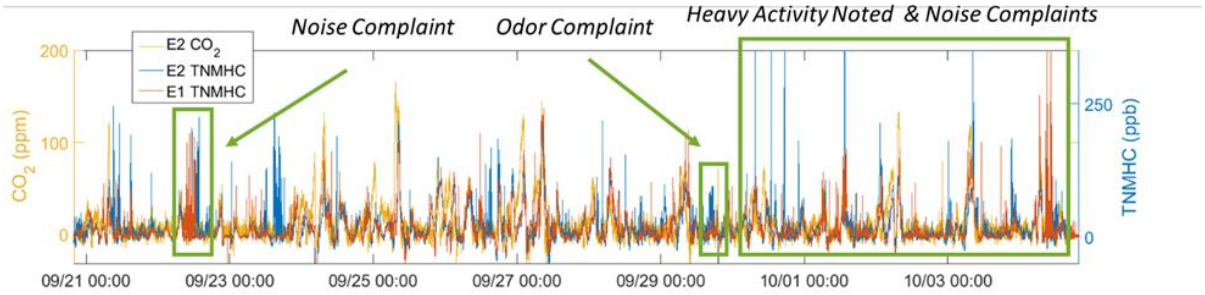


Figure 20: Baseline removed CO_2 and TNMHC data from Site E2 and Site E1, annotated with noise and odor complaints as well as observations of residents of heavy activity at the drill site.

In an effort to examine this data in a more systematic way, for each day the presence of a noise complaint, odor complaint, or observation of activity at the drill site were added together producing a number of observations per day. This data was then plotted against the 50th and 95th percentile data for each day for the TNMHC data (Figure 21). The results include a similar and positive relationship for the 95th percentile data points from both sites. Also, there were no low 95th percentile values on days with two or more observations. These examples suggest that it is possible that low-cost sensors are capable of recording emission events associated with the types of events reported by community members. However, this is a very limited and preliminary attempt at this type of analysis, and a more ideal dataset would include more observations from the entire study area and ideally, non-observations or confirmations of periods with no odors, or observable activity.

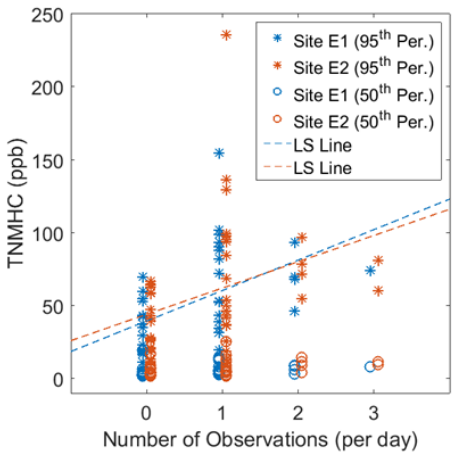


Figure 21: The daily 50th and 95th percentile values plotted against the number of observations per day (i.e., summed activity reports, odor complaints, and/or noise complaints).

3.4.2 A Model for Future Community-based Air Quality Projects

The analysis presented here provides an example of how sensors could support community-based air quality investigations in the future. As quantification for existing and new sensors continues to improve, this will increase the variety of sensors available for these studies. The inclusion of all these different sensor types will aid in mitigating cross-sensitivities as well as determining probable compositions of observed emissions. In addition to this sensor data collected across a network, other streams of data such as meteorological information will further help to identify sources and possibly even quantify the impact in a preliminary sense. This first step to determining what actions can reduce exposure.

The incorporation of community observations will provide further information regarding the context and effect of observed emissions, which may be especially valuable in areas with complex emissions mixtures like Los Angeles. It's also possible that the lived experiences of residents could help researchers to recognize patterns or trends in the data that otherwise may have been missed. However additional considerations will be needed to determine the best ways to collect this qualitative data and to merge it with the more quantitative, air quality data streams. One way of expanding the collection of this observational data could be through a mobile app or through physical log sheets utilized in a targeted campaign. It's possible that leveraging these two different types of data could facilitate more participatory and inclusive studies, and also could enhance data interpretation leading to more locally relevant and actionable results for a community.

4. CONCLUSION

In line with previous studies utilizing the same VOC sensors, we found that the results of sensor performance quantification were similar to expected results and that the converted sensor data reflected expected environmental trends – specifically regional and diurnal trends works (Collier-Oxandale et al., 2018a & 2018c). The unique information provided by this network of sensors includes the short-term enhancements in the CH_4 and TNMHC concentrations that are likely the result of local emission events given the time of day at which many of them occur. Furthermore, examining the CO_2 and CO data from the same time periods seems to suggest that at least some of these enhancements were the result of volatilized or vented hydrocarbons, as opposed to a combustion source such as motor vehicles. The inclusion of wind data also seems to suggest the source of these emissions may lie in the direction of the drill site. Finally, some of these enhancements also line up with observations made by nearby residents concerning odors or activity at the site of interest. While this data is not conclusive in terms of identifying all of the sources impacting concentrations at the site of interest, we can say that traffic and combustion emissions are not the only source in that area whose emissions are reaching the homes. Furthermore, given that the other source is comprised of volatilizing or venting hydrocarbons, the drill site provides a likely explanation for the other source. Similar to previous research, this study also illustrates that while individual low-cost sensors may lack the detection limits and accuracy to provide reliable information in ambient situations, when we leverage the information of multiple sensors we can better interpret the data from complex environments, leading to more useful information.

Acknowledgments

Funding provided through MetaSense NSF Grant CNS-1446912, NSF CyberSEES (1442971), NSF AWG SRN (CBET-1240584). We would also like to thank all project partners at the University of Southern California Keck School of Medicine, Redeemer Community Partnership (Nicole Wong), Sandy Navarro, William Flores, Esperanza Community Housing, and Occidental College. Additional acknowledgements to all community residents who assisted by hosting a Y-Pod sensor system. We would also like to thank regulatory partners at South Coast Air Quality Management District and San Joaquin Valley Air District/California Air Resources Board for this assistance with co-locations. Thanks to Christine Wiedinmyer (CIRES) for the deployment map included in this paper (Section 2.1). And a final thanks to current and former members of the Hannigan Research Lab, especially Nicholas Masson and Drew Meyers for their work on the U-Pod/Y-Pod hardware and software, and Evan Coffey and Kira Sadighi for their

assistance with the deployment in LA. (note, regarding reference air quality data: this data has not passed through the normal review process, and is therefore not QA'd and is unofficial data).

Chapter 9: Conclusions

SCIENTIFIC CONTRIBUTIONS

The contributions of my work speak to issues and needs identified by the low-cost sensor community in Chapter 2. Furthermore, these contributions begin to address different facets of the same problem, supporting the advancement of this technology as a whole. They include:

- Advances in the quantification of the performance of low-cost VOC sensors
- Recommendations and guidelines for the deployment and use of sensors systems, particularly in networks on small-spatial scales and in partnership with communities
- Strategies for leveraging multi-sensor devices to better understand not only sensor behavior, but also air quality trends and potential sources of pollution

The contributions of this thesis expand the type of gases available to be measured by the sensor community as well as pave the way for more use of this technology by communities. One contribution includes techniques that can be used to quantify metal oxide VOC sensor signals to estimate methane concentrations in complex environments, such as an area with high density natural gas and oil production activity. This work also demonstrated the different choices that are available to the sensor user in terms of calibration model selection, or the processing that can be applied to sensor or reference data and discussed what choices might be appropriate for a given situation. Another contribution includes techniques to quantify multiple metal-oxide sensor signals for individual and grouped hydrocarbons, including TNMHCs. This work provided a preliminary assessment of sensor selectivity and the reliability of performance amid changing VOC compositions. In both instances, the techniques described include specific multiple linear regression calibration models, examples of approaches to measurement validation, and results that sensor users can compare to. These techniques are unique for methane and non-methane hydrocarbons quantification as they are intended to be utilized in field calibration and field deployments of the sensors for ambient monitoring, as opposed to laboratory studies. These techniques have also been assessed in both rural and urban environments, and this feature, along with the difference in approaches, makes the work complementary to that of my colleague, which explored the utility of advanced machine learning techniques to achieve similar aims (Casey et al., 2018c).

Applying these quantification techniques to field data led to several case studies of sensor applications that demonstrate the potential information available, providing models for future studies. The quantification of sensor

signals for methane levels and their deployment in rural Colorado illustrated how sensor networks are capable of providing data leading to observations that correspond to observations resulting from studies using higher quality instrumentation and measurement methods. Specifically, the low-cost sensors captured the same large night-time enhancements in methane at the Platteville site, amid high density oil and gas activity, as another study (Collier-Oxandale et al., 2018; Halliday et al., 2016). Also, the data from the sensor network reflected a gradient of decreasing nighttime enhancements in methane, similar to the gradient of decreasing levels of alkanes observed by another team in the same area (Collier-Oxandale et al., 2018; Thompson et al., 2014). This case study, described in Chapter 6, further supports the assertion that sensors can provide useful preliminary information or data that supplements more conventional methods by, for example, illustrating the persistence of spatial and temporal trends across a region. The use of sensors in Los Angeles demonstrated their utility in a complex environment. There may quite a lot of uncertainty around the information from a single sensor, especially given the limitation of cross-sensitivities. However, when the signals from multiple sensor types are leveraged together and multiple system are examined across a network, some of the uncertainty regarding cross-sensitivities is minimized and interesting patterns can emerge. This is the type of data can provide insight into the multiple sources impacting a neighborhood. In our study, leveraging multiple sensors helped to clarify whether the source of enhancements in VOCs was more likely combustion or volatilization, which along with additional information, helps to indicate the origins of emission events. This case study, described in Chapter 7, provides a model for how low-cost sensors could be used to provide preliminary data in a community impacted by a complex mixture of emissions and possibly either inform more comprehensive monitoring plans. This study also reinforced the usefulness of utilizing existing analysis techniques, such as baseline removal, with sensor data. This analytical method was shown to be particularly useful for isolating short-term events that occur during larger diurnal swings and high levels of background pollutants. The work in Chapter 7 also provides a preliminary example of merging qualitative observational information and quantitative air quality sensor data, for which there currently are no best practices.

The analysis of the building-scale variability of pollutants and its impact on sensor data, presented in Chapter 5, provides an example that could inform sensor system placement for other users. Granted this example has many limitations as it was conducted at a single building in an urban location, it may still contain useful lessons. Additionally, the conclusions of this study agree with those of another study considering sensor siting and placement in a very different environment (Miskell et al., 2017); although the spatial scales vary between that study and the one

described here, the agreement suggests some broader conclusions about sensor siting. The examples of quantification techniques, provided in Chapters 6 and 7, may also support the development of best practices for field calibrations, especially if they are considered in conjunction with the other studies utilizing field calibrations (Sadighi et al., 2018; Zimmerman et al., 2018; Hagen et al., 2018).

RELATED RESOURCES AND LESSONS LEARNED

Though these are not scientific contributions, I would like to provide a list of resources developed or discovered throughout this work that I think are particularly well-suited to support community-based air quality research. It is not uncommon for researchers initiating these projects to have backgrounds primarily in air quality or engineering and little experience engaging with communities. Considering resources such as these might lead to more successful partnerships.

Related to this Thesis:

- *Air Quality Inquiry (AQ-IQ) Curriculum* – This project-based learning curriculum provides background information on air quality, sensors use, and help planning and conducting air quality research. The curriculum was developed for a high school audience but could be adapted for younger audiences or the public. Additionally, this curriculum has undergone multiple reviews and is publicly available on the TeachEngineering Digital Library.

  https://www.teachengineering.org/curricularunits/view/cub_airquality_unit
- *Memorandum of Understanding, Template* – This template was modeled after an MOU developed with community partners in Los Angeles. This document aids in ensuring that important issues are discussed at the outset of the project, including the capabilities and limitations of sensor systems, and the community's expectations regarding final products (e.g., reports) and data dissemination. (available in Appendix I)
- *Air Quality Sensors Project 1-Pager, Template* – This template provides a helpful tool for making a 1-page document to share with regulatory partners and the public. This type of document provides regulatory partners with the needed information to assist with facilitating co-locations at reference stations. This document can also answer questions from the community about the study and provide relevant information to community members that may consider hosting a monitor at their home or on their property. It is helpful to have a document such as this available early on in sensor projects. (available in Appendix II)

- *Lessons Learned & Other Observations*:
    - Taking Neighborhood Health to Heart has a published overview of their CBPR methods and there is one aspect that I believe would be particularly beneficial to low-cost sensor projects – the Data Review and Dissemination Committee (Main et al., 2012). The function of this committee is to review the data, ensure the protection of community member privacy throughout data dissemination, and maintain a copy of the final data (Main et al., 2012). Having this committee in place at the start of the project also leads to more accountability and dialogue between researchers and their community partners as the expectations are clear that the final reports and dissemination of data will be iterative and involve the participation of all partners.
    - Another important lesson from my partnership with TNH2H, described in Chapter 1, was the importance of discussing possible results and the appropriate next steps before beginning the project. At the outset of the partnership, we discussed the potentiality of finding high levels of radon or perchloroethylene in homes. For radon, we wanted to ensure that options were in place for low-income residents to receive financial support to mitigate their homes. For perchloroethylene, I outlined a plan for additional sampling, including sampling methods and labs that could conduct the analysis. We also contacted local public health officials to share our plan and receive feedback. This action had the duel benefit of these officials sharing their expertise and other resources as well as making them aware of our project. A worksheet with guiding questions has also been included in the appendix (Appendix III) to lead these discussions with partners. These discussions ensure that all partners are on the same page regarding possible next steps, after the completion of the initial data collection.

External Resources:

- *Resources from the American Geophysical Union's Thriving Earth Exchange (TEX)* – TEX works to facilitate and support community-based science, pairing scientists from the AGU membership with communities and local governments with questions about their environment. They support projects across a wide range of disciplines and have a wealth of knowledge on this topic. Their website includes some great resources including readings and project planning templates, these would be especially valuable for sensor

researchers and scientists without and prior experience in CBPR looking to begin a community-based air quality project. https://thrivingearthexchange.org/resources/

- *'Four questions to ask before buying an air quality sensor' infographic* - Another helpful tool is an infographic, developed by Tim Dye, this tool is intended to guide the selection of sensors for a project and includes helpful questions and links. http://tdenviro.com/wp-content/uploads/2017/09/4-questions-when-buying-air-sensors.pdf
- *Resources for those new to community-based research in general*: Often researchers participating in sensor projects come from an atmospheric chemistry or engineering background and may have little experience working with communities. There is already a wealth of information and resources to support successful academic/community partnerships. Below are several links to this work.
    - *Living Knowledge Toolbox* – The Living Knowledge Network is an international network working to support engaged work. Section 3 of the Toolbox includes links to a variety of helpful manuals, tools, and guides. http://www.livingknowledge.org/resources/toolbox/
    - *National Coordinating Center for Public Engagement's Guide to working with Local Communities* – This is a brief, but simple guide to getting partnerships started and includes helpful tips for both researchers and communities. http://www.publicengagement.ac.uk/do-engagement/partnership-working/working-with-local-communities
    - *Reading*: This article, entitled a "Review of community-based research: assessing partnership approaches to improve public health", is also a good resource, reviewing the benefits and challenges of CBPR. https://www.ncbi.nlm.nih.gov/pubmed/9611617

- Finally, I believe it is worth mentioning the resources from Chapter 2 once more as many of these tools are specifically geared to supporting air quality sensor projects. Table from Chapter 2 (Clements et al., 2017):

Category	Tool	Description	Website
Government	EPA's Air Sensor Toolbox	Resources to support communities and citizens selecting and using low-cost sensors	epa.gov/air-sensor-toolbox
	AirNow	Aggregated national data from regulatory grade monitoring stations and information on potential health impacts	airnow.gov
	EJSCREEN	Environmental justice screening and mapping tool	epa.gov/ejscreen
	RETIGO	EPA software for mapping collected air quality data	epa.gov/retigo
	C-FERST	Community-focused exposure and risk screening tool	epa.gov/c-ferst
	AQ-SPEC	Sensor evaluation program	aqmd.gov/aq-spec
	US EIA Energy Mapping System	Mapping tool to explore energy use and development across the United States	eia.gov/state/maps.php
	CalEnviroscreen	California specific environmental justice screening tool	oehha.ca.gov/calenviroscreen
Non-Profit and Academic	IVAN Monitoring System	Community-based monitoring system for environmental concerns	ivan-imperial.org/air
	FrackTracker	App for logging reports and observations around oil and gas activity	fractracker.org
	Institute for Health and Metrics Evaluation	Resource for exploring health data and statistics	healthdata.org
	EDF Air Sensor Workgroup Data Standards	Guidelines for sensor data formatting	edf.org/health/data-standards-date-and-timestamp-guidelines
	AQ-IQ Air Quality and Sensor Curriculum	K-12 curriculum supporting the use of low-cost sensors by students	teachengineering.org/curricularunits/view/cub_airquality_unit
	OpenAQ	Aggregated global data from regulatory grade monitoring stations	openaq.org
	SensorThingsAPI	Open source API for connecting IoT sensing devices, data, and applications over the web	opengeospatial.org/standards/sensorthings
Commercial	Envirosuite	Software for environmental management and planning	envirosuite.com
	Breezometer	Commercial API for displaying air quality indicators	breezometer.com

FUTURE WORK

Research will need to continue along the current trajectories to reach a point where sensor systems can be integrated into existing monitoring plans. Sensor calibration procedures and techniques for quantifying uncertainty will eventually need to be standardized, which may be challenging as these methods will need to be adaptable to the wide variety of potential sensor uses and applications. Before standardization, researchers will need to explore the use of more advanced calibration techniques (i.e., neural networks and other machine learning approaches) for methane and non-methane VOCs to improve the utility of sensors for assessing hazardous air pollutants. While there has been some research into advanced calibration techniques, specifically for several criteria pollutants (Cross et al., 2017; Zimmerman et al., 2018), methane (Casey et al., 2018c) and benzene (De Vito et al., 2008), more studies will help to identify the most reliable technique overall or for a given situation. Factor analysis methods are worth exploring as these conventional analysis tools could be applied to sensor data to produce useful information, example tools might be positive matrix factorization. Finally, researchers need to gain a better understanding of calibration transferability; simply stated, how robust are calibrations developed at one reference site when sensors are moved to their field sites.

This question has started to be explored as well (Castell et al., 2017; Hagan et al., 2018; Casey et al., 2018b), including a study by my colleague examining the transferability of carbon dioxide and ozone calibrations.

Another next step, not directly related to the technical capacity of sensors, but very important for community-based air quality research is the development of best practices regarding how the results of these sensor projects are presented and determining what potential actions may follow certain results. Communities often engage in projects utilizing air quality sensors because they would like to investigate a particular local concern, therefore, the results of these projects should provide either an answer to the initial research question or data to inform appropriate next steps. It would be helpful for the low-cost sensor community to consider what types of reports could be produced for community partners that could then be used to pursue additional research, more funding, or a conversation with local regulatory and public health representatives. This is a task that will require all stakeholders (researchers, community members, public health/regulatory representatives, and if possible, industry representatives), and decisions will need to be made regarding what needs to be included in these reports so the data can be properly assessed and acted on. A model for a roadmap that would lead to a potential mitigation action, that could be developed at the start of a project would also be beneficial. This roadmap model is likely much longer-term and will require many more case studies and iterations.

CLOSING REMARKS

My work and that of other researchers has demonstrated low-cost air quality sensors are capable of providing us with new and unique information about our environment. Sensors technologies are also versatile, these technologies may develop into tools to supplement conventional air quality research or become an established method for regulatory agencies and communities to work together to identify environmental inequities. Sensors may also find purpose as educational tools, teaching users about air quality, monitoring, and how our personal choices impact our exposure. While low-cost sensors may not lead to the kind of discoveries that change the thinking of the academic community, they will support bottom-up, localized research, research that takes on difficult questions, such as “is my child safe, breathing this air?”. Though they may never be able to provide a final answer to these questions, sensors can provide a starting point and create momentum. The niche capabilities of low-cost sensing systems and the potential to contribute to reduced exposure and improved environmental protections, despite the ongoing challenges and limitations, make them worthy of continued development and consideration.

References

Adgate, J. L., Goldstein, B. D., & McKenzie, L. M.: Potential public health hazards, exposures and health effects from unconventional natural gas development, Environ. Sci. Technol., 48(15), 8307–8320, doi:10.1021/es404621d, 2014.

Allen, D. T.: Methane emissions from natural gas pro-duction and use: reconciling bottom-up and top-down measurements, Curr. Opin. Chem. Eng., 5, 78–83, https://doi.org/10.1016/j.coche.2014.05.004, 2014.

Apte, J.S.; Messier, K.P.; Gani, S.; Brauer, M.; Kirchstetter, T.W.; Lunden, M.M.; Marshall, J.D.; Portier, C.J.; Vermeulen, R.C.H.; Hamburg, S.P. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. Environ. Sci. Technol., 51, 6999–7008. 2017.

Arfire, A., Marjovi, A., and Martinoli, A.: Mitigating slow dynamics of low-cost chemical sensors for mobile air quality monitoring sensor networks. ESWN ’16 Proceedings of the 2016 International Conference on Embedded Wireless Systems and Networks, Conference Proceedings, 159–167, ISBN: 978-0-9949886-0-7, 2016.

Arsenault, J., S. Godsoe, C. Holden and J. Vetelino, "Integration of Sensors into Secondary Classrooms," in 35th ASEE/IEEE Frontiers in Education Conference, Indianapolis, 2005.

Austin, E.; Novosselov, I.; Seto, E.; Yost, M.G. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor. PLoS ONE 2015, 10. 2015.

Bamberger, I., Stieger, J., Buchmann, N., and Eugster, W.: Spatial variability of methane: attributing atmospheric concentrations to emissions, Environ. Pollut., 190, 65–74, https://doi.org/10.1016/j.envpol.2014.03.028, 2014.

Bart, M.;Williams, D.E.; Ainslie, B.; Mckendry, I.; Salmond, J.; Grange, S.K.; Alavi-Shoshtari, M.; Steyn, D.; Henshaw, G.S. High density ozone monitoring using gas sensitive semi-conductor sensors in the lower Fraser Valley, British Columbia. Environ. Sci. Technol., 48, 3970–3977. 2014.

Basu, R.; Woodruff, T.J.; Parker, J.D.; Saulnier, L.; Schoendorf, K.C. Comparing exposure metrics in the relationship between PM2.5 and birth weight in California. J. Expo. Anal. Environ. Epidemiol., 14, 391–396. 2004.

Beloff, B. R., Beaver, E. R., & Massin, H. Assessing societal costs associated with environmental impacts. Environmental Quality Management, 10(2), 67–82. http://doi.org/10.1002/1520-6483(200024)10:2<67::AID-TQEM8>3.0.CO;2-B, 2000.

Berrocal, V.J.; Gelfand, A.E.; Holland, D.M. A Spatio-Temporal Downscaler for Output From Numerical Models. J. Agric. Biol. Environ. Stat., 15, 176–197. 2010.

Biaggi-Labiosa, A., Sola, F., Lebron-Colon, M., Evans, L. J., Xu, J. C., Hunter, G., Berger, G. M., and Gonzalez, J. M.: A novel methane sensor based on porous SnO2 nanorods: room temperature to high temperature detection, Nanotechnology, 23, 45, https://doi.org/10.1088/0957-4484/23/45/455501, 2012.

Bickerstaff, K. Risk perception research: Socio-cultural perspectives on the public experience of air pollution. Environ. Int. 30, 827–840. 2004.

Borrego, C.; Costa, A.M.; Ginja, J.; Amorim, M.; Coutinho, M.; Karatzas, K.; Sioumis, T.; Katsifarakis, N.; Konstantinidis, K.; De Vito, S.; et al. Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. Atmos. Environ., 147, 246–263. 2016.

Boulder County: Past review of oil and gas development regulations, available at: https://www.bouldercounty.org/ property-and-land/land-use/planning/oil-gas-development/ history/, last access: 8 November 2017. 2016.

Brown, P. Popular Epidemiology and Toxic Waste Contamination: Lay and Professional Ways of Knowing. Journal of Health and Social Behavior, 33(3), 267–281.1992.

Brown, P. Race, Class, and Environmental Health: A Review and Systemization of the Literature. Environmental Research, 69, 15–30. 1995.

Brugge, D., Durant, J. L., & Rioux, C. Near-highway pollutants in motor vehicle exhaust: A review of epidemiologic evidence of cardiac and pulmonary health risks. Environmental Health: A Global Access Science Source, 6(23), 1–12. http://doi.org/10.1186/1476-069X-6-23, 2007.

California Department of Public Health (CDPH), P.H.I. California Environmental Health Tracking Program: Improving Public Health with Better Information. Available online: http://www.cehtp.org/page/main (accessed on 20 October 2017). 2015.

Casey J. G., Ortega, J., Coffey, E., and Hannigan, M.: Low-cost measurement techniques to characterize the influence of home heating fuel on carbon monoxide in Navajo homes, Sci. Total Environ., 625, 608–618, https://doi.org/10.1016/j.scitotenv.2017.12.312, 2018a.

Casey, J. G. and Hannigan, M. P.: Testing the performance of field calibration techniques for low-cost gas sensors in new deployment locations: across a county line and across Colorado, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2018-81, in review, 2018b.

Casey, J.G., Collier-Oxandale, A., Hannigan, M. Performance of Artificial Neural Networks and Linear Models to Quantify 4 Trace Gas Species in an Oil and Gas Production Region with Low-Cost Sensors, [manuscript submitted] 2018c.

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., … Bartonova, A. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? Environment International, 99, 293–302. http://doi.org/10.1016/j.envint.2016.12.007, 2017.

Cheadle, L., Deanes, L., Sadighi, K., Casey, J. G., Collier-Oxandale, A., and Hannigan, M.: Quantifying neighborhood-scale spatial variations of ozone at open space and urban sites in Boulder, Colorado using low-cost sensor technology, Sensors, 17, 9, https://doi.org/10.3390/s17092072, 2017.

Chen, S., Liao, W., Chang, C., Su, Y., Tong, Y., Chang, J. S., & Wang, J. Network monitoring of speciated vs. total non-methane hydrocarbon measurements. Atmospheric Environment, 90, 33–42. http://doi.org/10.1016/j.atmosenv.2014.03.020, 2014.

Chilingar, G. V., & Endres, B. Environmental hazards posed by the Los Angeles Basin urban oilfields: An historical perspective of lessons learned. Environmental Geology, 47(2), 302–317. http://doi.org/10.1007/s00254-004-1159-0, 2005.

Clements, A., Griswold, W., RS, A., Johnston, J. E., Herting, M. M., Thorson, J., Collier-Oxandale, A., and Hannigan, M.: Low-cost air quality monitoring tools: from research to practice (a workshop summary), Sensors, 17, 2478, https://doi.org/10.3390/s17112478, 2017.

Clougherty, J. E., Kheirbek, I., Eisl, H. M., Ross, Z., Pezeshki, G., Gorczynski, J. E., … Matte, T. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants : The New York City Community Air Survey (NYCCAS), 23(3), 232–240. http://doi.org/10.1038/jes.2012.125. 2013.

Cole, L.W.; Foster, S.R. From the Ground up: Environmental Racism and the Rise of the Environmental Justice Movement; NYU Press: New York, NY, USA, 2001.

Collier, A.M., D. Knight, K. Hafich, M. P. Hannigan, B. Graves and M. Polmear, "The North Fork Valley Project - a Project-based Learning Curriculum to Support the use of Next-generation Monitoring Technologies in Rural Communities," in ASEE Rocky Mountain Section Conference, Denver, 2015a.

Collier, A., Knight, D., Hafich, K., Hannigan, M., Polmear, M., and Graves, B.: On the development and implementation of a project-based learning curriculum for air quality in K-12 schools, 2015 IEEE: Frontiers in Education Conference (FIE), Conference Proceedings, https://doi.org/10.1109/FIE.2015.7344246, 2015b.

Collier, J. D., & Mahoney, D. B. The First Cortical Implant of a Semiconductor Multielectrode Array: Continued Data Analysis (No. AFIT/GE/GSO/ENG/84D-20). AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING. 1984.

Collier-Oxandale, A., Hannigan, M. P., Casey, J. G., Piedrahita, R., Halliday, H. S., & Johnston, J. Assessing a low-cost methane sensor quantification system for use in complex rural and urban environments. Atmospheric Measurement Techniques, 11, 3569–3594. http://doi.org/10.5194/amt-2017-421, 2018a.

Collier-Oxandale, A., Coffey, E., Thorson, J., Johnston, J., & Hannigan, M. Comparing Building and Neighborhood-Scale Variability of CO2 and O3 to Inform Deployment Considerations for Low-Cost Sensor System Use. Sensors, 18(5), 1349. http://doi.org/10.3390/s18051349, 2018b.

Collier-Oxandale, A. M., Thorson, J., Halliday, H., and Hannigan, M. Understanding the ability of low-cost MOx sensors to quantify ambient VOCs, [manuscript in preparation] 2018c.

Colorado Department of Public Health and the Environment (CDPHE), https://www.colorado.gov/pacific/coepht/radon-data [last accessed June 2018]. 2018.

Colorado Department of Education (CDE), "Colorado Graduation Data 2014," 3 February 2015. [Online]. Available: http://www2.cde.state.co.us/cdereval/graduationdatamap2014.asp. [Accessed 1 March 2015]. 2015.

Colorado Oil and Gas Conservation Commission (COGCC), Online GIS Tool, tool available here: https://cogccmap.state.co.us/cogcc_gis_online/, (last access April 2018), 2017.

Colorado Oil and Gas Conservation Commission (COGCC), Wells, data layer, available at: http://cogcc.state.co.us/documents/data/ downloads/gis/WELLS_SHP.ZIP (last access: 4 October 2017), 2015.

Comite Civico del Valle, IVAN Imperial: Welcome to IVAN; Comite Civico del Valle, Inc.: Brawley, CA, USA, 2017a.

Comite Civico del Valle, IVAN Imperial: Air Monitoring Q&A. Available online: https://www.ivan-imperial.org/resources/airfaqs (accessed on 20 October 2017). 2017b.

Cooperative Institute for Research in Environmental Sciences, "Climate Literacy Education and Awareness Network," University of Colorado, Boulder, [Online]. Available: http://cires.colorado.edu/education-outreach/projects/current-projects/clean/. [Accessed 1 March 2015]. 2015.

Corburn, J. Environmental Justice, Local Knowledge, and Risk: The Discourse of a Community-Based Cumulative Exposure Assessment. Environ. Manag. 29, 451–466. 2002.

Crilley, L.R.; Shaw, M.; Pound, R.; Kramer, L.J.; Price, R.; Young, S.; Lewis, A.C.; Pope, F.D. Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. Atmos. Meas. Tech. Discuss., 1–24. 2017.

Cross, E. S., Williams, L. R., Lewis, D. K., Magoon, G. R., Onasch, T. B., Kaminsky, M. L., Worsnop, D. R., and Jayne, J. T.: Use of electrochemical sensors for measurement of air pollution: correcting interference response

and validating measurements, At-mos. Meas. Tech., 10, 3575–3588, https://doi.org/10.5194/amt-10-3575-2017, 2017.

Croxford, B.; Penn, A. Siting considerations for urban pollution monitors. Atmos. Environ., 32, 1049–1057. 1998.

Davis K. J., Deng A., Lauvaux T., Miles N. L., Richardson S. J., Sarmiento D. P., Kevin R., Gurney, R., Hardesty, M., Bonin, T. A., Brewer, W. A., Lamb, B. K., Shepson, P. B., Harvey, R. M., Cambaliza, M. O., Sweeney, C., Turnbull, J. C., Whetstone, J., and Karion, A.: The Indianapolis flux ex-periment (INFLUX): a test-bed for developing urban green-house gas emission measurements, Elem. Sci. Anth. 5, 21, https://doi.org/10.1525/elementa.188, 2017.

De Vito, S., Massera, E., Piga, M., Martinotto, L., and Di Francia, G.: On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, Sensor. Actuat. B-Chem., 129, 750–57, https://doi.org/10.1016/j.snb.2007.09.060, 2008.

NASA: Discover-AQ Database, available at: https://www-air.larc.nasa.gov/cgi-bin/ArcView/discover-aq.co-2014?GROUND-PLATTEVILLE=1?GROUND-PLATTEVILLE=1 [last accessed: May 2018], 2015.

Dunn, S., "Oil, gas industry has positive impact on Weld governments," Greeley Tribune, Greeley, 2014.

Duvall, R.M.; Long, R.W.; Beaver, M.R.; Kronmiller, K.G.; Wheeler, M.L.; Szykman, J.J. Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide. Sensors, 16. 2016.

Eisele, A. P., Mukerjee, S., Smith, L. A., Thoma, E. D., Whitaker, D. A., Oliver, K. D., … Stallings, C. Volatile organic compounds at two oil and natural gas production well pads in Colorado and Texas using passive samplers. Journal of the Air & Waste Management Association, 66(4), 412–419. http://doi.org/10.1080/10962247.2016.1141808, 2016.

English, P. B., Olmedo, L., Bejarano, E., Lugo, H., Murillo, E., Seto, E., … Northcross, A. The Imperial County Community Air Monitoring Network: A Model for Community-based Environmental Monitoring for Public Health Action, 1–5, 2016.

Environmental Defense Fund (EDF): Air Sensor Workgroup: Community-Driven Standards and Infrastructure for Air Quality Data. Available. Available online: https://www.edf.org/health/air-sensor-workgroup (accessed on 9 March 2018). 2018.

Esposito, E.; De Vito, S.; Salvato, M.; Bright, V.; Jones, R.L.; Popoola, O. Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems. Sens. Actuators B-Chem., 231, 701–713. 2016.

Etheridge, D. M., Steele, L. P., Francey, R. J., and Langenfelds, R. L.: Atmospheric methane between 1000 A.D. and present: evidence of anthropogenic emissions and climatic variability, J. Geophys. Res-Atmos., 103, 15979–15993, https://doi.org/10.1029/98JD00923, 1998.

Eugster, W., & Kling, G. W. Performance of a low-cost methane sensor for ambient concentration measurements in preliminary studies. Atmospheric Measurement Techniques, 5(8), 1925–1934. http://doi.org/10.5194/amt-5-1925-2012, 2012.

Fang, Y., Naik, V., Horowitz, L. W., and Mauzerall, D. L.: Air pollution and associated human mortality: the role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present, Atmos. Chem. Phys., 13, 1377–1394, https://doi.org/10.5194/acp-13-1377-2013, 2013.

Figaro: TGS 2600 – for the detection of air contaminants, On-line product data sheet, available at: http://www.figarosensor.com/products/2600pdf.pdf (last access: 4 October 2017), 2005a.

Figaro: TGS 2602 – for the detection of air contaminants, On-line product data sheet, http://www.figarosensor.com/products/2602pdf.pdf (last access: April 2018), 2005b.

Fiore, A. M., West, J. J., Horowitz, L. W., Naik, V., and Schwarzkopf, M. D.: Characterizing the tropospheric ozone response to methane emission controls and the benefits to climate and air quality, J. Geophys. Res., 113, 2156–2202, https://doi.org/10.1029/2007JD009162, 2008.

Fischer, P. H., Hoek, G., Van Reeuwijk, H., Briggs, D. J., Lebret, E., Van Wijnen, J. H., … Elliott, P. E. Traffic-related differences in outdoor and indoor concentrations of particles and volatile organic compounds in Amsterdam. Atmospheric Environment, 34(22), 3713–3722. http://doi.org/10.1016/S1352-2310(00)00067-4, 2000.

Gålfalk, M., Olofsson, G., Crill, P., & Bastviken, D.: Making methane visible. Nature, 6, 426–430. http://doi.org/10.1038/NCLIMATE2877, 2016.

Gao, M.L.; Cao, J.J.; Seto, E. A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. Environ. Pollut., 199, 56–65. 2015.

Gouw, J. De, & Warneke, C. Measurements of volatile organic compounds in the earth's atmosphere using proton-transfer-reaction mass spectrometry. Mass Spectrometry Reviews, 26, 223–257. http://doi.org/10.1002/mas.2007.

Griswold, W.; Hannigan, M.; Rosing, T.; Dasgupta, S.; Patrick, K.; Menarini, M.; Chan, C.; Vikram, S.; Dhanaraj, R.; Collier, A. MetaSense: Calibration of Personal Air Quality Sensors in the Field. In Proceedings of the NSF Cyber-Physical Systems Principal Investigators' Meeting, Arlington, VA, USA, 16–17 November 2015.

Hagan, D. H., Isaacman-vanwertz, G., Franklin, J. P., Wallace, L. M. M., Kocar, B. D., Heald, C. L., & Kroll, J. H. Calibration and assessment of electrochemical air quality sensors by co-location with regulatory-grade instruments, 315–328. 2018.

Hagler, G. S. W., Yelverton T. L. B., Vedantham, R., Hansen, A. D. A., and Turner, J. R.: Post-processing method to reduce noise while preserving high time resolution in aethalometer real-time black carbon data, Aerosol Air Qual. Res., 11, 539–546, https://doi.org/10.4209/aaqr.2011.05.0055, 2011.

Hagler, G.; Solomon, P.A.; Hunt, S.W. New Technology for Low-Cost, Real-Time Air Monitoring. In EM: Air and Waste Management Association's Magazine for Environmental Managers; Air & Waste Management Association: Pittsburgh, PA, USA, 2014.

Hall, E.S.; Kaushik, S.M.; Vanderpool, R.W.; Duvall, R.M.; Beaver, M.R.; Long, R.W.; Solomon, P.A. Integrating Sensor Monitoring Technology into the Current Air Pollution Regulatory Support Paradigm: Practical Considerations. Am. J. Environ. Eng., 4, 147–154. 2014.

Halliday, H. S., Thompson, A. M., Wisthaler, A., Blake, D., Hornbrook, R. S., Mikoviny, T., Muller, M., Eicher, P., Apel, E. C., Hills, A. J.: Atmospheric benzene observations from oil and gas production in the Denver Julesburg basin in July and August 2014, Journal of Geophysical Research: Atmospheres, 121(18), 11055-11074, doi: 10.1002/2016JD025327, 2016.

Hang, J., Li, Y., Sandberg, M., Buccolieri, R., & Di, S. The influence of building height variability on pollutant dispersion and pedestrian ventilation in idealized high-rise urban areas. *Building and Environment*, *56*, 346–360. http://doi.org/10.1016/j.buildenv.2012.03.023, 2012.

Heimann, I., Bright, V. B., McLeod, M. W., Mead, M. I., Popoola, O. A. M., Stewart, G. B., and Jones, R. L.: Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors, Atmos. Environ., 113, 10–19, https://doi.org/10.1016/j.atmosenv.2015.04.057, 2015.

Helmig, D., Thompson, C. R., Evans, J., Boylan, P., Hueber, J., and Park, J. H.: Highly elevated atmospheric levels of volatile organic compounds in the Uintah basin, Utah, Environ. Sci. Technol., 48, 4707–4715, https://doi.org/10.1021/es405046r, 2014.

Hew K.F., and T. Brush, "Integrating Technology into K-12 Teaching and Learning: Current Knowledge Gaps and Recommendations for Future Research," Educational Technology Research and Development,, vol. 55, no. 3, pp. 223-252, 2007.

Holm, M., "Project-based instruction: A review of the literature on effectiveness in prekindergarten through 12th grade classrooms," Insight: Rivier Academic Journal, vol. 7, no. 2, 2011.

Huyberechts, G., Szeco, P., Roggen, J., and Licznerski, B. W.: Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artifi-cial neural network, Sensor. Actuat. B-Chem., 45, 123–30 https://doi.org/10.1016/S0925-4005(97)00283-9, 1997.

Intergovernmental Panel on Climate Change (IPCC): Climate Change 2014 Synthesis Report, available at: https://www.ipcc. ch/pdf/assessment-report/ar5/syr/SYR_AR5_FINAL_full.pdf (last access: 4 October 2017), ISBN: 978-92-9169-143-2, 2015.

Israel, B.A.; Schulz, A.J.; Parker, E.A.; Becker, A.B. Review of community-based research: Assessing partnership approaches to improve public health. Annu. Rev. Public Health, 19, 173–202. 2012.

Jerrett, M., Donaire-gonzalez, D., Popoola, O., Jones, R., Cohen, R. C., Almanza, E., … Nieuwenhuijsen, M. Validating novel air pollution sensors to improve exposure estimates for epidemiological analyses and citizen science. Environmental Research, 158, 286–294. http://doi.org/10.1016/j.envres.2017.04.023, 2017.

Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; Giovis, C. A review and evaluation of intraurban air pollution exposure models. J. Expo. Anal. Environ. Epidemiol. 15, 185–204. 2005.

Jiao, W.; Hagler, G.; Williams, R.; Sharpe, R.; Brown, R.; Garver, D.; Judge, R.; Caudill, M.; Rickard, J.; Davis, M.; et al. Community Air Sensor Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. Atmos. Meas. Tech., 9, 5281–5292. 2016.

Johnson, K. K., Bergin, M. H., Russell, A. G., & Hagler, G. S. W. Field Test of Several Low-Cost Particulate Matter Sensors in High and Low Concentration Urban Environments. Aerosol and Air Quality Research, 18, 565–578. http://doi.org/10.4209/aaqr.2017.10.0418, 2018.

Johnston, J. E., & Gibson, J. M. Spatiotemporal variability of tetrachloroethylene in residential indoor air due to vapor intrusion: a longitudinal, community-based study. Journal of Exposure Science & Environmental Epidemiology, 24(6), 1–8. http://doi.org/10.1038/jes.2013.13, 2013.

Jovasevic-Stojanovic, M.; Bartonova, A.; Topalovic, D.; Lazovic, I.; Pokric, B.; Ristovski, Z. On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter. Environ. Pollut., 206, 696–704., 2015.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Hardesty, R. M., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R. S., Tans, P., Trainer, M., Zamora, R., and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophys. Res. Lett., 40, 4393– 4397, https://doi.org/10.1002/grl.50811, 2013.

Karp, T., and R. Gale, "Sustainability of K-12 Engineering Outreach Activities Beyond the Grant Period," in 39th ASEE/IEEE Frontiers in Education Conference, San Antonio, 2009.

Kaufman, A.; Williams, R.; Barzyk, T.; Greenberg, M.; O'Shea, M.; Sheridan, P.; Hoang, A.; Ash, C.; Teitz, A.; Mustafa, M. A Citizen Science and Government Collaboration: Developing Tools to Facilitate Community Air Monitoring. Environ. Justice, 10, 51–61. 2017.

Kelly, K. E., Whitaker, J., Petty, A., Widmer, C., Dybwad, A., Sleeth, D., … Butter, A. Ambient and laboratory evaluation of a low-cost particulate matter. Environmental Pollution, 221, 491–500. http://doi.org/10.1016/j.envpol.2016.12.039, 2017.

Kim, J., Shusterman, A. A., Lieschke, K. J., Newman, C., and Cohen, R. C.: The BErkeley Atmospheric CO2 Observation Network: field calibration and evaluation of low-cost air quality sensors, Atmos. Meas. Tech., 11, 1937–1946, https://doi.org/10.5194/amt-11-1937-2018, 2018.

Kim, J.; Shusterman, A.A.; Lieschke, K.J.; Newman, C.; Cohen, R.C. The berkeley atmospheric CO2 observation network: Field calibration and evaluation of low-cost air quality sensors. Atmos. Meas. Tech. Discuss., 2017.

Knight D.W., and J. F. Sullivan, "Integrating an assessment plan into a k-12/university engineering partnership," in Assessment in science: practical experiences and educational research, Arlington, National Science Teachers Association, pp. 149-158., 2006.

Knight, D.W., M. P. Hannigan, M. Polmear, L. Gardiner, K. Hafich and A. M. Collier, "Integrating a K-12 Education and Outreach Initiative into a Sustainability Research Network," in 122nd ASEE Annual Conference, Seattle, 2015.

Kort, E. A., Frankenberg, C., Costigan, K. R., Lindenmaier, R., Dubey, M. K., and Wunch, D.: Four corners: the largest US methane anomaly viewed from space, Geophys. Res. Lett., 41, 6898–6903, https://doi.org/10.1002/2014GL061503, 2014.

Król, S., Zabiegała, B., & Namieśnik, J. Monitoring VOCs in atmospheric air I. On-line gas analyzers. TrAC - Trends in Analytical Chemistry, 29(9), 1092–1100. http://doi.org/10.1016/j.trac.2010.05.007, 2010.

Kumar, A., & Víden, I. Volatile organic compounds: Sampling methods and their worldwide profile in ambient air. Environmental Monitoring and Assessment, 131(1–3), 301–321. http://doi.org/10.1007/s10661-006-9477-1, 2007.

Künzli, N., McConnell, R., Bates, D., Bastain, T., Hricko, A., Lurmann, F., ... & Peters, J. Breathless in Los Angeles: the exhausting search for clean air. American Journal of Public Health, 93(9), 1494-1499. 2003.

Kwon, J., Weisel, C. P., Turpin, B. J., Zhang, J., Korn, L. R., Morandi, M. T., … Colome, S. Source Proximity and Outdoor-Residential VOC Concentrations: Results from the RIOPA Study. Environmental Science and Technology, 40(13), 4074–4082. http://doi.org/10.1021/es051828u, 2006.

Lambert, T.W.; Guyn, L.; Lane, S.E. Development of local knowledge of environmental contamination in Sydney, Nova Scotia: Environmental health practice from an environmental justice perspective. Sci. Total Environ., 368, 471–484. 2006.

Laursen, S., C. Liston, H. Thiry and J. Graf, "What good Is a scientist in the classroom? Participant outcomes and program design features for a short-duration science outreach intervention in K–12 classrooms," Life Sciences Education, vol. 6, pp. 49-64, 2007.

Leidinger, M., Baur, T., Sauerwald, T., Schütze, A., Reimringer, W., Spinelle, L., & Gerboles, M. Highly sensitive benzene detection with MOS gas sensors. Proceedings Sensor 2017, 92–97. http://doi.org/10.5162/sensor2017/A4.3, 2017.

Leidinger, M., Rieger, M., Sauerwald, T., Alépée, C., & Schütze, A. Sensors and Actuators B: Chemical Integrated pre-concentrator gas sensor microsystem for ppb level benzene detection. Sensors & Actuators: B. Chemical, 236, 988–996. http://doi.org/10.1016/j.snb.2016.04.064, 2016.

Leidinger, M., Sauerwald, T., Reimringer, W., Ventura, G., & Schütze, A. Selective detection of hazardous VOCs for indoor air quality applications using a virtual gas sensor array. Journal of Sensors and Sensor Systems, 3(2), 253–263. http://doi.org/10.5194/jsss-3-253-2014, 2014.

Lewis, A. C., Lee, J., Edwards, P. M., Shaw, M. D., Evans, M. J., Moller, S. J., Smith, K. R., Buckley, J. W., Ellis, M., Gillot, S. R., and White, A.: Evaluating the performance of low cost chemical sensors for air pollution research, Faraday Discuss., 189, 85–103, https://doi.org/10.1039/C5FD00201J, 2016.

Lin, C.; Gillespie, J.; Schuder, M.D.; Duberstein, W.; Beverland, I.J.; Heal, M.R. Evaluation and calibration of Aeroqual series 500 portable gas sensors for accurate measurement of ambient ozone and nitrogen dioxide. Atmos. Environ. 2015, 100, 111–116.

Loh, M. M., Levy, J. I., Spengler, J. D., Houseman, E. A., & Bennett, D. H. Ranking cancer risks of organic hazardous air pollutants in the United States. Environmental Health Perspectives, 115(8), 1160–1168. http://doi.org/10.1289/ehp.9884, 2007.

M. Holm, "Project-based instruction: A review of the literature on effectiveness in prekindergarten through 12th grade classrooms," Insight: Rivier Academic Journal, vol. 7, no. 2, 2011.

Main, D.S.; Ware, G.; Iwasaki, P.G.; Burry, M.; Steiner, E.; Fedde, K.; Haverhals, L.M. Taking Neighborhood Health to Heart (TNH2H): Building a community-based participatory data system. Prev. Chronic Dis. 2012, 9.

Manikonda, A.; Zíková, N.; Hopke, P.K.; Ferro, A.R. Laboratory assessment of low-cost PM monitors. J. Aerosol Sci. 2016, 102, 29–40.

Marchese, A. J., Vaughn, T. L., Zimmerle, D. J., Martinez, D. M., Williams, L. L., Robinson, A. L., … Herndon, S. C. (2015). Methane Emissions from United States Natural Gas Gathering and Processing. http://doi.org/10.1021/acs.est.5b02275, 2015.

Markus, B., How a mine closure could transform the North Fork Valley, Colorado Public Radio, 2013.

Marshall, J. D. (2008). Environmental inequality: Air pollution exposures in California's South Coast Air Basin. Atmospheric Environment, 42, 5499–5503. http://doi.org/10.1016/j.atmosenv.2008.02.005, 2008.

Masson, N., Piedrahita, R., and Hannigan, M.: Approach for quantification of metal oxide type semiconductor gas sensors used for ambient air quality monitoring, Sensor. Actuat. B-Chem., 208, 339–345, https://doi.org/10.1016/j.snb.2014.11.032, 2015a.

Masson, N., Piedrahita, R., and Hannigan, M.: Quantification method for electrolytic sensors in long-term monitoring of ambient air quality, Sensors, 15, 27283–27302, https://doi.org/10.3390/s151027283, 2015b.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., ... & Cui, Y. Y. Volatile chemical products emerging as largest petrochemical source of urban organic emissions. Science, 359(6377), 760-764. 2018.

McKercher, G.R.; Salmond, J.A.; Vanos, J.K. Characteristics and applications of small, portable gaseous air pollution monitors. Environ. Pollut., 223, 102–110. 2017.

Mead, M. I., Popoola, O. A. M., Stewart, G. B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J. J., McLeod, M. W., Hodgson, T. F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J. R., and Jones, R. L.: The use of electro-

chemical sensors for monitoring urban air quality in low-cost, high-density networks, Atmos. Environ., 70, 186–203, https://doi.org/10.1016/j.atmosenv.2012.11.060, 2013.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, P. Natl. Acad. Sci. USA, 110, 20018–20022, https://doi.org/10.1073/pnas.1314392110, 2013.

Miller-Rushing, A., R. Primack and R. Boney, "The history of public participation in ecological research," Frontiers in Ecology, vol. 10, no. 6, pp. 285-290, 2012.

Mills J.E., and D. F. Treagust, "Engineering education - is problem-based learning or project-based learning the answer," Australasian Journal for Engineering Education, 2003.

Minkler, M. Linking Science and Policy Through Community-Based Participatory Research to Study and Address Health Disparities. Am. J. Public Health, 100, S81–S87. 2010.

Miskell, G., Salmond, J., & Williams, D. E. Low-cost sensors and crowd-sourced data: Observations of siting impacts on a network of air-quality instruments. Science of the Total Environment, The, 575, 1119–1129. http://doi.org/10.1016/j.scitotenv.2016.09.177. 2017.

Mobile Sensing Technology: U-Pod open-source construction and parts information, available at http://mobilesensingtechnology.com/ (last access: October 2017), n.d.

Mocon, Inc – Baseline: Series 9000 Total Hydrocarbon Analyzer, available at: www.baseline-mocon.com (last access: 7 May 2018), 2017.

Moltchanov, S., Levy, I., Etzion, Y., Lerner, U., Broday, D. M., and Fishbain, B.: On the feasibility of measuring urban air pollution by wireless distributed sensor networks, Sci. Total Environ., 502, 537–547, https://doi.org/10.1016/j.scitotenv.2014.09.059, 2015.

Moore, C. W., Zielinska, B., Petron, G., and Jackson, R. B.: Air impacts of increased natural gas acquisition, processing, and use: a critical review, Environ. Sci. Technol., 48, 8349–8359, https://doi.org/10.1021/es4053472, 2014.

Moskal, B.M., C. Skokan, L. Kosbar, A. Dean, C. Westland, H. Barker, Q. N. Nguyen and J. Tafoya, "K-12 Outreach: identifying the broader impacts of four outreach projects," Journal of Engineering Education, pp. 173-189, 2007.

Mukerjee, S., Smith, L. A., Thoma, E. D., Oliver, K. D., Whitaker, D. A., Wu, T., … Stallings, C. Spatial analysis of volatile organic compounds in South Philadelphia using passive samplers. Journal of the Air & Waste Management Association, 66(5), 492–498. http://doi.org/10.1080/10962247.2016.1147505, 2016.

Nam, E. K., Jensen, T. E., and Wallington, T. J.: Methane Emissions from Vehicles, Environ. Sci. Technol., 38, 2005–2010, https://doi.org/10.1021/es034837g, 2004.

National Oceanic and Atmospheric Administration (NOAA). Global Greenhouse Gas Reference Network, (Global Monitoring Division). Available online: https://www.esrl.noaa.gov/gmd/ccgg/trends/ (accessed on 9 March 2018). 2018.

Nikzad, N.; Verma, N.; Ziftci, C.; Bales, E.; Quick, N.; Zappi, P.; Patrick, K.; Dasgupta, S.; Krueger, I.; Rosing, T.Š. Citisense: Improving geospatial environmental assessment of air quality using a wireless personal exposure monitoring system. In Proceedings of the Conference on Wireless Health, La Jolla, CA, USA, 22–25 October 2012; ACM: New York, NY, USA, 2012; p. 11.

Peischl, J., Ryerson T. B., Brioude, J., Aikin, K. C., Andrews, A. E., Atlas, E., Blake, D., Daube, B.C., De Gouw, J. A., Dlugokencky, E., Frost, G. J., Gentner, D. R., Gilman, J. B., Goldstein, A. H., Harley, R. A., Holloway, J. S., Kofler, J., Kuster, W. C., Lang, P. M., Novelli, P.C., Santoni, G. W., Trainer, M., Wofsy, S. S. C., and Parrish, D. D.: Quantifying sources of methane using light alkanes in the Los Angeles Basin, California, J. Geophys. Res.-Atmos., 118, 4974–90, https://doi.org/10.1002/jgrd.50413, 2013.

Penza, M., Suriano, D., Cassano, G., Pfister, V., Amodio, M., Trizio, L., Brattoli, M., and De Gennaro, G.: A case-study of microsen-sors for landfill air-pollution monitoring applications, Urban Cli-mate, 14, 351–369, https://doi.org/10.1016/j.uclim.2014.09.002, 2015.

Perrenet, J., B. P.A.J. and S. J.G.M.M., "The suitability of problem-based learning for engineering education: theory and practice," Teaching in Higher Education, vol. 5, no. 3, pp. 345-358, 2000.

Pétron, G., Karion, A., Sweeney, C., Miller, B. R., Montzka, S. A., Frost, G. J., Trainer, M., Tans, P., Andrews, A., Kofler, J., Helmig, D., Guenther, D., Dlugokencky, E., Lang, P., New-berger, T., Wolter, S., Hall, B., Novelli, P., Brewer, A., Conley, S., Hardesty, M., Banta, R., White, A., Noone, D., Wolfe, D., and Schnell, R.: A new look at methane and nonmethane hydro-carbon emissions from oil and natural gas operations in the Col-orado Denver-Julesburg Basin, J. Geophys. Res.-Atmos., 119, 6836–6852, https://doi.org/10.1002/2013JD021272, 2014.

Pfister G., Flocke, F., Hornbrook, R., Orlando, J., Lee, S., Schroeder, J.: Final Report: Process-Based and Regional Source Impact Analysis for FRAPPÉ and DISCOVER-AQ 2014, National Center for Atmospheric Research, report available at: https://www.colorado.gov/airquality/tech_doc_repository.aspx?action=open&file=FRAPPE-NCAR_Final_Report_July2017.pdf, 2017.

Phillips, N. G., Ackley, R., Crosson, E. R., Down, A., Hutyra, L. R., Brond, M., … Jackson, R. B. (2013). Mapping urban pipeline leaks : Methane leaks across Boston, 173, 11–14. http://doi.org/10.1016/j.envpol.2012.11.003

Picarro, Inc: Picarro G2301 CRDS Analyzer CO2 CH4 H2O Mea-surements in Air, available at: www.picarro.com (last access: 7 May 2018), 2015.

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, Atmos. Meas. Tech., 7, 3325–3336, https://doi.org/10.5194/amt-7-3325-2014, 2014.

Pilot Project with Taking Neighborhood Health to Heart. Investigating Indoor Air Quality in Northeast Denver (TNH2H), https://experiment.com/projects/investigating-indoor-air-quality-in-northeast-denver [last accessed: June 2018] 2015.

Popa, M. E., Vollmer, M. K., Jordan, A., Brand, W. A., Pathirana, S. L., Rothe, M., and Röckmann, T.: Vehicle emissions of green-house gases and related tracers from a tunnel study: CO:CO2, N2O:CO2, CH4:CO2, O2:CO2 ratios, and the stable isotopes 13C and 18O in CO2 and CO, Atmos. Chem. Phys., 14, 2105– 2123, https://doi.org/10.5194/acp-14-2105-2014, 2014.

Quaranta, F., Rella, R., Siciliano, P., Capone, S., Epifani, M., Vasanelli, L., Licciullic, A., and Zoccod, A.: A novel gas sen-sor based on SnO2/Os thin film for the detection of methane at low temperature, Sensor. Actuat. B-Chem., 58, 350–355, https://doi.org/10.1016/S0925-4005(99)00095-7, 1999.

Ruckstuhl, A. F., Henne, S., Reimann, S., Steinbacher, M., Vollmer, M. K., Doherty, S. O., & Buchmann, B. Robust extraction of baseline signal of atmospheric trace species using local regression, 2613–2624. http://doi.org/10.5194/amt-5-2613-2012, 2012.

Rumchev, K., Brown, H., & Spickett, J. Volatile Organic Compounds: Do the present a risk to our health? Reviews on Environmental Health, 22(1), 39–55. 2007.

Sadd, J. and Shamasunder, B.: Oil extraction in Los Angeles: health, land use, and environmental justice consequences, from the Drilling Down Report, by the Liberty Hill Foundation, avail-able at: https://www.libertyhill.org/sites/libertyhillfoundation/ files/DrillingDownReport_1.pdf (last access: 4 October 2017), 2015.

Sadighi, K., Coffey, E., Polidori, A., Feenstra, B., Lv, Q., Henze, D. K., and Hannigan, M.: Intra-urban spatial variability of surface ozone in Riverside, CA: viability and validation of low-cost sensors, Atmos. Meas. Tech., 11, 1777–1792, https://doi.org/10.5194/amt-11-1777-2018, 2018.

Sahagun, L. Chemical odor, kids' nosebleeds, few answers in South L.A. Neighborhood. Los Angeles Times, 2013.

Sahagun, L. South L.A.'s Allenco fined $99,000 over toxic emissions. Los Angeles Times, 30 July 2014.

Sauerwald, T., Baur, T., Leidinger, M., Reimringer, W., Spinelle, L., Gerboles, M., … Schütze, A. Highly sensitive benzene detection with metal oxide semiconductor gas sensors – An inter-laboratory comparison. Journal of Sensors and Sensor Systems, 7(1), 235–243. http://doi.org/10.5194/jsss-7-235-2018, 2018.

Schneider, P., Castell, N., Vogt, M., Dauge, F. R., Lahoz, W. A., & Bartonova, A.. Mapping urban air quality in near real-time using observations from low- cost sensors and model information. Environment International, 106(May), 234–247. http://doi.org/10.1016/j.envint.2017.05.005, 2017.

Schütze, A., Baur, T., Leidinger, M., Reimringer, W., Jung, R., Conrad, T., & Sauerwald, T. Highly Sensitive and Selective VOC Sensor Systems Based on Semiconductor Gas Sensors: How to? Environments, 4(1), 20. http://doi.org/10.3390/environments4010020, 2017.

Shamasunder, B., Collier-Oxandale, A., Blickley, J., Sadd, J., Chan, M., Navarro, S., Hannigan, M., and Wong, N. J.: Community-based health and exposure study around urban oil developments in south Los Angeles, Int. J. Environ. Res. Public Health, 15, 1, https://doi.org/10.3390/ijerph15010138, 2018.

Shen, Jialei, and Zhi Gao. Ozone removal on building material surface: A literature review. *Building and Environment, 134, 205-217,* 2018.

Shen, Z., Zhang, X., Ma, X., Mi, R., Chen, Y., & Ruan, S. The significant improvement for BTX (benzene, toluene and xylene) sensing performance based on Au-decorated hierarchical ZnO porous rose-like architectures. Sensors & Actuators: B. Chemical, 262, 86–94. http://doi.org/10.1016/j.snb.2018.01.205, 2018.

Shusterman, A. A., Teige, V. E., Turner, A. J., Newman, C., Kim, J., & Cohen, R. C. The BErkeley Atmospheric CO2 Observation Network: initial evaluation, 13449–13463. http://doi.org/10.5194/acp-16-13449-2016, 2016.

Simmons, A., & Colbeck, I. Resistance of various building materials to ozone deposition. *Environmental Technology*, *11*(10), 973-978, 1990.

Smith, Steven W. The scientist and engineer's guide to digital signal processing. Available at: http://www.dspguide.com/CH28.PDF, (last accessed 08/18) 1997.

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin D. A., Kilaru, V. J., and Preuss, P. W.: The changing paradigm of air pollution monitoring, Environ. Sci. Technol., 47, 11369–11377 https://doi.org/10.1021/es4022602, 2013.

South Coast Air Quality Management District (SCAQMD): Air Quality Sensor Performance Evaluation Center (AQ-SPEC). Available online: http://www. aqmd.gov/aq-spec/home (accessed on 9 March 2018). 2015

Souza, J. C. D., Jia, C., Mukherjee, B., & Batterman, S. Ethnicity, housing and personal factors as determinants of VOC exposures. Atmospheric Environment, 43(18), 2884–2892. http://doi.org/10.1016/j.atmosenv.2009.03.017, 2009.

Spinelle, L.; Aleixandre, M.; Gerboles, M. Protocol of Evaluation and Calibration of Low-Cost Gas Sensors for the Monitoring of Air Pollution; EUR 26112 EN; European Commission Joint Research Centre: Brussels, Belgium, 2013.

Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., and Bonavitacola, F.: Field calibration of a cluster of low-cost available sensors for air quality monitoring. part A: ozone and nitrogen dioxide, Sensor. Actuat. B-Chem., 215, 249–257, https://doi.org/10.1016/j.snb.2015.03.031, 2015a.

Spinelle, L.; Gerboles, M.; Aleixandre, M. Performance evaluation of amperometric sensors for the monitoring of O3 and NO2 in ambient air at ppb level. Procedia Eng. 2015, 120, 480–483. 2015b.

Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., and Bonavitacola, F.: Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. part B: NO, CO and CO2, Sensor. Actuat. B-Chem., 238, 706–715, https://doi.org/10.1016/j.snb.2016.07.036, 2017a.

Spinelle, L., Gerboles, M., Kok, G., Persijn, S., & Sauerwald, T. Performance Evaluation of Low-Cost BTEX Sensors and Devices within the EURAMET Key-VOCs Project. Proceedings, 1(4), 425. http://doi.org/10.3390/proceedings1040425, 2017b.

Spinelle, L., Gerboles, M., Kok, G., Persijn, S., & Sauerwald, T. Review of Portable and Low-Cost Sensors for the Volatile Organic Compounds. Sensors, 17. http://doi.org/10.3390/s17071520, 2017c.

Subramanian, R., Williams, L. L., Vaughn, T. L., Zimmerle, D., Roscioli, J. R., Herndon, S. C., Yacovitch, T. I., Floerchinger, C., Tkacik, D. S., Mitchell, A. L., Sullivan, M. R., Dallman, T. R., and Robinson, A. L.: Methane emissions from natural gas compressor stations in the transmission and storage sector: measurements and comparisons with the EPA greenhouse gas reporting program protocol, Environ. Sci. Technol., 49, 3252–3261, https://doi.org/10.1021/es5060258, 2015.

Suh, H. H., Bahadori, T., Vallarino, J., & Spengler, J. D. Criteria Air Pollutants and Toxic Air Pollutants. Environmental Health Perspectives, 108(4), 625–633. 2000.

Sullivan, J.F., S. E. Davis, J. L. deGrazia and D. W. Carlson, "Beyond the Pipeline: Building a K-12 Engineering Outreach Program," in 29th Annual Frontiers in Education Conference, San Juan, 1999.

Sun, L.; Wong, K.C.; Wei, P.; Ye, S.; Huang, H.; Yang, F.H.; Westerdahl, D.; Louie, P.K.K.; Luk, C.W.Y.; Ning, Z. Development and Application of a Next Generation Air Sensor Network for the Hong Kong Marathon 2015 Air Quality Monitoring. Sensors, 16, 211. 2016.

Sun, Y. F., Liu, S. B., Meng, F. L., Liu, J. Y., Jin, Z., Kong, L. T., and Liu, J. H.: Metal oxide nanostructures and their gas sensing properties: a review, Sensors 12, 2610–2631, https://doi.org/10.3390/s120302610, 2012.

Szulczynski, B., & Gebicki, J. Currently Commercially Available Chemical Sensors Employed for Detection of Volatile Organic Compounds in Outdoor and Indoor Air. Environments, 4(21). http://doi.org/10.3390/environments4010021, 2017.

The Buck Institute for Education, "What is Project Based Learning (PBL)?," The Buck Institute for Education, [Online]. Available: http://bie.org/about/what_pbl. [Accessed 5 March 2015]. 2018.

Thoma, E. D., Brantley, H. L., Oliver, K. D., Whitaker, D. A., Mukerjee, S., Mitchell, B., … Philadelphia, S. (2016). South Philadelphia passive sampler and sensor study. Journal of the Air & Waste Management Association, 66(10), 959–970. http://doi.org/10.1080/10962247.2016.1184724, 2016.

Thompson, C. R., Hueber, J., and Helmig, D.: Influence of oil and gas emissions on ambient atmospheric non-methane hydrocarbons in residential areas of Northeastern Colorado, Elementa: Science of the Anthropocene, 3, 35, https://doi.org/10.12952/journal.elementa.000035, 2014.

Trabue, S., Scoggin, K., Mcconnell, L. L., Li, H., Turner, A., Burns, R., … Hatfield, J. Performance of commercial nonmethane hydrocarbon analyzers in monitoring oxygenated volatile organic compounds emitted from animal feeding operations. Journal of the Air & Waste Management Association, 63(10), 1163–1172. http://doi.org/10.1080/10962247.2013.804464, 2013.

Trumbull, D.J.; Bonney, R.; Bascom, D.; Cabral, A. Thinking scientifically during participation in a citizen-science project. Sci. Educ. 84, 265–275. 2000.

Tsujita, W.; Yoshino, A.; Ishida, H.; Moriizumi, T. Gas sensor network for air-pollution monitoring. Sens. Actuators B-Chem., 110, 304–311. 2005.

U.S. Energy Information Administration (US EIA): States Profiles and Energy Estimates, Colorado, available at: https: //www.eia.gov/state/analysis.php?sid=CO (last access: 4 October 2017), 2016.

U.S. Energy Information Administration (US EIA): States Profiles and Energy Estimates, California, available at: https:// www.eia.gov/state/analysis.php?sid=CA (last access: 4 October 2017), 2016.

U.S. Energy Information Administration (US EIA): Top 100 oil fields, available at: https://www.eia.gov/naturalgas/ crudeoilreserves/top100/pdf/top100.pdf (last access: 4 October 2017), 2015.

U.S. Environmental Protection Agency (US EPA), "EJ View (Colorado search)," 2 March 2015. [Online]. Available: http://epamap14.epa.gov/ejmap/. [Accessed 1 March 2015]. 2015.

U.S. Environmental Protection Agency (US EPA): Technical Note – Reporting Negative Values for Criteria Pollutant Gaseous Monitors to AQS, available at: https://www.epa.gov/sites/production/files/2017-02/documents/ negative_values_reporting_to_aqs_10_6_16.pdf (last access: 15 April 2018), 2016.

U.S. Environmental Protection Agency (US EPA): Air Sensor Toolbox for Citizen Scientists, Researchers and Developers. Available online: https://www.epa.gov/air-sensor-toolbox (accessed: October 2017). 2017a.

U.S. Environmental Protection Agency (US EPA): Air Sensor Toolbox: Evaluation of Emerging Air Pollution Sensor Performance. Available online: https://www.epa.gov/air-sensor-toolbox/evaluation-emerging-air-pollution-sensor-performance (accessed on 20 October 2017). 2017b.

U.S.Environmental Protection Agency (US EPA): EJ Small Grant Program recipients, https://www.epa.gov/sites/production/files/2017-11/documents/ejsg_2017_project_summaries_0.pdf, 2017c.

U.S. Environmental Protection Agency (US EPA), Quality Assurance Handbook for Air Pollution Measurement Systems Volume II Ambient Air Quality Monitoring Program Chapter 6, Monitor Network Design, available at: https://www3.epa.gov/ttn/amtic/files/ambient/pm25/qa/Final%20Handbook%20Document%201_17.pdf (last accessed 08/2018) 2017d.

U.S. Environmental Protection Agency (US EPA): Overview of greenhouse gases: methane, available at: https: //www.epa.gov/ghgemissions/overview-greenhouse-gases# CH4reference, last access: 4 October 2017. 2017e.

U.S. Environmental Protection Agency (US EPA), Health Risks of Radon, https://www.epa.gov/radon/health-risk-radon (last accessed: 06/2018). 2018.

Van den Bossche, M., Rose, N. T., and De Wekker, S. F. J.: Potential of a low-cost gas sensor for atmospheric methane monitoring, Sensor. Actuat. B-Chem., 238, 501–509, https://doi.org/10.1016/j.snb.2016.07.092, 2017.

Vardoulakis, S., Gonzalez-flesca, N., & Fisher, B. E. A. (2005). Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris, *39*, 2725–2736. http://doi.org/10.1016/j.atmosenv.2004.05.067, 2005.

Vardoulakis, Sotiris, et al. Modelling air quality in street canyons: a review. *Atmospheric environment* 37(2), 155-182. 2003.

Verma, N.; Zappi, P.; Rosing, T. Latent variables based data estimation for sensing applications. In Proceedings of the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Adelaide, Australia, 6–9 December 2011; pp. 335–340.

Wang, C., Yin, L., Zhang, L., Xiang, D., and Gao, R.: Metal oxide gas sensors: sensitivity and influencing factors, Sensors, 10, 2088–2106, https://doi.org/10.3390/s100302088, 2010.

Wang, Y.; Li, J.Y.; Jing, H.; Zhang, Q.; Jiang, J.K.; Biswas, P. Laboratory Evaluation and Calibration of Three Low-Cost Particle Sensors for Particulate Matter Measurement. Aerosol Sci. Technol., 49, 1063–1077. 2015.

Warneke, C., Geiger, F., Edwards, P. M., Dube, W., Pétron, G., Kofler, J., … Brown, S. S. (2014). Volatile organic compound emissions from the oil and natural gas industry in the Uintah Basin , Utah : oil and gas well pad emissions compared to ambient air composition, (x), 10977–10988. http://doi.org/10.5194/acp-14-10977-2014.

West, J. J., Arlene, F. M., Horowitz, L. W., and Mauzerall, D. L.: Global health benefits of mitigating ozone pollution with methane emission controls, P. Natl. Acad. Sci. USA., 103, 3988– 3993, https://doi.org/10.1073/pnas.0600201103, 2006.

Wilcox, J., Gopstein, A. M., Arent, D., Wofsy, S., Brown, N. J., Bradley, R., and Stucky, G. D.: Methane leaks from north american natural gas systems, Science, 343, 733–735, https://doi.org/10.1126/science.1247045, 2014.

Williams, R., & Kaufman, A. Next Generation Air Monitoring (NGAM) VOC Sensor Evaluation Report VOC Sensor Evaluation Report, (May). 2015.

Williams, R.; Kaufman, A.; Hanley, T.; Rice, J.; Garvey, S. Evaluation of Field-deployed Low-Cost PM Sensors; EPA/600/R-14/464; U.S. Environmental Protection Agency: Washington, DC, USA, 2014.

Williams, R.; Kilaru, V.; Snyder, E.; Kaufman, A.; Dye, T.; Rutter, A.; Russell, A.; Hafner, H. Air Sensor Guidebook; EPA/600/R-14/159 (NTIS PB2015-100610); U.S. Environmental Protection Agency: Washington, DC, USA, 2014.

Wilson, J. G., Kingham, S., Pearce, J., & Sturman, A. P. A review of intraurban variations in particulate air pollution: Implications for epidemiological research, *39*, 6444–6462. http://doi.org/10.1016/j.atmosenv.2005.07.030, 2005.

Wing, S. Environmental justice, science and public health. Environ. Health Perspect., 113, 54–63. 2005.

Wolfrum, E. J., Meglen, R. M., Peterson, D., & Sluiter, J. Metal oxide sensor arrays for the detection, differentiation, and quantification of volatile organic compounds at sub-parts-per-million concentration levels. Sensors and Actuators, B: Chemical, 115, 322–329. http://doi.org/10.1016/j.snb.2005.09.026, 2006.

Wong, K.W., Fu, D., Pongetti, T. J., Newman, S., Kort, E. A., Duren, R., Hsu Y.K., Miller, C. E., Yung, Y. L., Sander, S. P.: Mapping CH4:CO2 ratios in Los Angeles with CLARS-FTS from Mount Wilson, California. Atmos. Chem. Phys. 15(1), 241-252, doi:10.5194/acp-15-241-2015, 2015.

Wong, M., Bejarano, E., Carvlin, G., Id, K. F., King, G., Lugo, H., … English, P. (2018). Combining Community Engagement and Scientific Approaches in Next-Generation Monitor Siting: The Case of the Imperial County Community Air Network. http://doi.org/10.3390/ijerph15030523, 2018.

Woodruff, T. J., Axelrad, D. A., Caldwell, J., Morello-Frosch, R., & Rosenbaum, A. Public health implications of 1990 air toxics concentrations across the United States. Environmental Health Perspectives, 106(5), 245–251., 1998.

World Health Organization (WHO), Air Pollution, available at http://www.who.int/airpollution/en/ (last accessed: 08/2018). 2018.

Wu, X., Fan, Z., Zhu, X., Jung, K. H., Ohman-strickland, P., Weisel, C. P., & Lioy, P. J. Exposures to volatile organic compounds (VOCs) and associated health risks of socio-economically disadvantaged population in a "hot spot" in Camden, New Jersey. Atmospheric Environment, 57, 72–79. http://doi.org/10.1016/j.atmosenv.2012.04.029, 2012.

Yacovitch T. I., Herndon, S. C., Pétron, G., Kofler, J., Lyon, D., Zahniser, M. S., and Kolb C. E.: Mobile laboratory observations of methane emissions in the barnett shale region, Environ. Sci. Technol., 49, 7889–7895, https://doi.org/10.1021/es506352j, 2015.

Zarske, M., J. Yowell, J. Sullivan and D. Knight, "The TEAMS Program: A Study of a Grades 3-12 Engineering Contiunuum," in ASEE 2007 Annual Conference and Exposition, Honolulu, 2007.

Zavala-araiza, D., Lyon, D. R., Alvarez, R. A., Davis, K. J., Harriss, R., Herndon, S. C., Karion, A., Kort, E. A., Lamb, B. K., Lan, X. Marchese, A. J., Pacala, S. W., Robinson A. L., Shepson, P. B., Sweeny, C., Talbot, R., Townsend-Small, A., Yacovitch, T. I., Zimmerle D. J., and Hamburg, S. P.: Reconciling divergent estimates of oil and gas methane emissions, P. Natl. Acad. Sci. USA., 112, 15597–15602, https://doi.org/10.1073/pnas.1522126112, 2015.

Zhang, Z., Tong, J., Chen, D., and Yu-bin Lan, Y.: Electronic nose with an air sensor matrix for detecting beef freshness, J. Bionic Eng., 5, 67–73, https://doi.org/10.1016/S1672-6529(08)60008-6, 2008.

Zikova, N.; Hopke, P.K.; Ferro, A.R. Evaluation of new low-cost particle monitors for PM2.5 concentrations measurements. J. Aerosol Sci., 105, 24–34. 2017.

Zimmerman, N., Presto, A. A., Kumar, S. P. N., Gu, J., Hauryliuk, A., Robinson, E. S., Robinson, A. L., and R. Subramanian: A machine learning calibration model using random forests to im-prove sensor performance for lower-cost air quality monitoring, Atmos. Meas. Tech., 11, 291–313, https://doi.org/10.5194/amt-11-291-2018, 2018.

Appendix I

MEMORANDUM OF UNDERSTANDING TEMPLATE

List of Project Partners and Affiliations

Partner 1

Affiliation 2

Partner 2

Affiliation 1

Project Overview

Include here

- A description of the project and the objectives
- A general timeframe and where the project will occur, including a brief description of the community and sampling sites
- An overview of the methods (e.g., number of sensor systems, sensor types, and any important notes regarding the limitations of these sensors as it relates to the project)
- An overview of each partner's background, skills, and role in the project

Responsibilities and Rights of the Partners

Partner # (list responsibilities and rights)

Include here:

- Responsibilities of partner, and associated deadlines
- Rights of each partner (for example to the data)
- Other expectations (e.g., a number of participants to be recruited, or a description of the final report that will be delivered to a community partner)
- Be sure to include expectations around analyzing, interpreting, and sharing the data (for example, you may want to include the expectation that no partner will publicly share the data without consulting all partners, or that data will only be shared after the data is de-identified with respect to participants or sampling locations)
- **Complete for each partner*

Partner Signatures

Appendix II

PROJECT ONE-PAGER TEMPLATE

Note, this may be tailored for either regulatory partners assisting with sensor co-locations or community members hosting stations sensor systems.

Project Overview

- Where sampling will occur and what pollutants are of interest
- Project objectives, location, and general timeframe
- Plans for data or where/how it will be shared (e.g., a journal article)

Sensor System Specifications

- Size, weight, and power requirements (for those hosting monitors, translate this into a total electricity cost)
- Requirements for deployment locations (e.g., preferred placement at site), and how they are secured at a site
- Maintenance and access to system required
- List sensors and include a photo

Contact Information

- Include appropriate contacts of there are issues during the co-location or field deployment

Additional Optional Sections

- Detailed schedule of sensor system deployment and access needed (for residents)
- Details of project participant compensations or gift cards that will be provided as a thank you

Appendix III

INITIAL QUESTIONS FOR PROJECT PARTNERS

These questions are intended to support initial conversations among project partners ahead of the development of a complete plan for data collection and analysis.

1. List the all potential results of the data collection.

2. What are the follow-up actions that your team agrees would be best to conduct for each potential outcome?

3. What would you need to conduct these actions (e.g., additional funding? More resources? a longer project timeline?)

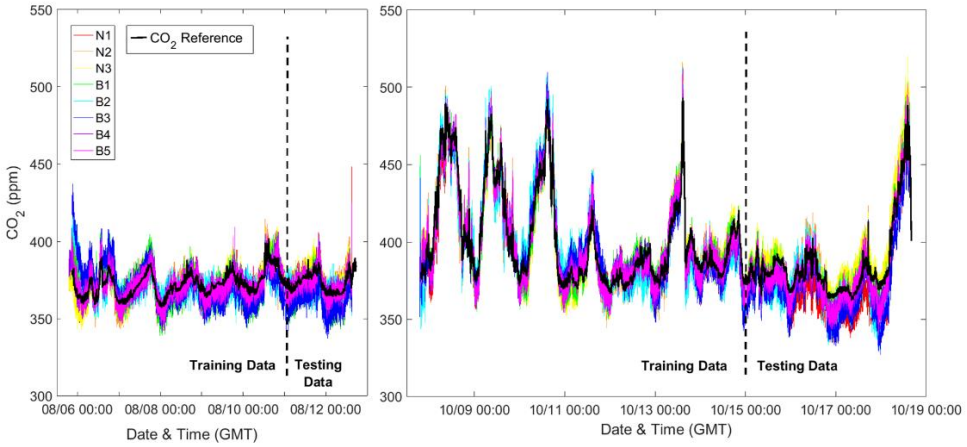
At this point develop a plan with your project partners regarding how you will proceed in each of these cases and determine if any additional resources need to be in place prior to the beginning of data collection. It's also good to incorporate these thoughts into the project plan, then an actionable and agreed upon plan is already in place.

Appendix IV

SUPPLEMENTAL FOR CHAPTER 5

Table 1: Performance Statistics for each Y-Pod; RMSE - CO_2 (ppm) & O_3 (ppb); MB – Mean Bias

	Carbon Dioxide						Ozone					
	Training Data			Testing Data			Training Data			Testing Data		
Pod	R^2	RMSE	MB	R^2	RMSE	MB	R^2	RMSE	MB	R^2	RMSE	MB
B2	0.93	7.32	-0.01	0.94	9.69	-6.35	NA	NA	NA	NA	NA	NA
B3	0.93	7.85	-0.04	0.93	7.39	2.69	0.97	3.45	-0.11	0.94	5.16	-1.88
B4	0.86	11.2	0.01	0.83	11.3	5.84	0.96	3.88	-0.07	0.96	4.73	-2.18
B5	0.94	7.35	-0.03	0.92	7.25	1.90	0.96	4.00	-0.11	0.92	6.51	-3.07
B7	0.94	7.32	0.01	0.77	14.7	10.4	0.97	3.71	-0.13	0.94	5.30	-1.99
B8	0.91	9.11	-0.04	0.88	9.00	0.39	0.98	2.78	-0.08	0.97	3.96	-0.94
C9	0.88	10.3	-0.04	0.89	14.6	12.8	0.97	3.74	-0.08	0.95	5.10	-2.96
D2	0.95	6.19	-0.02	0.93	6.83	3.49	0.96	3.95	-0.08	0.93	6.21	-3.08
Ave	0.92	8.33	-0.02	0.89	10.1	3.89	0.97	3.65	-0.09	0.94	5.28	-2.30
SD	0.03	1.71	0.02	0.06	3.16	5.95	0.01	0.42	0.02	0.02	0.86	0.79



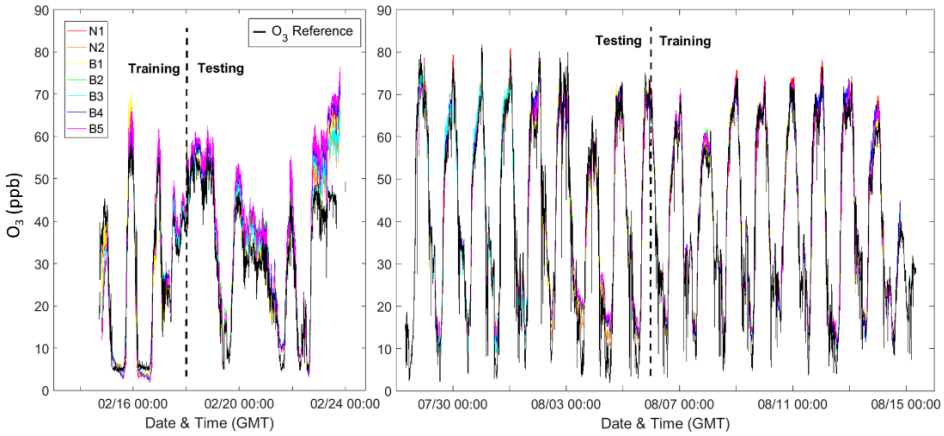


Figure 1: Time series of Y-Pod data (shown in colors), and reference data (shown in black), for CO_2 (top) and O_3 (bottom).

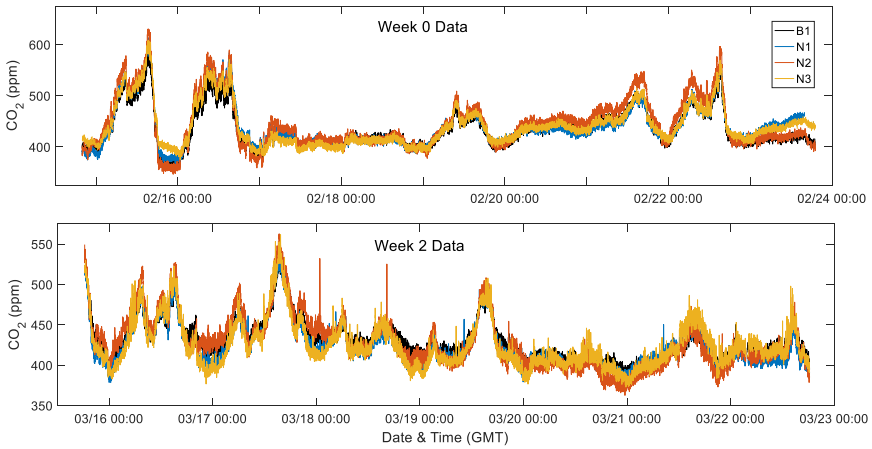


Figure 2: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining neighborhood-scale variability of CO_2.

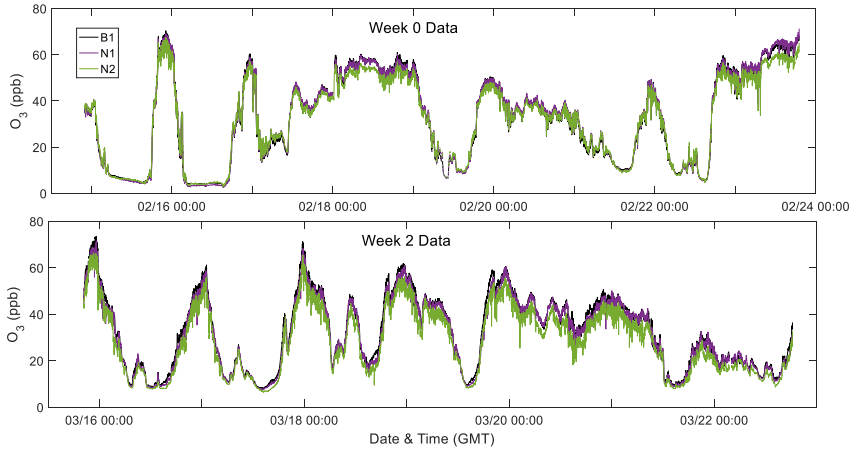


Figure 3: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining neighborhood-scale variability of O_3.

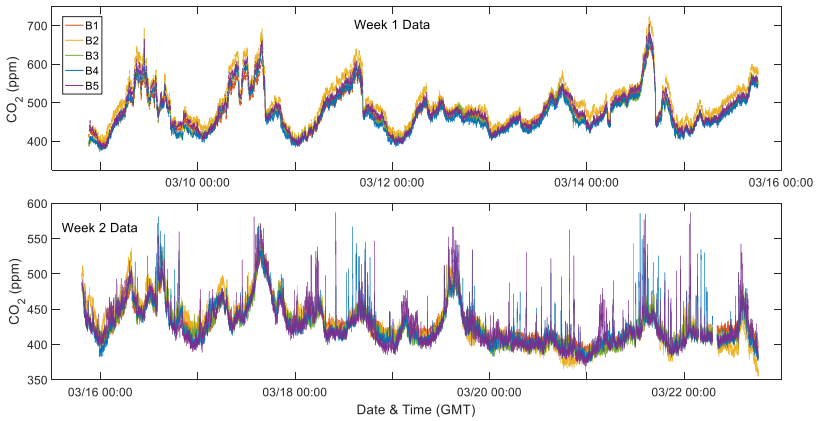


Figure 4: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining building-scale variability of CO_2.

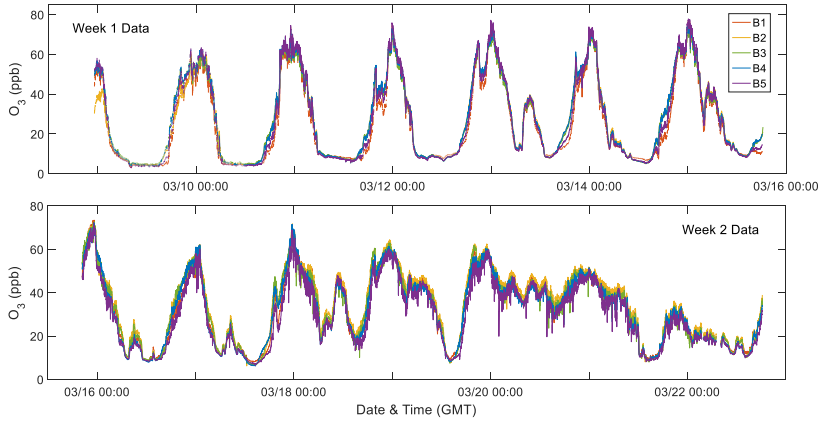


Figure 5: Time series of co-located (top) and deployed (bottom) Y-Pods, for examining building-scale variability of O_3.

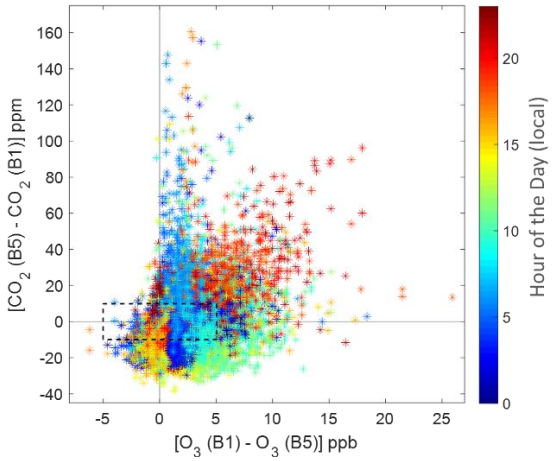


Figure 6: Scatter plot of differences between B1 and B5 (on the x-axis, for O_3, B1 – B5; and on the y-axis, for CO_2, B5 – B1), colored by the hour of the day. The dotted lines around the center axes indicate the respective RMSEs for CO_2 and O_3.

Appendix V

SUPPLEMENTAL FOR CHAPTER 6

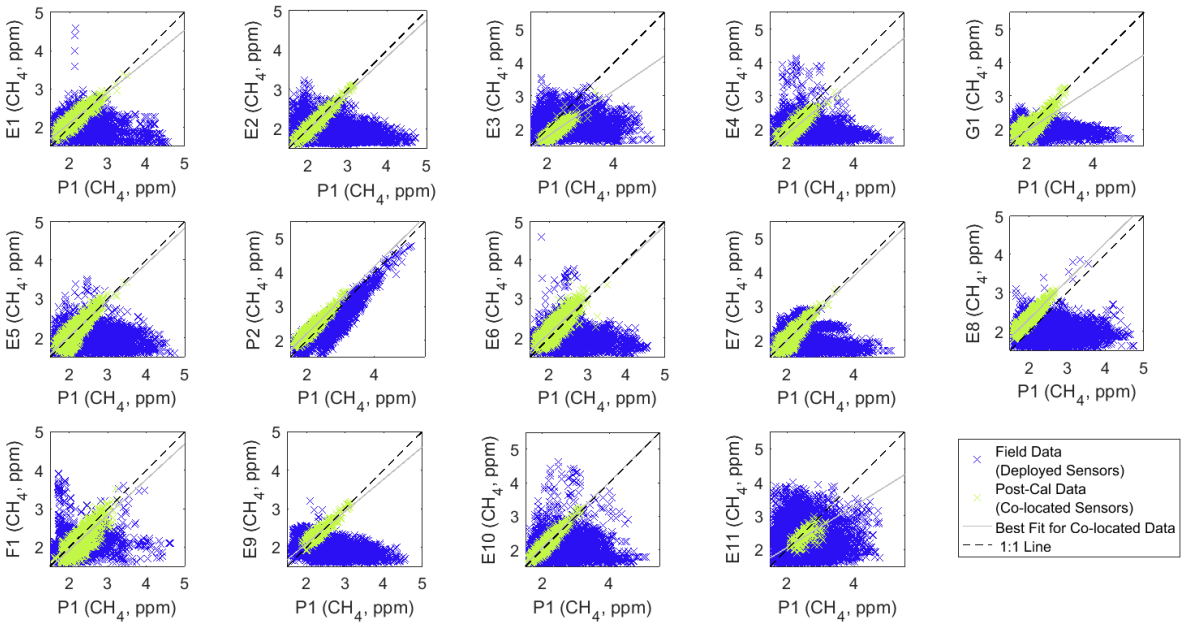


Figure 1. Scatter plots of P1 versus every other Y-Pod including both co-located data (in green) and deployed data (in blue).

Table 1: Statistics for Colorado data converted using the universal model method, including data for each individual Y-Pod and statistics for each Y-Pod versus Y-Pod P1.

	Co-located		Deployed	
	R	RMSE	*R*	RMSE
E1	0.914	0.117	−0.159	0.581
E2	0.940	0.084	−0.095	0.594
E3	0.909	0.284	0.168	0.619
E4	0.863	0.201	0.050	0.591
G1	0.737	0.200	−0.179	0.565
E5	0.849	0.132	0.054	0.529
P2	0.961	0.210	0.904	0.241
E6	0.847	0.166	−0.061	0.535
E7	0.864	0.124	−0.067	0.572
E8	0.931	0.336	−0.186	0.599
F1	0.819	0.182	0.221	0.550
E9	0.931	0.095	−0.454	0.655
E10	0.929	0.107	0.053	0.585
E11	0.624	0.208	0.104	0.571
Average	0.866	0.175	0.025	0.556
SD	0.092	0.073	0.306	0.096

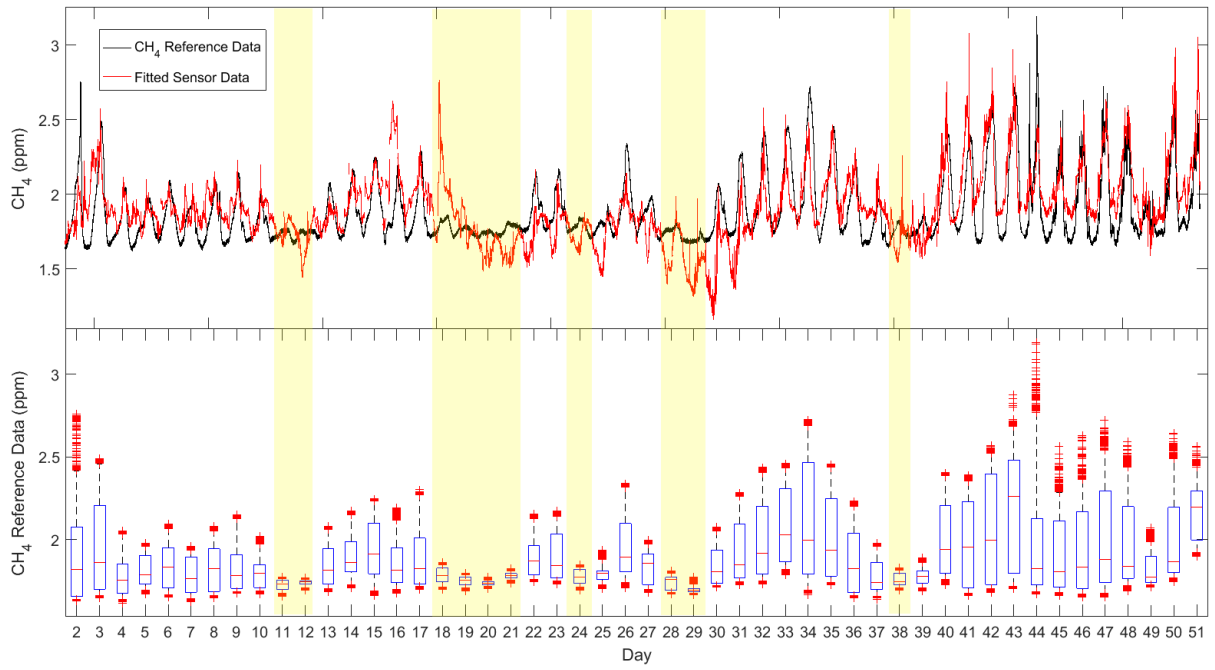


Figure 2: Panel (a) includes time series of methane reference data and fitted sensor data. Panel (b) includes the methane data for the Validation 2 dataset binned by days. The yellow segments highlight periods when underestimations below atmospheric background were removed, which coincide with days where the dynamic range is less than the expected uncertainty (RMSE is 0.18 ppm).

Appendix VI

SUPPLEMENTAL FOR CHAPTER 7

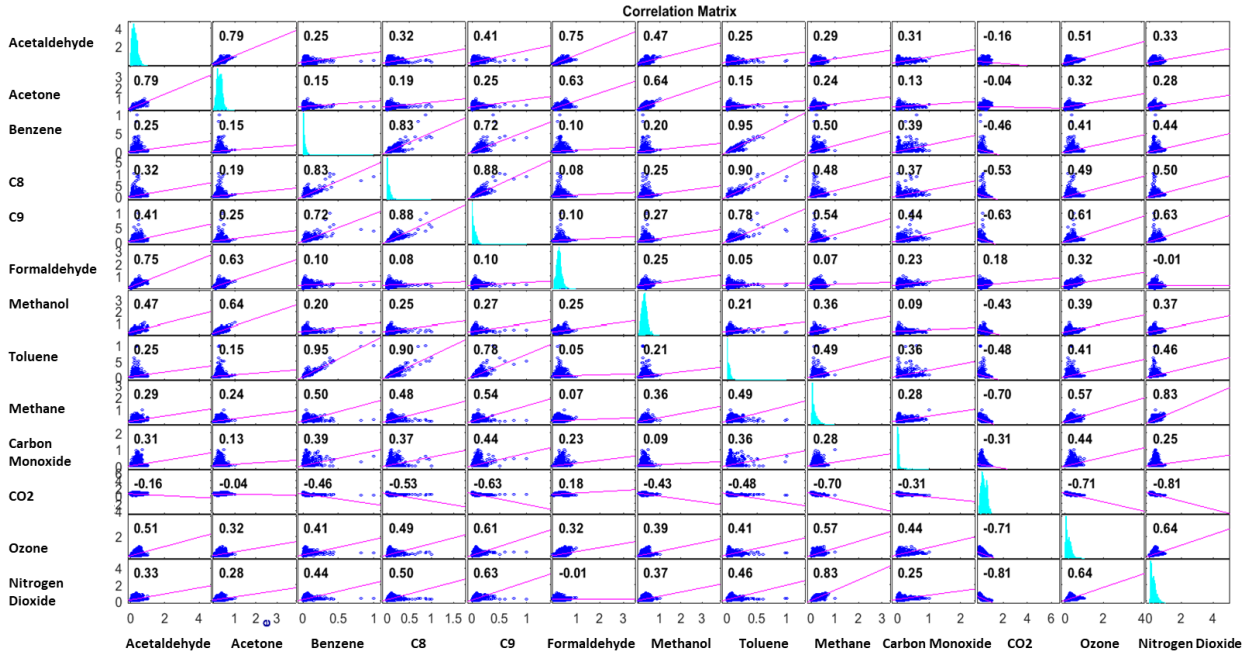
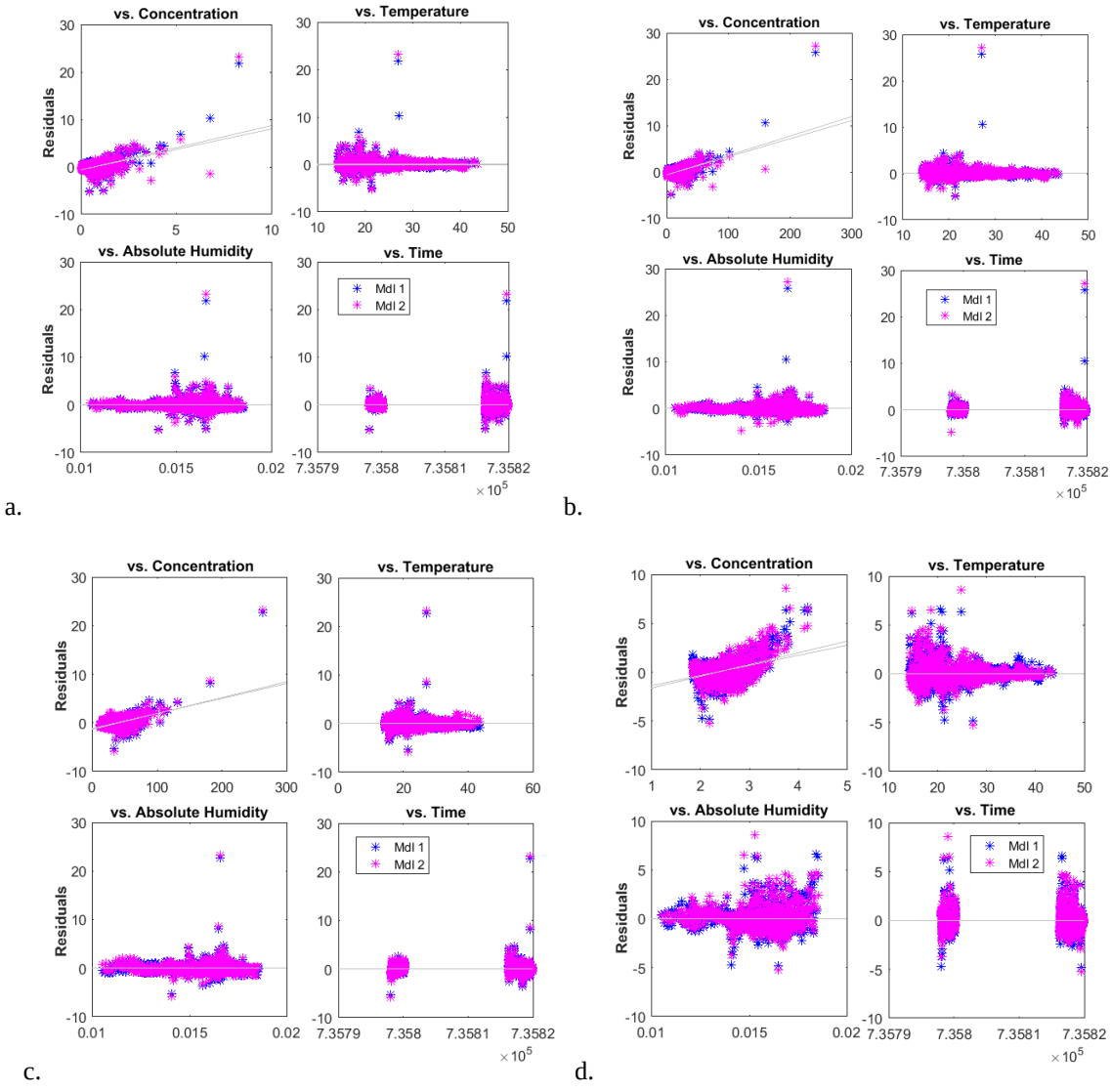
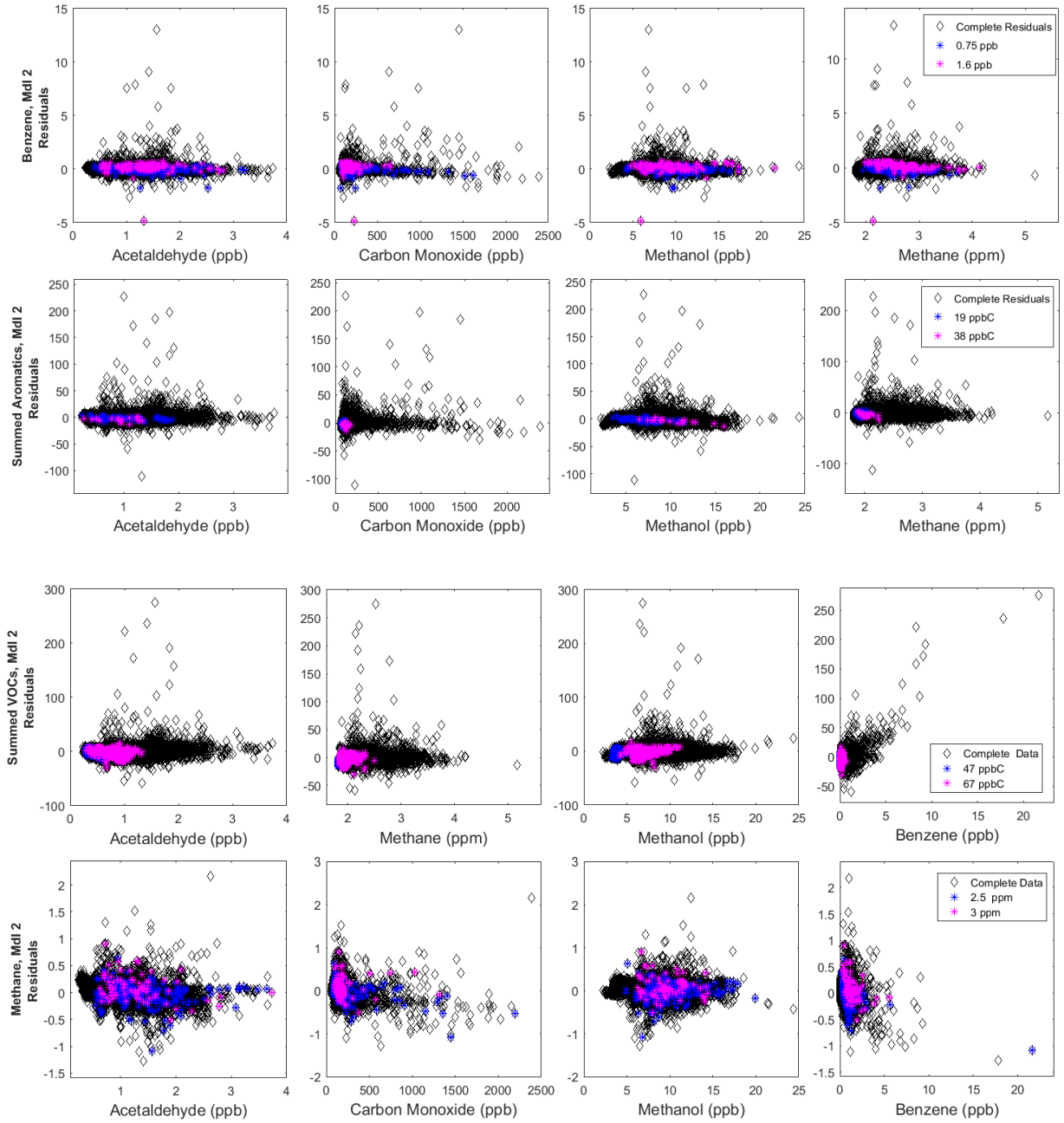


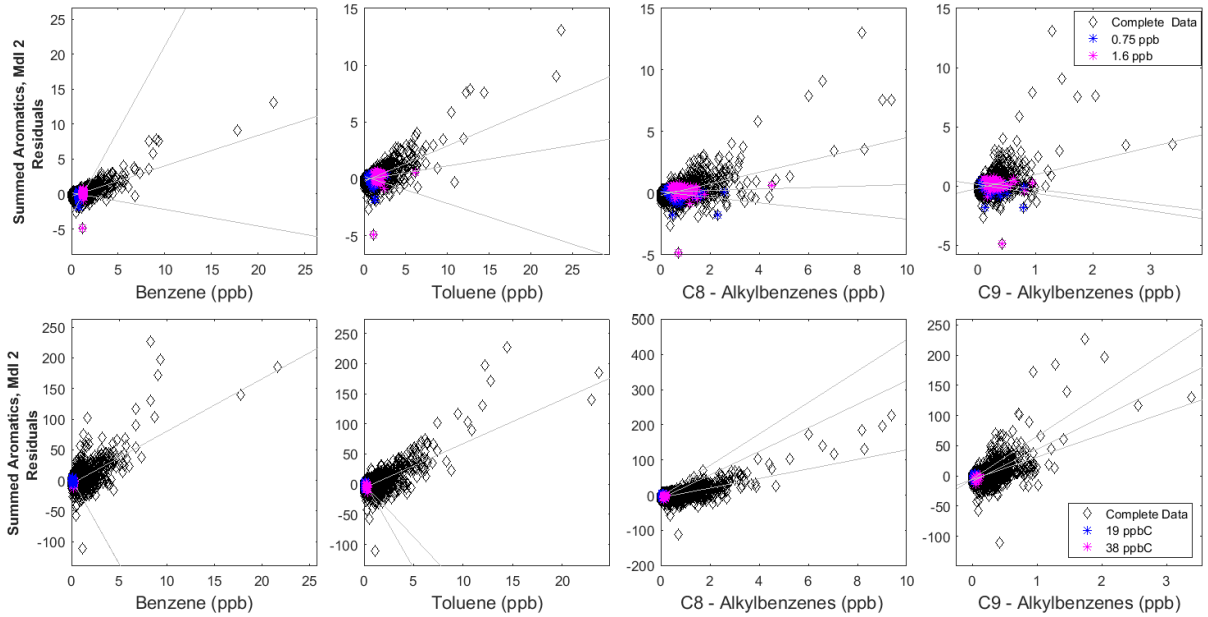
Figure 1 – Correlation plot for reference gases



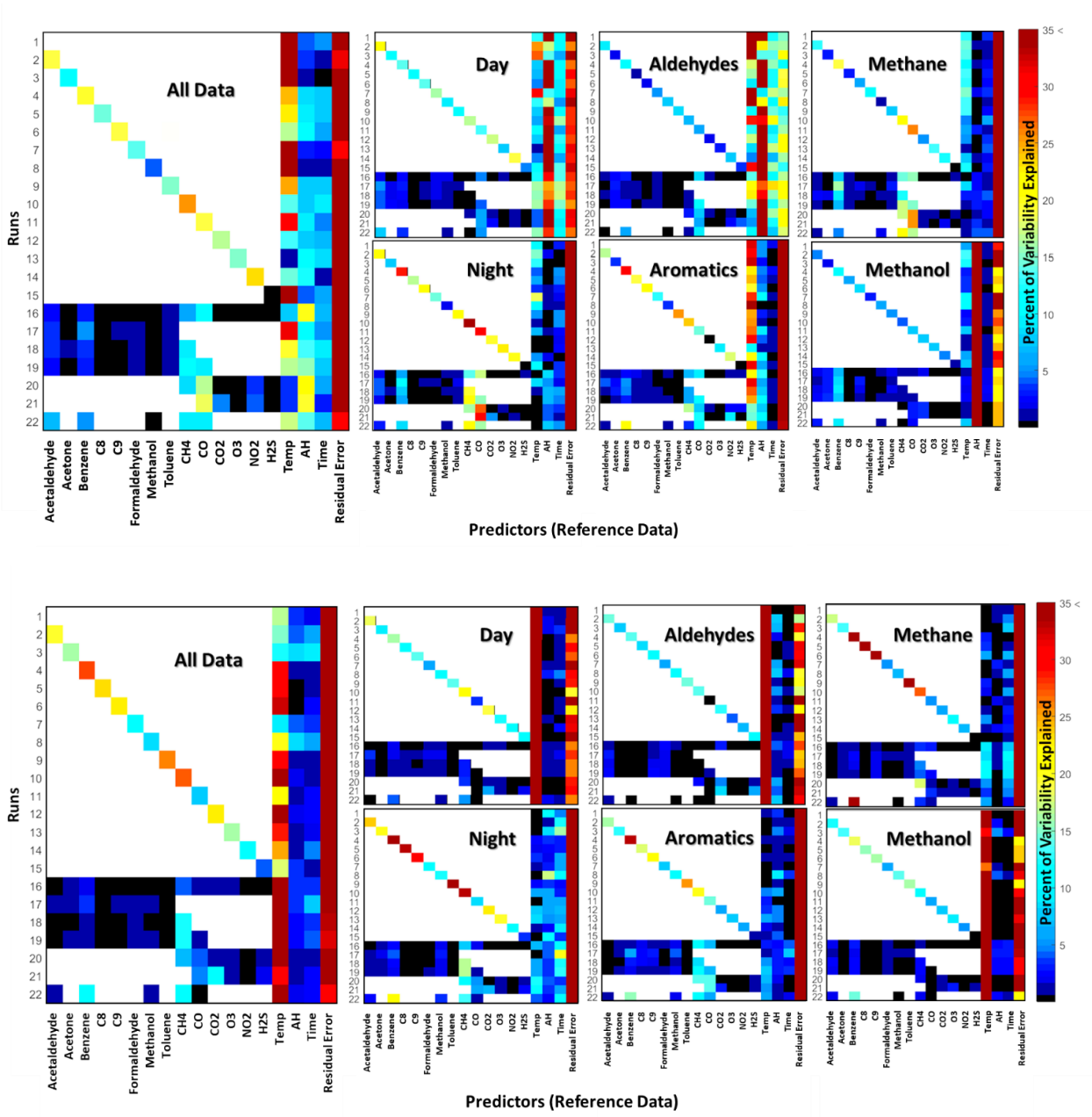
Figures 2 - Regression model residuals (from Section 3.1)

a (top left four panels) – models for benzene
b (top right four panels) – models for summed aromatics
c (bottom left four panels) – models for summed VOCs
d (bottom right four panels) – models for methane


Benzene, Mdl 2 Residuals
Acetaldehyde (ppb)
Carbon Monoxide (ppb)
Methanol (ppb)
Methane (ppm)
Complete Residuals
0.75 ppb
1.6 ppb
Summed Aromatics, Mdl 2 Residuals
Acetaldehyde (ppb)
Carbon Monoxide (ppb)
Methanol (ppb)
Methane (ppm)
Complete Residuals
19 ppbC
38 ppbC
Summed VOCs, Mdl 2 Residuals
Acetaldehyde (ppb)
Methane (ppm)
Methanol (ppb)
Benzene (ppb)
Complete Data
47 ppbC
67 ppbC
Methane, Mdl 2 Residuals
Acetaldehyde (ppb)
Carbon Monoxide (ppb)
Methanol (ppb)
Benzene (ppb)
Complete Data
2.5 ppm
3 ppm

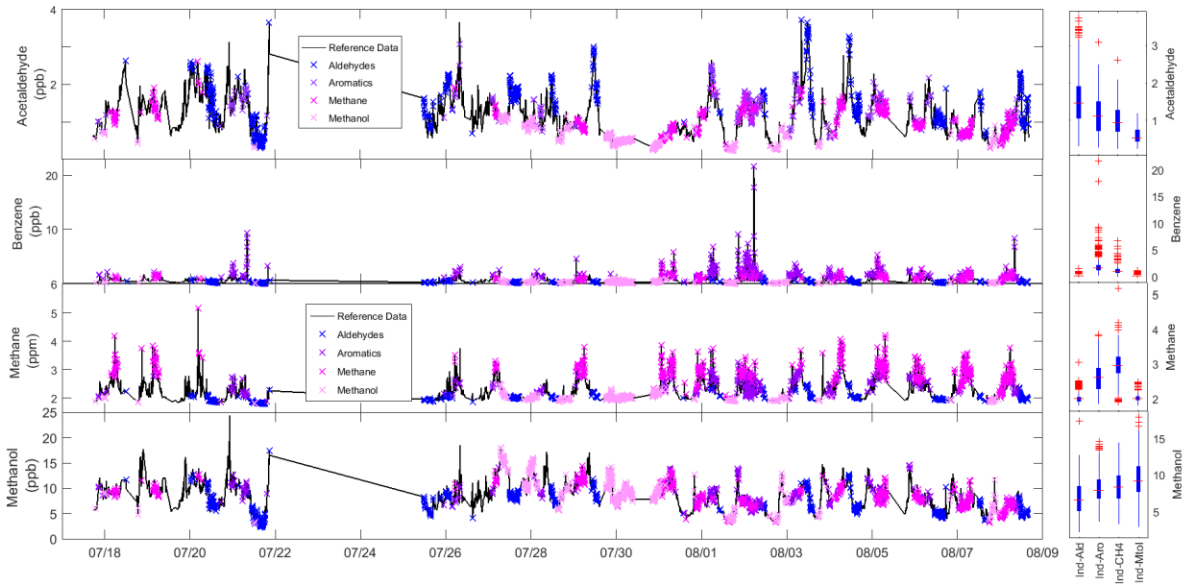


Figures 3 – Additional residuals for models vs. target and non-target VOCs



Figures 4 – ANOVA results illustrating the percentage of variance in the sensor signal explained by various predictors for complete data and subsets of data (top – Figaro 2600, bottom – Figaro 2602)

(White indicates that a predictor was not included in a run, and the VOC species subsets are defined in Figure VI5)



Figures 5 – Illustration of selected periods of different relative composition

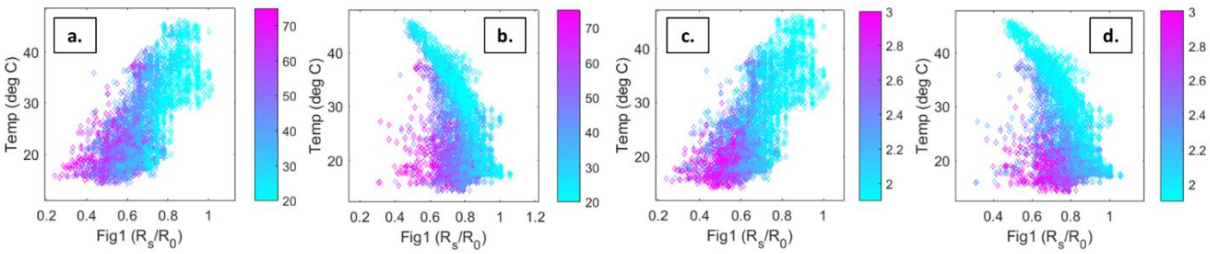


Figure 6: Sensor signal vs. temperature and pollutant concentrations

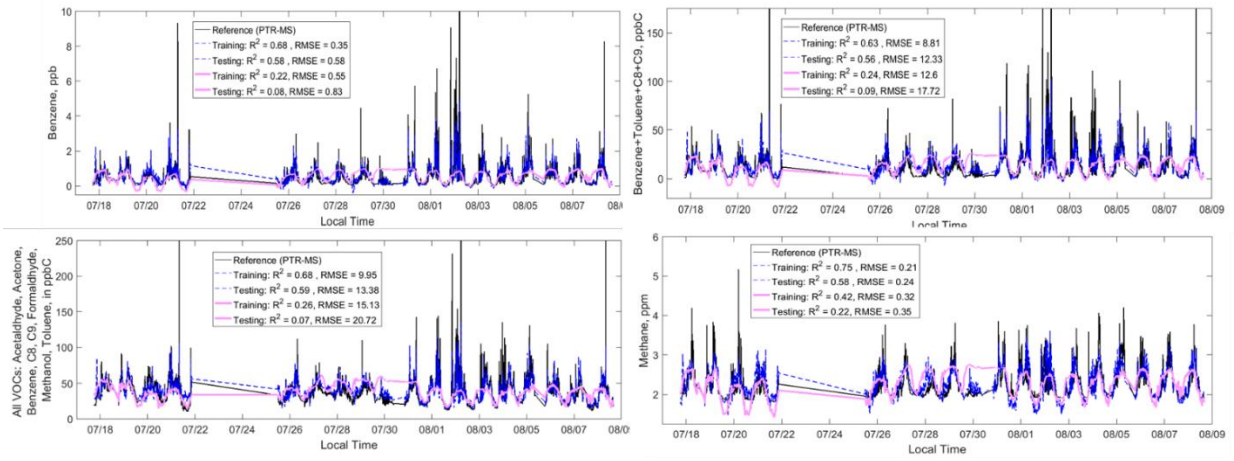


Figure 7: Regression analysis results, excluding VOC sensor signals

(training data before 7/20 & after 8/5, remaining data is testing)

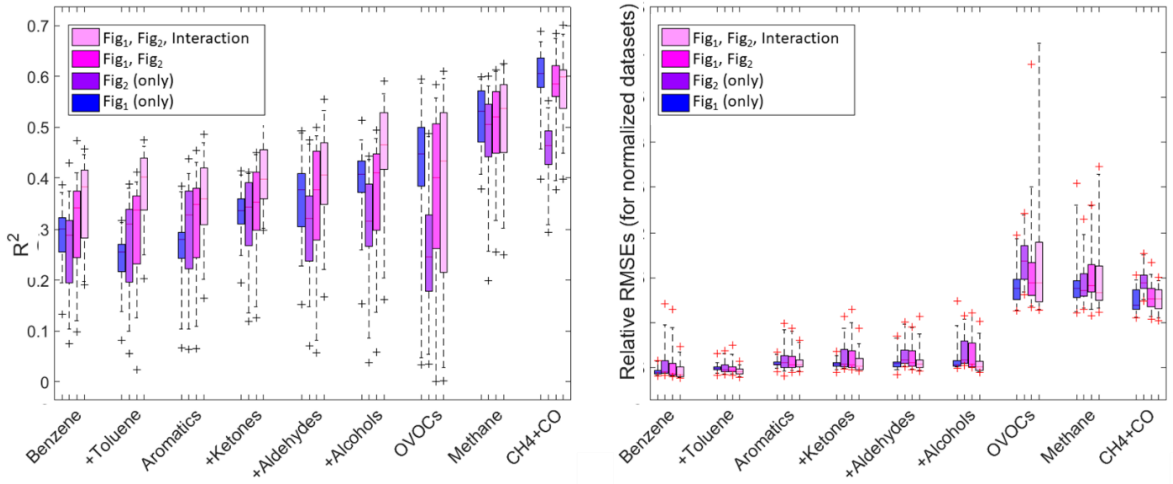


Figure 8: Bootstrap analysis for sensor set in secondary U-Pod

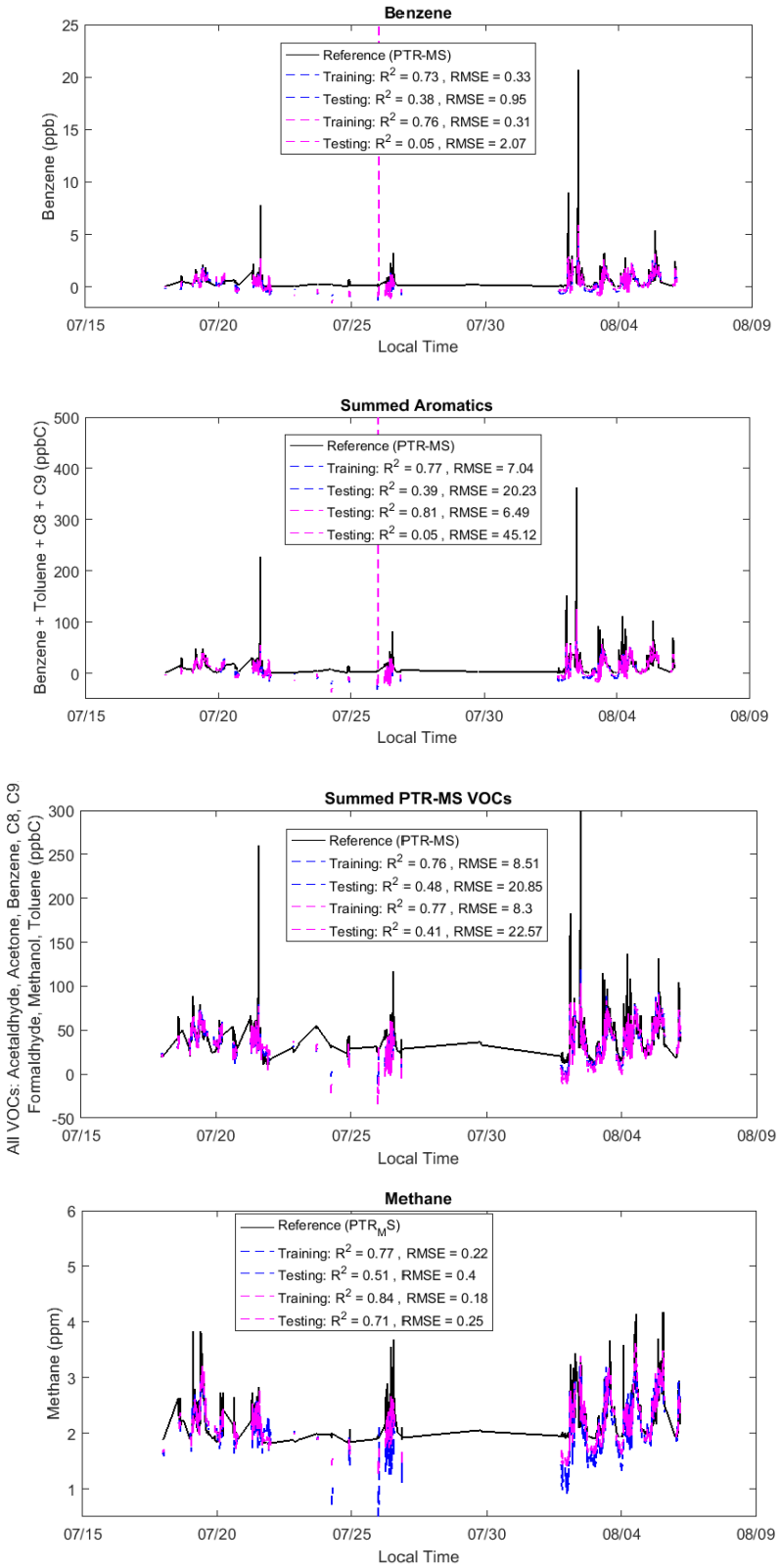
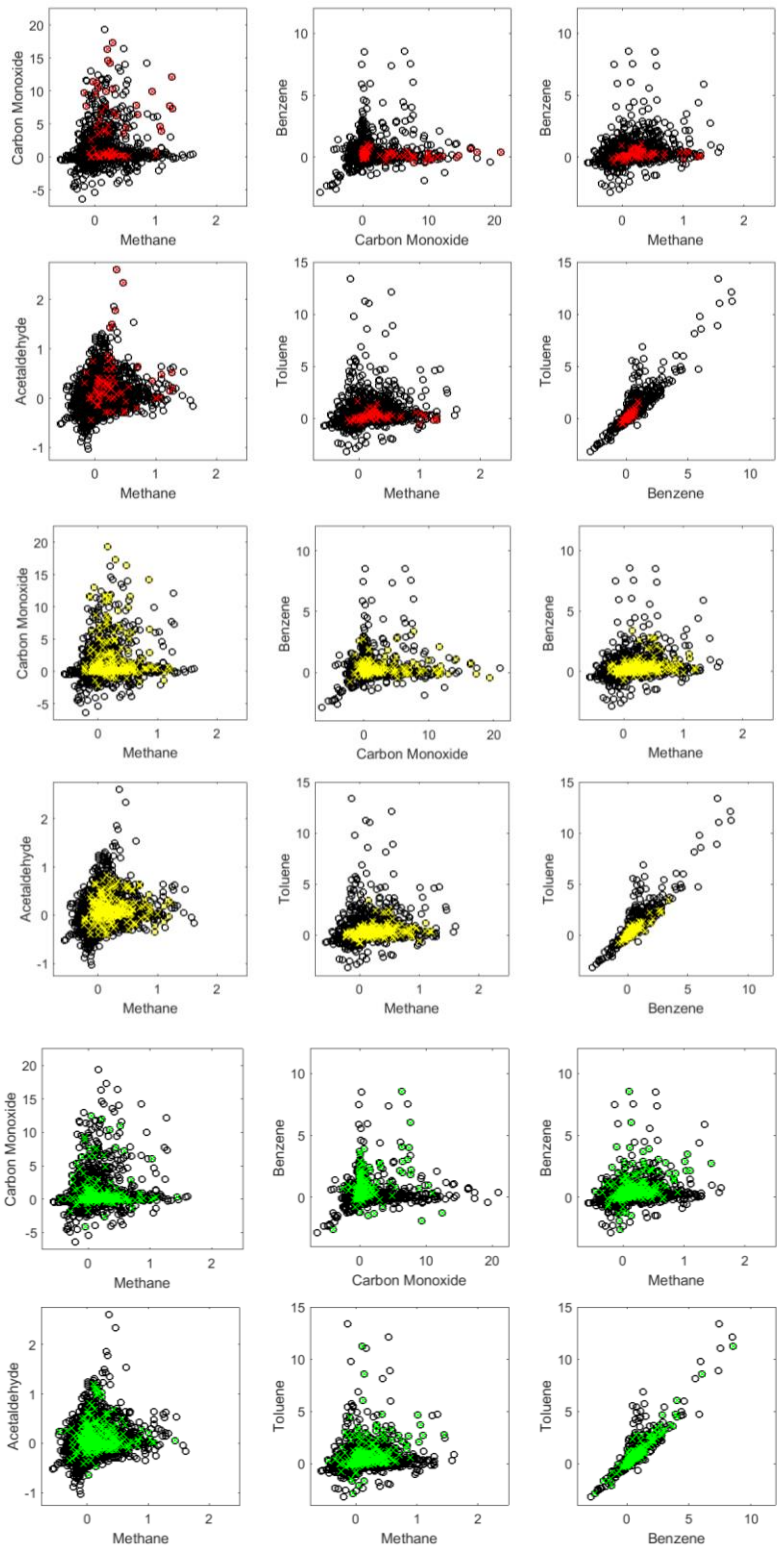
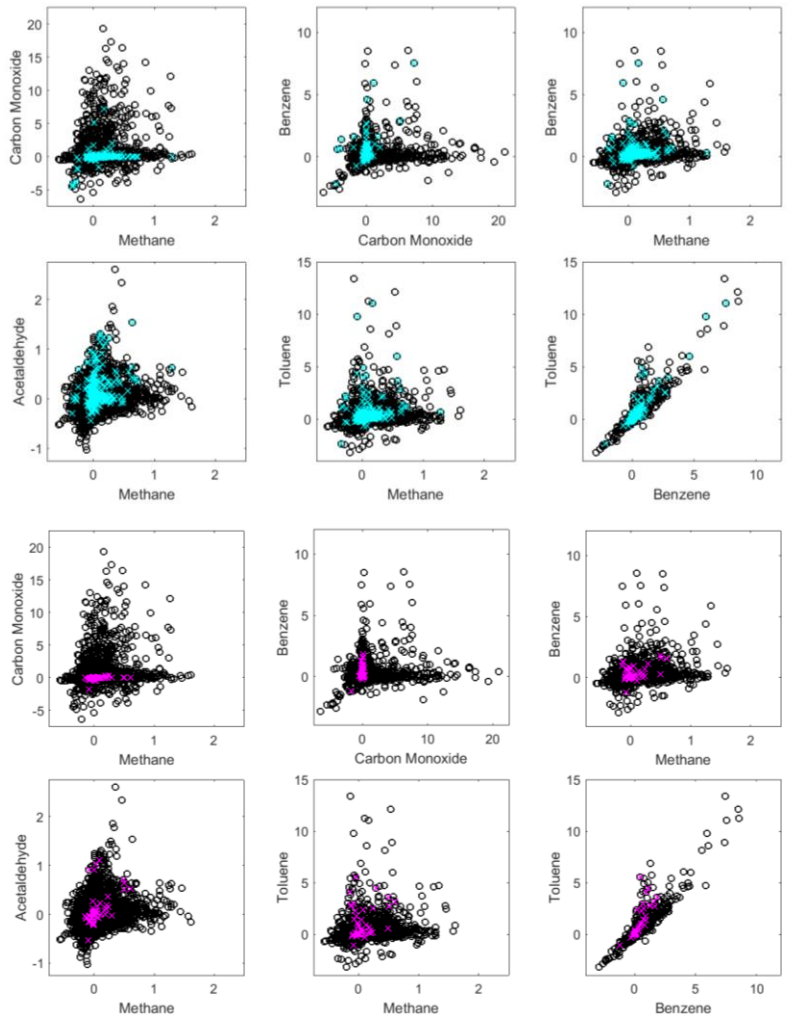


Figure 9: Regression analysis results for sensor set in secondary U-Pod

(training data before 7/20 & after 8/5, remaining data is testing

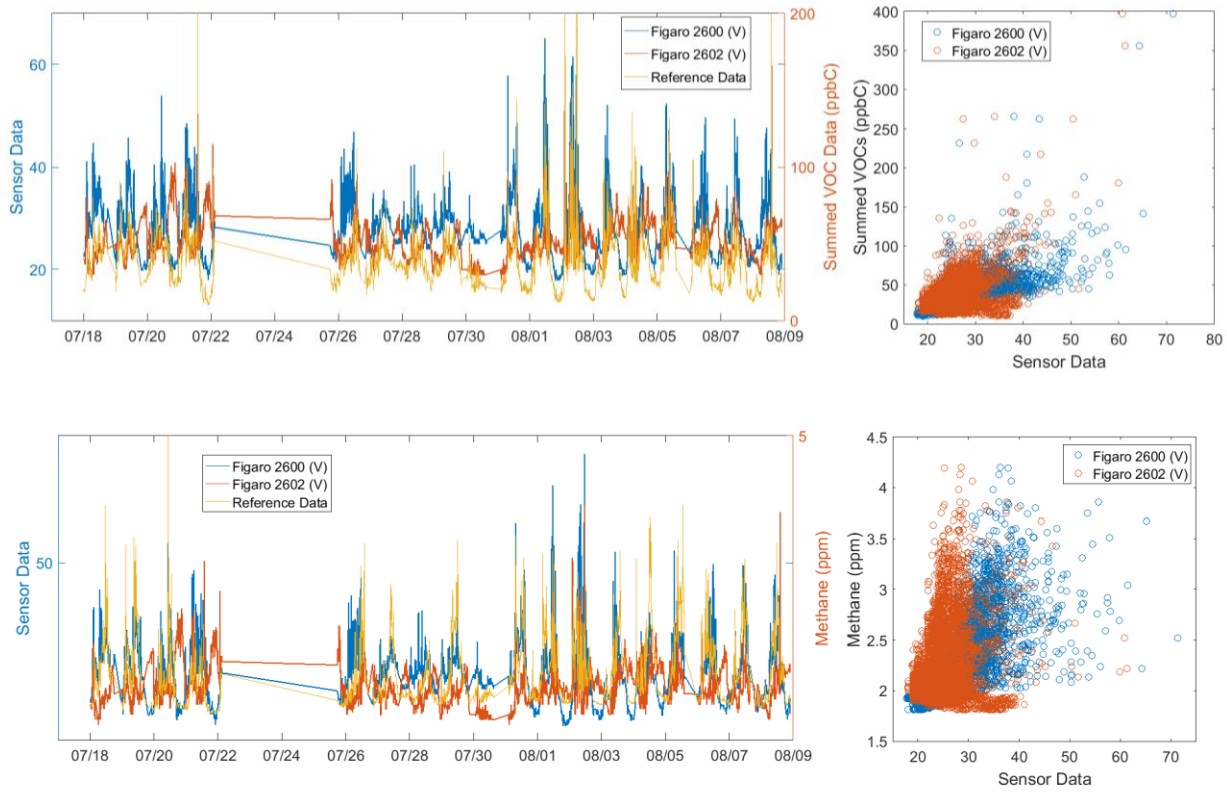

Carbon Monoxide
Methane
Benzene
Carbon Monoxide
Benzene
Methane
Acetaldehyde
Methane
Toluene
Methane
Toluene
Benzene
Carbon Monoxide
Methane
Benzene
Carbon Monoxide
Benzene
Methane
Acetaldehyde
Methane
Toluene
Methane
Toluene
Benzene
Carbon Monoxide
Methane
Benzene
Carbon Monoxide
Benzene
Methane
Acetaldehyde
Methane
Toluene
Methane
Toluene
Benzene



Figures 10: Complete sensor ratio plots

Fig1/Fig2 Ratio
(R/R0, w/
baseline removed)

- 1.75 < r
- 1.15 < r < 1.75
- 0.65 < r < 0.95
- 0.40 < r < 0.65
- 0.00 < r < 0.40



Figures 11: Raw sensor data versus reference data, top – summed non-methane VOCs, bottom – methane

Appendix VII

SUPPLEMENTAL FOR CHAPTER 8

Table 1: Methane Statistics – both methods

		Single-Sensor: CH_4						Multi-Sensor: CH_4					
		Training			Testing			Training			Testing		
		R^2	RMSE (ppm)	MB	R^2	RMSE (ppm)	MB	R^2	RMSE (ppb)	MB	R^2	RMSE (ppb)	MB
C1		0.821	0.149	0.002	0.755	0.161	0.063	0.881	0.11	0.000	0.802	0.163	0.086
C2		0.852	0.132	0.001	0.733	0.181	0.021	0.882	0.109	0.000	0.8	0.163	0.084
E1		0.834	0.141	0.002	0.719	0.183	0.023	0.882	0.109	0.000	0.802	0.152	0.062
E2		0.804	0.156	0.001	0.761	0.191	-0.01	0.879	0.11	0.000	0.806	0.152	0.067
N1		0.805	0.155	0.001	0.71	0.194	0.007	0.876	0.112	0.000	0.809	0.154	0.07
N2		0.794	0.16	0.001	0.764	0.156	0.061	0.878	0.111	0.000	0.811	0.149	0.061
N3		0.803	0.156	0.001	0.728	0.183	0.018	0.88	0.11	0.000	0.807	0.148	0.059
N4		0.82	0.147	0.001	0.753	0.176	0.019	0.879	0.11	0.000	0.808	0.151	0.062
N5		0.804	0.156	0.002	0.694	0.196	0.007	0.881	0.109	0.000	0.799	0.155	0.067
N6		0.831	0.143	0.002	0.712	0.19	0.011	0.885	0.108	0.000	0.79	0.167	0.088
R1		0.796	0.176	0.001	0.754	0.163	0.053	0.879	0.121	0.000	0.797	0.177	0.108
R2		0.789	0.163	0.001	0.739	0.172	0.036	0.883	0.108	0.000	0.8	0.158	0.073
R3		0.8	0.157	0.001	0.757	0.169	0.033	0.873	0.113	0.000	0.8	0.16	0.077
R4		0.818	0.149	0.001	0.762	0.171	0.025	0.883	0.109	0.000	0.807	0.154	0.071
R5		0.806	0.155	0.001	0.713	0.192	0.008	0.878	0.111	0.000	0.792	0.161	0.076
Ave.		0.812	0.153	0.001	0.737	0.179	0.025	0.880	0.111	0.000	0.802	0.158	0.074

Table 2: Total Non-Methane Hydrocarbons

			Multi-Sensor: TNMHC				
Y-Pod		Training			Testing		
		R^2	RMSE (ppb)	MB	R^2	RMSE (ppb)	MB (ppb)
C1		0.62	30.9	0.00	0.48	46.4	9.3
C2		0.6	31.6	0.00	0.46	47.6	11.6
E1		0.59	32.1	0.00	0.45	48	10.5
E2		0.6	31.8	0.00	0.46	48.2	12.8
N1		0.59	32.1	0.00	0.47	47.9	13.7
N2		0.59	31.9	0.00	0.47	47.3	11.6
N3		0.58	32.6	0.00	0.43	47.6	4.6
N4		0.6	31.4	0.00	0.46	47.4	8.1
N5		0.62	30.8	0.00	0.48	46.6	9.7
N6		0.57	33	0.00	0.43	48.6	10.8
R1		0.58	34.7	0.00	0.48	46.1	5.4

R2		0.6	31.5	0.00	0.47	46.2	6.5
R3		0.61	19.5	0.00	0.4	32.7	-17.8
R4		0.59	32.1	0.00	0.45	47.3	6.9
R5		0.61	31.4	0.00	0.47	47.2	10.6
Ave.		0.597	31.160	0.000	0.457	46.340	7.620

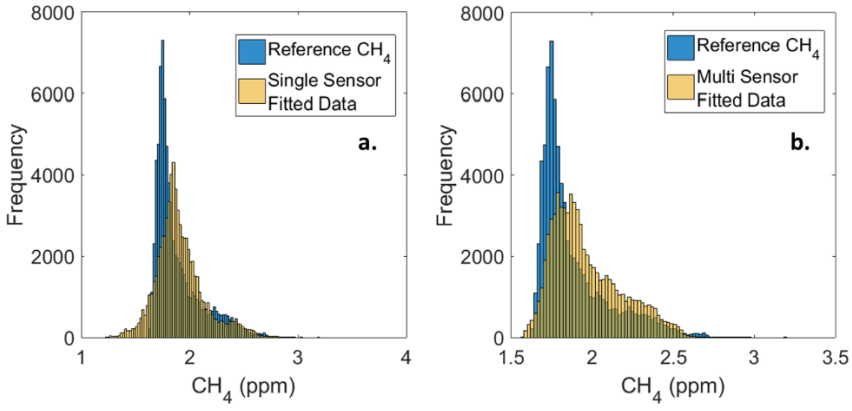


Figure 1: Histogram illustrating the accuracy of the distribution from the single-sensor method versus the multi sensor method

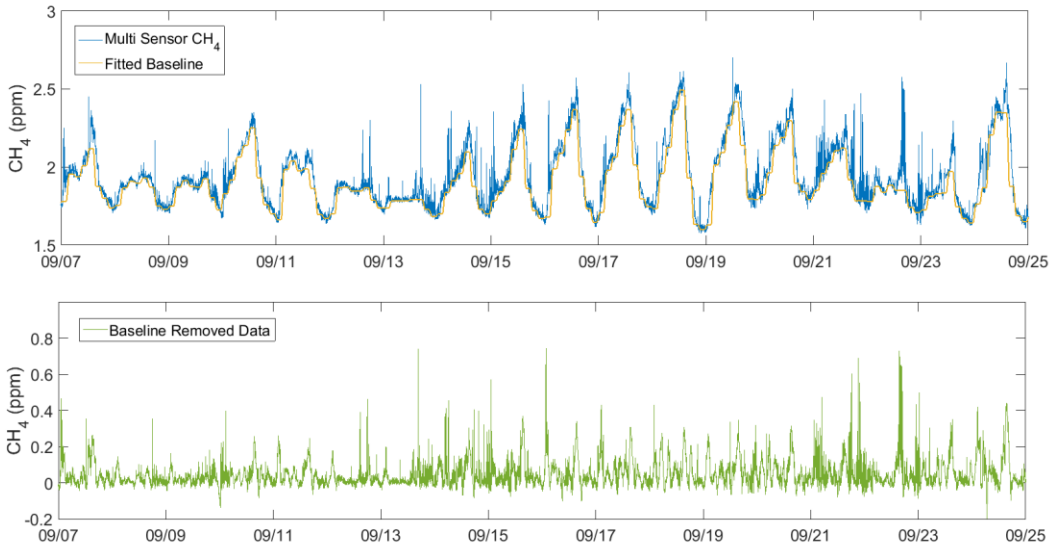


Figure 2: Baseline Extraction Methodology

Following the method utilized in:

Heimann, I., Bright, V. B., McLeod, M. W., Mead, M. I., Popoola, O. A. M., Stewart, G. B., & Jones, R. L. (2015). Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors. Atmospheric Environment, 113, 10–19. http://doi.org/10.1016/j.atmosenv.2015.04.057

The steps include:

1. Breaking complete data into equal sized pieces (3 hours long), 3 hours found to be appropriate for removing diurnal variability but not short-term enhancements (i.e., local emissions)
2. Binning the distributions using discrete concentration intervals
3. Then bi was determined to be the 25th percentile value of the applied smooth kernel fit
4. B(t) determined though interpolation between the baseline values (*Heiman et al.*,)

$$\boldsymbol{S(t) = B(t) + L(t)}$$

Carbon Dioxide Sensor Quantification Details

The data used here was from a co-location at a SCAQMD reference site, where a Licor CO2 analyzer was added and maintained by our team. A model utilizing, concentration, temperature, humidity, and time as predictor to solve for raw sensor voltage was utilized. Below are plots to illustrate the performance as well as the results statistics (in ppm) from four Y-Pods.

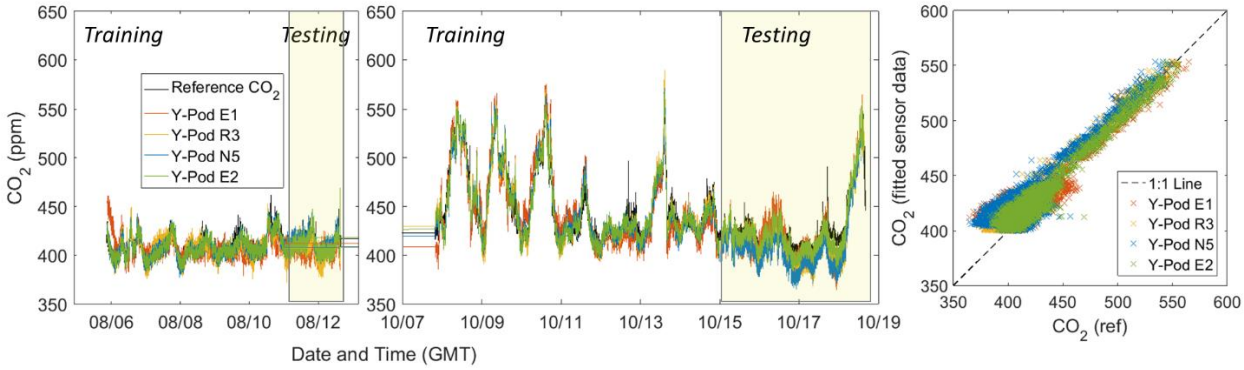


Figure 3: Carbon Dioxide Sensor Quantification Details

Table 3: Carbon Dioxide Sensor Quantification Results

	Training			Testing		
	R^2	RMSE	MB	R^2	RMSE	MB
E1	0.876	13.220	0.136	0.835	13.286	-6.864
R3	0.933	9.377	0.247	0.876	14.378	-10.787
N5	0.953	7.739	0.072	0.729	18.382	-12.020
E2	0.958	7.317	0.050	0.789	13.130	-4.302
Average	0.930	9.413	0.126	0.807	14.794	-8.494

Carbon Monoxide Sensor Quantification Details

The data used here was from a co-location at a SCAQMD reference site that included an FRM/FEM carbon monoxide monitor. A model utilizing, concentration, temperature, humidity, and time as predictor to solve for raw sensor voltage was utilized. Below are plots to illustrate the performance as well as the results statistics from one U-Pod. This U-Pod was co-located at the E2 site during the field deployment and was utilized for it's capacity to run the added CO electrochemical sensor.

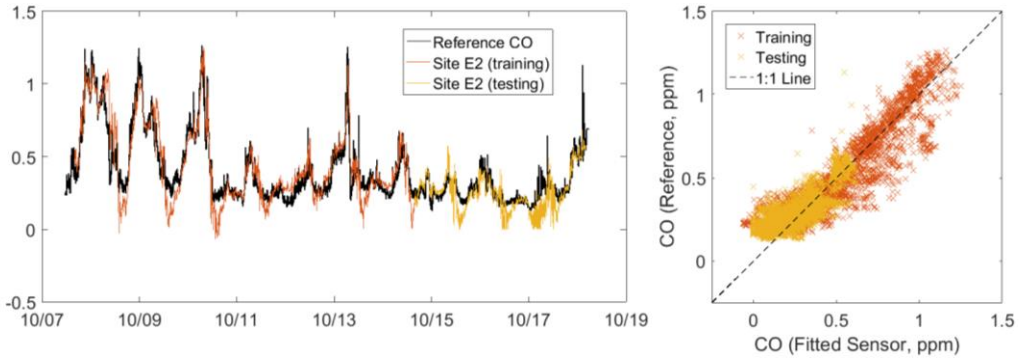


Figure 4: Carbon Monoxide Sensor Quantification Details

Statistics (R^2, RMSE in ppm, Mean Bias in ppm)

- Training: 0.841, 0.11, 0.002
- Testing: 0.603, 0.09, -0.036

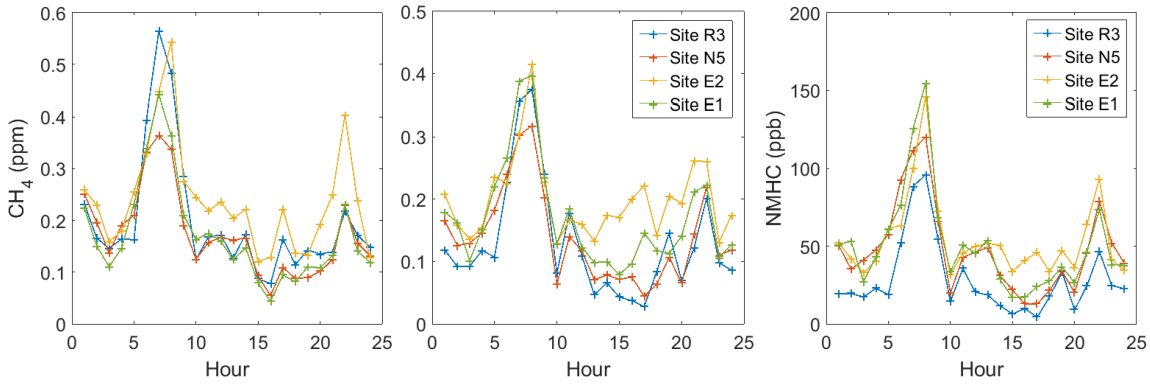


Figure 5: 95th percentile plots (from hourly grouped data shown in paper)

Figure 6: Box plots of grouped hourly data for all sites