THESIS

AUTONOMOUS LOW-COST NETWORK OF OZONE SENSORS: TO STUDY THE SPATIAL DISTRIBUTION AND EXPOSURE IN UNDERSERVED AGRICULTURAL COMMUNITIES IN CENTRAL CALIFORNIA

Submitted by

Akshay Kumar Gunniah Vijayakumar

Department of Mechanical Engineering

In partial fulfillment of requirements

For the Degree Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2025

Master's Committee:

Advisor: Shantanu Jathar

Christian L'Orange
Sheryl Magzamen
Casey Quinn
Ellison Carter

ABSTRACT

AUTONOMOUS LOW-COST NETWORK OF OZONE SENSORS: TO STUDY THE SPATIAL DISTRIBUTION AND EXPOSURE IN UNDERSERVED AGRICULTURAL COMMUNITIES IN CENTRAL CALIFORNIA

Ozone (O_3), a criteria air pollutant, is often overlooked in rural and remote regions, leaving the spatial distribution and exposure levels poorly understood, particularly in underserved communities. In this study, we developed and deployed a network of 12 autonomous, low-cost, and solar-powered air quality monitoring units (VOZboxes) in California's San Joaquin Valley (SJV). Co-located with a reference monitor in Fresno, CA, the VOZboxes underwent calibration before and after field deployment in June and November 2023, respectively, to measure O_3 concentrations over a dynamic range of 20 to 100 ppbv with an RMSE < 5 ppbv. Deployed across 11 unique locations in SJV from July to November 2023 at varying periods, the VOZboxes revealed modest spatial variability in O_3 mixing ratios, with elevated concentrations recorded in bigger cities and smaller eastern townships, while lower concentrations were found in smaller western regions. By leveraging multivariate regression models for data calibration, the VOZboxes effectively assessed compliance with the national ambient air quality standard (NAAQS) for O_3 (maximum daily 8-hour average of 70 ppbv) across locations. This study underscores the potential of low-cost environmental sensors for characterizing air quality and O_3 exposure in rural and remote environments. Additionally, it emphasizes their utility as tools for

addressing the monitoring needs of underserved communities and acts as a tool for environmental justice.

ACKNOWLEDGEMENTS

I want to convey my heartfelt thanks and gratitude to various individuals who have helped me strive and achieve in completing my thesis.

I want to express my gratitude to my advisor, Dr. Shantanu Jathar, who has been a valuable professional and personal resource for me. He has given me invaluable advice on this project and has helped me stay on track and focused. He has been an excellent source of support for me throughout difficult times, and he is always available to answer my questions. Thank you so much for your continued support and belief in me.

I would like to thank Dr. Casey Quinn for being a tremendous resource and support for me during this project, as well as a very competent technical guide. I'd like to thank my committee member, Dr. Christian L'Orange for his assistance in providing lab support and direction in utilizing the appropriate apparatus and helping me learn about the equipment.

Dr. Sheryl Magzamen has been a valuable mentor to me in terms of community outreach and understanding the effects of O_3 pollution in those places. Dr.Ellison Carter has been an excellent mentor in assisting me with the statistical aspects of the research, as well as an excellent instructor in teaching me a variety of new methodologies.

I'd like to thank my friends, lab mates, co-workers, and family for their constant support, aid, and belief in me to pursue my aspirations; without them, it wouldn't have been possible.

DEDICATION

To my mother, Devisri, and my sister Sushmitha,

For their unwavering love, sacrifices, and constant encouragement. Mom, your strength and compassion inspires me every day.

And in loving memory of my father, Vijayakumar,

Whose guidance, wisdom, and enduring spirit remains a source of strength in my life. Though you are no longer here, your legacy lives on in every step I take forward.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1 Ozone -A Global Concern for Air Quality

In recent times, there has been an increasing global concern regarding air quality and the impact of tropospheric ozone on it. The effect is most pronounced in low- and medium-income countries (LMICs) with substantial populations (Rentschler & Leonova, 2023). The spatial correlation in urban areas is studied well, whereas the effects on rural areas remain uncertain (Strosnider et al., 2017).

Ozone (O_3) is a molecule produced through the reaction of a free oxygen atom (O) with molecular oxygen (O_2) and is present in both the stratosphere and troposphere. Tropospheric ozone is a reactive gas, and a criteria pollutant formed by the interaction of sunlight with Volatile Organic Compounds (VOCs) and Nitrogen oxides (NOx) emitted from various sources such as vehicle exhaust and industrial emissions (Finlayson-Pitts & Pitts, 1997) and can be emitted from biogenic sources as well. O_3 is a key indicator of poor air quality and poses significant risks to human health and the environment (Donzelli & Suarez-Varela, 2024).

Acute exposure to elevated levels of O_3 leads to negative health effects, especially on the respiratory system (Jerrett et al., 2009). Inhaled ozone is a highly reactive gas that causes irritation and inflammation of the respiratory tract. Short-term exposure may provoke coughing, throat irritation, chest tightness, and difficulty breathing (Filippidou & Koukouliata, 2011; Keller, 1992). Acute ozone exposure may aggravate asthma, Chronic Pulmonary Obstructive Disease (COPD), and other respiratory ailments, leading to heightened medical drug requirements, emergency department visits, and, in severe instances, hospitalization (Malig et al.,

2016). Elevated ozone concentrations can negatively impact even healthy persons, especially during daylight hours, reducing lung function and outdoor physical activity (Kim et al., 2011).

Acute ozone exposure triggers systemic inflammatory responses. This may provoke oxidative stress, linked to cardiovascular impacts, acute effects are as thoroughly understood as chronic ones. Increased exposure or reduced physiological resilience jeopardizes children, the elderly, and outdoor workers. Children are more vulnerable because of their developing lungs and heightened outdoor activities, where ozone levels are elevated (Ito et al., 2005). Moreover, ground-level O_3 pollution can damage vegetation and ecosystems (Emberson, 2020). The health implications highlight the need for extensive ozone monitoring and public health notifications to reduce exposure during high ozone occurrences, especially in areas with poor air quality or during pollution incidents.

Conventional methods for monitoring ozone levels typically involve expensive instrumentation deployed in a limited number of fixed locations, often failing to capture the spatial and temporal variability of ozone concentrations in both urban and rural environments (Schultz et al., 2017). Furthermore, the high cost of such systems poses a barrier to their widespread deployment, particularly in resource-constrained regions where an air quality monitoring infrastructure is lacking (Cheadle et al., 2017; Miskell et al., 2019).

1.2 Need for ozone monitoring and effects

Human health concerns and environmental justice primarily drive the need for O_3 measurements in less monitored areas. Air quality management often overlooks rural and economically disadvantaged areas, leaving vulnerable populations disproportionately exposed to unknown O_3 concentrations. Rural areas and LMICs significantly contribute to global ozone dynamics due to

their significant emissions, including biomass burning, agriculture, and unique atmospheric chemistry (Bourgeois et al., 2021; Levine et al., 1995). Adequate monitoring is key in these areas to understand the trends and complex patterns found in these locations, and also to analyze the extent of ozone pollution in the field of agriculture, as ozone is a powerful oxidation agent capable of reducing the yield. The effect of ozone on agriculture affects the economic stability in the LMICs, agriculture being the backbone of the economy in these countries (Dewan et al., 2024). Monitoring and knowing the concentrations of ozone can help understand the impacts of ozone pollution on the food and crop industry and measures can be taken to lower them.

Ozone is a short-lived, potent greenhouse gas that absorbs the thermal infrared radiations emitted by the earth's surface thus increasing the atmosphere temperature and thus increasing the surface temperature (Ledley et al., 1999). Ozone is the third most potent greenhouse gas after CO_2 and CH_4. Ozone can have its say in the overall radiative forcing in both regional and global atmospheres (Naik et al., 2005). Longer exposure to ozone can accelerate the degradation of materials such as rubber and fabrics (Cho et al., 2011).

Additionally, ozone negatively impacts forests and vegetation, which in turn reduces biodiversity and alters the ecosystem. Chronic exposure to higher concentrations of ozone impairs the reproduction capabilities of plant species, negatively impacting the quality of wildlife (Weigel et al., 2015). By affecting the health of the vegetation, they influence the hydrological cycles by altering the evapotranspiration rates (Sun et al., 2012).

Good spatial coverage for ozone monitoring is necessary for several reasons, including the high degree of variation in ozone formation, which is influenced by local precursors, geographical location, and ambient conditions. Therefore, deploying monitors throughout a city can aid in

capturing the variations inside a densely populated location. Understanding hotspots, vulnerable locations, and a comprehensive set of data for each location is crucial for effective air quality management. Enhancing the spatial coverage allows for the execution of model validations by comparing against the value measured on-site for new locations, which was previously predicted, thus enhancing satellite-based observations (Di et al., 2017). This also paves the way for conducting health risk assessments, evaluating population exposure, and providing information on daily exposure. Finally, understanding the spatial coverage enables us to track the transport of ozone between regions, thereby facilitating the study of complex mechanisms (Sadighi et al., 2018).

In the United States, nearly 83% of the population resides in urban areas, where dense air quality monitoring networks are in place due to high population density, economic activities, and hotspots. However, the remaining 17% of the population, residing in 97% of the country's landmass, lack adequate monitoring, leading to significant spatial gaps as densely populated areas could be monitored easily but sparsely populated areas need a lot of resources due to a lot of resources. These locations, exposed to unknown concentrations, remain unaccounted for in exposure assessments. Globally, 43% of the population resides in rural areas, often receiving insufficient monitoring. By understanding their exposures, we can effectively implement regulations and control measures (Miranda et al., 2011; Steinle et al., 2013).

1.3 Involvement of low-cost sensors

Conventional monitoring techniques use regulatory-grade ozone monitors, such as the Thermo Scientific 49i, and models from 2B Technologies that comply with U.S. EPA standards. They often range in price from $5,000 to $25,000 per unit. Their yearly operational expenses vary

between $5,000 and $10,000 per device, comprising calibration, quality control evaluations, routine maintenance (such as filter replacement and optical component cleaning), utility costs, and personnel expenses for equipment management. These monitors exhibit exceptional accuracy and are predominantly utilized for regulatory compliance, research-grade investigations, and long-term monitoring networks. Nonetheless, their higher expense limits their implementation to urban locations and regions with considerably more pollution issues, resulting in many rural and resource-constrained environments being inadequately monitored.

To address these challenges, there is a growing interest in developing low-cost, low-power ozone sensors and integrated devices that are of smaller weight and size and are capable of providing real-time, high-resolution data on ambient ozone levels across rural and urban areas. These new ozone sensors and a network of sensors leverage advances in sensor technology, data analytics, machine learning, and wireless communication to create distributed monitoring systems that offer better spatial coverage and accessibility at a fraction of the cost of traditional monitoring approaches (Snyder et al., 2013).

Regulatory monitoring, supplementary monitoring using low-cost sensors, and satellite data can be combined to provide a comprehensive approach to estimating the spatiotemporal distribution of ozone (O_3) (Di et al., 2017). Although model-based or satellite-derived ozone values are useful, they frequently predict values that are far from the relative physical measurements, highlighting the necessity of integrating these datasets with direct, on-site measurements (de Nazelle & Serre, 2006).

Low-cost sensors are essential for addressing spatial gaps in regulatory networks, especially in inadequately monitored areas, and they improve satellite observations by providing localized,

high-resolution data. The strategy of using both in-situ data and satellite data leverages the ozone level assessment by using the strengths of each approach to address their respective limitations and provide a robust framework for understanding and addressing ozone pollution (Miskell et al., 2019; Weissert et al., 2020).

Their portability facilitates high spatial and temporal resolution, enabling real-time applications such as exposure mapping and community participation. Moreover, these low-cost sensors have facilitated public engagement via people science projects, enhancing awareness and enabling communities to advocate policy reforms. Nonetheless, constraints persist, encompassing data accuracy challenges such as biases, drift, and restricted detection capabilities compared to reference-grade instruments.

Environmental variables such as temperature, humidity, and cross-sensitivities to other gases increase the challenges of calibration and data interpretation. Numerous sensors encounter calibration issues, as precise models customized for unique deployment conditions are crucial for reliable measurements. Durability is a challenge, as several sensors lack the durability required for prolonged deployments in severe conditions (Nalakurthi et al., 2024).

Opportunities for enhancement encompass the utilization of machine learning calibration models that integrate environmental covariates, the establishment of standardized QA/QC (Quality Assurance/ Quality Control) protocols for deployment and data processing, the amalgamation of sensor networks with satellite data to augment spatial coverage, and the advancement of sensor design incorporating features such as active cooling mechanisms and enhanced radiation shields to mitigate environmental interferences. By tackling these issues, affordable ozone sensors can significantly enhance air quality monitoring and accessibility.

Literature suggests that low-cost ozone sensors face several challenges, including calibration issues and performance deterioration over time. Some examples are, electrochemical sensors like Sensoric O_3 3E 1 showed good initial performance but experienced a decline in accuracy after nine months. Semiconductor sensors such as SM50 OZU demonstrated high linear relationships but required regular recalibration. Environmental factors like temperature and humidity can affect sensor accuracy, necessitating the inclusion of these variables in calibration models (Badura et al., 2022). Despite these challenges, low-cost sensors can effectively capture spatial variations in ozone concentrations at neighborhood scales (Cheadle et al., 2017) and meet EU (European Union) requirements for indicative measurements (Badura et al., 2022). When properly calibrated and operated, some low-cost sensors can provide consistent measurements comparable to reference monitors (Duvall et al., 2016).

1.4 Ozone Pollution in Central Valley, California

One of the most ozone-polluted regions in the United States is Central California, particularly the San Joaquin Valley (SJV). Its climate, geography, and emission sources create a unique and challenging environment for air quality management (Pusede et al., 2014). Despite years of regulatory efforts and decades of environmental justice efforts, this region frequently exceeds the air quality standards, leading to problems for both public health and the ecosystem (White, 2020). The Fresno-Madera-Hanford region ranks in the top 5 metropolitan areas for the most polluted cities (*Fresno-Madera-Hanford, CA*, n.d.).

Locality and occupation both influence the sources of emissions. Urban areas like Fresno, Bakersfield, and Modesto primarily contribute to ozone precursors through transportation, industrial operations, and residential activities. They emit significant ozone precursors, including

nitrogen oxides (NOx) and volatile oxide compounds (VOCs). The SJV is a hub for agriculture, and as such, farming operations play a major role in contributing to ozone precursors through fertilizer applications, pesticide usage, and the use of agricultural equipment. Additionally, farm workers burn a significant amount of biomass, a major contributor (Kumar et al., 2011; Pfannerstill et al., 2023).

Extensive oil and gas activities in this region result in the release of numerous VOCs and NOx gases, which can further contribute to the formation of ozone. Wildfires, which are becoming more frequent in this region, also impact ozone concentrations by releasing volatile organic compounds (VOCs), which can also lead to secondary ozone formation.

The next significant factor contributing to the formation of ozone is the combination of meteorological conditions and geographical location (Yates et al., 2020). The Sierra Nevada Mountains and the Coastal Ranges surround the bowl-shaped SJV. This geography is capable of trapping the ozone and has very limited dispersion. The temperatures are warmer, and there is abundant sunlight available, which increases the photochemical reactions and creates more ozone. Inversion which occurs during the summer worsens ozone pollution by trapping ozone close to the surface.

The transportation phenomenon is a key factor when considering ozone pollution, as both the ozone and its precursor transportation occur in parallel. The primary sources of ozone precursors in the region include the emissions from San Fransico Bay Area, and Sacramento Valley and also consist of the local sources in the valley. Bay Area contributes up to 50% of the Volatile Organic Compounds (VOCs) and Nitrogen Oxide emissions. Also, Industrial point sources release pollutants above the mixed layer and this allows for long-range transport of intact plumes into

the downwind regions. The transportation occurs through well-defined wind patterns, where the sea breeze through the Carquinez Strait and Pacheco Pass allows the movement of the Bay area emissions into the SJV. Long-range transports which are induced by the nocturnal jets of air, transport air southward and lead to high ozone concentrations in the early morning.

In recent years, Central California has recorded the highest number of non-compliance days in the country. The imposed air quality regulations have shown progress in lowering ozone concentrations, but this progress has been slower than in other regions due to the complex dynamics involved in its formation.

1.5 Aim, Motivation and Objectives

This study will design, implement, and evaluate a network of low-cost, low-power ozone sensors-named VOZboxes-that can be used to increase the resolution of ozone-monitoring networks. Particular emphasis will be placed on determining the accuracy, reliability, and long-term stability of the sensors in field experiments by comparative analyses with reference-grade monitors. Capturing the spatial and temporal variability of ozone concentrations across deployed locations is a key deliverable.

This research will help decision-makers, rural planners, and public health officials advance the cause of air quality monitoring by demonstrating the efficacy of low-cost ozone sensor networks. Accomplished through the insights gained, this will allow such networks to be more widely implemented in both urban and rural areas. Understanding the limitations of low-cost ozone sensors is crucial for their effective application in supplementing the existing monitoring infrastructure.

The motivation of this research is to establish the need to measure ground-level ozone concentrations in rural and remote areas, which are disproportionately affected. The health impacts in terms of respiratory and cardiovascular health from long-term exposure to levels above regulatory limits may culminate in premature deaths. These regions' residents usually do not have air quality data to support claims of health risks due to pollution. This is of particular concern to agricultural communities who spend extensive hours outdoors when ozone concentrations are typically at their peak.

This study represents the critical need to analyze urban-rural trends and ozone concentrations across the San Joaquin Valley. In order to capture the gaps in the current monitoring network, which essentially captures urban trends, VOZboxes have been strategically placed. These sensors are low-cost, giving values in very remote areas, enabling a better understanding of the spatial distribution and variability of ozone.

In line with the objective, the following are the specific objectives for this study:

Development and Calibration: Design and build the VOZbox fleet, developing a robust sensor calibration strategy using multivariate regression and machine learning models to optimize accuracy.

Deployment and Data Collection: Deploy sensors across multiple locations within the Central Valley, collecting data in a systematic manner with regard to comprehensive spatial coverage.

Post-Deployment Analysis: Post-field deployment analysis, to quantify sensor drift over time, will be performed using established methodologies for calibration to correct and mitigate drift-related inaccuracies.

Performance Evaluation: Collected data from VOZboxes will be evaluated against reference-grade monitors at deployment locations.

Spatial and Temporal Analysis: Assess the variability of ozone concentrations across several sites in the Central Valley, showing how adding the VOZbox network increases spatial coverage and provides important information on trends in ozone pollution.

2. Materials and Methods

2.1 VOZBox Description

In this work, we built a fleet of autonomous (which can measure, transmit, and store information with a self-sustained power system), low-cost, and low-power units called VOZbox to measure ground-level O_3 concentrations. The VOZbox consists of three primary parts: a radiation shield (housing the O_3 sensor), a control box, and the solar panel, all attached to a tripod stand with metal brackets and clamps.

The VOZbox radiation shield is equipped with a primary metal oxide sensor (Aeroqual SM50) for O_3 detection and an environmental sensor (Adafruit SHT30) to measure the relative humidity and temperature of the ambient sampled air. The radiation shield is custom-designed and 3D printed using Polyethylene Terephthalate Glycol (PETG) with high durability for long field deployment periods. PETG is resistant to ultraviolet light and chemically inert thus reducing the interaction with the sampled O_3 and minimizing wall losses. The radiation shield is equipped with two fans that are responsible for sampling air and reducing the impact of temperature by cooling the SM50 and SHT30 and ensuring that the temperature within the radiation shield is kept within 1℃ of ambient temperature as shown in Figure S7.

To minimize chemical interactions with the sampled air, the air is channeled through a Teflon tube. PETG exhibits more durable properties, while Teflon offers enhanced chemical resistance. Teflon is challenging to use as a 3D print filament alone; therefore, we selected PETG for the 3D printing of the radiation shield and utilized a Teflon tube for sampling where great chemical resistance is essential.

The control box is made up of a weatherproof polycarbonate box (WQ-57; Polycase, OH), that consists of a 12V 9Ah battery (BLF-1209WS; Bioenno Tech, LLC, CA), charge controller (Renogy Wanderer), low-cost PM sensors (SEN55; Sensirion and PMS5003T; Plantower), temperature and humidity inside the box (SHT30; Adafruit Inc), a microcontroller unit (MCU) with cellular connectivity (BRN404X; Particle Boron LTE), custom Printed Circuit Board (PCB) shown in Figure S5, Adalogger (2922; Adafruit Inc) and a GPS (15210; Sparkfun). The MCU integrates and reads the sensor values and stores them locally in an SD card using the adalogger. The data is measured at a frequency of 1Hz and averaged across 1 minute before uploading and it is transmitted wirelessly through the mobile network as 10-minute averages using the Particle Boron IOT device to the Particle Cloud server. The work was more focused on O_3 so the sensors for PM were not analyzed and would be a future prospect for the VOZbox.

Table 2.1 Sensors in the VOZbox and their characteristics.

Sensor	**Cost**	**Measurements (units)**	**Characteristics**
Sensirion SEN55	$35	PM 1.0 (µg/m³, #/cm³), PM 2.5 (µg/m³, #/cm³), PM 4 (µg/m³, #/m³), PM 10 (µg/m³, #/cm³), VOC Index (#), NOx Index (#), Humidity(%), Temperature(℃).	Mass concentration precision (PM) – ± 10% Typ. relative humidity accuracy – 4.5 %RH Operating (RH) range – 0 - 90 %RH Response time (RH) (τ63%) – 20 s Response time (VOC) (τ63%) < 10s Device-to-device var. (VOC) – ±15 % m.v. Limit of detection (VOC) – < 0.05 ppm of ethanol equivalents or < 10 % of concentration setpoint, whichever is larger Operating temperature range [°C] – -10 - 50°C Response time (Temp) (τ63%) < 60s Typ. Temp accuracy – 0.45 °C

Plantower PMS5003T	$31	PM 1.0 (μg/m³, #/cm³), PM 2.5 (μg/m³, #/cm³), PM 10 (μg/m³, #/cm³), Temperature(℃).	Resolution – 1 μg/m³ Single Response Time＜1 s Total Response Time ≤10 s Maximum Consistency Error (PM2.5 standard data)* – ±10%@100~500μg/m³, ±10μg/m³@0~100μg/m³
Adafruit SHT30	$25	Temperature (℃), Humidity (%).	Accuracy tolerance – ±0.3 °C Resolution Typ. 0.05 °C Response time >2 s Accuracy tolerance Typ. ±3 %RH Resolution Typ. 0.05 %RH Response time – 8s
Aeroqual SM50	$435	Ozone(ppbv).	Limit of detection – 5 ppbv Resolution – 1 ppbv
Sparkfun GPS	$43	Latitude (deg), Longitude (deg), Altitude (m).	Heading Accuracy: 0.3 degrees 18Hz Max Update Rate Time-To-First-Fix: Cold: 26s Hot: 1s

Table 2.2 Simplified Bill of Materials for a single unit VOZbox

Components	**Manufacturer**	**Description**	**Price**
Fasteners	Multiple	Linking the components together	$15
Aluminum Components	8020	Mounts for the solar panels and control box to attach to the stand	$40
Enclosure and Stand	Polycase and Pyle	Waterproof enclosure, system stand	$95
Power System - Battery, Solar panel, Charge controller, wires	Bioenno, Wanderer, Renogy	Powers the system	$187
Electronics - PCB	Multiple	Integrating sensor outputs	$20

Electronics - Sensors, Data logger, wires	Multiple	Sensing elements	$665
Radiation Shield	Multiple	Housing for the ozone sensor	$75
		Total	**$1097**

Table 2.3 Power Calculations for the VOZbox

Battery Voltage: 12 V	Solar Panel rating: 30 W
Battery Capacity: 9 Ah	Solar Output (With direct sunlight) : 30 x 6 x 0.75 = 135 Wh (Passive charging not included)
Energy Storage Capacity : 12V x 9Ah = 108 Wh	Charging Time: 108/135 = 0.8 days (Actual charging is faster and closer to 2 hours)
Power consumed by device = 5 -10 W	Total run time on full battery = 108/5 = 21.6 hours

The power system consists of a solar panel (30 W), a charge controller, and a battery. The VOZbox power system is capable of running the system effectively day and night with the solar panel and battery all season indicated in Table 2.3. This enables the VOZbox to be located at any site without any dependencies, making it fully autonomous. The entire setup is mounted on a tripod stand that is >2 m above ground level for measuring ozone. The setup is secured using guy wires and counterweights. Table 2.2 shows a simplified bill of materials with all the components and the total cost for one VOZbox, Figure S1 gives a detailed bill of materials.

(a)  (b) 

Figure 2.1.1 a) Full standing image of the VOZbox on the rooftop of Powerhouse, CSU. b) Image of the internal components of the control box, which houses the battery, charge controller, PM sensors, GPS unit, Boron, and PCB.

2.2 Sensors and Measurement

The sensor (SM50; Aeroqual Ltd, NZ) is a metal oxide ozone sensing sensor that is capable of detecting in the range of 0 - 150 ppbv which is in the sufficient range to measure ozone concentrations that can cause harm to people even when exposed for 1 hour (Hubbell et al., 2005). This sensor showed a very high linear performance with r-squared values exceeding 0.94 for the electrochemical sensors with the reference monitor from a study conducted in Wroclaw, Poland by (Badura et al., 2022) it also indicates the need for regular recalibration every 6 months as there is the scaling of the sensor mesh thus leading to irregular measurements(*Sensors - Ozone*, 2022). The twelve sensors used for the fleets were purchased in the year 2022 for the cost of USD 435 per unit. SM50s are factory-calibrated, and their output is read as analog voltage signals (0 to 5 V) and used for reading the ozone concentration values. The data is then stored

locally in the SD card, which are 1-minute averages. The data is then averaged for 10 minutes and transmitted to the cloud via the Particle Boron LTE device. The data is stored in the CSU cloud services and a daily summary of the data is pulled into a public GitHub repository for easy access. The analysis was done in Python by pulling the data from the repository and running them locally. The finalized results and the spatial maps are published on a Tableau Public Dashboard shown in Figure S6 for easy access to data visualization for the outreach activities.

A fleet of 12 VOZbox units was used for this research study. All the sensors and the electronic components used are identical. The 12 Aeroqual SM50s were obtained from the same manufacturing batch. The fleet was pre-tested and calibrated at the CSU Energy Campus (Fort Collins, CO) before their deployment at the Central Valley. This calibration is then compared with the localized collocated calibration to understand the significance of the region-specific calibration approach.

2.3 CCEJN and CARB

The Central California Economic Justice Network (CCEJN), an environmental justice organization, is involved in this project as a partner to deploy a low-cost network of sensors to measure pollutants, specifically ozone, at rural and agricultural sites. The project aims to comprehend exposure and pollutant levels, empowering the community to make informed and strategic decisions. Their mission is "*to empower our communities and secure our children's future by eliminating negative environmental impacts in low-income communities and communities of color*". They were involved in everything from the development of the VOZbox to selecting the sites and deploying them, troubleshooting them whenever needed, and communicating with community members about the results obtained. The California Air

Research Board (CARB) provided funding for this project through AB 617 and also due to their interest in knowing about ozone exposure in agricultural communities, as they provided us the required support by providing the location for the collocation measurement phase and allowed us to utilize their reference-grade instrument measurements, which served as the base for the calibration that we built for the sensors; the pre- and post-field calibration were done at their Fresno Garland site.

2.4 Collocation Site and Calibration Study

A short study to analyze sensor performance was conducted on the CSU Powerhouse campus, and the Thermo Environmental Instruments Gas Analyzer (Model 49C, Franklin, MA; Thermo hereafter) was used as a calibrated reference for the VOZbox. This study was carried out for 3 days (May 29, 2023, to May 31, 2023) to test the sensors where the ozone values are considerably high, and the sensors showed a high linear relationship with an R-squared value higher than 0.9 as seen in FigureS10 that illustrates the data for 3 of the units. The VOZboxes (referring to many VOZbox units) were then packed and shipped to Fresno, California.

The measurement campaign in the San Joaquin Valley (SJV), California was conducted from June 2023 to November 2023. A pre-calibration was required before the deployment which was done at the Fresno-Garland site a CARB facility to ensure high reproducibility of the data in comparison to the reference monitor and amongst them as well. Ambient conditions are a major factor in determining the value of ozone measured, thus a geographically localized calibration is required.

The collocation approach is a method where the VOZbox sensors are placed adjacent to each other next to a reference-grade instrument and operated simultaneously and under the same

ambient conditions for a fixed period of at least 2 weeks to obtain optimal results. This way both datasets are comparable, and the reference data can be used as the base on which the calibration for the VOZbox can be built.

The calibration period for the VOZ Box fleet was divided into segments, pre-field and post-field deployment at the same reference site using the collocation approach. This strategy was adopted as the sensors could have sufficient high ozone exposure that the calibration model could capture for better performance of the model in the early ozone season and the post-field calibration allowed for the drift correction of the sensor and to capture ozone concentrations at the lower range. This co-location strategy is less labor-intensive than a deployment-specific calibration at the deployment site.

For the pre-calibration, all the 12 sensors of the fleet were placed adjacent to each other and along a reference grade FEM (Federal Equivalence Monitor) at the California Air Research Board (CARB) Fresno Garland site (36.78, -119.77). The VOZboxes collocation was set up in a way by having two sets of VOZboxes in a single tripod to ensure it was space efficient and also very close to the reference monitor to ensure maximum reproducibility (Figure 2.4.1b).

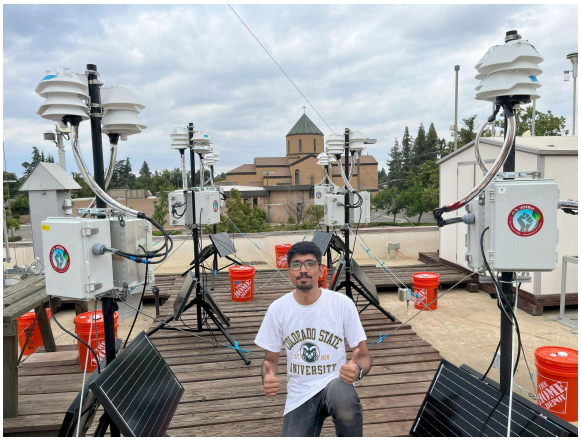
(a)  (b) 

Figure 2.4.1 a) Author along with the deployed sensors during the pre-field calibration phase b) Pre-field deployment at the Fresno-Garland site with two sensors paired together and firmly put in place with the buckets and the guy wires.

The sensors underwent the pre-field calibration for a period of 3 weeks from June 07, 2023, to June 21, 2023, in common with all the units, the data for the reference monitors were obtained from the California Air Research Board (CARB) AQMIS2 website (https://www.arb.ca.gov/aqmis2/aqdselect.php), the sensors were then moved to their designated deployment sites that are distributed throughout San Joaquin Valley.

To assess drift and to validate how the current calibration holds, all of the units were brought back to the Fresno Garland site for a post-field calibration after the deployment which captured the low ozone values after the ozone season—the post-calibration phase spanned from Nov 2, 2023, to Nov 20, 2023. The setup for the post-field calibration was similar to the pre-field calibration in terms of positioning the sensors.

2.5 Field Deployment

The deployment was done in partnership with the Central California Economic Justice Network (CCEJN, Fresno, California) to study the spatial patterns in the agricultural communities in the Central Valley and inform residents about ground-level ozone pollution where regulatory measurements are unavailable. The deployment was done at two different types of sites: the Type A site where the VOZbox is located near a FEM (Federal Equivalence Monitor) /FRM (Federal Reference Monitor) regulatory monitor and the other where there is no regulatory monitor in the proximity of the VOZbox. For type A, the identified locations were Fresno-Garland (Figure 2.5.2.b)) and Tranquility sites (Figure S8). This type of site enables continuous monitoring alongside the regulatory monitor to evaluate the performance of the sensor, measure compliance

with the NAAQS (National Ambient Air Quality Standard) standard, and analyze the drift of the sensor with high temporal resolution.

For the type B locations, our partner CCEJN was able to identify key deployment locations where the ozone pollution is less understood, these locations included sites with no federal regulatory monitors present in a 10-mile radius and where there is a decent population and some activities like oil and gas or high agricultural activity prone areas. Some of these sites are essential to understanding the urban vs rural trends.

The deployment for the type A site was done immediately after the pre-field calibration phase on the rooftop (~10m above ground) of the California Air Research Board site at the Fresno Garland site (VOZbox 1) and on the rooftop (~12 m above ground) of the Air Quality Management District (AQMD) site at the Tranquility site (VOZbox 12). The height from the ground is a key parameter as it reduces interference and the preferable location will be a rooftop for accurate measurement(Weissert et al., 2020).

The deployment of Type B sites was based on the availability of the people at the site and it was done in batches. The Type B VOZboxes have been installed in a variety of sites, including community center rooftops (~10 m above ground), community member house backyards (~1.75 m above ground), and agricultural fields (~1.75 m to 3 m above ground). All of the VOZboxes were either secured with guy wires or linked to an existing structural member.

The deployed sites are listed in Table 2.4 alongside their calibration and deployment dates. The sensors were brought back from their assigned location for the post-field calibration to the Fresno/Garland site which marked the end of the deployment research study. Sensor 12 was left

at the Tranquility site and the local FEM monitor reference was utilized for its post-field deployment calibration.

(a)  (b) 

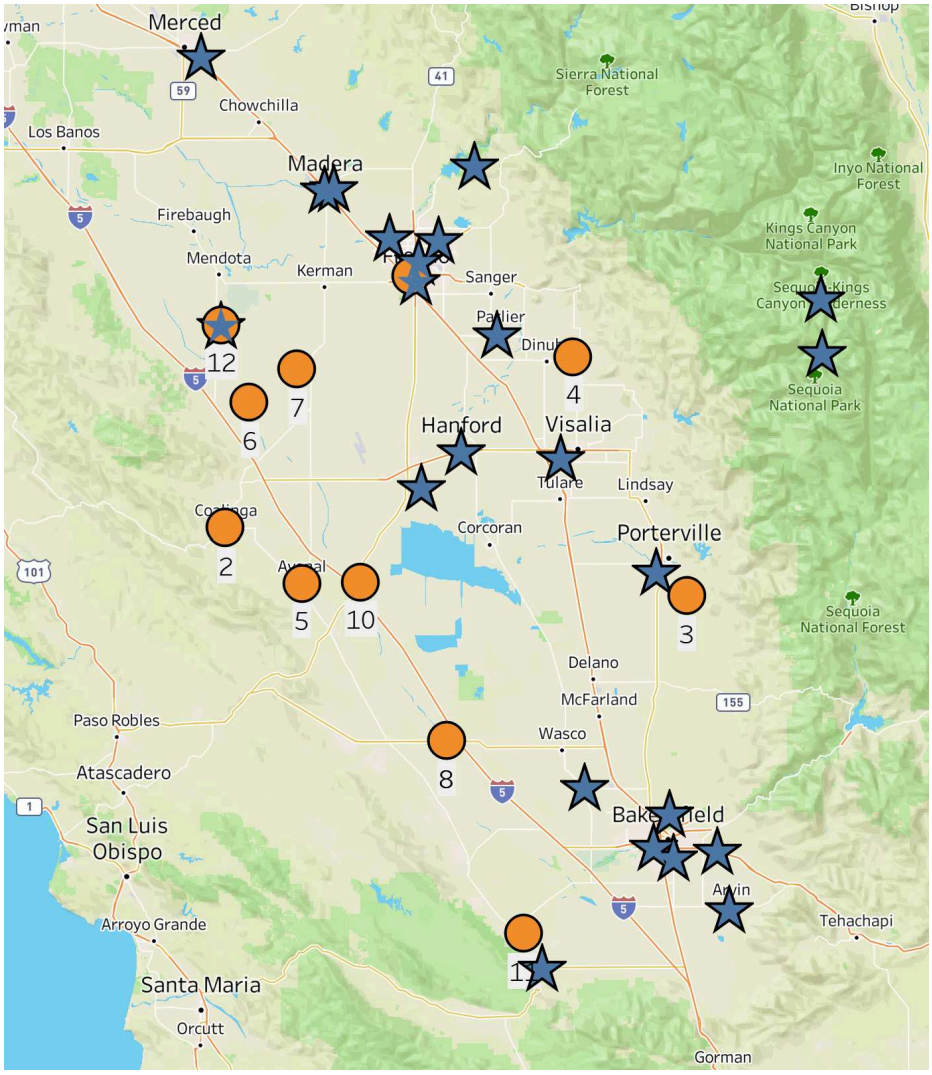
Figure 2.5.1 a) The image of the established site in the backyard of a community member in Coalinga. b) Photograph of the established location in the community member's backyard in Avenal.

(a)  (b) 

Figure 2.5.2 a) Depicts an image of the designated site in the orchid area of a community member in Lost Hills. b) Image of the designated site on the rooftop of the Fresno-Garland CARB facility.

Table 2.4 Overview of the VOZbox deployment sites and duration for calibration and deployment phases.

Unit	Location Type	Site Location	Pre-field Calibration Start	Deployment Date	Post-Field Calibration Start
VOZBox_1	Type A	Fresno Garland	6/7/2023	6/22/2023	11/01/2023
VOZBox_2	Type B	Coalinga	6/7/2023	7/3/2023	10/27/2023
VOZBox_3	Type B	Terra Bella	6/7/2023	8/2/2023	11/2/2023
VOZBox_4	Type B	Cutler	6/7/2023	7/6/2023	11/2/2023
VOZBox_5	Type B	Avenal	6/7/2023	7/10/2023	11/2/2023
VOZBox_6	Type B	Cantua Creek	6/7/2023	7/22/2023	10/27/2023
VOZBox_7	Type B	Lanare	6/7/2023	6/22/2023	10/27/2023
VOZBox_8	Type B	Lost Hills	6/7/2023	6/23/2023	11/2/2023
VOZBox_9	N/A	Not Deployed	6/7/2023	N/A	10/27/2023
VOZBox_10	Type B	Kettleman City	6/7/2023	6/29/2023	10/27/2023
VOZBox_11	Type B	Taft	6/7/2023	8/12/2023	11/14/2023
VOZBox_12	Type A	Tranquility	6/7/2023	6/22/2023	11/01/2023

(a) 
Merced
Chowchilla
Los Banos
Madera
Firebaugh
Mendota
Kerman
Sanger
Hanford
Visalia
Tulare
Lindsay
Corcoran
Porterville
Delano
McFarland
Wasco
Paso Robles
Atascadero
San Luis Obispo
Arroyo Grande
Santa Maria
Orcutt
Tehachapi
Gorman
Sierra National Forest
Kings Canyon National Park
Sequoia National Park
Sequoia National Forest
Inyo National Forest
12
7
6
4
2
5
10
3
8

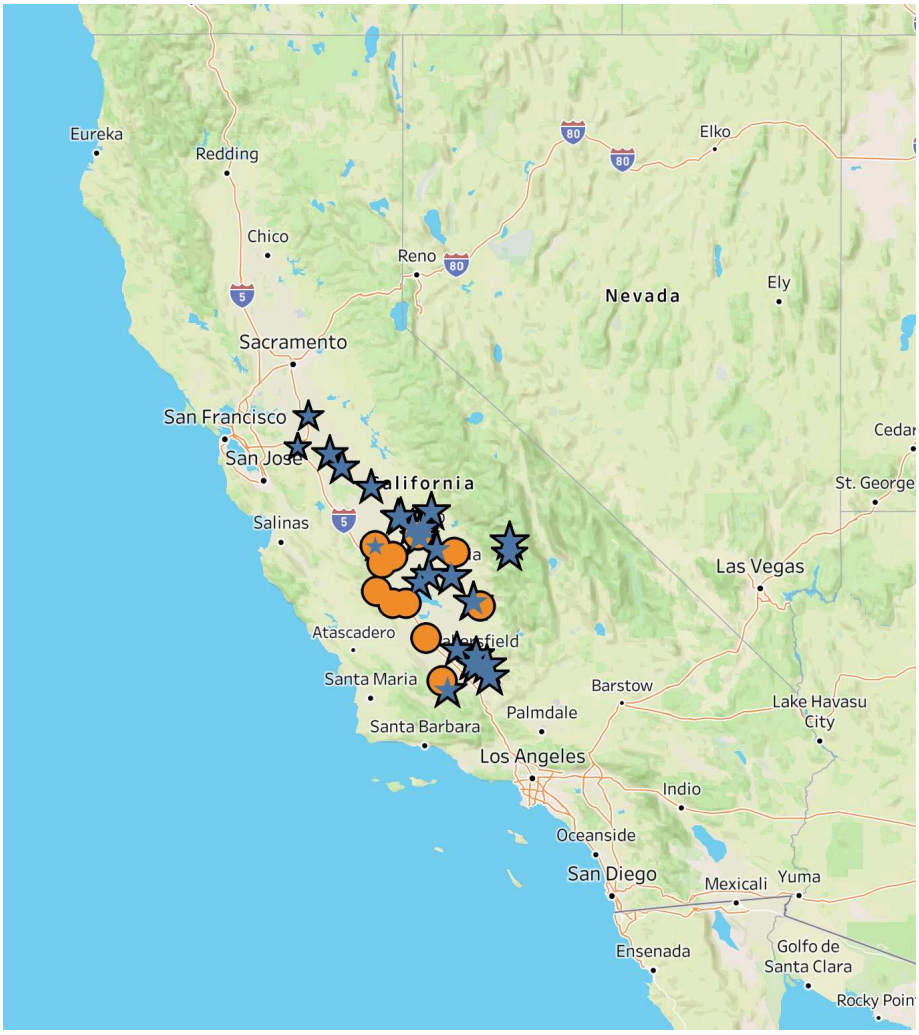
(b) 

Figure 2.5.3 a) Field deployment of the VOZboxes across the central California region. This map shows the existing reference sites (represented by stars) and the newly deployed VOZbox sites (represented by circles) with their numbers. b) Exploded view of the network of sensors in the state of California.

2.6 Calibration Model

A pre and post-field calibration data set is used for the calibration model to calibrate the sensor's raw data. The manufacturer's calibration works well when the units are out of the box with some units but lags considerably with others (Giardina, n.d.); this necessitates calibrating the sensors based on their deployment location while taking into account the ambient factors that influence pollutant concentration; in our case, temperature and humidity play an important role in

determining the value of ozone to be measured. The sensor's response displayed linear variability, implying that a linear relationship model would function well with these sensors. Calibration of the sensors was accomplished using a multivariate linear regression (MVLR) model. The formula for MVLR calibration is as follows.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots\ldots + \beta_n x_n + \epsilon$$

Where:

- y is the dependent variable (the outcome after prediction/calibration).
- β_0 is the intercept (the value of y when all x values are 0).
- $\beta_1, \beta_2, \ldots\ldots, \beta_n$ are the coefficients (weights) of the independent variables.
- $x_1, x_2, \ldots\ldots, x_n$ are the independent variables (predictors).
- ϵ Is the error term (the difference between the predicted and actual values of y).

The sensors are calibrated using the MVLR model that combines the data from the pre-and post-field calibration phases along with the ambient variables to produce a calibration model for the units with a high linear correlation between the reference data and the calibrated data that had high r-squared values and low Root Mean Square Error (RMSE) values for the calibration phases indicating the high performance of the model that was developed.

Data for this model is prepared by combining the data from pre- and post-field calibration phases, the data is cleaned and segregated for the respective units and the time zones are unified to the Pacific Time zone to ensure they are in line with the reference data time zone. The outliers over the pre and post-field calibration phase are removed using the z-score approach as they

skew the data and will affect the overall results. Data is resampled from 10-minute averages to an hourly data frequency to match the reference data frequency, and the missing data was not considered for the calibration as imputation would affect the data quality. Also, readings less than 10 ppbv are eliminated because part of the sensor's downtime had values close to zero and the values below 10 ppbv are not above the confidence range of the sensor, affecting the low-range calibration.

The MVLR model employs ambient parameters such as temperature and relative humidity. To account for sensor drift, a new computed parameter called week is introduced, representing the number of weeks the sensor was active since the pre-field calibration phase. The PM2.5 and PM10 measurements are also employed as they had a good correlation with the predicted variable; several independent variables that were measured by the VOZbox were selected for the model based on the correlation score between the reference value and the independent variable. The independent variables that have a strong Spearman coefficient(which can weigh values that are not linear) between them and the reference values are effectively used in the model to calibrate the values from the sensor. The model is trained throughout the calibration period and tested at the deployment site for units 1 and 12.

To further investigate the calibration model of this MVLR sensor we used three other models with regularization, a Least Absolute Shrinkage and Selection Operator (LASSO) regression technique was employed to evaluate its performance relative to the MVLR model, particularly in analyzing the reliance on independent variables for calibrating the actual value. The LASSO model incorporates L1 regularization, resulting in the reduction of less dependent, independent variables to zero. Therefore, effectively executing a variable selection.

All evaluated independent variables are utilized as features, allowing the LASSO model to determine their respective weights. Likewise, Ridge Regression was employed to evaluate the efficacy of the MVLR model; this regression technique utilizes L2 regularization, which compresses the independent variables towards zero without eliminating any, ensuring that no variables are excluded.

The Elastic Net Regression served as the final model for testing, including L1 and L2 regularization, hence offering stable regularization to the model. This model was also employed to validate the parameters specified for the MVLR model. The parameters that were selected in common with all three models and using Spearman's correlation coefficient were used to build the final MVLR model for calibration, this model was preferred for its simplicity and robustness against failure.

The MVLR model performance is evaluated based on the Square of Pearson correlation factor (R^2), Root Mean Square Error (RMSE), and Mean Bias Error (MBE) values in accordance with the reference values.

$$1.)\ r = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \Sigma(y_i - \bar{y})^2}}$$

$$2.)\ RMSE = \sqrt{\frac{\Sigma(x_i - y_i)^2}{n}}$$

$$3.)\ MBE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)$$

In terms of the sensor evaluation the pre-field calibration data was missing in VOZbox 6 for the 6/07/23 to 6/21/23 period where the plot was plotted; the issue was then fixed, and sensor 6 had data after that specified period, which was used for the pre-field calibration data. VOZbox 8 had a sensor failure and it had to be replaced and VOZbox 9 sensor radiation shield was replaced, thus leaving us with only 11 VOZbox sensors than the planned 12 sensors. Hence we have data only for 11 different locations in the Central Valley of California.

The Final Equation that was used for the calibration of the sensors included the independent variables represented below:

$$Calibrated\ O_3 = \beta_1(T_C) + \beta_2(RH) + \beta_3(O_3) + \beta_4(week) + \beta_5(PM_{2.5}) + \beta_6(PM_{10}) + \epsilon$$

Where:

- $Calibrated\ O_3$ Is the dependent variable.
- O_3 is the raw value of ozone measured by the SM50
- T_C, RH, O_3, $week$, $PM_{2.5}$, PM_{10} are the independent variables used for the calibration.
- β_1, β_2, ……., β_n are the coefficients (weights) of the independent variables.
- ϵ is the intercept term (the difference between the predicted and actual values of y).

The values for the weights can be found in Table S1, which summarizes the weight values for each individual sensor after the pre and post-calibration data calibrations.

2.7 Kriging

Kriging is a geostatistical technique that employs values from fixed measurement sites to infer spatial values at an unmeasured location without the need for measuring tools. This method is beneficial in regions lacking the resources for a measurement device (Candiani et al., 2010). This technique relies on the constructed variogram. A variogram is a statistical tool that assesses the dependency between two places concerning their distance and direction. Various types of variograms can be computed based on multiple criteria. When the variogram exhibits shortcomings, it could adversely impact the values displayed on the spatial maps. The availability of data influences the quality of the resulting maps; a greater number of stationary points correlates with enhanced map quality. Furthermore, there is an excessive dependence on data correlation, resulting in the created map potentially lacking accuracy compared to actual maps due to the dynamic interactions involved. This map offers a broad overview of the potential values of locations depending on interactions with other locations.

A measurement station for pollutants must meet specific standards, including minimum population density or objectives such as assessing background pollution or identifying hotspots. Locations that do not meet the criteria for a measurement station lack the availability of data regarding local pollutants. Kriging effectively addresses this by estimating values for these locations using data from other cities. The incorporation of VOZboxes significantly enhances the range of these maps.

Kriging can effectively predict ozone distribution with high spatial resolution, providing valuable information for environmental and health impact assessments (Araki et al., 2015).

The data from the reference sites and VOZboxes are extracted and then used to develop the spatial gradient map using universal kriging which is developed based on the universal variogram that assumes there is a inverse linear variation in the amount of ozone concentrations, the prediction of each site extends to a radius of 20 miles, which in turn when combined with all the other sensors that are available in the region makes the map more effective. The maps are developed only with the existing reference sites and then with the addition of the VOZbox data to compare the improvement in the range of the map.

3. Results

3.1 Data Availability and Quality

Figure 3.1.1(a) gives us an overall picture of the measured environmental parameters and the pollutants data that were available for the VOZbox 7, which has a modest amount of data for the measured period. Figure 3.1.1(b) shows the overall picture of a VOZbox 10, which had less data and a lot of data discontinuity and data losses. Figure 3.1.1(c) provides an overview of data availability for all units and Table 3.1 provides the data availability across the 8 am to 8 pm where the ozone values are prominent. The table shows that we typically have data above 80%, with some units experiencing downtime due to battery issues and temperature issues. This gives us the confidence that the VOZboxes were able to capture the data more often than not. This gives us an idea about the data stability and availability of these low-cost pollutant measuring devices.

All the measured variables in the VOZbox have the same amount of data variability across the devices, thus concluding that whenever the device was online, all the sensors were providing data and the data was getting uploaded to the cloud. The loss of data in the devices was majorly caused by an outage due to the temperature rise inside the box.

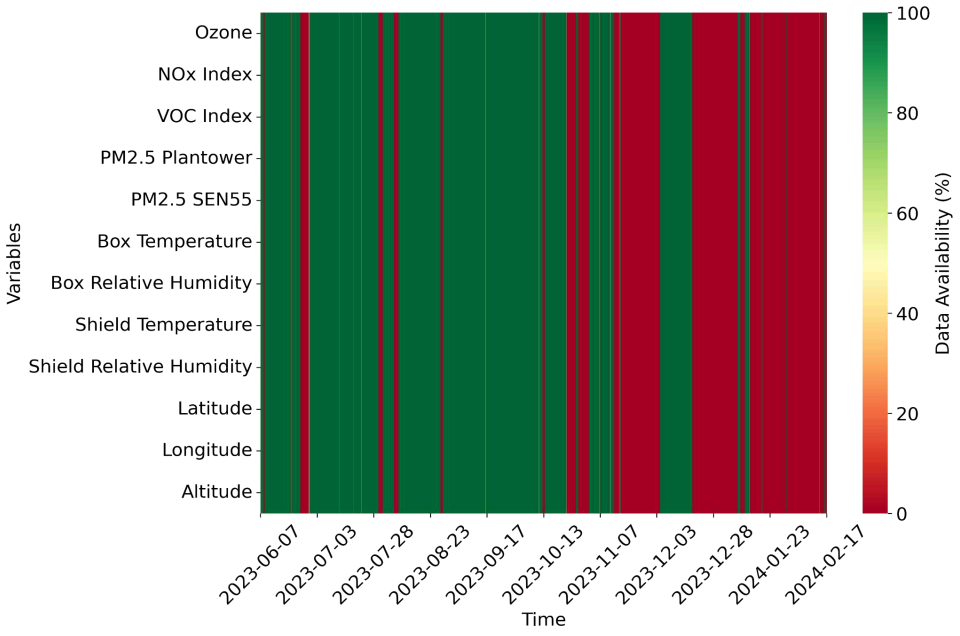
Table 3.1 Average data availability (%) across units across June to November of 2023 for the period of 8 am to 8 pm daily.

Unit	June	July	August	September	October	November
VOZBox_1	97	95	88	100	68	85
VOZBox_2	86	87	88	100	49	86
VOZBox_3	99	81	87	99	73	46

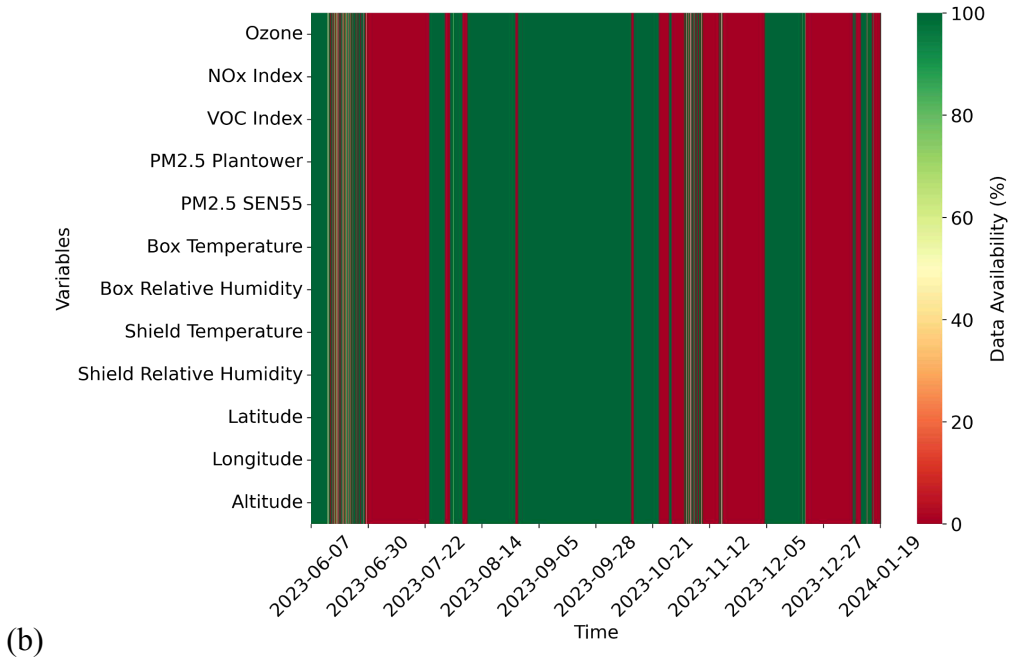
VOZBox_4	89	91	88	99	70	75
VOZBox_5	78	65	88	100	76	13
VOZBox_6	0	86	82	93	74	05
VOZBox_7	79	95	88	100	71	39
VOZBox_8	89	95	88	100	73	10
VOZBox_9	99	35	44	70	76	19
VOZBox_10	80	20	88	100	72	23
VOZBox_11	34	35	84	100	76	32
VOZBox_12	96	95	87	99	75	86

The LiPoFe4 battery has an incorporated safety mechanism that prevents it from charging or discharging beyond a temperature point, which is around 45 °C, to preserve the health and chemistry of the battery. This led to a lot of outages in the data. As ambient temperatures exceeded 40 °C, the control box, containing substantial hardware and a battery that generated heat, was entirely enclosed due to its weatherproof design, resulting in internal temperatures surpassing 45 °C, thereby activating the safety trigger. The electronics kept running with the active electricity generated by the solar panel. The battery, upon cooling, began to discharge; nonetheless, little energy remained to sustain the device overnight, resulting in its shutdown. The devices began to reactivate once the solar panel was operational. The power cycling problem resulted in data loss throughout the nighttime hours. The charge controller requires a functional battery to operate power from either the solar panel or the battery itself.

Some of the units had the batteries discharge below the critical limit, thus inducing an error code, which stopped the unit from functioning, thus a loss in data completely. The battery had to be manually charged and then had to be put back for the units to kick in and measure data again. This flaw in the design of the units where the continuous data flow wasn't available, was not found in the initial testing, as these were tested in Fort Collins, Colorado, which is relatively cooler than the Central Valley, and never faced such issues as the temperatures never soared so high. However, when the units were put to test in the field, this issue was identified. This issue caused some outage in data but we were able to capture most of the ozone activity in that area as the ozone concentrations are higher in the presence of sunlight. The batteries started failing at the latter half of the deployment due to its high battery cycle count thus failing the entire system.



(a)



(b)

(c)

Figure 3.1.1 a) Data availability heatmap in VOZbox 7 across all measured environmental parameters and pollutants. b) Data availability heatmap in VOZbox 10 across all measured environmental parameters and pollutants. c) Overall data availability heatmap across all VOZboxes throughout calibration and deployment. (white squares indicate that they weren't turned on)

3.2 Calibration

3.2.1 VOZbox O_3 Calibration

The calibration results obtained from the PowerHouse CSU(Colorado State University) campus yielded poor results when applied to the values observed at the Fresno Garland site, as the R^2 values were much lower than 0.6, which shows a significant effect of ambient conditions on the measured tropospheric ozone concentrations. The calibration that was built was a simple linear regression without any variables and it performed poorly than what was expected.

Figure 3.2.1.1 compares the raw and calibrated sensor outputs from two VOZboxes against reference measurements made at the Fresno/Garland site during the pre-field calibration period. Results are plotted both as a time series (Figure 3.2.1.1 panels a and c) and as scatter plots (panels b and d). While the calibrated outputs were developed by using both the pre-and post-field data, we only show comparisons for ~2 weeks during the pre-field data (June 07 to 29, 2024) to keep the visual presentation clean but representative of the overall performance. For VOZbox 7, the raw (or manufacturer's) calibration underestimates the O_3 concentrations but the calibrated output accurately reproduces the absolute O_3 concentrations and the diurnal profile over this ~2-week period (Figure 3.2.1.1(a)). The raw calibration does relatively much better for VOZbox 1 in reproducing the peak O_3 concentrations but fails in capturing the low O_3 concentrations (Figure 3.2.1.1(c)). Similar calibrations and comparisons were performed for all other VOZboxes and these are shown in Figure 3.2.2.2, which includes all data used for calibration.

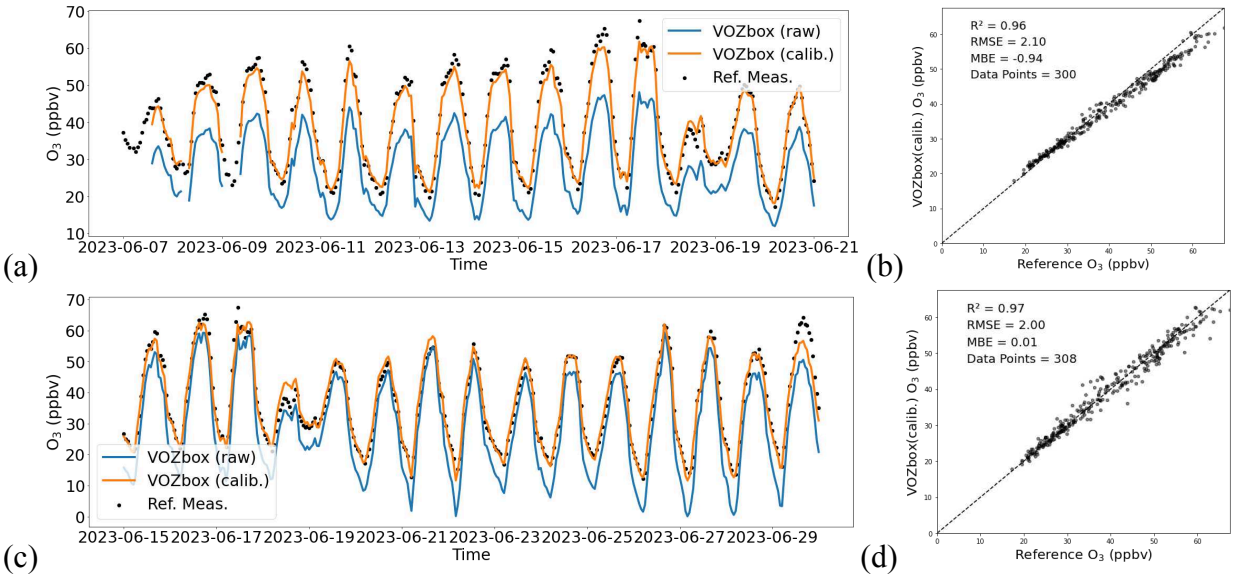


Figure 3.2.1.1 a) Time series scatter-line plot for VOZbox 7 over the period of June 07 to June 21, 2023, showcasing the sensor's reference, raw, and calibrated values. b) Scatter plot for VOZbox 7 calibrated data against the reference ozone value at the Fresno Garland site during pre-field calibration. c) Time series scatter-line plot for VOZbox 1 over the period of June 15 to June 29, 2023, showcasing the sensor's reference, raw, and calibrated values. d) Scatter plot for VOZbox 1 calibrated data against the reference ozone value at the Fresno Garland site during pre-field calibration.

3.2.2 Pre and Post Calibration

The need for the pre-and post-calibration is important to capture the sensor-specific drift and to reduce the bias of the calibration model on high ozone values, this is clear where the usage of only pre-field calibration results in hourly ozone concentration values that are scattered above the 1:1 line indicating overprediction of the values and have a lower linear correlation with the reference measurements yielding an R^2 value close to 0.58 for the VOZbox 1 at Fresno Garland and when the pre-field calibration is applied to the VOZbox 12 as well and yields an R^2 value of 0.47.

When the post-field data is included in the data that is used to train the model, the deployment data looks very good, with the VOZbox 1 yielding an r-squared value of 0.96 and the VOZbox

12 yielding a value of 0.96 as well; thus, this indicates the success of the post-field calibration that was carried out.

Conclusively calibrations were performed using both the pre- and post-field data as calibrations performed using only the pre- or the post-field data did not effectively account for unit-specific sensor drift (e.g., see Figure 3.2.2.1). The data was observed to have a linear drift and the data collected was post-processed totally after the measurement period. Overall, Figures 3.2.2.1a) and 3.2.2.1 b) suggest that multivariable regression with post-field data was able to accurately estimate O_3 over a sufficiently large dynamic range (20 to 90 ppbv) which covers the exposure range that can cause health issues with a very small error (RMSE ~ 4 ppbv).

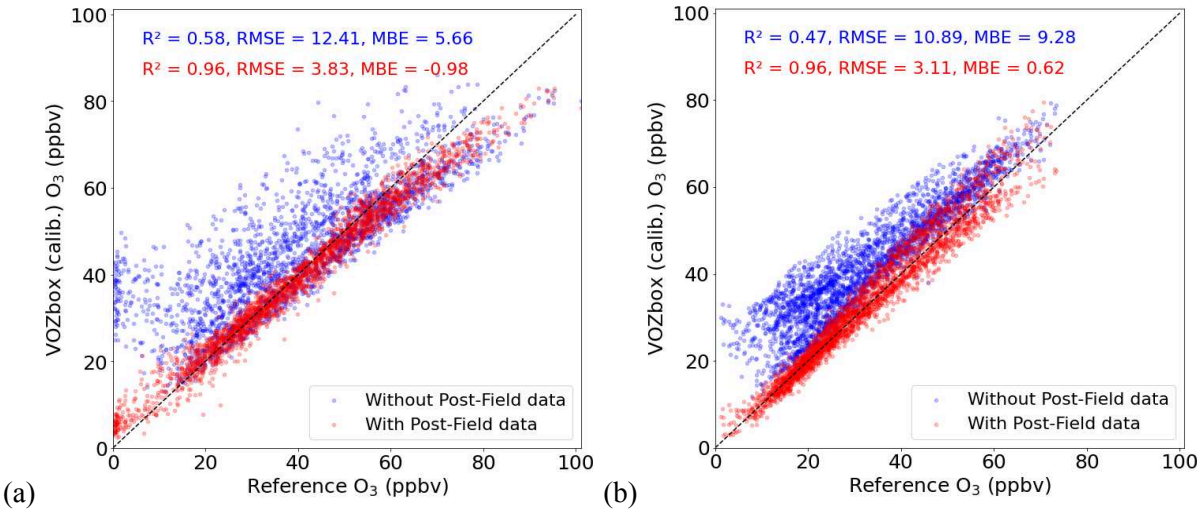


Figure 3.2.2.1 a) VOZbox 1 deployment scatter plot with and without the post-field calibration values. b)VOZbox 12 deployment scatter plot with and without post-field calibration values.

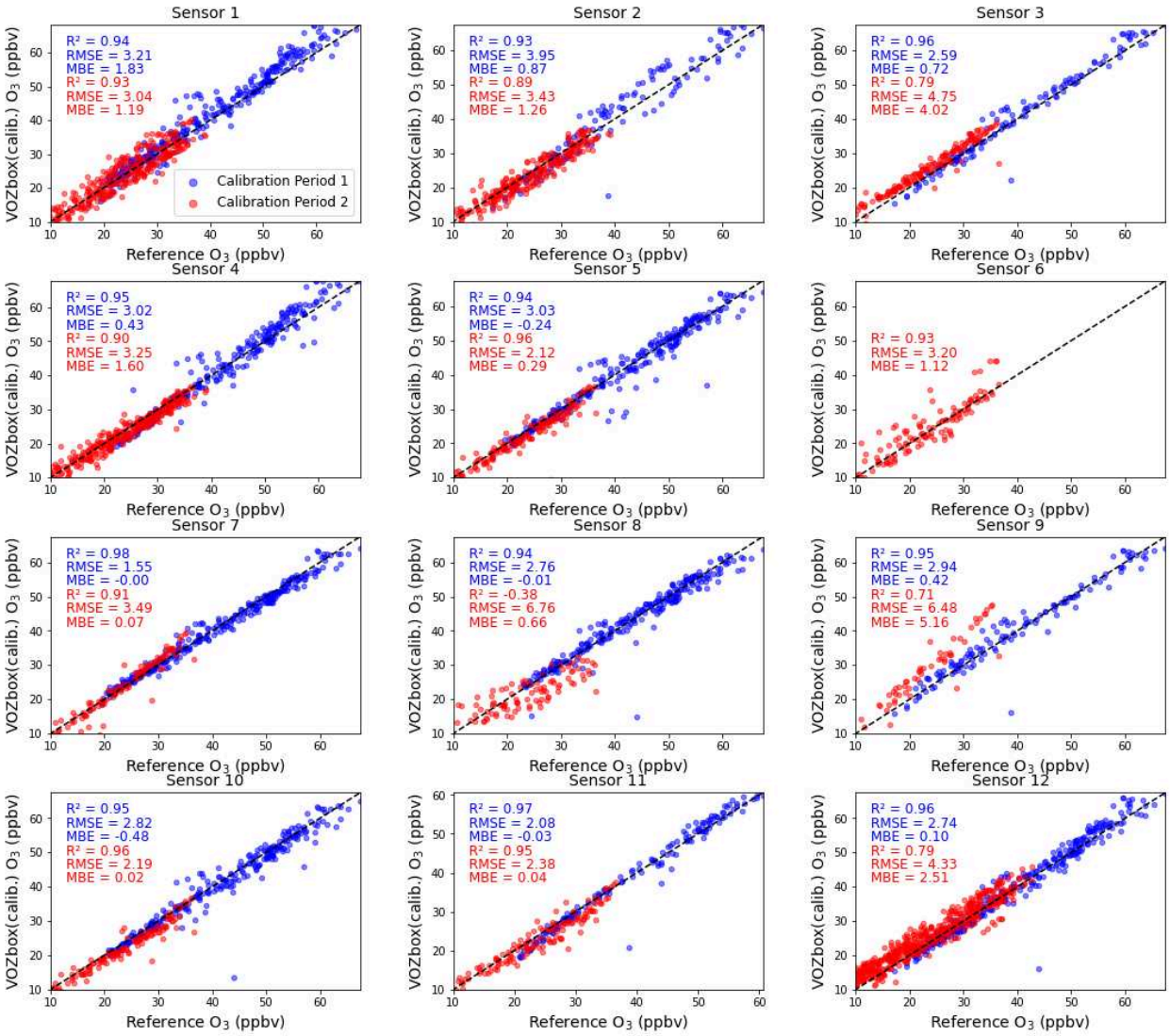


Figure 3.2.2.2 Pre (calibration period 1) and Post-field (calibration period 2) calibrated data against the reference data for all the 12 VOZboxes. All the sensors have high linear correlation after calibration indicating a high performance. Sensor 6 was down during this time frame.

Figure 3.2.2.2 is a consolidated plot with all the VOZboxes with their pre-field (calibration period 1) and post-field (calibration period 2) phases calibrated data against the reference values, which indicate the high correlation that can be established for the sensor when they are deployed in the field.

The sensors were tested with an 80-20 split for the train and a test split for pre-field deployment, and the calibration was successful in yielding values of R^2 higher than 0.95. Then the entire data was considered for the sensors' training to ensure that we have sufficient data to train the model that can capture the values accurately in the field.

The plot clearly indicates that the sensors were trained well in both calibration periods, with most of the sensors having an r-squared value higher than 0.9. Table S1 shows the values for the final calibration that was obtained with the intercept values as well.

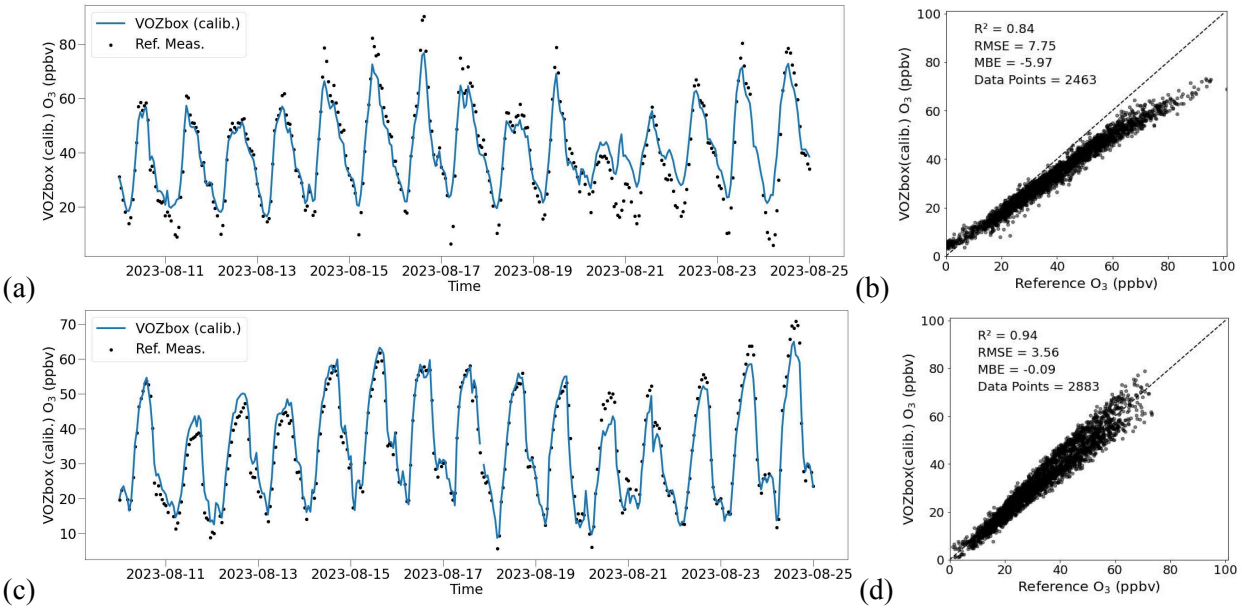


Figure 3.2.2.3 a) Time series scatter of deployment data for VOZbox 1 at Fresno Garland for the time period (08/10/23 to 08/25/23) indicating the reduced performance by the sensor due to the inability to capture high hourly ozone values above 60 ppbv b) Reference versus calibrated ozone scatter plot for VOZbox 1at Fresno Garland for the time period (08/10/23 to 08/25/23) indicating the clear decline at higher ozone concentrations, c) Time series scatter of deployment data for VOZbox 12 at Tranquility for the time period (08/10/23 to 08/25/23) indicating good performance of the sensor by capturing most of the high ozone points.d) Reference versus calibrated ozone scatter plot for VOZbox 12 at Tranquility for the time period (08/10/23 to 08/25/23)indicating a good performance for a range of ozone concentrations.

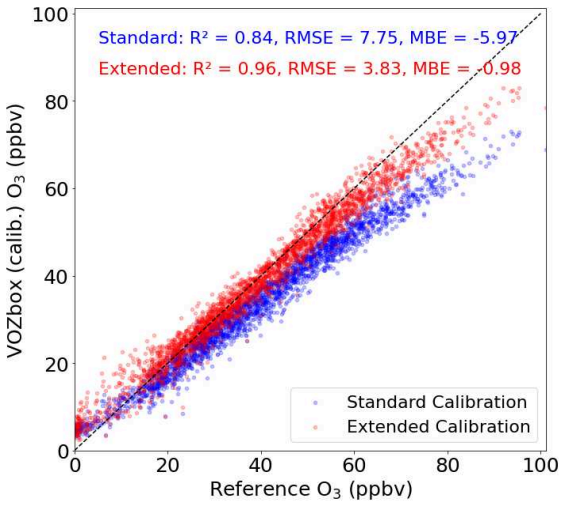
Figure 3.2.2.3 illustrates the contrasting performance of the sensors at the type A site, where the VOZbox 1, located in Fresno, a metropolitan area, encountered elevated ozone concentrations as denoted by the reference values. However, the model did not accurately forecast values in proximity to the actual measurements, raising significant concerns as the expected level of performance was not achieved, evidenced by an R^2 value of 0.84. This was inferior to what we saw during the calibration period. The VOZbox 12 effectively measured the ozone concentrations to which it was subjected, achieving a performance characterized by an R^2 value of 0.94, closely aligning with calibration results and meeting expectations. The inference drawn is that the calibration model exhibits superior performance within the range of its calibration. This leads us to another association with the necessity for an increased number of days during the calibration phase.

The calibration period plays a key role in training the calibration model as the number of high ozone days available in the calibration phase data dictates how the performance of the model pans out in the deployment phase, this was tested out in the type 1 site units that were placed adjacent to the reference monitors. The unit at the Fresno-Garland site did better than the originally intended calibration when extra periods of data were added to the pre-deployment calibration phase data, as the R^2 value improved from 0.73 to 0.87 with 5 additional days, as the extra periods had high ozone data in adequate which improved the overall performance of the unit, this highlights the importance of the high ozone days captured into the data when the deployment phase has higher ozone values than the calibration phase then the model struggles to capture that band of values.

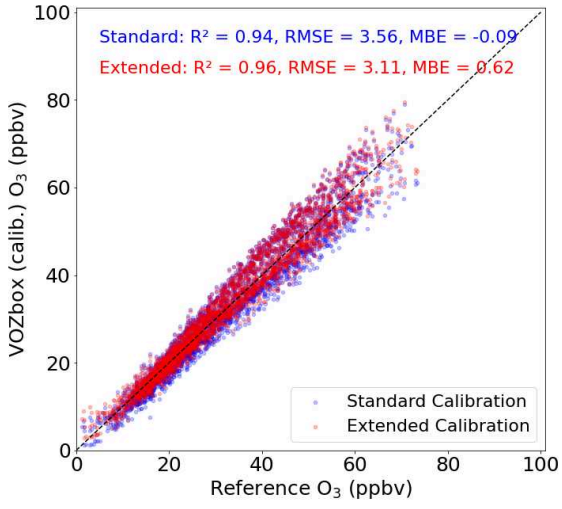
With 25 days of additional data in the pre-deployment calibration, the R^2 value increased to 0.95 indicating that maximum performance can be extracted with sufficient data. There was an

improvement of 20% in the efficacy of the sensor when an additional 10% of the actual data set was provided to the training model as depicted in Figure 3.2.2.4 c) and d). This truly helps in improving the capability of the sensor in predicting the high ozone non-compliance days.

We can see that the unit at Tranquility did very well even without the extra days added to the calibration phase data, the extra days added to the model were from the Tranquility reference site as it was deployed immediately. This emphasizes that the model does well when the calibration phase data can capture the range of values the unit is about to measure in the deployment phase. This also cements that the units measure values accurately to the actual values when the model is well equipped.



(a)



(b)

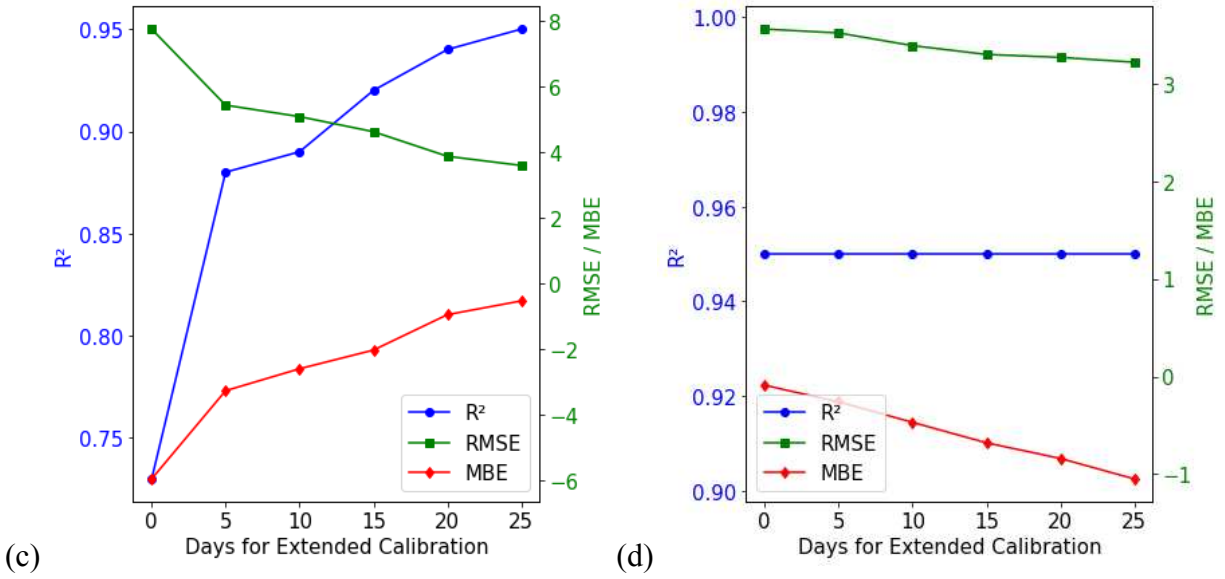


Figure 3.2.2.4 Extended Calibration a) VOZbox 1 and 12 O_3 hourly field data compared to reference hourly data at the Fresno Garland site and Tranquility site respectively with 0 days extended calibration (deployment data) b) VOZbox 1 and 12 O_3 hourly field data compared to reference hourly data at the Fresno Garland site and Tranquility site respectively with 25 days extended calibration (deployment data) c) VOZbox 1 R-square and RMSE/MBE of deployment VOZ/Reference correlation vs the number of extended calibration days. d) VOZbox 12 R-square and RMSE/MBE of deployment VOZ/Reference correlation vs the number of extended calibration days.

3.3.Drift Analysis and Correction

The requirement for pre-field and post-field data is essential for recording the long-term drift related to the sensor. The Type A site enables the study of continuous deployment data at the Fresno-Garland and Tranquility locations, facilitating the integration of a temporal factor into the calibration model to address short and long-term temporal sensor drift.

Figure 3.3.1 a) illustrates data from the Fresno Garland site, which reflects calibration based solely on pre-field data without extended calibration. It is evident that drift occurs, as values at first align closely with the 1:1 line during the calibration phase but subsequently diverge at higher concentrations, significantly underpredicting actual values by up to 25%. This

discrepancy raises concerns regarding the sensor's reliability for supplementary monitoring. Over time, the drift appears to improve slightly, suggesting enhanced performance at lower concentrations. On the contrary, figure 3.3.1 b) illustrates a linear decline in the sensor's performance, characterized by a gradual decrease in output and an underestimation of values; however, this behavior is consistent and predictable. In contrast, the drift observed in the VOZbox 1 is unique to that specific unit. These two units emphasize the necessity of implementing drift correction by the addition of post-field calibration, as well as incorporating a temporal variable to mitigate drift, in addition to training the model for high ozone days to enhance the sensor's overall efficiency.

The VOZboxes at the Fresno Garland site and the Tranquility site exhibit distinct drifts from each other and they are corrected. We assessed the drift after the correction using hourly data spanning almost 170 days. Figure 3.3.1 c) displays the performance of the Fresno Garland sensor which was exceptional throughout the field period with an R^2 value of 0.96 and the RMSE value of 3.52 ppbv owing to the extended calibration data that was incorporated. Still, the drift was particularly noticeable in the higher ozone range (above 70 ppbv) where the sensor was underpredicting the actual value being reduced to 10% of the actual value which can be clearly seen in Figure 3.3.2 a) which is a residual plot showing the difference between the calibrated value and the reference value, suggesting that the sensor’s ability to capture higher concentrations in the ozone season (June to August) is compromised slightly and can capture NAAQS compliance. The model has a variable accounting for the temporal change (number of active weeks) and also included the post-field calibration data to capture the lower range of ozone concentrations.

In contrast, Figure 3.3.1 d) shows that the VOZbox at the Tranquility site exhibited a linear drift while maintaining good performance with an R^2 value of 0.95 and an RMSE value of 3.36 ppbv, the linear drift pattern which was noticeable with the monthly drift downward has reduced and the values are concentrated close to the 1:1 line, which suggests a strong improvement in the sensor's performance. This slightly contrasting drift pattern after the drift correction illustrates the varied characteristics of the sensors in terms of performance over the deployment period. Figure S12 - S14 provides detailed performance charts for both sensors.

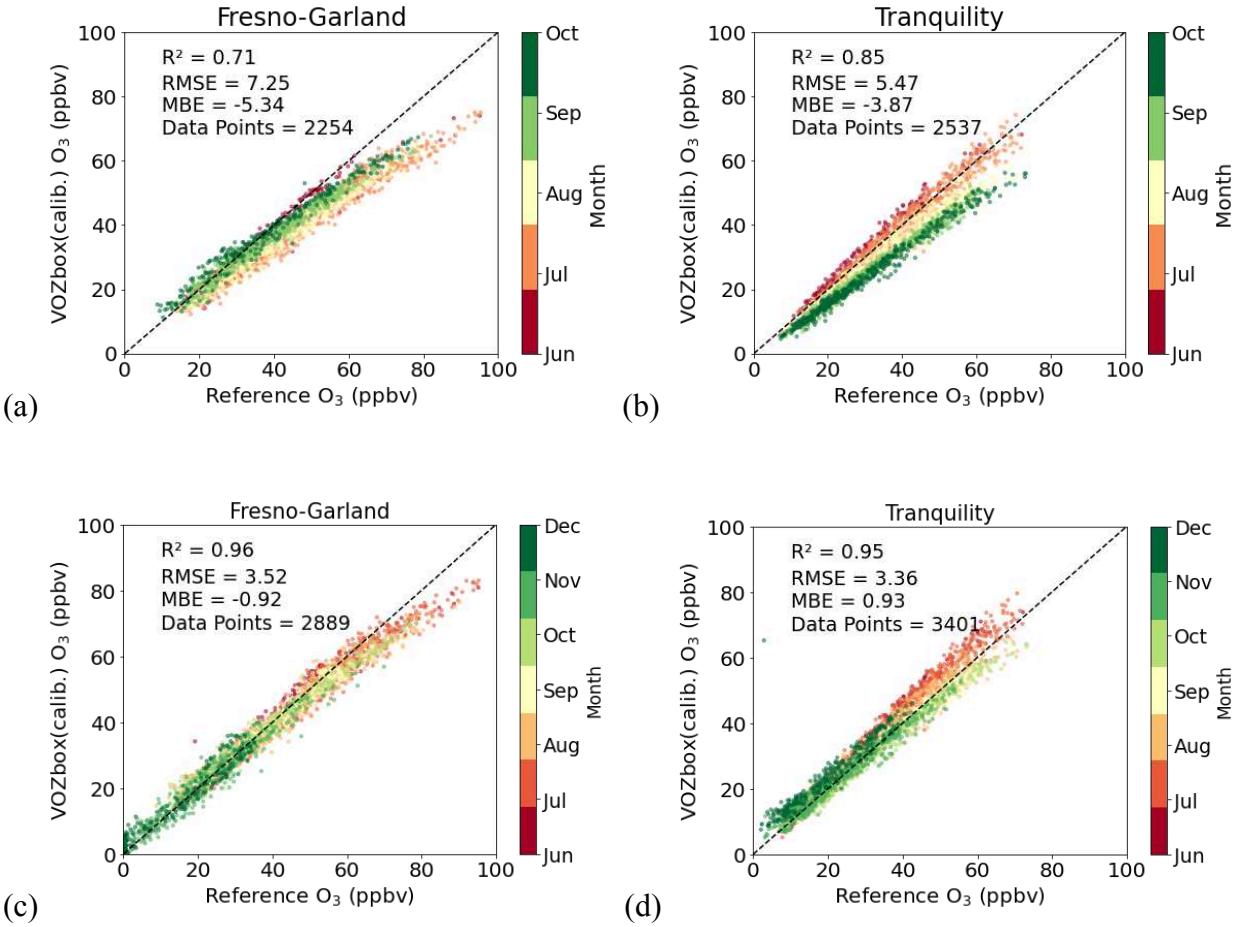


Figure 3.3.1 Drift analysis a)VOZbox 1 O_3 hourly value compared to reference hourly measurements at the Fresno Garland site with pre-field data calibration. b)VOZbox 12 O_3 hourly value compared to reference hourly measurements at the Tranquility site with pre-field data calibration. c)VOZbox 1 O_3 hourly value compared to reference hourly measurements at the Fresno Garland site with extended calibration days and post-field data d) VOZbox 12 O_3 hourly

value compared to reference hourly measurements at the Tranquility site with extended calibration days and post-field data.RMSE - Root Means Square Error, MBE - Mean Bias Error.

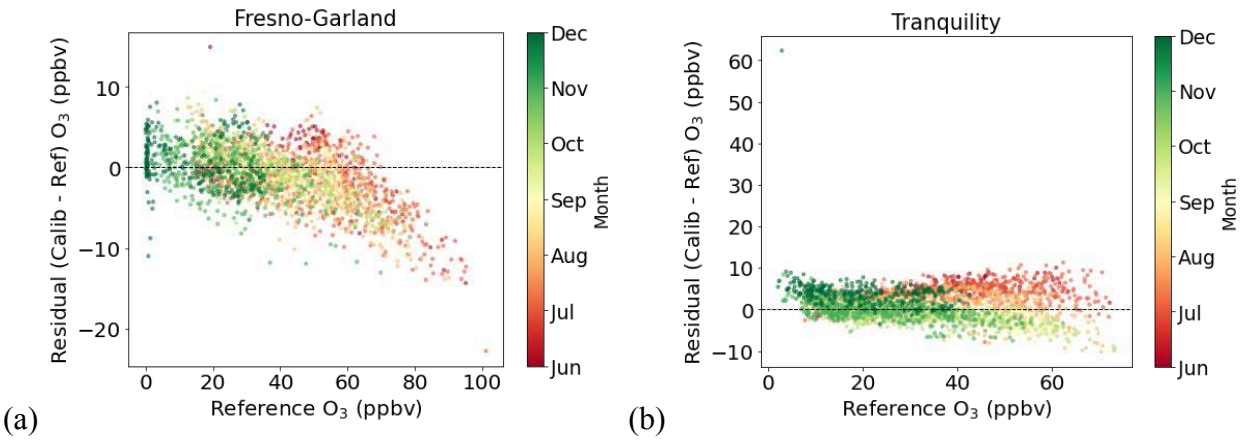


Figure 3.3.2 a)VOZbox 1 O_3 hourly residuals compared to reference hourly measurements at the Fresno Garland site with extended calibration days and post-field data. b)VOZbox 1 O_3 hourly residuals compared to reference hourly measurements at the Fresno Garland site with extended calibration days and post-field data.

3.4 NAAQS Capability and Performance Analysis

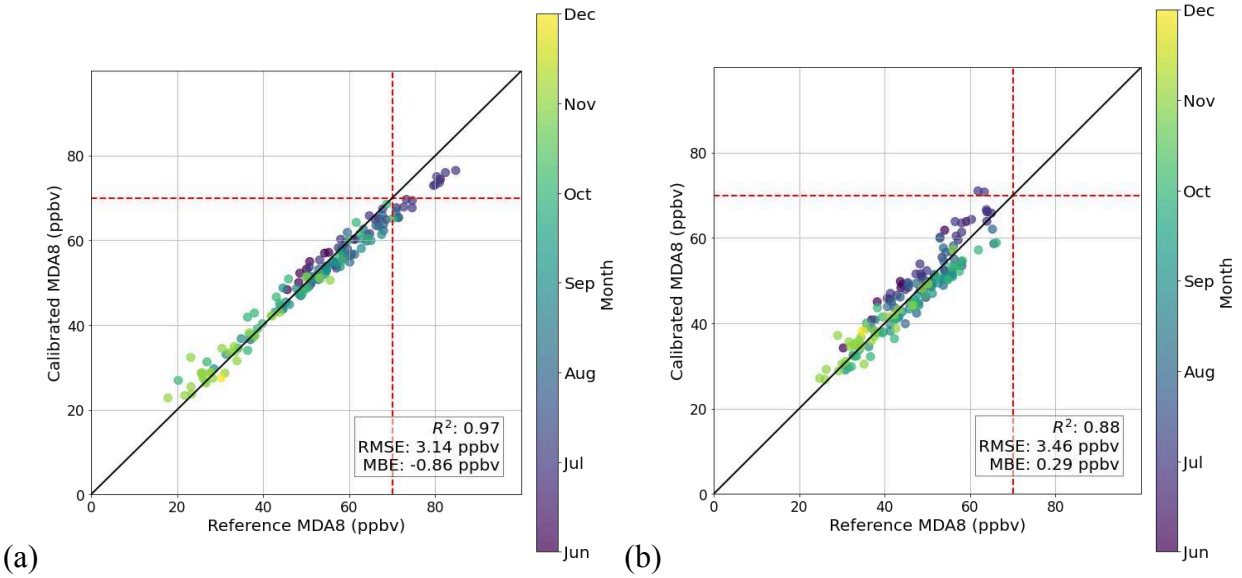


Figure 3.4.1 NAAQS detection capability a) VOZbox 1 O_3 MDA8 estimates compared to reference MDA8 measurements at the Fresno/Garland site b) VOZbox 12 O_3 MDA8 estimates compared to reference MDA8 measurements at the Tranquility site. True positives (TP) represent >70 ppbv O_3 events (the NAAQS standard for O_3 exposure) measured by the collocated reference monitor where VOZbox estimates also exceed 70 ppbv. True negatives (TN) represent

<70 ppbv O_3 events where VOZbox estimates are also below 70 ppbv. False positives (PN) represent <70 ppbv O_3 events where VOZbox estimates exceeded 70 ppbv and false negatives (FN) represent >70 ppbv O_3 events where VOZbox estimates below 70 ppbv. RMSE - Root Means Square Error, MBE - Mean Bias Error.

We assessed the capability of the sensor to measure O_3 compliance, which could predict the Maximum Diurnal Average 8-hour average O_3 (MDA8) value. This metric is used to evaluate the air quality, National Ambient Air Quality Standards (NAAQS) for tropospheric ozone are established at 70 ppbv, and any exceedance is flagged as non-compliance.

The two sensors deployed at the Type A site were utilized to assess the sensors' capability to predict the MDA8 value after the drift correction. Figure 3.4.1 shows the calibrated VOZbox compared to the reference value at two different sites. The figures include both the training and the test data where the training data constitutes about 8 - 12% of the full data set.

The VOZbox 1 at the Fresno Garland site recorded 7 of the 16 non-compliance days which were the true positives out of the 153 days measured and observed during the deployment period. The Type I (false positives) error indicates where the VOZbox predicted a value greater than 70 ppbv for MDA8 values but the true value is lower than 70 ppbv. The Type II (false negatives) error indicates where the VOZbox predicted a value lower than 70 ppbv for MDA8 values whereas the true value is higher than 70 ppbv. The VOZbox 1 has zero Type I errors and 9 Type II errors. The Type II errors that were present in the VOZbox 1 were very close to 70 ppbv thus indicating a strong correlation between the VOZbox sensor and the reference measurement, evidenced by an r-squared value of 0.97 and an RMSE value of 3.34 ppbv.

The VOZbox 12 at Tranquility forecasted two non-compliance days during the deployment period of 153 days; however, the reference grade instrument recorded no non-compliance days,

suggesting a slight overestimation of values at the higher concentrations while still adhering to the prediction range. The VOZbox 12 had two instances of Type I errors and 0 Type II errors. The Type I errors that were present in the VOZbox 1 were very close to 70 ppbv thus indicating a strong correlation between the VOZbox sensor and the reference measurement, evidenced by an r-squared value of 0.88 and an RMSE value of 3.46 ppbv. Both locations exhibited commendable performance relative to the reference monitors, offering a clear assessment of sensor efficacy and establishing a benchmark for the performance of other sensors. Table 3.2 summarizes the performance metrics for the two units, with the standard deviation calculated using the bootstrap approach. They exhibit a high accuracy of approximately 88%, assuming an accuracy of 100% when the difference is within a range of 5 ppbv. The cost-to-performance ratio achieved by the sensors, coupled with advanced calibration processes, is commendable, encouraging confidence in its deployment across various sites.

Table 3.2 Final consolidated performance metrics for the VOZboxes.

VOZbox 1 - Fresno-Garland	**VOZbox 12 - Tranquility**
R^2: 0.97 ± 0.00	R^2: 0.88 ± 0.01
RMSE: 3.13 ± 0.18	RMSE: 3.45 ± 0.17
MBE: -0.87 ± 0.24	MBE: 0.28 ± 0.29
Accuracy: 88.21 ± 2.61	Accuracy: 87.78 ± 2.67

3.5.Spatial Analysis

We assess the spatial disparities between various sites, with some having nearby reference sites and others lacking such monitors. We majorly do two analyses, one being the cluster analysis and the other being the paired analysis. In the first method, where reference values are available, we

directly compare the VOZbox unit-calibrated ozone concentrations against the reference monitor in the closest proximity (paired analysis) and in the cluster analysis, we compare a cluster of devices with the reference monitor in the same proximity to identify spatial trends. The second method, in the absence of reference monitors, involves comparing calibrated ozone concentration in the clustered VOZboxes within the same cluster to identify spatial trends.

3.5.1 Cluster Analysis

When we place the monitors close to each other, we consider them to form a cluster of VOZboxes. The analysis of these clusters reveals trends among those locations. We anticipate a change in the spatial variability of approximately 5-15 ppbv (Sadighi et al., 2018). We have identified two clusters in the VOZbox deployment: cluster 1(VOZboxes 6, 7, 12) is centered around Tranquility, where we have a reference FEM-grade instrument; we would be using method one for this cluster, and the second cluster is centered around the westernmost region of the Central Valley, where the values remain unknown due to a lack of a reference grade instrument, thus utilizing the second method both these clusters are visually showcased on a map in figure 3.5.1. We plotted the ozone concentrations on a time series for the clustered sensors in Figures 3.5.1.1.1 and 3.5.1.1.2, matching the dates with high ozone days observed at the Fresno Garland site in August when all the sensors were in the field. We then conducted a comparison among the sensors in that cluster.

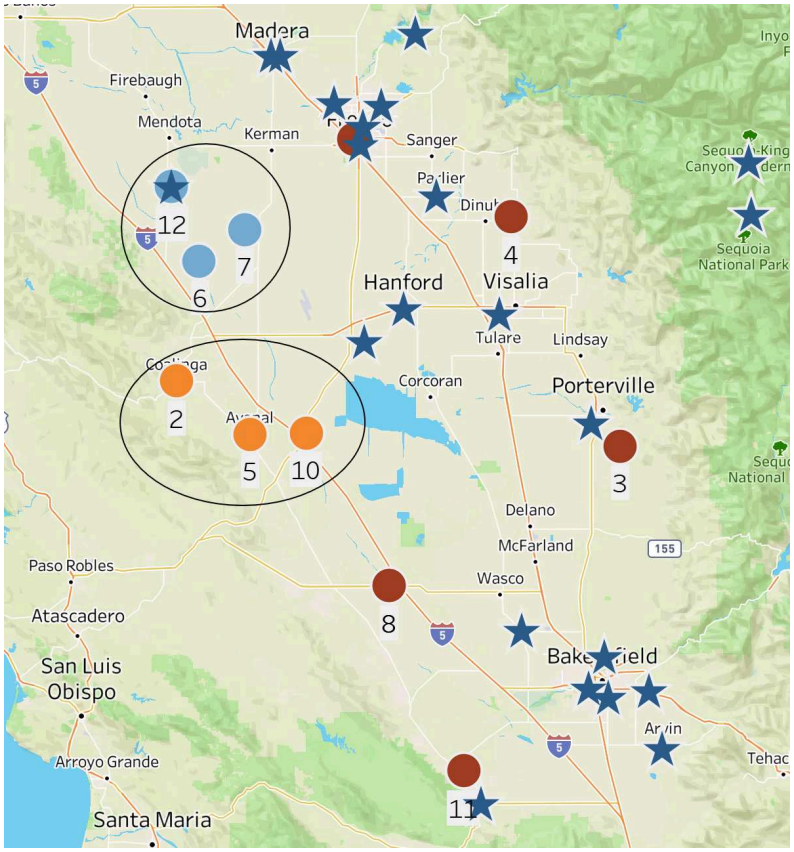


Figure 3.5.1 illustrates the location of the clusters considered, whereas the blue circles indicate the cluster 1 and the orange circles indicate the cluster 2.

3.5.1.1 Cluster 1: (Tranquility, Lanare, Cantua Creek)

This cluster has the luxury of comparing the sensor values to the reference-grade instrument in Tranquility. The VOZbox 12 (Tranquility) calibration was built partly on the values from this reference instrument; hence, that sensor provided the best ozone concentrations when compared to the reference monitor. On the contrary, VOZbox 6 (Cantua Creek) and VOXbox 7 (Lanare) were calibrated at the Fresno Garland site, and they were detecting values closer to values from the Tranquility reference site; hence, we can see that the spatial variability amongst these cluster of sensors is low and the variation is within 15 ppbv from the Figure 3.6.1.1.1 which gives us an hourly time series plot between the dates of Aug 09 to Aug 25, 2023, we picked these dates as they are after the deployment of all three sensors and the Fresno site had elevated ozone

concentrations in this period. The values from Lanare had lower values during nights compared to the other two sites indicating that the activity levels in the morning are high compared to the other two sites.

The other two sites measured higher peak ozone concentrations than the reference site having a variation of ~10 ppbv. The MDA8 values for these sites are very comparable as they have each picked up 2 non-compliance days in terms of high ozone days. The population of all these three sites is comparable and less than 1000 as per the census in 2010.

The differences in values are in the range of ~10 ppbv, and the three sites measured 60 ppbv plus MDA8 values: Lanare – 33 times, Cantua Creek – 20 times, and Tranquility – 40 times, respectively. This brings us to the conclusion that the compliance monitoring done at the Tranquility site with the reference grade instrument is effective for the regions surrounding it as well owing to its higher values being measured and values comparable at high ozone days.

The VOZ monitors that have been placed are giving a good indication and have helped to see the trends amongst these zones and a case that the values in these agricultural communities part of the Central Valley are well below the compliance limit in the ozone season of 2023 as the Tranquility site had zero non-compliance days registered in the regulatory monitor.

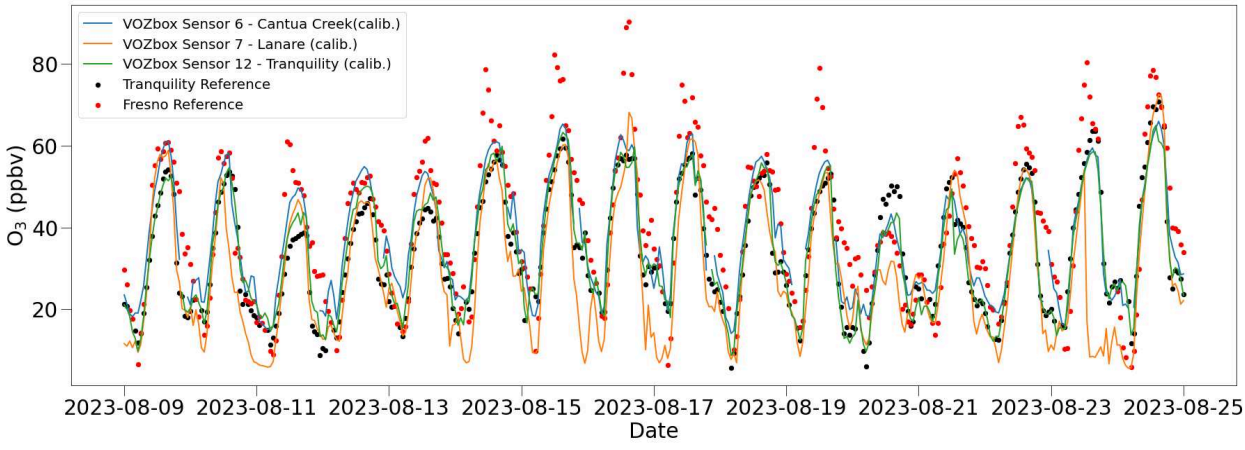


Figure 3.5.1.1.1 Time series scatter for cluster 1 comprising the Tranquility reference monitors and VOZboxes 6 (Cantua Creek), 7 (Lanare), 12 (Tranquility) for the time period of Aug 09, 2023 to Aug 25, 2023.

3.5.1.2. Cluster 2: (Coalinga, Avenal, Kettleman City)

This cluster is centered around the westernmost region of the Central Valley, a zone that is prone to forest fires but there were no fires during the measurement period. This cluster lies in the northerly wind downstream direction, potentially impacting ozone concentrations in these specific locations. VOZbox 2 (Coalinga) is the closest to the mountains, VOZbox 5 (Avenal) is located south of Coalinga, and VOZbox 10 (Kettleman City) is located east of Avenal thus creating a three VOZbox cluster. Figure 3.5.1.1.2, which displays the hourly ozone concentrations for the period from August 11 to August 22, 2023, clearly demonstrates that the ozone concentrations in Coalinga are significantly lower than those in the other two VOZboxes. The measurements at the other two sites also show a significant spatial gradient, with the peak hourly ozone concentrations at Coalinga significantly lower than those at the other two sites, varying by nearly 20 ppbv, and the Kettleman City values marginally higher than the Avenal values, indicating an increase in ozone concentrations towards the south. The number of non-compliance days recorded is a good indicator. Coalinga and Kettleman City had zero noncompliance days, whereas Avenal had 3 non-compliance days. Whereas comparing the number of days above 60 ppbv, Coalinga had only 1 day, and Kettleman City had 9 days, whereas Avenal had 17 days. This gives us a good case to state that the wind patterns observed in the Central Valley are having an impact on the concentration of ozone while traveling down south.

When predicting MDA8 values below 70 ppbv, the VOZbox performs well as the calibration holds good in that range of values, as showcased by the VOZbox 1 and 12. The values in this

cluster significantly fall below the compliance limit for the most part, as evidenced by the spatial plots in Figures 3.6.3.1 and 3.6.3.2 for the period when Fresno recorded high ozone days. Thus we can confidently say that this cluster had lower ozone exposure than the metropolitan locations where we expect higher ozone days. The absence of a reference monitor at this location complicates our ability to ensure the actual values this location encounters. Considering the range of values the cluster encounters and with the confidence of our VOZboxes, we can conclude that the predicted values closely match the actual values.

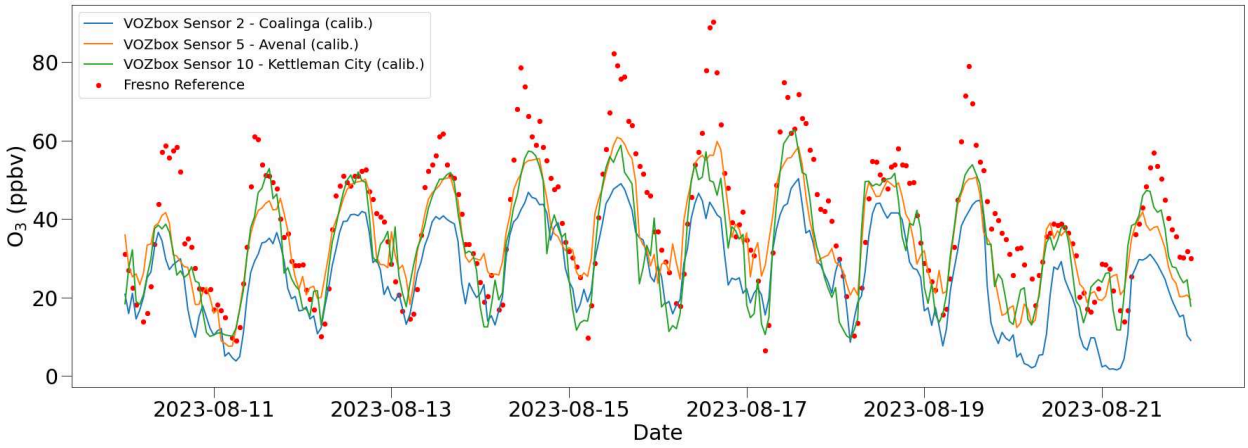


Figure 3.5.1.1.2 Time series scatter for cluster 2 comprising of the VOZboxes 2 (Coalinga), 5 (Avenal), and 10 (Kettleman City) for the time period of Aug 11, 2023 to Aug 22, 2023.

3.5.2 Paired Analysis

Units 1 and 12 are located at type A sites; therefore, they are excluded from evaluation in this section. The units that are located with the nearest reference monitor within a 5–20 mile radius are units 3 (Terra Bella), 4 (Cutler), 7 (Lanare), 6 (Cantua Creek), and 11 (Taft). The calibrated ozone results are compared with those from the nearest monitor to assess the monitor's prediction and to see if there is any spatial variation, as ozone is a spatial pollutant that exhibits variation over a specific distance during a defined timeframe. A link exists between the device's hourly data and the reference values. A distinct divergence in elevated ozone concentrations is apparent

in this paired analysis, evident across all tested units and confirmed by the sensor evaluations conducted for units 1 and 12. The difficulty of this evaluation arose when there was the absence of reference sites in the western rural areas, making the assessment considerably more challenging than for other units. However, given that the other units consistently predict ozone values and that the theoretically obtained values closely align with the measured values, there is substantial evidence that the sensors accurately capture values near the actual measurements.

The units at Terra Bella and Cutler, located in the eastern part of the Central Valley, experienced high ozone concentrations, as indicated by the reference monitors available there and the recorded 10 and 12 non-compliance days respectively. The nearest reference for Terra Bella VOZbox was Porterville (distance between ~6 miles), and for the VOZbox at Cutler was Parlier and Visalia. Figure 3.5.2.2 a) shows, the scatter plot between the VOZbox 3 and the reference data from Porterville; we can see there is a linear relationship that exists between them, but there is a deviation at higher ozone values.

This indicates two things: one, there is spatial variability between the two sites, and two, the VOZbox might be underpredicting some ozone values or a mix of both. Figure 3.5.2.1 c)shows the scatter plot between the VOZbox 4 and the reference measurements at the Parlier site (distance between ~ 12 miles) and the Visalia site (distance between ~ 15 miles); this also had a similar characteristic as the VOZbox 3, but the linear correlation here is lower, indicating that there is a higher varied spatial variability amongst these measuring locations. Both the comparisons yielded Visalia (R^2 = 0.18) having less correlation than Parlier (R^2 = 0.35) but the data spread is similar for both sites.

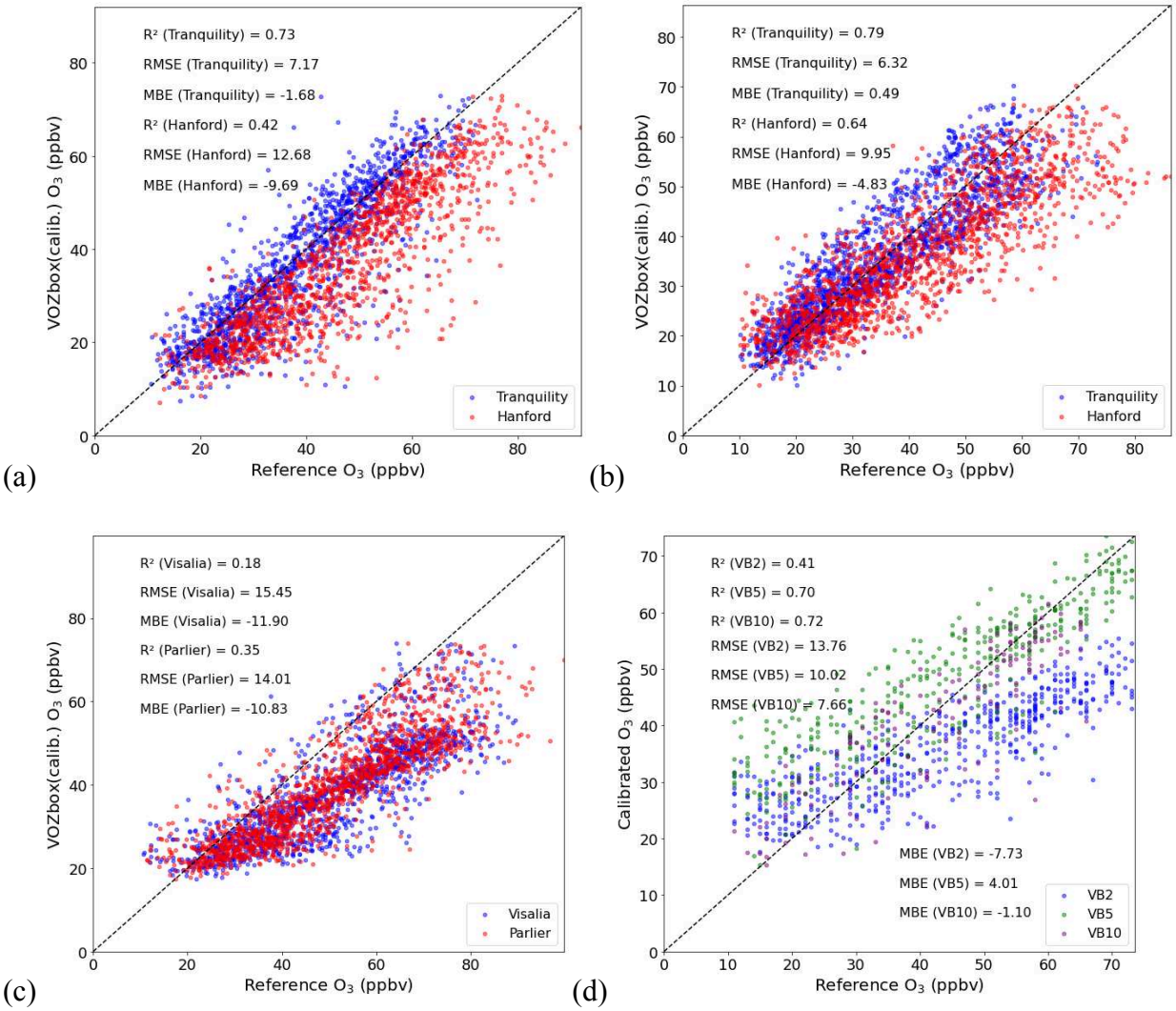


Figure 3.5.2.1 a) Scatter plot between VOZbox 6 and reference monitor at Tranquility/Hanford. b) Scatter plot between VOZbox 7 and reference monitor at Tranquility/Hanford. c) Scatter plot between VOZbox 4 and reference monitor at Visalia/Parlier. d) Scatter plot between VOZbox 2,5,10 and reference monitor at Santa Rosa.

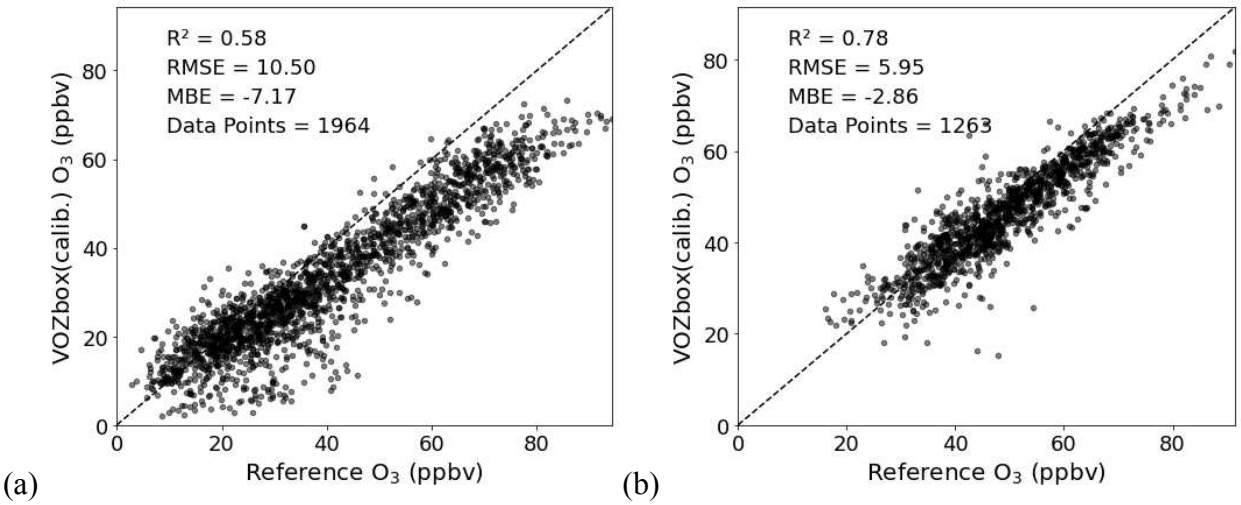


Figure 3.5.2.2 a) Scatter plot between VOZbox 3 and reference monitor at Porterville. a) Scatter plot between VOZbox 11 and reference monitor at Taft.

The VOZbox in the Tranquility clusters was compared individually with the reference monitors, where Figure 3.5.2.1 a) is the scatter plot between the VOZbox 6 at Cantua Creek against the Tranquility reference site (distance between ~12 miles) and Hanford reference site (distance between ~ 32 miles) with Tranquility being the closer site; we can see that the linear correlation is very high for Tranquility (R^2 = 0.73) but low for Hanford (R^2 = 0.42), confirming that there is spatial variability and Hanford has higher concentrations which is located in the south of VOZbox 6.

Similar patterns can be found for the VOZbox 7 at Lanare with the Tranquility reference monitor (distance between ~ 13 miles) and Hanford reference monitor(distance between ~ 27 miles) as well in Figure 3.6.2.1 b). The results demonstrate a linear correlation, with a high R^2 value of 0.79 for Tranquility and at Hanford, which is distant compared to the Tranquility, from VOZbox 7. They have a lower linear correlation of 0.64.

The monitors located in the southern part of the Central Valley include the VOZbox 11 at Taft and the nearest reference monitor at Maricopa (distance from ~ 6 miles). Due to its late

deployment, this VOZbox has less data than the other units. The scatter for ozone concentrations between these two sites is depicted in Figure 3.5.2.2 b). The plot indicates a strong correlation between the sensor and the reference, with an R^2 value of 0.78, and a minimal deviation at higher ozone values. This suggests that the proximity of the two sites is a crucial factor in understanding the spatial gradient.

The monitors situated in the westernmost part of the Central Valley were not in proximity to any reference monitors; the nearest monitor is approximately 30 miles away from the VOZbox 2. Figures 3.6.2.5 d) illustrate the scatter between the VOZbox 2, VOZbox 5, and VOZbox 10 with the reference measurement at Santa Rosa. The values from Coalinga are significantly lower than those from Avenal, which is closer to Santa Rosa. The spatial gradient is evident in these two plots, providing tangible proof of the developed spatial trends and their significance concerning distances between the VOZbox and the reference monitor.

3.5.3. Spatial Maps - Existing Network

In the San Joaquin Valley, a network of regulatory sensors is placed to capture the temporal and spatial variability of various pollutants. These are mostly located on the CA -99 highway which has densely populated cities all along, thus the sensor locations make sense to capture the urban pollutants such as ozone and the emissions from the vehicles that travel on the highway, as they are a major precursor for O_3 pollution. However, these sensors fail to capture trends in the rural-agricultural areas that are located on the far west side of the valley and other interior remote regions.

With the current set of reference sensors, we observe that months from June to August have higher ozone concentrations, and from Figure 3.6.3.1 a) we can see that July has recorded the

highest ozone monthly average of 76 ppbv, which is due to the peak summer temperature which has accelerated the formation of ozone. The cities of Modesto, Merced, Fresno, Visalia, and Bakersfield all lie on the same highway and they consistently have monthly-averaged values of about 65 ppbv in July indicating a high urban value due to the presence of precursors that are produced internally. Whereas locations such as Tranquility have lower values compared to the other cities, indicating that the ozone concentrations are lower in the western part of the Valley.

The far east valley near the mountains where there is a high chance of ozone getting captured is seen vividly in our data as in Figure 3.6.3.1 a) Sequoia National Park located far east has recorded the maximum number of non-compliance days among all the locations in the Central Valley from the regulatory monitors. The northern winds transport the precursors of the ozone into the eastern and southern regions by entering into the valley from the western zone near the Bay Area and moving down south thus increasing the ozone concentrations significantly (Jin et al., 2013; Pun et al., 2000). This trend can be seen in Figure 3.6.3.1 a), where all the months have similar spatial trends to support this movement of the precursors.

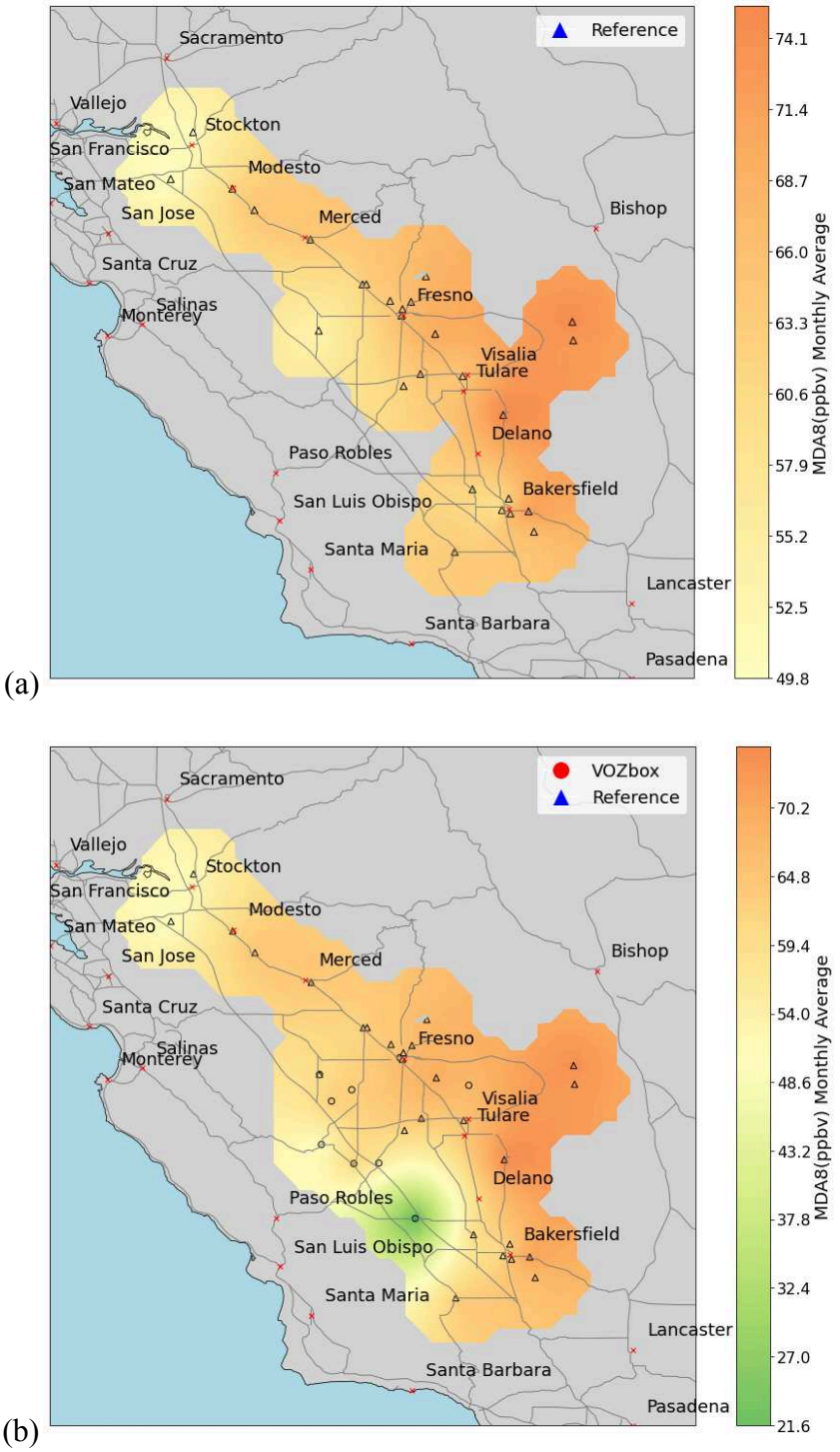
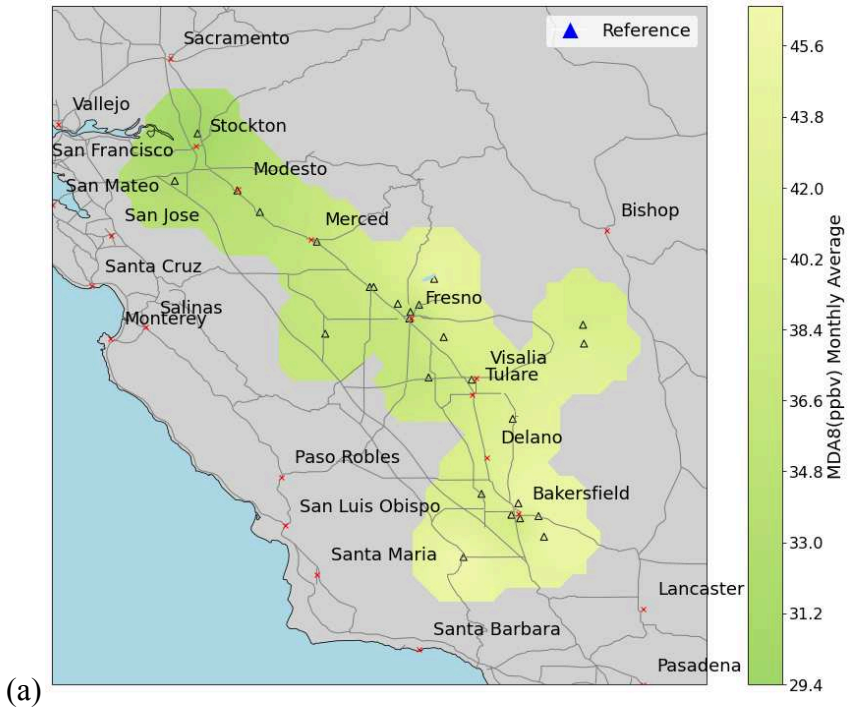


Figure 3.6.3.1 a) Kriged map of the Central Valley using the reference monitors(represented by triangles) available in the locality for July 2023. b) Kriged map of the Central Valley using the reference monitors and the newly deployed VOZboxes (represented by circles) available in the locality for July 2023.

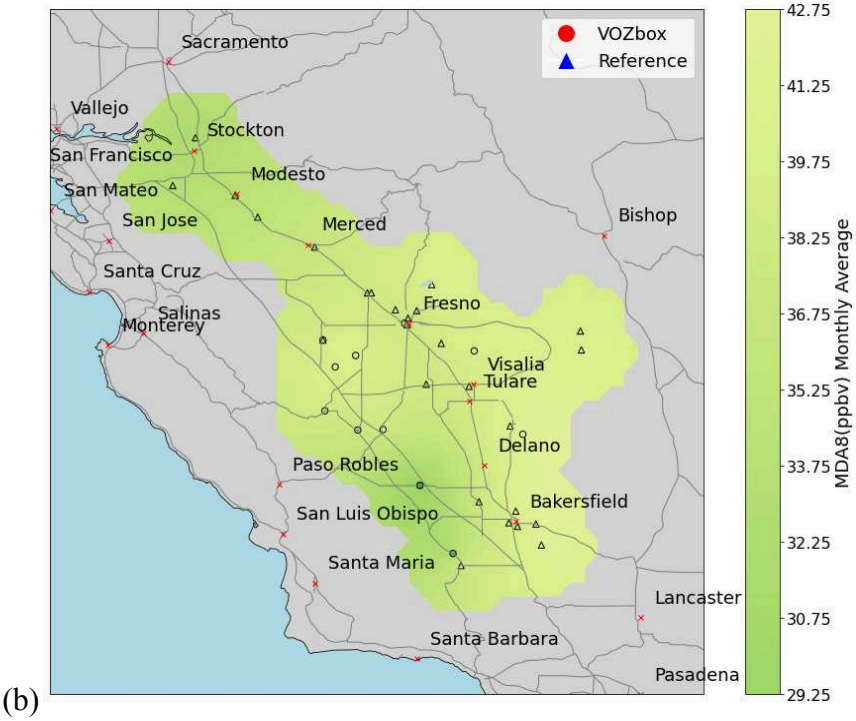
3.5.4 Spatial Maps - Enhanced Network with VOZboxes

With the addition of 11 new VOZboxes which are spread across the SJV Central Valley, we enhance the spatial gradient map, as there are new points of data that were previously not available, thus giving us an idea of how the urban-rural gradient is present in the valley. The rural parts of the valley never had access to monitoring equipment thus it was assumed that they were exposed to unknown amounts of ozone pollution, From this monitoring effort the Figure 3.6.3.1 b) showcasing the month with the highest ozone concentrations in the Central Valley has made it clear that the ozone concentration in these regions is lower than in the urban zones that were previously monitored and they have fewer non-compliance days compared to the nearest urban city which was Fresno in the north and Bakersfield in the south.

The VOZboxes indicate that cities like Coalinga, Avenal, and Kettleman City have lower concentrations and the average MDA8 value of each month is lower than Fresno or the nearest monitor which was Tranquility. The sensor down south at Taft was reading values higher than these and the monitors in the east which were Cutler and Terra Bella consistently had higher ozone values than the sensor at the Fresno Garland indicating the higher ozone trend at the eastern region due to the trapping of ozone, the metropolitan regions which had a consistent production of NOx were having the ozone concentration in control due to the ozone titration, but the areas which had the movement of the precursors, had exposure to higher ozone pollution.



(a)



(b)

Figure 3.6.3.2 a) Kriged map of the Central Valley using the reference monitors(represented by triangles) available in the locality for November 2023. b) Kriged map of the Central Valley using the reference monitors and the newly deployed VOZboxes (represented by circles) available in the locality for November 2023.

Figure 3.6.3.3 indicates the spatial distribution on four different days with non-compliance days recorded at the Fresno- Garland site; these four days were picked based on recording very high ozone concentrations in the valley. All these days were in the month of July and they were subsequent days indicating the effect of ozone season peaks in the month of July also the ozone concentrations are present a lot more in the eastern parts of the valley than the western parts as they have lower concentrations than them. Also, there is a clear view of the hotspots for ozone from these maps like Fresno, Bakersfield, and Delano which are big cities along the CA99 highway.

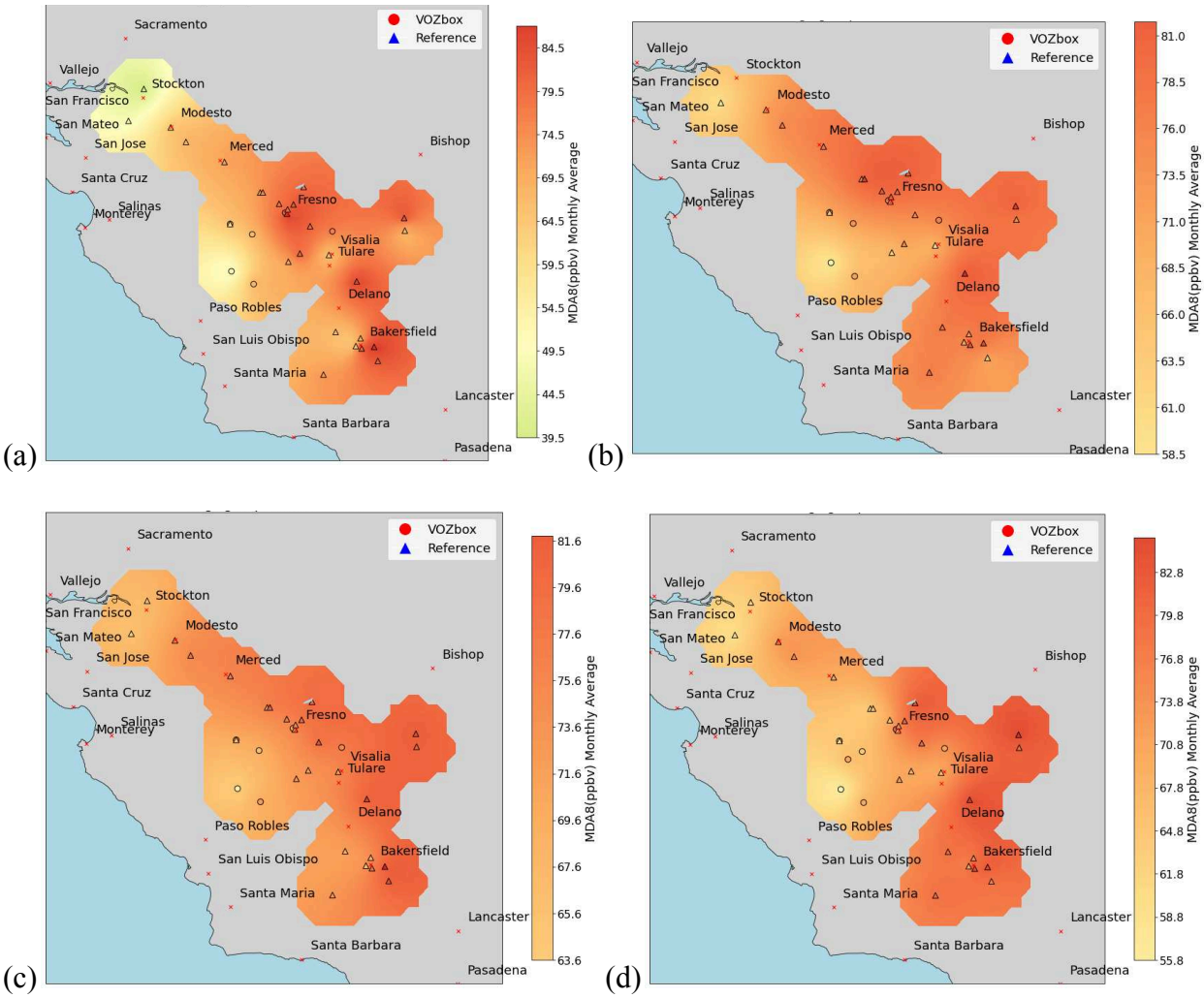


Figure 3.6.3.3 Ozone concentrations over the SJV for four different days with the non-compliance recorded in Fresno-Garland site on the dates a) 07/15/2023 b) 07/16/2023 c)

07/21/2023 d) 7/22/2023 these were the highest concentration in the measurement campaign carried out.

3.5.5 Temporal Variations Observed

The variation of ozone concentration variations was discussed in detail spatially, but they also vary significantly in terms of time. Over the period of observation, we could clearly see that the values were on the rise from June as the average ozone value of the valley went from 53 ppbv to 60 ppbv in the month of July where it peaked, it recorded 10-plus days of noncompliance across the central valley through the VOZbox network that was in place. The 16th of July was the most noncompliant day of all, as the VOZboxes recorded 7 noncompliant locations from the monitors that were placed. The month of August also has its fair share of noncompliance, with it recording about 5 days of noncompliance across the valley. The month of June was mostly under pre-field calibration, thus the variation was not captured accurately by the network. The values started dropping post-August, and the average value across the valley dropped to 40–50 ppbv, which is better than the high 60s that were visible in the previous months at several locations. Figures S16 and S17 show us the variation of the ozone concentrations in the valley and across the months, thus indicating the change over time with and without the VOZbox network and we can clearly see how VOZboxes influence understanding the ozone concentrations at different locations.

3.6 Location-based observations

The VOZboxes were deployed in various communities and monitored over an extended duration. Some experienced data loss, others encountered sensor malfunctions and were subsequently replaced, and some faced delayed deployment. All factors were considered, resulting in summarized outcomes based on their deployment locations.

Table 3.3 Location-based observation summary.

Site	Observations
Fresno Garland (VOZbox 1)	The sensor stayed at the same location from day 1 of the pre-field calibration phase till the end of the post-field calibration phase. Out of the 16 recorded non-compliance days, the sensor accurately predicted 7, with the remaining days falling within a 5% margin, indicating good performance at this location. Our monitor captured 55 days with an MDA8 value above 60 ppbv, placing it in a very high ozone-polluted zone.
Coalinga (VOZbox 2)	Since there are no nearby reference monitors in the westernmost city where the VOZbox is located, the VOZboxes significantly contribute to community monitoring at this site. The VOZboxes recorded the lowest ozone levels here across the entire Central Valley VOZbox network. The VOZboxes recorded values above 50 ppbv 14 times and once above 60 ppbv, indicating that the city is under compliance and its exposure to ozone is significantly lower than that of other cities.
Terra Bella (VOZbox 3)	Another city in the eastern part of the city has high exposure to higher ozone concentrations, which are revealed by the nearest reference monitors to the VOZboxes. The VOZboxes recorded 15 days had a MDA8 value above 60 ppbv and out of it 10 were non-compliance days recorded. This indicates it is one of the highly ozone-polluted cities in the Central Valley.

Cutler (VOZbox 4)	The ozone levels in this city, which is in the northeast region of the Central Valley, were on par with those in Fresno. The atmosphere primarily traps the ozone, causing this phenomenon. Our device recorded 12 non-compliance days and 24 days above the MDA8 value of 60 ppbv. This region of the Central Valley is exposed to higher values of ozone.
Avenal (VOZbox 5)	This site, located in the west, exhibits values higher than those of Coalinga. Situated south of Coalinga, the VOZboxes recorded 3 non-compliance days and 17 days recording above an MDA8 value of 60 ppbv. This VOZbox adds data to locations that are not monitored, thereby quantifying the ozone concentration in these zones.
Cantua Creek (VOZbox 6)	This location was in close proximity to the Tranquility city. The values from the VOZbox at this location, when compared to the reference in Tranquility, were close and within range. The VOZbox here recorded 2 non-compliance days and 20 days above the MDA8 value of 60 ppbv.
Lanare (VOZbox 7)	This location was also in close proximity to the Tranquility. The VOZbox values at this location were close to the Tranquility reference. The hourly lows at this location were significantly lower than those at the Tranquility site, and the maximum values were in proximity to the reference, suggesting the presence of localized precursors in this city. The VOZbox here recorded 3 non-compliance days and 33 days above the MDA8 value of 60 ppbv.

Lost Hills (VOZbox 8)	The sensors in this city initially experienced malfunctions, yielding either no data or low readings, which led to inaccurate information and considerable downtime even after a change in sensors.
Kettleman City (VOZbox 10)	This site, situated in the west, demonstrates values exceeding those of Coalinga. Located south of Coalinga and east of Avenal, its values vary from 50 to 60 ppbv MDA8 on average, signifying less ozone exposure relative to nearby cities. The VOZbox here recorded 0 non-compliance days and 9 days above the MDA8 value of 60 ppbv. This VOZbox supplements data for unmonitored places, thereby quantifying ozone concentration in these areas.
Taft (VOZbox 11)	This city's VOZbox serves as the southernmost point in the Central Valley within the VOZbox network. The recorded values in August aligned well with the closest reference monitor, Maricopa, despite the sensor's late deployment. The VOZbox here recorded 2 non-compliance days and 18 days above the MDA8 value of 60 ppbv. This, along with the city's close proximity to Bakersfield, implies that the primary source of ozone pollution in this area is the transportation of precursors.
Tranquility (VOZbox 12)	In one of the locations where the VOZbox and a reference monitor coexist, the reference monitor recorded zero noncompliance days in 2023. The VOZbox forecasted two instances of noncompliance, but these were within a 5% error margin. This suggests that the location is in compliance with the standards, and the monitoring efforts here benefit the surrounding

	communities. Additionally, the deployment of the VOZbox here aids in quantifying the sensor's performance in a compliant location.

Table 3.4 Summary of non-compliance days recorded by VOZboxes.

Unit	Location Type	Site Location	Number of non-compliance days
VOZBox_1	Type A	Fresno Garland	7
VOZBox_2	Type B	Coalinga	0
VOZBox_3	Type B	Terra Bella	10
VOZBox_4	Type B	Cutler	12
VOZBox_5	Type B	Avenal	3
VOZBox_6	Type B	Cantua Creek	2
VOZBox_7	Type B	Lanare	3
VOZBox_8	Type B	Lost Hills	0
VOZBox_9	N/A	Not Deployed	N/A
VOZBox_10	Type B	Kettleman City	0
VOZBox_11	Type B	Taft	2
VOZBox_12	Type A	Tranquility	2

3.7 Sensor Failures and Troubleshooting

The 12 VOZboxes that were built and shipped to Fresno had no issues in the initial testing that was carried out on the Powerhouse CSU campus. After we set them up for pre-field calibration, the heat caused some of the sensors to malfunction and shut down. We had to get to the location and charge the battery manually to crank the system up; this happened for 5 to 6 units as the pre-calibration phase was done in the summer and the temperatures had started shooting up and above 30 ℃, and thus the box temperature was reaching above the threshold temperatures. We didn't encounter these issues during our ambient testing in Colorado. We left the sensors that

experienced these failures at the location to collect additional data and then relocated the sensors that had collected data for approximately three weeks.

We deployed some of the sensors late because we were busy obtaining community members' permission to place them there and making sure the locations were safe and free from vandalism. Consequently, we deployed some of the sensors as late as August. After deployment, we continuously monitored the sensors for any abnormalities in the collected data, troubleshooting any issues immediately. VOZbox 6 had insufficient data initially due to a technical fault and it was rectified.

We deployed the VOZbox 8 in the city of Lost Hills, which recorded nearly zero ozone concentrations for several days. However, our troubleshooting efforts yielded no improvements.

The sensors in this city initially experienced malfunctions, yielding either no data or low readings, which led to inaccurate information and considerable downtime. We then exchanged the sensors to resolve the issue. Subsequently, after several days, the sensors once more had significant downtimes and generated readings that deviated from our expectations. This led to a significant failure rate at this site. We expected to observe variations in readings as we progressed southward and compared them with the nearest northern and southern sensors. Unfortunately, we could not perform this analysis because of the inadequate data from this unit. Consequently, we replaced the radiation shield with a VOZbox 9 radiation shield and turned off the faulty sensor. As a result, we reduced the total number of locations from 12 to 11. The replaced sensor functioned well for a few days before the same issue resurfaced; as a result, the data from that site was of poor quality, and the maps revealed a discrepancy with the data displayed by other units nearby.

4. Conclusion, Discussion and Future work

In this work, we developed a low-cost, low-power, autonomous network of sensors (VOZboxes) to measure ozone using the Aeroqual SM50 ozone sensors. We evaluated the long-term performance of the sensor by building a calibration algorithm using a pre-and post-field calibration phase strategy, thereby accounting for drifts that are visible in the sensors and adding corrections to them, also we were able to check the NAAQS capturing capability and assess the spatial variability in the central valley using this network of sensors. We tested the sensors' performance by putting them in type 1 sites and constantly comparing them against reference monitors. The sensors work well when they are properly trained for a sufficient period of high ozone concentrations and allowing them to capture high ozone days for the calibration model, they were able to detect non-compliance with a 10% error margin. This suggests that they are high potential low-cost sensor units that can be used in places where there isn't enough monitoring to improve the spatial data and help with community monitoring to keep people informed. We also analyzed the spatial variability with the presence of these monitors, which broadened the scope of spatial maps, pinpointing locations with unknown concentrations and generating data through statistical techniques such as kriging. We were able to quantify the number of non-compliance days for all the sites and see how ozone has an effect over the course of summer in the SJV.

4.1 Future scope and work

This study shows great potential for future applications, especially in ozone monitoring and the integration of more sensors to assess a wide range of pollutants. In low- and middle-income countries where monitoring costs are beyond their budgets, VOZboxes can fill in the data gap

and offer important information about the level of pollution. Such devices will allow people to better understand their exposure to the environment and will also enable governments to make evidence-based policy decisions.

Generally, VOZboxes can report ozone concentrations within an error margin of 10% and are, therefore, an affordable solution for spatial monitoring. Future improvements involve refining the calibration techniques and the cooling system to maintain sensor performance optimally. An active cooling mechanism could be used to regulate the temperature of the control boxes, which in turn would minimize battery cycling and improve the accuracy of the data. Integration of GPS and particulate matter sensors enables year-round monitoring by shifting the focus to PM concentrations outside of the ozone season.

The aging assessment is needed for the SM50 sensors to achieve long-term reliability. The current VOZboxes within the western cluster are showing inconsistent readings and will need further testing and possible repositioning. Working with CCEJN, a reference-grade monitor is being placed in the area to validate sensor performance. A side-by-side comparison of new and aging sensors will be conducted to determine their depreciation over time and inform future sensor deployment. Finally, ozone monitoring in high-exposure areas will provide information on long-term trends and mitigation strategies.

4.2 Environmental Justice Works

Other concerns about environmental justice include air quality monitoring site location, which continues to place a disproportionate amount of sites in rural areas, underrepresenting both urban and rural communities. We partnered with CCEJN to better communicate ozone exposure to the public with VOZbox data from the agricultural communities of Kettleman City, Terra Bella,

Lanare, and Cantua Creek. Community-specific visuals and outreach events were performed to communicate air quality results and suggest ways of reducing exposure.

In an effort to further increase access, we developed a Tableau Public dashboard of the campaign data visualized through interactive visuals and maps showing ozone pollution trends. The platform provides residents with real-time data, enabling awareness and motivating local actions toward the risk of exposure.

4.3 Discussion

One of the important objectives of this study is the integration of monitoring techniques by including regulatory-grade monitors, low-cost sensors, and satellite data into one robust air quality surveillance network. The regulatory monitors serve as a benchmark for sensor comparison due to precise calibration and frequent maintenance. These networks are supplemented by low-cost sensors that fill in spatial gaps and enhance the assessment of pollutant distribution. VOZboxes, with their mobile network connectivity, support real-time data streaming to allow for immediate policy and public health responses.

Despite the many advantages, low-cost sensors also suffer from environmental susceptibility, durability concerns, and limitations in data accuracy. Improvement in sensor design and machine learning for calibration will continue to improve these issues and long-term performance. Forecasting ozone levels from existing datasets is a valuable next step in extending the utility of these sensors.

The price disparity between regulatory monitors and low-cost sensors stands as a barrier, especially in rural areas. Their integration into established monitoring frameworks must not come at the cost of data integrity for wide-scale adoption.

This study underlines how networks of low-cost ozone sensors may support decisions by policymakers, urban planners, and authorities responsible for the protection of public health. Advanced tools are realized for the implementation of ozone pollution control, healthier environments, and enabling sustainable policies about the natural environment. Standardization of the QA/QC process is further required to reach higher reliability from sensors, at which point mass deployment will also be possible.

In the future, interdisciplinary collaboration and technological innovation will continue to be at the forefront of solving air quality challenges. Harmonization of various monitoring techniques, together with real-time data analytics, allows a holistic and cost-effective approach to mitigating air pollution. This study opens up avenues for the use of low-cost ozone sensors in global environmental policy to ensure that current and future generations have wider access to air quality data for decision-making.

5. References

Araki, S., Yamamoto, K., & Kondo, A. (2015). Application of Regression Kriging to Air Pollutant Concentrations in Japan with High Spatial Resolution. *Aerosol and Air Quality Research*, *15*(1), 234–241. https://doi.org/10.4209/aaqr.2014.01.0011

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., & Modzel, P. (2022). Low- and Medium-Cost Sensors for Tropospheric Ozone Monitoring—Results of an Evaluation Study in Wrocław, Poland. *Atmosphere*, *13*(4), Article 4. https://doi.org/10.3390/atmos13040542

Bourgeois, I., Peischl, J., Neuman, J. A., Brown, S. S., Thompson, C. R., Aikin, K. C., Allen, H. M., Angot, H., Apel, E. C., Baublitz, C. B., Brewer, J. F., Campuzano-Jost, P., Commane, R., Crounse, J. D., Daube, B. C., DiGangi, J. P., Diskin, G. S., Emmons, L. K., Fiore, A. M., … Ryerson, T. B. (2021). Large contribution of biomass burning emissions to ozone throughout the global remote troposphere. *Proceedings of the National Academy of Sciences*, *118*(52), e2109628118. https://doi.org/10.1073/pnas.2109628118

Candiani, G., Carnevale, C., Pisoni, E., & Volta, M. (2010). Assimilation of Chemical Ground Measurements in Air Quality Modeling. In I. Lirkov, S. Margenov, & J. Waśniewski (Eds.), *Large-Scale Scientific Computing* (pp. 157–164). Springer. https://doi.org/10.1007/978-3-642-12535-5_17

Cheadle, L., Deanes, L., Sadighi, K., Gordon Casey, J., Collier-Oxandale, A., & Hannigan, M. (2017). Quantifying Neighborhood-Scale Spatial Variations of Ozone at Open Space and Urban Sites in Boulder, Colorado Using Low-Cost Sensor Technology. *Sensors*, *17*(9), 2072. https://doi.org/10.3390/s17092072

Cho, K., Tiwari, S., Agrawal, S. B., Torres, N. L., Agrawal, M., Sarkar, A., Shibato, J., Agrawal, G. K., Kubo, A., & Rakwal, R. (2011). Tropospheric ozone and plants: Absorption,

responses, and consequences. *Reviews of Environmental Contamination and Toxicology*, *212*, 61–111. https://doi.org/10.1007/978-1-4419-8453-1_3

de Nazelle, A., & Serre, M. L. (2006). Ozone Exposure Assessment in North Carolina Using Bayesian Maximum Entropy Data Integration of Space Time Observations and Air Quality Model Prediction. *Epidemiology*, *17*(6), S189.

Dewan, S., Bamola, S., & Lakhani, A. (2024). Addressing ozone pollution to promote United Nations sustainable development goal 2: Ensuring global food security. *Chemosphere*, *347*, 140693. https://doi.org/10.1016/j.chemosphere.2023.140693

Di, Q., Rowland, S., Koutrakis, P., & Schwartz, J. (2017). A hybrid model for spatially and temporally resolved ozone exposures in the continental United States. *Journal of the Air & Waste Management Association*, *67*(1), 39–52. https://doi.org/10.1080/10962247.2016.1200159

Donzelli, G., & Suarez-Varela, M. M. (2024). Tropospheric Ozone: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. *Atmosphere*, *15*(7), Article 7. https://doi.org/10.3390/atmos15070779

Duvall, R., Long, R., Beaver, M., Kronmiller, K., Wheeler, M., & Szykman, J. (2016). Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide. *Sensors*, *16*(10), 1698. https://doi.org/10.3390/s16101698

Emberson, L. (2020). Effects of ozone on agriculture, forests and grasslands. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *378*(2183), 20190327. https://doi.org/10.1098/rsta.2019.0327

Filippidou, E., & Koukouliata, A. (2011). Ozone effects on the respiratory system. *Progress in Health Sciences*.

https://www.semanticscholar.org/paper/Ozone-effects-on-the-respiratory-system-Filippidou-Koukouliata/1b6f79a792c30920b5383934efd5c712e6a10946

Finlayson-Pitts, B. J., & Pitts, J. N. (1997). Tropospheric air pollution: Ozone, airborne toxics, polycyclic aromatic hydrocarbons, and particles. *Science (New York, N.Y.)*, *276*(5315), 1045–1052. https://doi.org/10.1126/science.276.5315.1045

Fresno-Madera-Hanford, CA. (n.d.). Retrieved December 5, 2024, from https://www.lung.org/research/sota/city-rankings/msas/fresno-madera-hanford-ca

Giardina, D. M. (n.d.). *AUTONOMOUS LOW-COST OZONE SENSORS: DEVELOPMENT, CALIBRATION, AND APPLICATION TO STUDY EXPOSURE AND SPATIAL GRADIENTS*.

Hubbell, B. J., Hallberg, A., McCubbin, D. R., & Post, E. (2005). Health-Related Benefits of Attaining the 8-Hr Ozone Standard. *Environmental Health Perspectives*, *113*(1), 73–82. https://doi.org/10.1289/ehp.7186

Ito, K., De Leon, S. F., & Lippmann, M. (2005). Associations between ozone and daily mortality: Analysis and meta-analysis. *Epidemiology (Cambridge, Mass.)*, *16*(4), 446–457. https://doi.org/10.1097/01.ede.0000165821.90114.7f

Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., & Thun, M. (2009). Long-term ozone exposure and mortality. *The New England Journal of Medicine*, *360*(11), 1085–1095. https://doi.org/10.1056/NEJMoa0803894

Jin, L., Loisy, A., & Brown, N. J. (2013). Role of meteorological processes in ozone responses to emission controls in California's San Joaquin Valley. *Journal of Geophysical Research: Atmospheres*, *118*(14), 8010–8022. https://doi.org/10.1002/jgrd.50559

Keller, R. (1992). [Effects of tropospheric ozone on human respiratory organs]. *Schweizerische*

Rundschau Fur Medizin Praxis = Revue Suisse de Medecine Praxis. https://www.semanticscholar.org/paper/%5BEffects-of-tropospheric-ozone-on-human-respiratory-Keller/f9e4df3c72bd77ba3c7c04a29b7416f41dc637a3

Kim, C. S., Alexis, N. E., Rappold, A. G., Kehrl, H., Hazucha, M. J., Lay, J. C., Schmitt, M. T., Case, M., Devlin, R. B., Peden, D. B., & Diaz-Sanchez, D. (2011). Lung function and inflammatory responses in healthy young adults exposed to 0.06 ppm ozone for 6.6 hours. *American Journal of Respiratory and Critical Care Medicine*, *183*(9), 1215–1221. https://doi.org/10.1164/rccm.201011-1813OC

Kumar, A., Howard, C. J., Derrick, D., Malkina, I. L., Mitloehner, F. M., Kleeman, M. J., Alaimo, C. P., Flocchini, R. G., & Green, P. G. (2011). Determination of Volatile Organic Compound Emissions and Ozone Formation from Spraying Solvent‐based Pesticides. *Journal of Environmental Quality*, *40*(5), 1423–1431. https://doi.org/10.2134/jeq2009.0495

Ledley, T. S., Sundquist, E. T., Schwartz, S. E., Hall, D. K., Fellows, J. D., & Killeen, T. L. (1999). Climate change and greenhouse gases. *Eos, Transactions American Geophysical Union*, *80*(39), 453–458. https://doi.org/10.1029/99EO00325

Levine, J., Cofer, I. W. R., Cahoon, J. D. R., & Winstead, E. (1995). Biomass Burning: A Driver for Global Change! *Environmental Science & Technology*. https://www.semanticscholar.org/paper/Biomass-Burning%3A-A-Driver-for-Global-Change!-Levine-Cofer/f8b13518bd548ea7e243cccb701933a5cf36c32b

Malig, B. J., Pearson, D. L., Chang, Y. B., Broadwin, R., Basu, R., Green, R. S., & Ostro, B. (2016). A Time-Stratified Case-Crossover Study of Ambient Ozone Exposure and Emergency Department Visits for Specific Respiratory Diagnoses in California

(2005–2008). *Environmental Health Perspectives*, *124*(6), 745–753. https://doi.org/10.1289/ehp.1409495

Miranda, M. L., Edwards, S. E., Keating, M. H., & Paul, C. J. (2011). Making the Environmental Justice Grade: The Relative Burden of Air Pollution Exposure in the United States. *International Journal of Environmental Research and Public Health*, *8*(6), 1755–1771. https://doi.org/10.3390/ijerph8061755

Miskell, G., Alberti, K., Feenstra, B., Henshaw, G. S., Papapostolou, V., Patel, H., Polidori, A., Salmond, J. A., Weissert, L., & Williams, D. E. (2019). Reliable data from low cost ozone sensors in a hierarchical network. *Atmospheric Environment*, *214*, 116870. https://doi.org/10.1016/j.atmosenv.2019.116870

Naik, V., Mauzerall, D., Horowitz, L., Schwarzkopf, M. D., Ramaswamy, V., & Oppenheimer, M. (2005). Net radiative forcing due to changes in regional emissions of tropospheric ozone precursors. *Journal of Geophysical Research: Atmospheres*, *110*(D24), 2005JD005908. https://doi.org/10.1029/2005JD005908

Nalakurthi, N. V. S. R., Abimbola, I., Ahmed, T., Anton, I., Riaz, K., Ibrahim, Q., Banerjee, A., Tiwari, A., & Gharbia, S. (2024). Challenges and Opportunities in Calibrating Low-Cost Environmental Sensors. *Sensors*, *24*(11), 3650. https://doi.org/10.3390/s24113650

Pfannerstill, E. Y., Arata, C., Zhu, Q., Schulze, B. C., Woods, R., Seinfeld, J. H., Bucholtz, A., Cohen, R. C., & Goldstein, A. H. (2023). Volatile organic compound fluxes in the agricultural San Joaquin Valley – spatial distribution, source attribution, and inventory comparison. *Atmospheric Chemistry and Physics*, *23*(19), 12753–12780. https://doi.org/10.5194/acp-23-12753-2023

Pun, B. K., Louis, J.-F., Pai, P., Seigneur, C., Altshuler, S., & Franco, G. (2000). Ozone

Formation in California's San Joaquin Valley: A Critical Assessment of Modeling and Data Needs. *Journal of the Air & Waste Management Association*. https://www.tandfonline.com/doi/abs/10.1080/10473289.2000.10464140

Pusede, S. E., Gentner, D. R., Wooldridge, P. J., Browne, E. C., Rollins, A. W., Min, K.-E., Russell, A. R., Thomas, J., Zhang, L., Brune, W. H., Henry, S. B., DiGangi, J. P., Keutsch, F. N., Harrold, S. A., Thornton, J. A., Beaver, M. R., St. Clair, J. M., Wennberg, P. O., Sanders, J., … Cohen, R. C. (2014). On the temperature dependence of organic reactivity, nitrogen oxides, ozone production, and the impact of emission controls in San Joaquin Valley, California. *Atmospheric Chemistry and Physics*, *14*(7), 3373–3395. https://doi.org/10.5194/acp-14-3373-2014

Rentschler, J., & Leonova, N. (2023). Global air pollution exposure and poverty. *Nature Communications*, *14*(1), 4432. https://doi.org/10.1038/s41467-023-39797-4

Sadighi, K., Coffey, E., Polidori, A., Feenstra, B., Lv, Q., Henze, D. K., & Hannigan, M. (2018). Intra-urban spatial variability of surface ozone in Riverside, CA: Viability and validation of low-cost sensors. *Atmospheric Measurement Techniques*, *11*(3), 1777–1792. https://doi.org/10.5194/amt-11-1777-2018

Schultz, M. G., Schröder, S., Lyapina, O., Cooper, O. R., Galbally, I., Petropavlovskikh, I., von Schneidemesser, E., Tanimoto, H., Elshorbany, Y., Naja, M., Seguel, R. J., Dauert, U., Eckhardt, P., Feigenspan, S., Fiebig, M., Hjellbrekke, A.-G., Hong, Y.-D., Kjeld, P. C., Koide, H., … Zhiqiang, M. (2017). Tropospheric Ozone Assessment Report: Database and metrics data of global surface ozone observations. *Elementa: Science of the Anthropocene*, *5*, 58. https://doi.org/10.1525/elementa.244

Sensors—Ozone. (2022, August 17). Aeroqual Support Center.

https://support.aeroqual.com/Wiki/O3_Sensor_Head_Information

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., & Preuss, P. W. (2013). The changing paradigm of air pollution monitoring. *Environmental Science & Technology*, *47*(20), 11369–11377. https://doi.org/10.1021/es4022602

Steinle, S., Reis, S., & Sabel, C. E. (2013). Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *The Science of the Total Environment*, *443*, 184–193. https://doi.org/10.1016/j.scitotenv.2012.10.098

Strosnider, H., Kennedy, C., Monti, M., & Yip, F. (2017). Rural and Urban Differences in Air Quality, 2008–2012, and Community Drinking Water Quality, 2010–2015—United States. *MMWR. Surveillance Summaries*, *66*(13), 1–10. https://doi.org/10.15585/mmwr.ss6613a1

Sun, G., McLaughlin, S. B., Porter, J. H., Uddling, J., Mulholland, P. J., Adams, M. B., & Pederson, N. (2012). Interactive influences of ozone and climate on streamflow of forested watersheds. *Global Change Biology*, *18*(11), 3395–3409. https://doi.org/10.1111/j.1365-2486.2012.02787.x

Weigel, H. J., Bergmann, E., & Bender, J. (2015). Plant-Mediated Ecosystem Effects of Tropospheric Ozone. In U. Lüttge & W. Beyschlag (Eds.), *Progress in Botany: Vol. 76* (pp. 395–438). Springer International Publishing. https://doi.org/10.1007/978-3-319-08807-5_15

Weissert, L., Alberti, K., Miles, E., Miskell, G., Feenstra, B., Henshaw, G. S., Papapostolou, V., Patel, H., Polidori, A., Salmond, J. A., & Williams, D. E. (2020). Low-cost sensor

networks and land-use regression: Interpolating nitrogen dioxide concentration at high temporal and spatial resolution in Southern California. *Atmospheric Environment*, *223*, 117287. https://doi.org/10.1016/j.atmosenv.2020.117287

White, C. G. (2020). Reframing Air Pollution as a Public Health Crisis in California's San Joaquin Valley. *Case Studies in the Environment*, *4*(1), 1–9. https://doi.org/10.1525/cse.2020.sc.965681

Yates, E. L., Iraci, L. T., Tarnay, L. W., Burley, J. D., Parworth, C., & Ryoo, J.-M. (2020). The effect of an upwind non-attainment area on ozone in California's Sierra Nevada Mountains. *Atmospheric Environment*, *230*, 117426. https://doi.org/10.1016/j.atmosenv.2020.117426

6. Appendix: Supplementary Information

(a)

Date:	6/12/2023			Number of Units:	1
	Total Cost:	1107.9856		Unit Price per MOOS A:	0

Component	Item	PN	Description	Quantity per VOZBox	Unit Cost
Fasteners	Bolt U 5/16" x 2" x 3.69	51614	Long U bolt (holding box/solar panel on to pole)	3	3.39
x	Bolt U 5/16" x 2" x 3"	51610	Short U bolt (holding solar panel to pole)	1	2.79
x	M6 - 1.00 x 8mm	91280A060	Mount solar panel on to 8020 components	4	0.23
x	M5 - 0.8 x 6mm	94500A296	Mount box to 8020 components	4	0.2
x	4-40 X 3/8 (uses 3/32" Hex Wrench)	92949A108	Bolt to mount charge controlled to backing panel	2	0.69
x	4-40 Stainless steel stop nut	94407A100	Lock nut for charge controller bolt	2	0.3
x	M4 X 35 0.7 pitch	92095A199	Bolt for battery holder	2	0.4
x	M4 0.7 pitch Stop nut	90576A103	Lock nut for battery holder	2	0.65
x	8mm O.D. 4.3 mm I.D 15mm L (fits over M4 screw)	99072A213	spacer for battery holder	2	0.47
x	1/4x4 1/2 Grade 5 Hex Cap Screws	91247A560	Long hex bolt for sensor shroud	2	0.89
x	1/4x4 Grade 5 Hex Cap Screws	91247A558	Short hex bolt for sensor shroud	1	0.75
x	1/4-20 Hex Nuts coarse thread	95462A029	Hex Nut for all 1/4" bolts	5	0.11
x	1/4x2 1/2 grade 5 Hex Cap Screws	91247A552	Bolt for PM sensor mount	2	
x	M4 - 0.7 x 16mm	92095A194	Bolt to mount SM50 to sensor mount	2	0.65
x	M4 - 0.7 Locknut	94645A101	Lock nut for SM50 mount	2	0.3
x	8mm O.D. 4.3 mm I.D 5mm L (fits over M4 screw)	93657A039	Spacer for bolt on SM50 mount	2	0.37
8020 Components	25 x 25mm Quarter round T slotted profile (13.39")	25-2527	Solar panel mount	2	6.67
	50 x 12.5mm T slotted profile w/2 adjacent open T slots (9.65")	25-5013	Solar Panel Mount	1	6.74
	M6 x 10mm blue BHCS with slide in economy T nut	75-3404	Solar Panel mount	6	0.72
	25 series "L" structual pivot arm	25-4149	Solar Panel Mount	2	7.46
	M6 slide in economy T nut	3859	Solar Panel Mount	4	0.42
	20 x 20mm T slotted profile 4 open T slots (6.81")	20-2020	Box mount	2	4.49
	M5 slide in economy T nut block	14122	Box mount	4	0.32
	25 series 3 hole - slotted inside corner bracket	25-4250	Mount shroud to pole (upper)	1	6.72
	25 series 4 hole - tall inside corner bracket	25-4115	mount shroud to pole (lower)	1	6.06
Enclosure + Stand	Polycase WQ-57	WQ-57-02	Waterproof enclosure for housing power system anc	1	54.57
	Polycase WQ-91	WQ-91	Mounting feet for connecting enclosure to mounting	1	8.25
	Polycase WQ-57-p01	WQ-57-p01	Backboard for electronics mounting	1	5.01
	Pyle Audio Stand	PSTND2	System stand	1	27.41
Power System	12V, 30 W Solar Panel	B07ML3PSD	Solar Panel	1	59.99
	12V, 9Ah LFP Battery	BLF-1209WS	Battery	1	99.99
	Wanderer 12V Charge Controller	RNG-CTRL-WND10-US	Solar Charge Controller	1	19.99
	Waterproof Polarized 4-wire cable set	744	Wiring to shroud	4	2.5
	12 - 10 AWG 0.25 in. Tab Female Fully Insulated Disconnect	15-155F	quick connect to batteries	2	3.75
	Single core 22 gage wire (Y, W, B, R)	78-12210	wiring electrical components		
Electronics	PCB		Printed circuit board for electrical components		
PCB	Plantower Connector	530470810	connects cable from plantower pm2.5 sensor	1	0.52
	SPS30 Connector	B5B-ZR(LF)(SN)	connects cable from sensyrion pm2.5 sensor	2	0.26
	Temp/RH Connector	282834-4	connects 4 wires from SHT30 t/rh probe	1	2.75
	Moos analog	282834-2	connects data/gnd for sm50, and power/gnd for both far	1	1.27
	RES SMD 10K OHM 1% 1/4W 1206	RC1206FR-0710KL	resistors for pcb	6	0.0301
	FEATHER STACKING HEADERS FML	2830-ND	stacking headers for connecting featherwing	1	1.25
	FEATHER HEADER KIT FML	2886	traditional header for featherwing	1	0.95
	Optional GPS/General I2C JST connector	B4B-PH-K-S(LF)(SN)	Optional GPS connector	1	0.24

(b)

(not on digikey)	SEN55 6 pin 2.54mm pitch terminal header	KF120-6P	SEN 55 header	1	2.15
	RES 39K OHM 1% 1/4W 1206	RC1206FR-0739KL	39kOhm resistor for voltage divider	1	0.1
	RES 68K OHM 1% 1/4W 1206	RC1206FR-0768KL	68 kOhm resistor for voltage divider	1	0.1
DNS	MCU/Cellular Chip	BRN402	cellular data transfer to cloud from sensor	1	59.71
	Temp/RH Probe (SHT30)	4099	collects temp/rh data from within the radiation shield	1	24.95
	ADALOGGER FEATHERWING - RTC + SD	2922	stores data locally on an SD card	1	8.95
	Plantower PM2.5 Sensor	PMS5003T	PM sensor (2)	1	29.95
	Plantower Cable	151340801	connects plantower PM sensor to the pcb	1	6.27
	SM50 Ozone Sensor	SM50 UZ AA	collects ozone data	1	435
	2 pack 40 mm x 10 mm DC 5V Brushless Cooling Fan	YDM4010B05	cools ozone sensor unit within the radiation shield	1	14.38
(not on digikey)	Particle Sensor	SEN55 or SEN50	PM sensor	1	34.65
(not on digikey)	SEN55 cable	CAB-18079	PM sensor connector cable	1	1.31
	GPS breakout	GPS-15210	GPS unit	1	42.95
	GPS/I2c 4 POS CABLE ASSY RECT SKT-SKT	4424	GPS cable	1	0.95
	FAN AXIAL 45X10MM 5VDC WIRE	259-2000-ND	Fan for cooling battery inside box	2	5.06
Shroud/PM Mount	PETG 3D printer filment, 1kg	OVPETG157	Filament for printing rediation shield and housing compo	0.5	21.99
	3/4 - 1-3/4 in. Stainless Steel Hose Clamp	6720595	mount sensor to pole	2	1.98
	Chemical Resistant Slippery PTFE Tube 1" OD 7/8"ID	8457K29	PTFE Sampling tube for inside the shroud	1	20.47
	waterproof conduit for wiring to shroud				
	waterproof cable gland				
3D printed Parts	rad_shield_upper			1	
	rad_shield_lower			1	
	PM sensor mount upper			1	
	PM sensor mount Lower			1	
	PCB mount			1	
	fan holder			1	
	rad shield bottom			1	
	sm50/sht 30 mount			1	
	rad shield top cap				

Figure S1 a) Bill of Materials listing all the required items and costs for the construction of a single VOZbox. b) Continuation of page 1 of the bill of materials.

Figure S2 The Author along side people who work for CARB and CCEJN involved in the pre-calibration collocation effort at the Fresno Garland site on June 07, 2023.

Figure S3 Author alongside a member of CCEJN for the deployment of the VOZbox at Lanare Community Center rooftop.

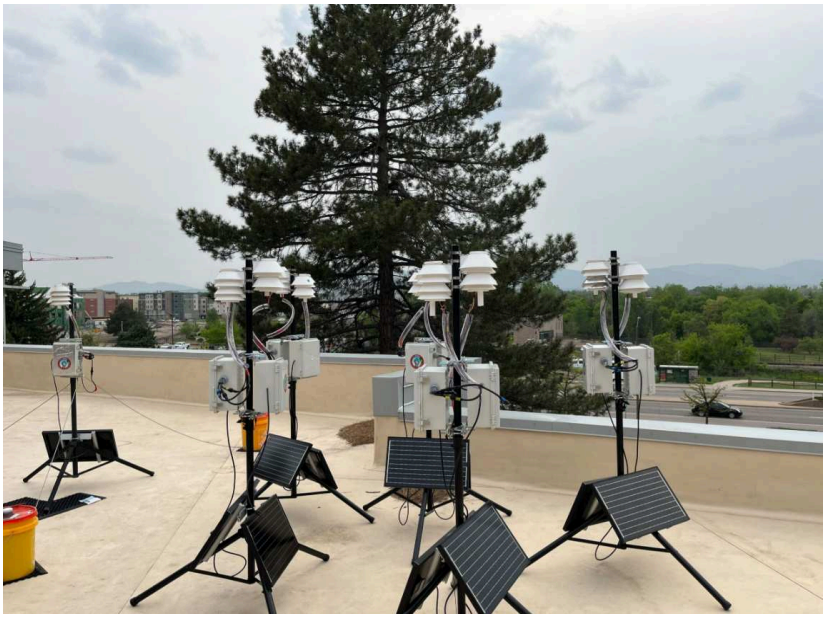
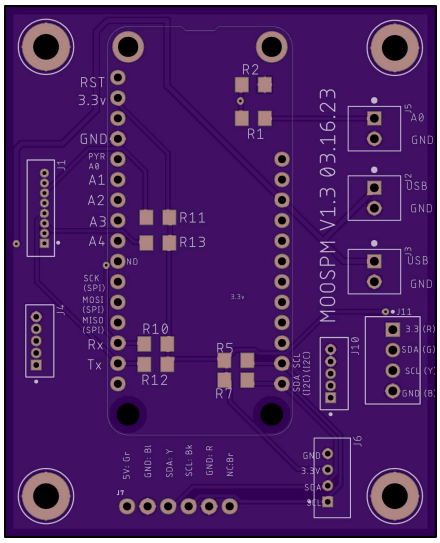


Figure S4 Image Illustrates the rooftop collocation effort done to check the capability of the sensors at the Powerhouse Campus, CSU, Colorado.

(a) 

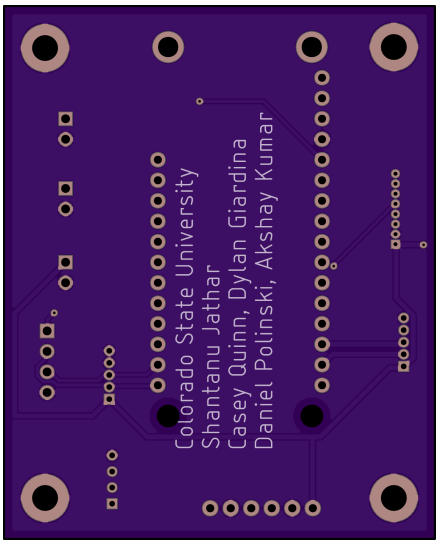
(b) 

Figure S5 a) and b) Illustrate the front and back panel of the custom PCB that has been used in the VOZbox to integrate sensor data collection with the Boron IOT device.

Figure S6 Snippet of the interactive Tableau dashboard used to effectively communicate information for community outreach activities.

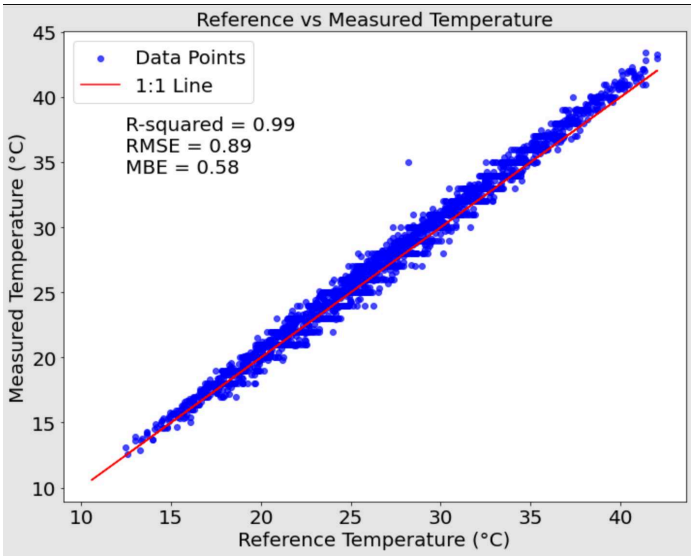


Figure S7 Temperature Scatter between the ambient temperature and the temperature inside the radiation shield.

Table S1: Final calibration values utilized for the VOZboxes

VOZboxes	Temp	RH	Ozone	Week	PM2.5	PM10	Intercept
Unit 1	0.943	0.096	0.713	-0.119	1.198	-1.107	-9.268
Unit 2	0.289	-0.059	0.959	0.370	-2.820	2.619	-0.269
Unit 3	0.492	0.074	0.843	0.018	-0.434	0.442	-9.977
Unit 4	0.778	0.043	0.952	0.213	-0.547	0.504	-11.651
Unit 5	0.529	-0.005	0.708	-0.004	-0.193	0.161	-3.230
Unit 6	1.661	0.238	0.601	-0.108	0.033	-0.001	-30.632
Unit 7	0.451	0.004	1.042	-0.368	0.243	-0.203	-0.710
Unit 8	0.349	-0.014	0.834	0.011	4.685	-4.425	5.258
Unit 9	0.929	0.128	1.591	0.169	1.181	-1.080	-24.502
Unit 10	0.593	0.034	0.766	-0.044	-0.089	0.069	-4.968
Unit 11	0.751	0.098	0.947	-0.220	1.292	-1.218	-7.042
Unit 12	0.521	0.037	0.835	0.262	-0.047	0.079	-8.117

Figure S8 (a) (b) Images of the Tranquility site where the VOZbox 12 was deployed and used as a performance analysis location.

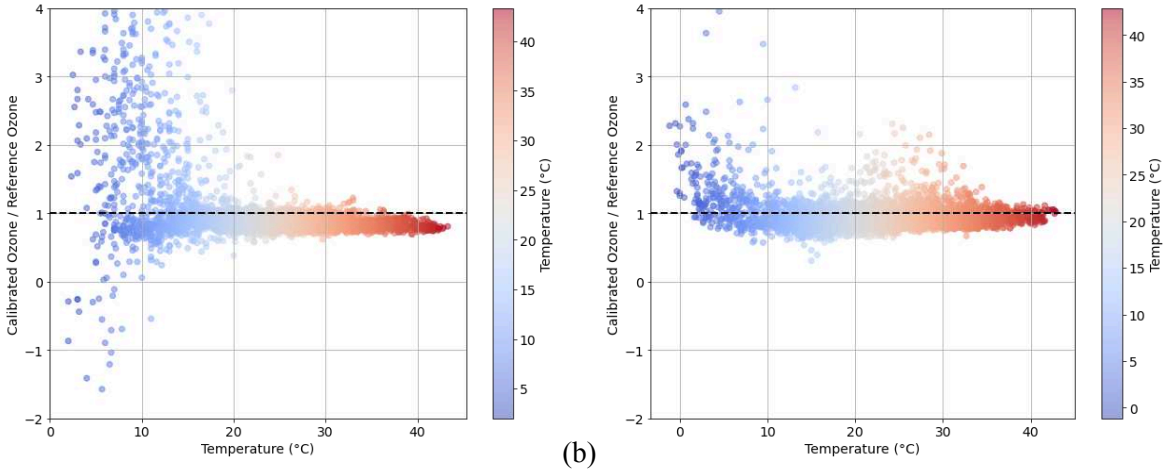


Figure S9 Temperature influence depiction at a) Fresno-Garland and b)Tranquility.

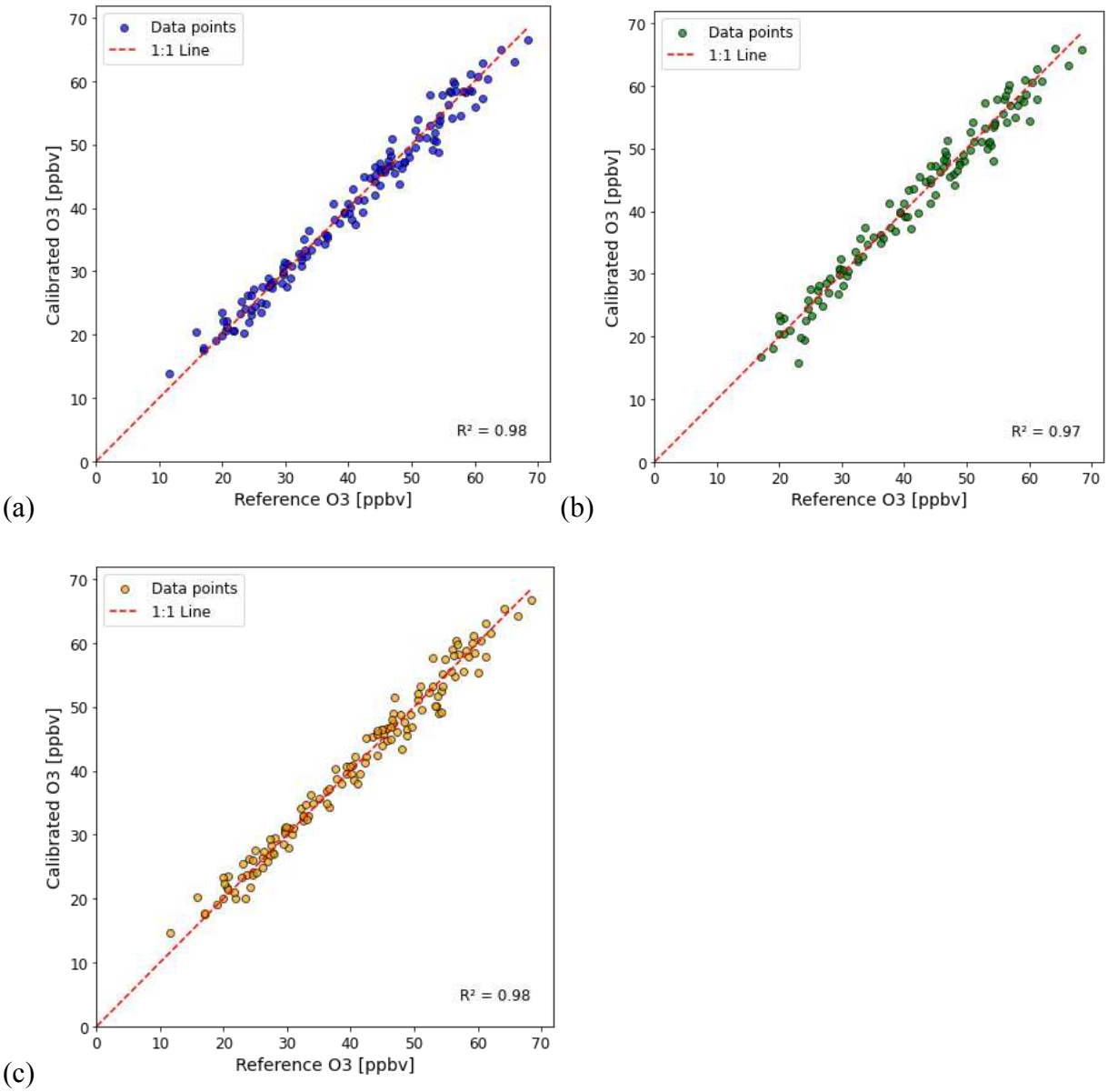
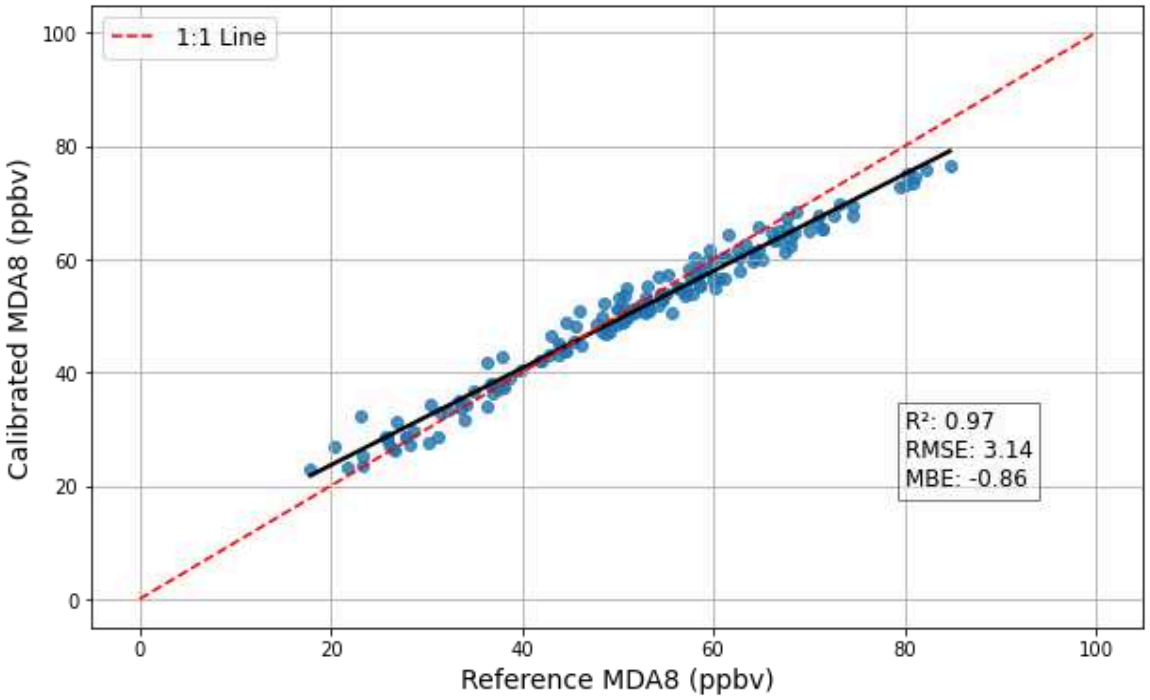
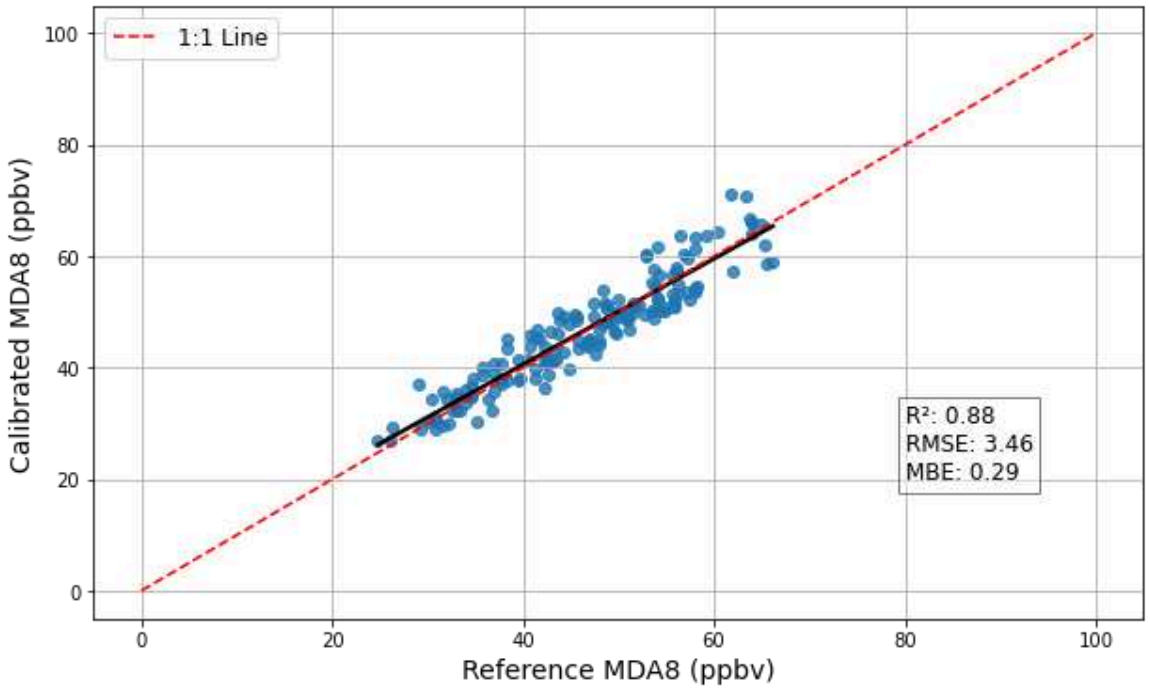


Figure S10 Scatter plot for Rooftop Performance of a)VOZboxes 10, b)VOZboxes 7, c)VOZboxes 8 for 3 days at the Powerhouse, CSU Campus, Colorado.

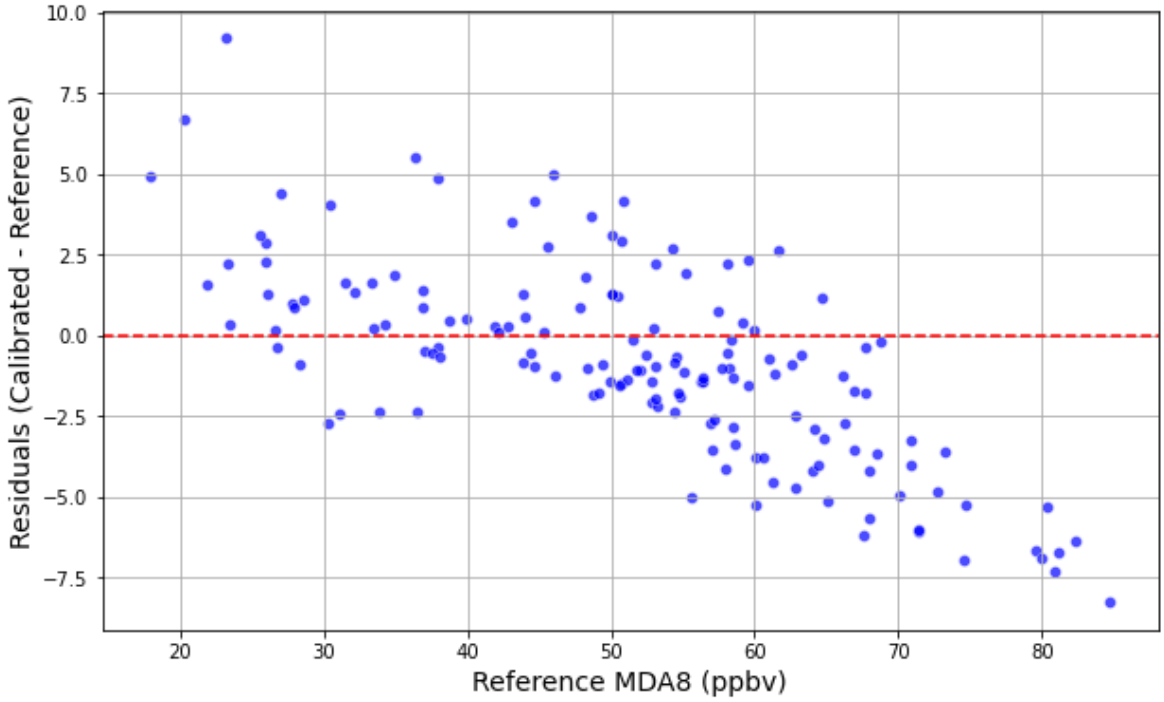


(a)

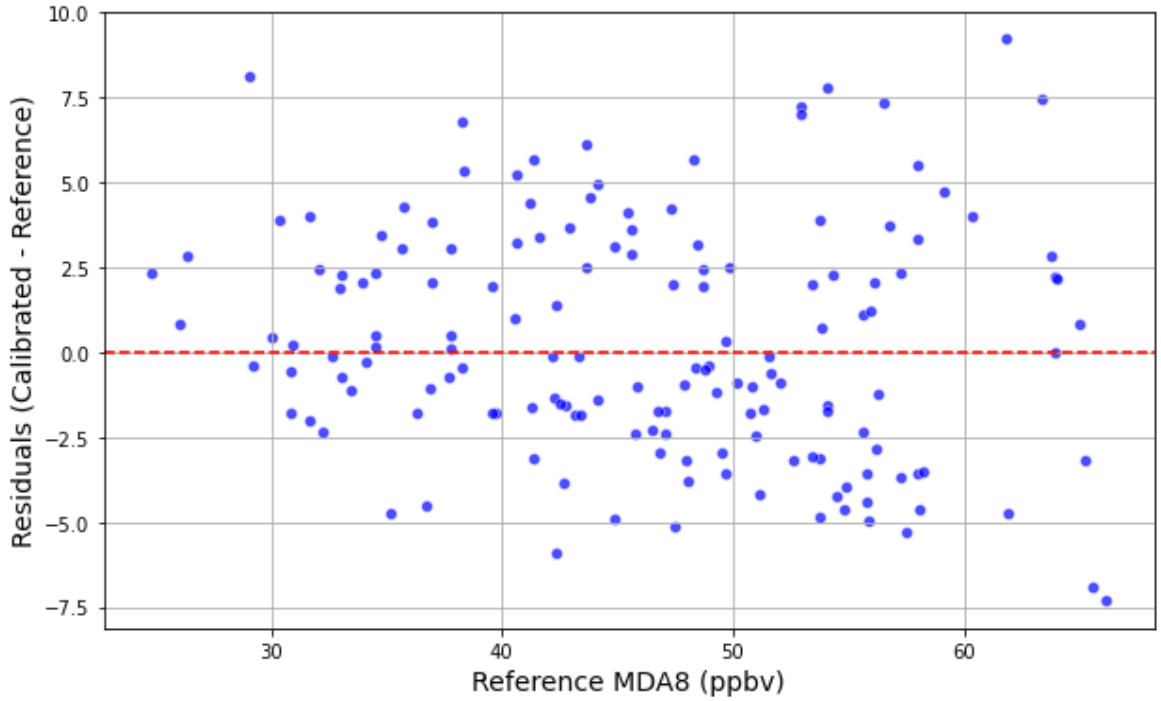


(b)

Figure S11 illustrates the performance of the VOZbox a) 1 and b) 12 by depicting the regression line and the correlation between the calibrated value and the reference value by a scatter plot.

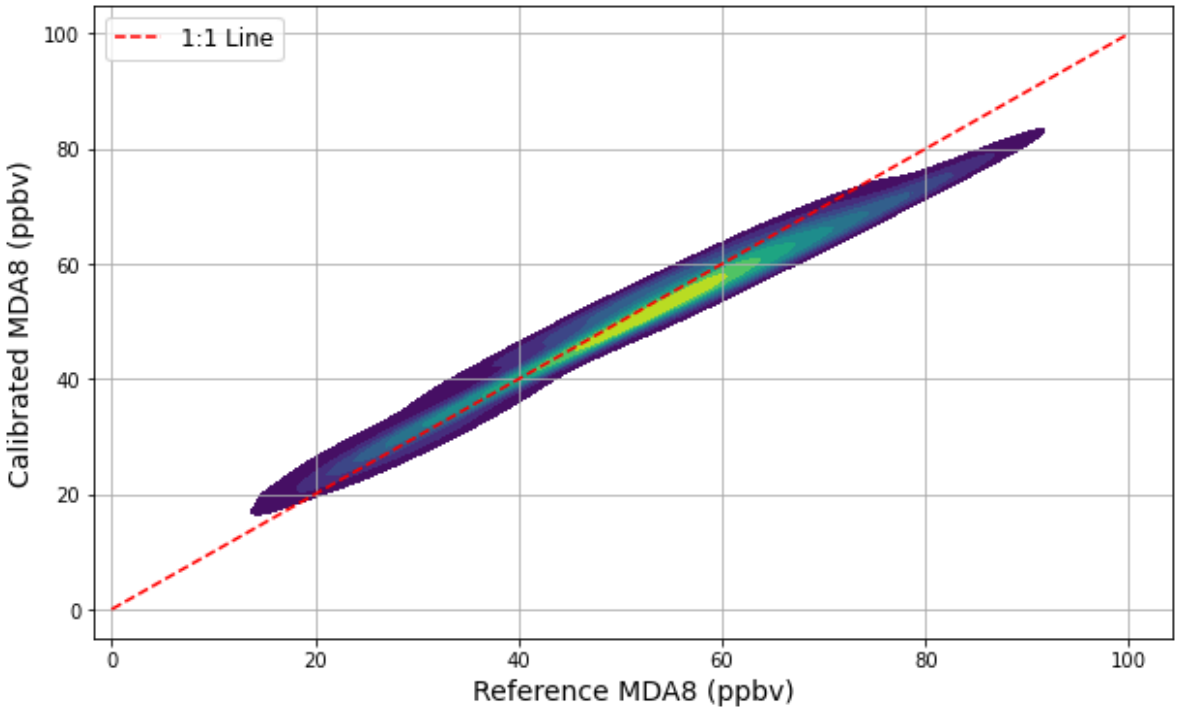


(a)

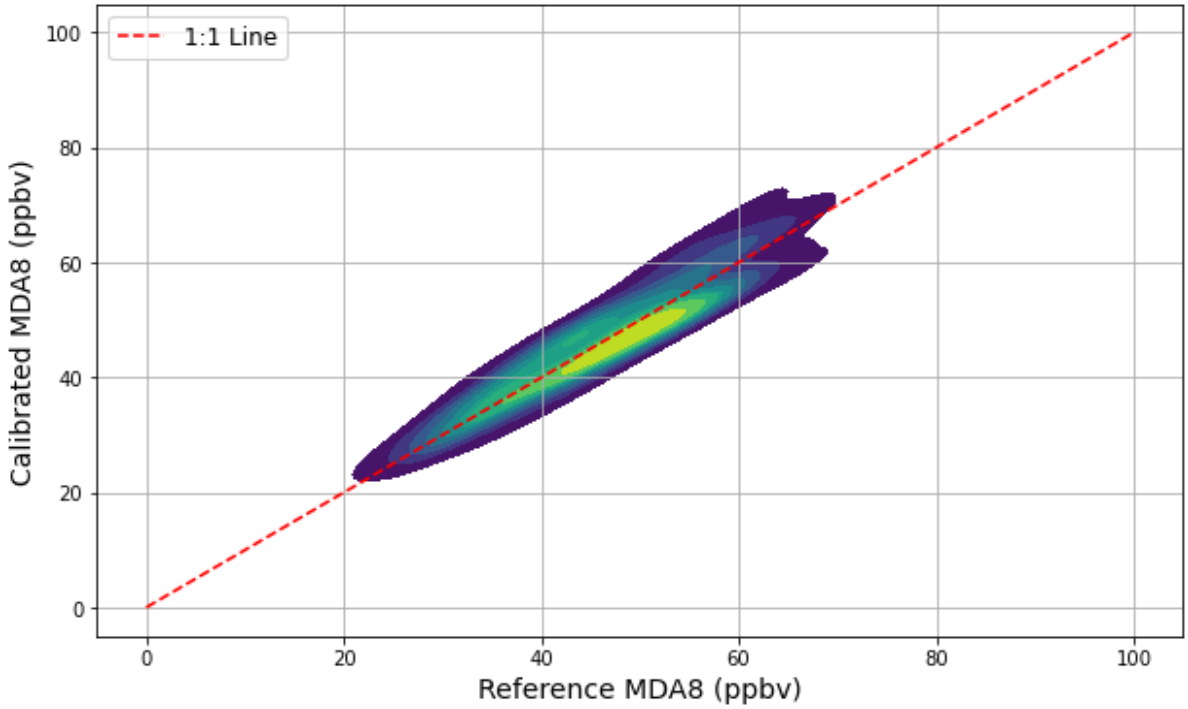


(b)

Figure S12 illustrates the performance of the VOZbox a) 1 and b) 12 by depicting the residuals between the calibrated value and the reference value by a scatter plot.

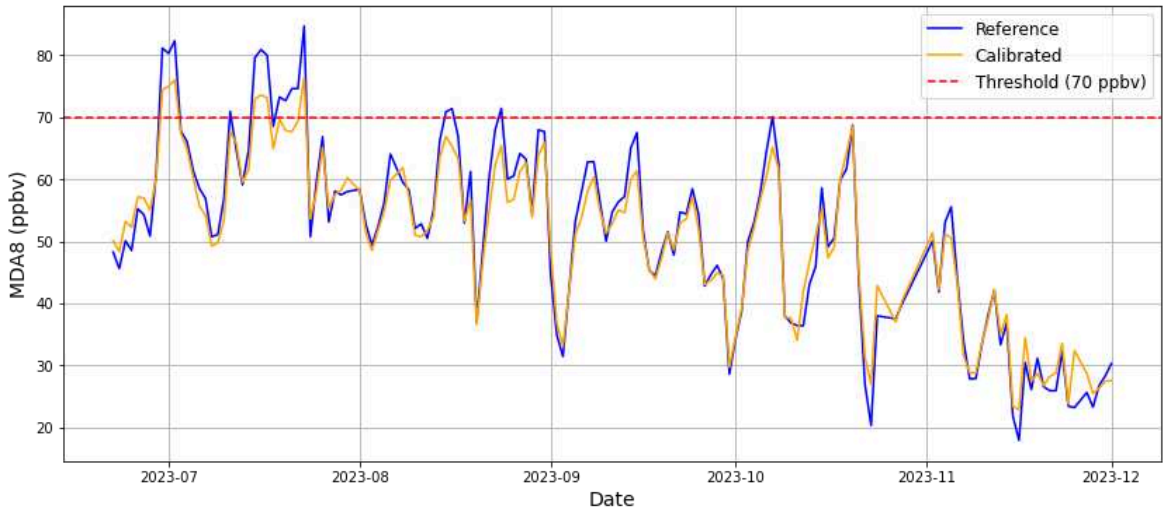


(a)

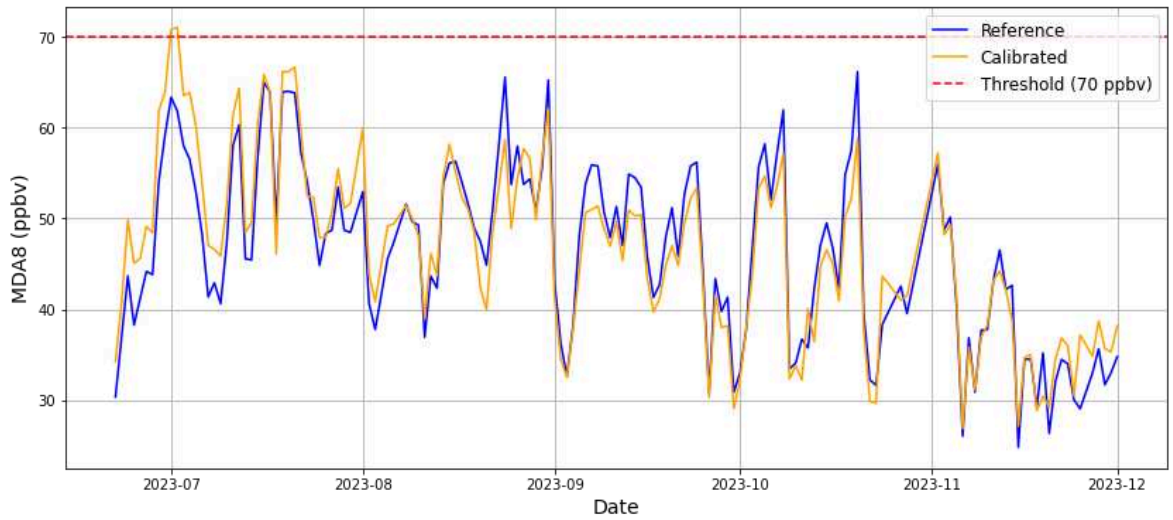


(b)

Figure S13 illustrates the performance of the VOZbox a) 1 and b) 12 by depicting the density of values between the calibrated value and the reference value using a density plot.

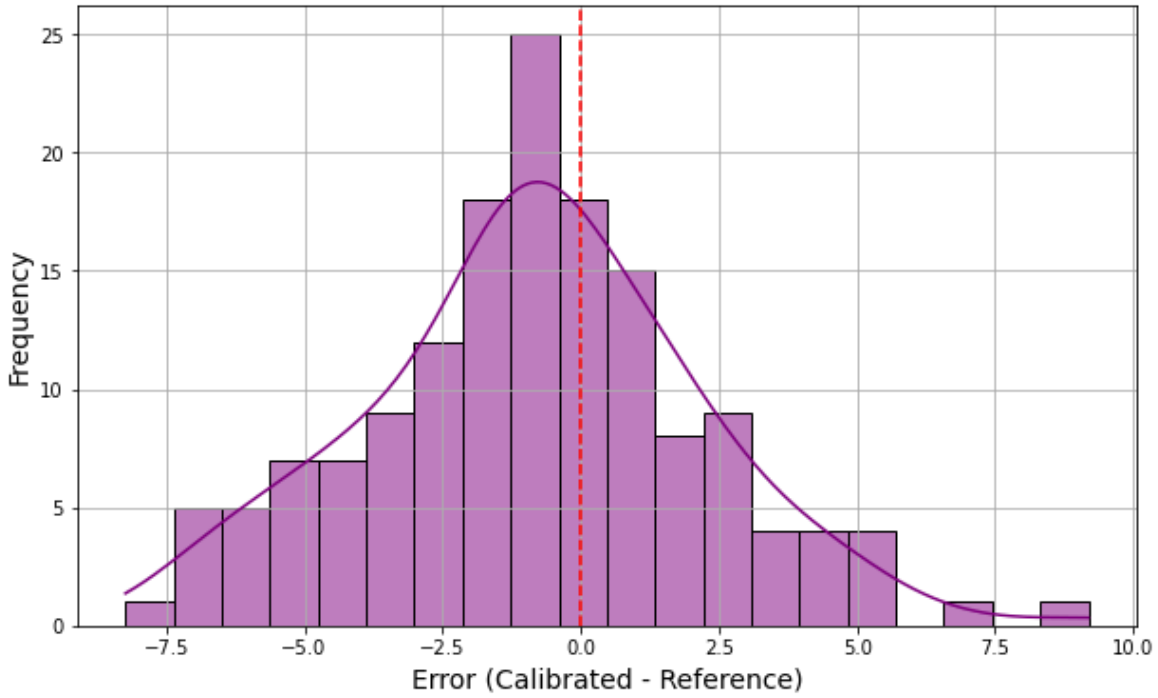


(a)

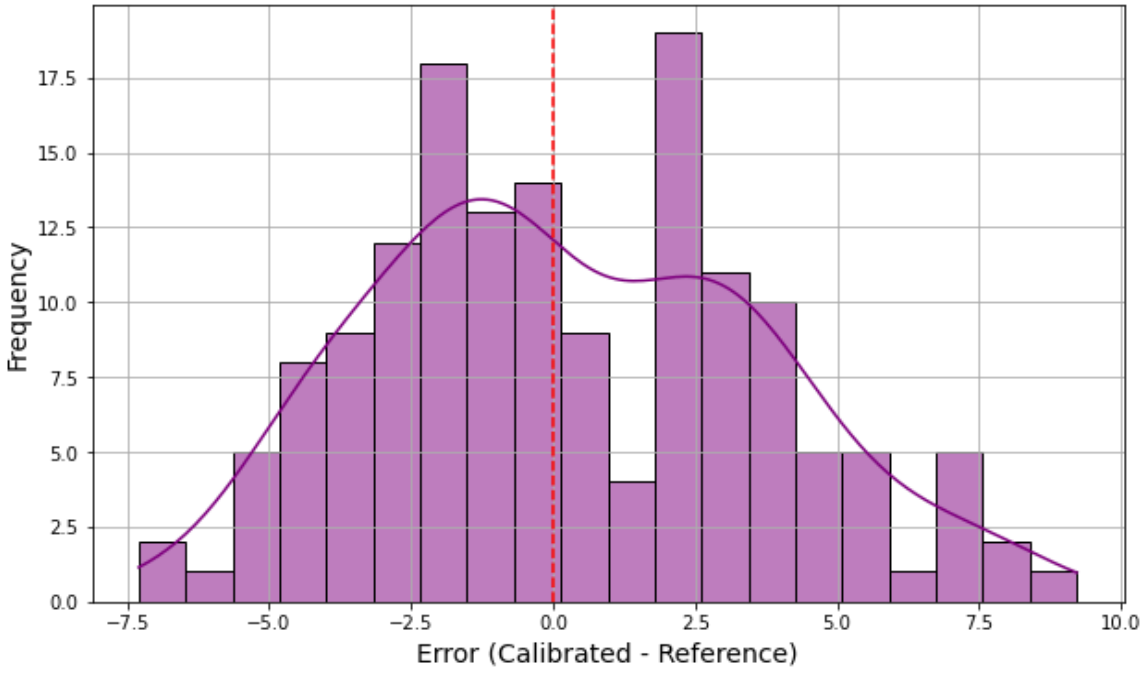


(b)

Figure S14 illustrates the performance of the VOZbox a) 1 and b) 12 by depicting the calibrated and the reference value in a time series line plot by how well it matches the calibration matches the actual values.



(a)



(b)

Figure S15 illustrates the performance of the VOZbox a) 1 and b) 12 by depicting the error bars between the calibrated value and the reference value by a histogram plot.

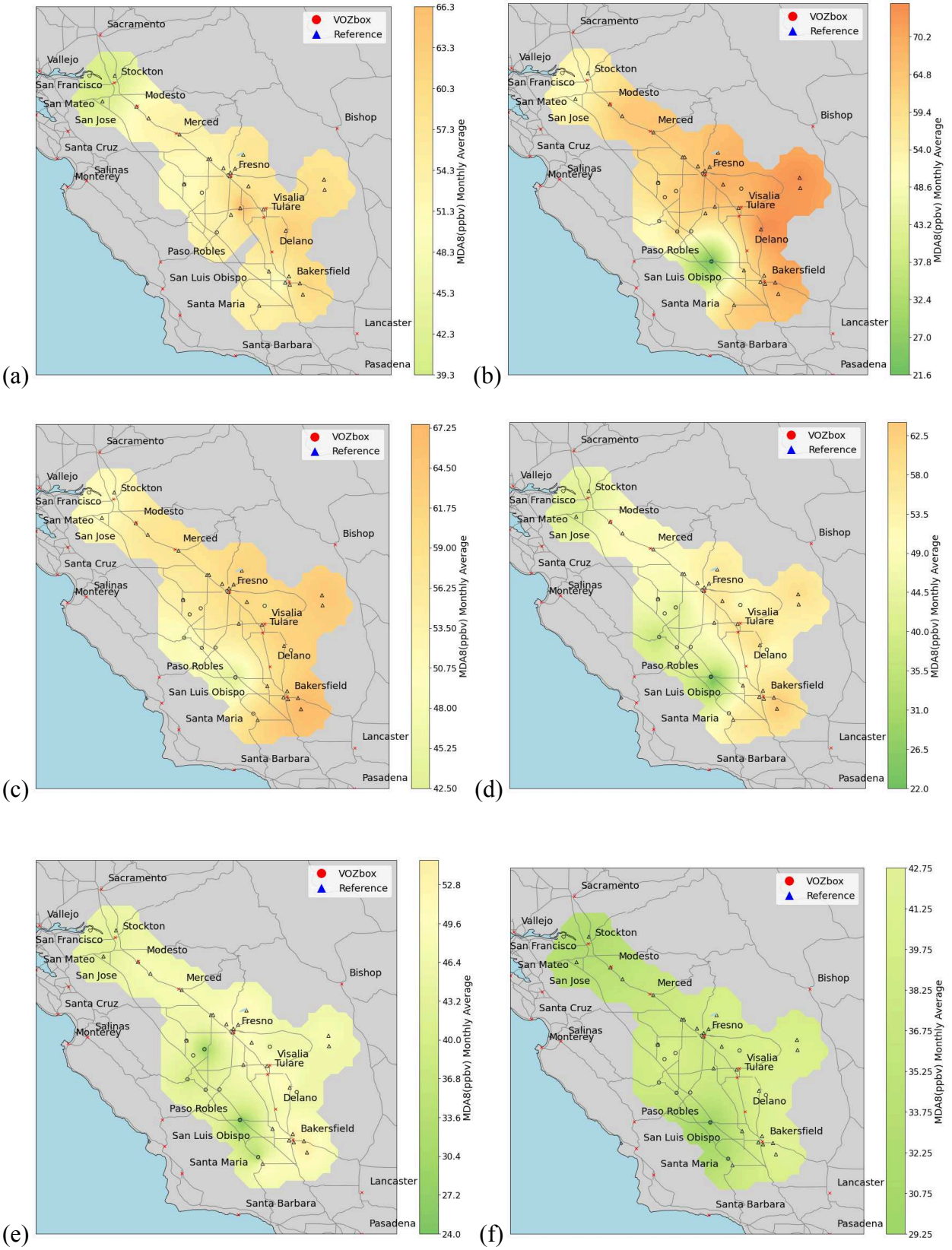
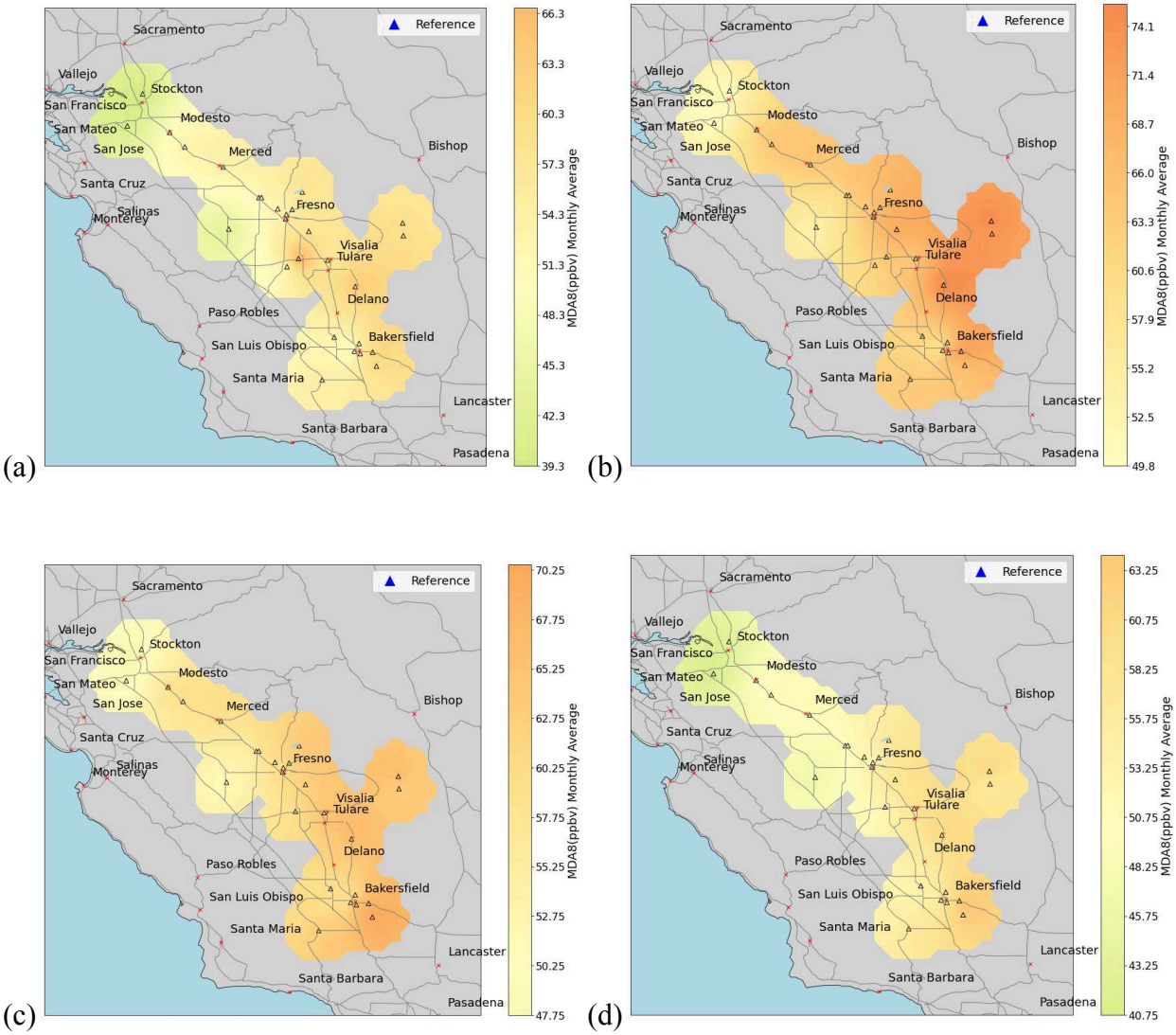

(a)
(b)
(c)
(d)
(e)
(f)
VOZbox
Reference
MDA8(ppbv) Monthly Average
Sacramento
Vallejo
Stockton
San Francisco
Modesto
San Mateo
San Jose
Merced
Bishop
Santa Cruz
Salinas
Monterey
Fresno
Visalia
Tulare
Delano
Paso Robles
Bakersfield
San Luis Obispo
Santa Maria
Lancaster
Santa Barbara
Pasadena
66.3
63.3
60.3
57.3
54.3
51.3
48.3
45.3
42.3
39.3
70.2
64.8
59.4
54.0
48.6
43.2
37.8
32.4
27.0
21.6
67.25
64.50
61.75
59.00
56.25
53.50
50.75
48.00
45.25
42.50
62.5
58.0
53.5
49.0
44.5
40.0
35.5
31.0
26.5
22.0
52.8
49.6
46.4
43.2
40.0
36.8
33.6
30.4
27.2
24.0
42.75
41.25
39.75
38.25
36.75
35.25
33.75
32.25
30.75
29.25

Figure S16 (a) - (f) Illustrates the Kriged Maps with Regulatory Monitors and VOZboxes for the months of June - November in the year 2023.



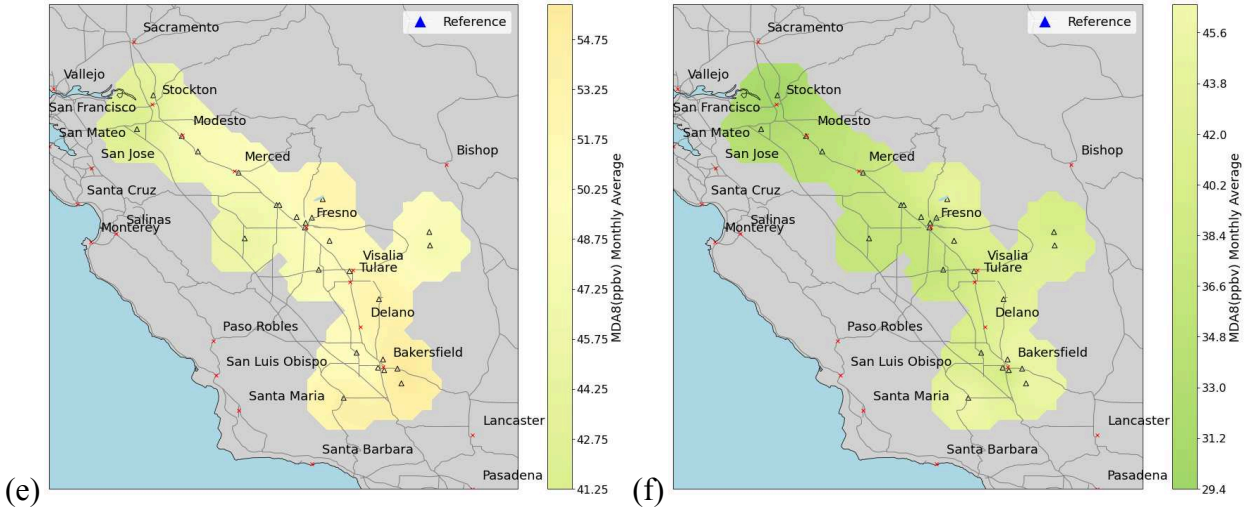


Figure S17 (a) - (f) Illustrates the Kriged Maps with Regulatory Monitors for the months of June - November in the year 2023.