DATA ANALYSIS FOR THE PORTSIDE AIR QUALITY IMPROVEMENT AND RELIEF (PAIR) PROGRAM

FINAL REPORT

Prepared By:
San Diego State University (SDSU)
Environmental Health Coalition (EHC)
San Diego County Air Pollution Control District (SDAPCD)

Prepared For:
California Air Resources Board (CARB)
Agreement Number: 22RD022

March 29, 2025

Acknowledgments

The authors wish to thank the California Air Resources Board (CARB) for sponsoring this study and especially CARB staff members Patrick Wong and Holly Jessop. The statements and conclusions in this Report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

Contents

List of Figures

List of Tables

Data Analysis for the Portside Air Quality Improvement and Relief (PAIR) Program

Execute Summary

The Portside Air Quality Improvement and Relief (PAIR) program was developed by the San Diego County Air Pollution Control District (SDAPCD) in collaboration with Environmental Health Coalition (EHC) for improving the indoor air quality (IAQ) in homes of Portside Communities that include the neighborhoods of Barrio Logan, Logan Heights, Sherman Heights, and West National City in the San Diego County. The two-year PAIR program started in 2022 and provided free delivery and installation of both an indoor air monitor (IQAir AirVisual Pro) and an air purifier (Blueair Blue Pure 311 Auto) to improve air quality in the homes of over 500 Portside Community participants. This document presents the work plan for the statistical analysis of the ambient and indoor air quality, meteorological, and household data obtained as part of the PAIR Data Analysis Project sponsored by the California Air Resources Board (CARB, project 22rd022) and developed by the San Diego State University (SDSU) in collaboration with SDAPCD and the EHC.

The goal of the data analysis project is to observe the IAQ impact of air purifiers in the homes of participants and investigate how factors such as outdoor air pollution, ambient meteorological conditions, and housing conditions can affect both the IAQ and the effectiveness of air purifiers. The sampling size and sampling period in the design for the statistical analysis process is delimited by the indoor data collected in five-minute intervals by the air monitor for Particulate Matter (PM) 10, 2.5, and 1 micrometers or less in diameter (PM10, PM2.5, and PM1, respectively), carbon dioxide, humidity, and temperature. Corresponding ambient air quality and meteorological datasets are obtained at one-hour intervals from the monitoring networks nearest to the Portside Communities for the same sampling period of the PAIR program. In addition, qualitative household data (e.g., house type, size, age, etc.) from participants residents are obtained by EHC. In here, we describe the analysis methods and statistical metrics used for assessing the impacts of the air purifier devices as revised by the PAIR project Advisory Committee to serve community needs for understanding the benefits of indoor air monitors and air purifiers in real-world working conditions.

The study results indicate that air purifiers effectively reduce indoor PM concentrations, with median reductions of 30–40% across different PM size fractions. The highest efficiency was observed in the 1st quantile, where some households experienced PM decreases exceeding 60%, particularly during the daytime. However, efficiency varied among households due to differences in environmental conditions, air circulation, and operational factors. Proper placement, regular maintenance, and continuous operation were key to maximizing efficiency.

The findings emphasize the importance of tracking installation dates and operational conditions to reduce uncertainty in evaluating air purifier impacts. Future studies should implement controlled interventions to isolate the effects of air purifiers and explore long-term health benefits. While air purifiers offer a valuable solution for reducing indoor air pollution, their effectiveness depends on optimal usage and environmental conditions.

Participants reported high levels of usage and perceived effectiveness of both air purifiers and air quality monitors. Most respondents noted improvements in air quality, respiratory health, and overall well-being, highlighting strong alignment between device functionality and user expectations. Concerns about indoor and outdoor air pollution were prevalent, reinforcing the need for ongoing efforts to improve air quality in vulnerable communities.

While most users encountered no major issues, some reported challenges with filter replacements, device connectivity, and data interpretation. Addressing these through better user support, improved accessibility, and enhanced educational resources could maximize the program's impact. Future research should focus on objective air quality measurements and long-term health outcomes to further validate these benefits. The results of this project are useful to understand the temporal and spatial characteristics of indoor air pollution levels and to identify additional resources needed to reduce the exposure to air pollutants in the Portside Communities. Findings from this analysis contribute to data-driven recommendations for environmental health interventions and community-based air quality improvement strategies.

Introduction

Indoor Air Pollution in Disadvantaged Communities

Household indoor air pollution arises from a complex interplay of indoor and outdoor sources, contributing to significant human exposure to harmful air contaminants. Key indoor sources include combustion activities such as cooking, smoking, and heating, which emit particulate matter (PM), carbon monoxide (CO), nitrogen oxides (NO_x), and volatile organic compounds (VOCs) (Naeher et al., 2007). In addition to combustion, a variety of evaporative sources—including emissions from building materials, furniture, cleaning solvents, and consumer and personal care products—release VOCs, semi-volatile organic compounds (SVOCs), and other hazardous air pollutants (Weschler & Nazaroff, 2008). Furthermore, outdoor pollution from industrial emissions, vehicular traffic from nearby roads, and other regional air pollution sources can infiltrate indoor environments, exacerbating air quality concerns of communities (Zhang & Smith, 2003).

Prolonged exposure to indoor air pollution has been linked to a range of adverse health effects, such as respiratory and cardiovascular illnesses, allergic symptoms, cancers, and premature mortality (Vardoulakis et al., 2020). Fine particulate matter (PM2.5) and ultrafine particles (UFPs) can penetrate deep into the respiratory system, leading to chronic respiratory conditions such as asthma, chronic obstructive pulmonary disease (COPD), and lung cancer (Pope & Dockery, 2006). Additionally, VOCs and other gaseous pollutants can cause acute symptoms such as eye, nose, and throat irritation, while long-term exposure has been linked to neurological disorders and endocrine disruption (Mendell, 2007). Vulnerable populations—including children, the elderly, and individuals with pre-existing health conditions—are particularly susceptible to the harmful effects of indoor air pollutants (Díaz et al., 2007). The World Health Organization (WHO) estimates that household air pollution contributes to approximately 3.2 million premature deaths annually, with over 237,000 of these deaths occurring in children under five years old (Puthumana et al., 2021).

Indoor air pollution is not only a health issue but also a matter of health equity. Vulnerable populations, including children, the elderly, and low-income communities, often experience higher exposure levels and bear a disproportionate burden of the associated health risks. Addressing these disparities requires integrated health and energy strategies aimed at reducing exposure to indoor air pollutants and improving the health of residents in communities with significant environmental and health inequities (Bruce et al., 2000; Balakrishnan et al., 2013). Comprehensive policies that promote the adoption of clean fuels and air filtration technologies, enhance ventilation, and reduce the use of products that emit harmful pollutants are essential for protecting public health and advancing health equity.

The use of filter-based commercial and non-commercial air purifiers in households has increased in recent years, primarily due to their ability to reduce indoor air pollution, particularly particulate matter and other airborne contaminants. These devices are widely adopted in response to growing concerns about indoor air quality, especially in urban and industrialized areas where outdoor pollution infiltrates homes. Additionally, climate change-related factors, such as increased wildfire smoke events and prolonged heatwaves, have further driven the demand for indoor air purification solutions (Chen et al., 2022).

Despite their widespread use, the effectiveness of air purifiers can vary significantly based on multiple factors, including household characteristics (e.g., ventilation type, room size, building materials), indoor emission sources (e.g., cooking activities, smoking habits, use of cleaning products), and operational settings (e.g., placement of the device, frequency and duration of use, filter replacement schedules). These variables can influence the overall filtration efficiency of the devices, as well as their ability to provide measurable health benefits to residents (Zhao et al., 2021). While controlled laboratory studies have demonstrated the ability of air purifiers to reduce airborne particulate matter and certain gaseous pollutants (Wang et al., 2020), real-world conditions introduce complex variables that may alter their performance. For example, air exchange rates, human activity patterns, and variations in pollutant sources all contribute to the dynamic nature of indoor air quality, which may lead to differing degrees of pollutant removal efficiency (Allen et al., 2019). Furthermore, the long-term health benefits of air purifier usage remain an area of active research, particularly in vulnerable populations, such as children, the elderly, and individuals with pre-existing respiratory conditions (Chen & Zhao, 2020).

Given these uncertainties, there is a critical need to assess the performance of air purifiers in real-world household environments. Evaluating their effectiveness in diverse settings will provide valuable insights into their role as a practical intervention for mitigating indoor air pollution exposure and reducing associated health risks, particularly for communities facing environmental justice challenges. Moreover, as the burden of indoor air pollution is disproportionately high in low-income and marginalized communities, residents need access to adequate ventilation and access to air filtration technologies, particularly those living in proximity to industrial facilities. Addressing these inequities requires comprehensive air quality control policies that integrate indoor pollution mitigation strategies with broader environmental justice efforts. Given the growing recognition of indoor air pollution as a critical public health issue, interdisciplinary research and policy efforts must focus on developing effective interventions to reduce exposure and improve indoor air quality. These efforts are essential for safeguarding the health of residents in environmentally burdened communities and achieving broader public health and environmental equity goals.

The Portside Community

The Portside Community includes the neighborhoods of Barrio Logan, Logan Heights, Sherman Heights, and West National City, located near the Port of San Diego, CA (see Figure 1). National City, Logan Heights, and Barrio Logan are historically marginalized communities in the San Diego County, facing substantial socioeconomic and environmental challenges. According to the CalEnviroScreen 4.0 tool, these communities rank among the most disadvantaged in California, with high exposure to air pollution from industrial sources, port operations, and major highways (OEHHA, 2023). Barrio Logan and Logan Heights are part of the Portside Environmental Justice Neighborhoods, identified by CARB for their exposure to emissions from freight, rail, and industrial activities. This exposure contributes to elevated rates of asthma and other respiratory illnesses (CARB, 2023). Socioeconomic challenges, such as high unemployment and limited access to healthcare, further exacerbate disparities in these neighborhoods. Educational attainment is also lower in these areas compared to countywide averages, with a significant portion of residents lacking a high school diploma (SDCH&HS, 2022).

National City, the county's most economically disadvantaged city, has a predominantly minority population, with high poverty rates (22%) and low educational attainment (35% lacking a high school diploma). The community is also predominantly Latino (63%) and Asian American/Pacific Islander (20%), with a relatively young demographic (26% under 18). Logan Heights shares similar challenges, with 86% of residents identifying as Latino and 50% lacking a high school diploma. Additionally, 30.8% live below the poverty line, and 34% are under 18. Barrio Logan, a historically significant but environmentally burdened neighborhood, was rezoned for mixed-use development following the construction of Interstate 5 (1963) and the San Diego-Coronado Bay Bridge (1969), disrupting its residential fabric. Today, the community experiences high poverty (78% low-income), linguistic isolation (32%), and low educational attainment (42% without a high school diploma). Environmental health concerns are severe, with asthma-related hospital visits among the highest in California (81 per 10,000 residents) and elevated cancer risks (80th–90th percentile nationally). These communities continue to face systemic inequities, necessitating targeted interventions to address socioeconomic and environmental disparities.

The PAIR program

The Portside Air Quality Improvement and Relief (PAIR) Program is a two-year initiative aimed at enhancing indoor air quality for residents of the Portside Communities, including Barrio Logan, Logan Heights, Sherman Heights, and West National City. The PAIR program provides participating households with air cleaning devices and indoor air monitors to reduce exposure to harmful air pollutants. The San Diego County Air Pollution Control District (SDAPCD) administers the project and has partnered with the Environmental Health Coalition (EHC) to conduct outreach and education, facilitate device installations, process applications, and provide troubleshooting support throughout the program's duration.

Residents of the Portside Communities face significant environmental and health justice challenges due to their proximity to major freeways, freight corridors, port activities, and industrial facilities, all of which contribute to elevated levels of air pollution. These exposure risks are further exacerbated by small housing units, high occupant density, inadequate ventilation, and older building construction, which can increase the infiltration of outdoor air pollutants into indoor spaces. By assessing air purifier performance under these conditions, the PAIR program aims to inform future air quality interventions and policy recommendations for communities disproportionately burdened by environmental pollution.

The process for enrolling participants in the PAIR program followed a structured protocol to ensure proper installation, education, and data collection. The EHC administered the 2-year PAIR Program grant from the APCD with funding from the County of San Diego and Port of San Diego to do the following:

- Conduct outreach and education to portside communities to recruit participants
- Assist participants with the application intake process
- Install air monitors during a home visit
- 30 days later install purifiers in up to 575 homes during two home visits to provide deep education about indoor air quality and the pollution in their neighborhoods
- Provide technical assistance and support remotely or via an additional home visit

Participants applied to the program directly through the PAIR website by completing an online application form. The application process was designed to be user-friendly and accessible, allowing interested individuals to provide relevant background information, outline their motivations for joining the program, and highlight any prior experience or qualifications. This streamlined approach ensured a consistent and equitable method for selecting participants. 556 portside participants received the complete set, including one air monitor, air purifier, and extra fabric filter with a 2-year supply of replacement filters. Participants also received a $125 stipend to help offset a potential increase in energy usage and as a thank you for participating in the two-year program. Of the 556 participants, 43 received an additional air purifier to help their families with respiratory illnesses.

Households applying through the PAIR application process first received a home visit by EHC personnel. During this visit, an air quality monitor (IQAir AirVisual Pro) was installed to measure indoor air pollution levels. EHC staff also provided information on the benefits of the program, shared practical tips to improve indoor air quality, and connected the monitor to a cloud-based IQAir Dashboard accessible to both EHC and the SDAPCD. In some cases, air quality data was stored and made accessible through the IQAir Dashboard. In many cases, baseline air quality data was collected for approximately 30 days before introducing the air purifier to establish indoor air pollution levels without filtration. After this initial monitoring period, the air cleaning device (Blueair Blue Pure 311 Auto) was activated and remained in use for the remainder of the program. During installation, EHC personnel also instructed participants on proper device operation, including filter replacement and routine maintenance, and provided each household with three additional replacement filters.

Through this process, the PAIR program has served over 500 households in the Portside Community. By comparing air quality levels before and after the activation of air purifiers, the collected data offers an opportunity to assess the effectiveness of air filtration under real-world conditions. However, variations in household environments, ventilation, and indoor pollution sources must be considered when interpreting results.

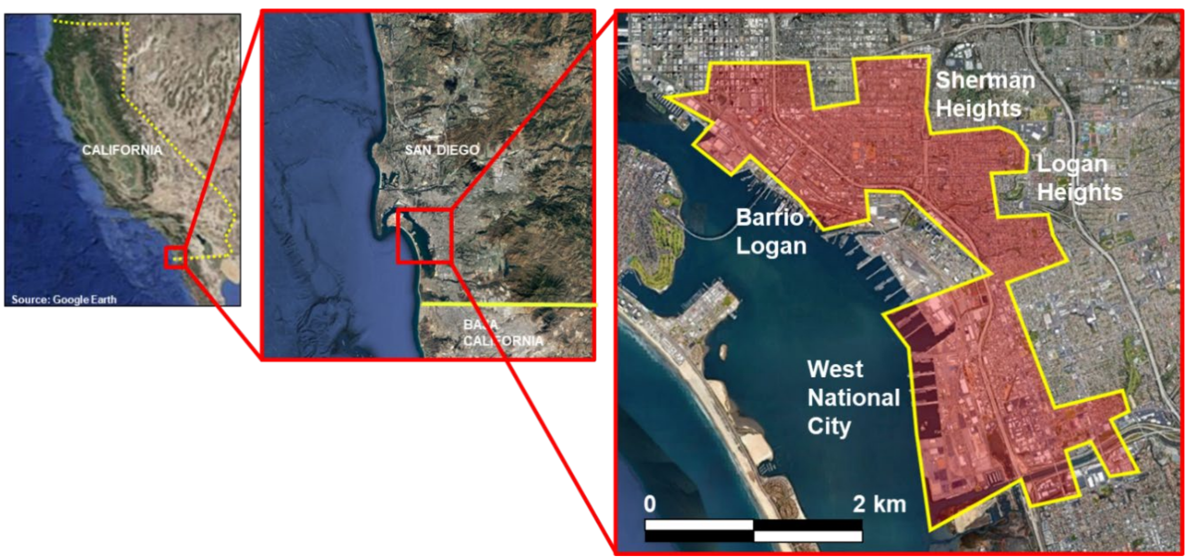

FIGURE 1. NEIGHBORHOODS PARTICIPATING IN THE PAIR PROGRAM (RED POLYGONS, RIGHT FIGURE) IN THE PORTSIDE COMMUNITY IN SAN DIEGO, CA.

The main objective of the PAIR Program was to provide new portable air purifiers and indoor air monitoring systems to selected residences in the Portside area, at no cost to participants. The data collected through the PAIR program presents a unique and valuable opportunity to assess the potential benefits of air filtration systems in real-world conditions within the Portside Communities. However, several important caveats must be considered when interpreting the findings. Notably, the PAIR program was not originally designed to evaluate the efficiency of air filtration devices, resulting in substantial data gaps in key areas of data collection. Critical parameters such as detailed housing conditions, ventilation rates, and indoor source emissions were not systematically recorded, limiting the ability to fully characterize the factors influencing air purifier performance. Nevertheless, as outlined in the Methodology Section, only a subset of relevant parameters was collected, allowing for a meaningful, though constrained, analysis. The limitations associated with these data gaps are explicitly acknowledged and considered throughout this document in the interpretation and discussion of the results.

Objective

The primary goal of this project is to analyze indoor and outdoor air quality data collected from communities surrounding the San Diego Port as part of the PAIR program. This analysis aims to evaluate the impact of air purifiers distributed to residents on indoor air quality and to examine how community-relevant factors—such as housing characteristics, ventilation conditions, and proximity to pollution sources—influence the effectiveness of these devices.

Methods

General Approach

The project was structured around eight key tasks, as illustrated in Figure 2, and incorporated the establishment of an Advisory Committee to guide data analysis. Each of the tasks listed below played a crucial role in ensuring the successful execution of the project, from initial planning, data collection and management, data analysis and reporting. Importantly, the Advisory Committee, composed of members from SDSU, SDAPCD, and EHC, was the initial task and played a central role in identifying community priority questions and shaping the strategies for data collection and management to ensure alignment with community needs and research objectives.

- **Task 1**. Establish an Advisory Committee
- **Task 2**. Set up an Indoor Air Pollution Database
- **Task 3**. Set up a Household Information Database
- **Task 4**. Set up a Meteorological Database
- **Task 5**. Set up an Outdoor Air Pollution Database
- **Task 6**. Literature Review of Methodologies and Develop a Work Plan
- **Task 7**. Perform Data Analysis
- **Task 8**. Project Reporting

This section provides a detailed description of the methodological approach used under Task 7: *Performing Data Analysis*, including the statistical analytical techniques, data processing procedures, hypothesis generation, and evaluation criteria applied to interpret the collected data. A comprehensive breakdown of the development and implementation of all the other tasks, including database management, data collection methodologies, literature review, timelines, and key outcomes, is provided in Appendix 1.

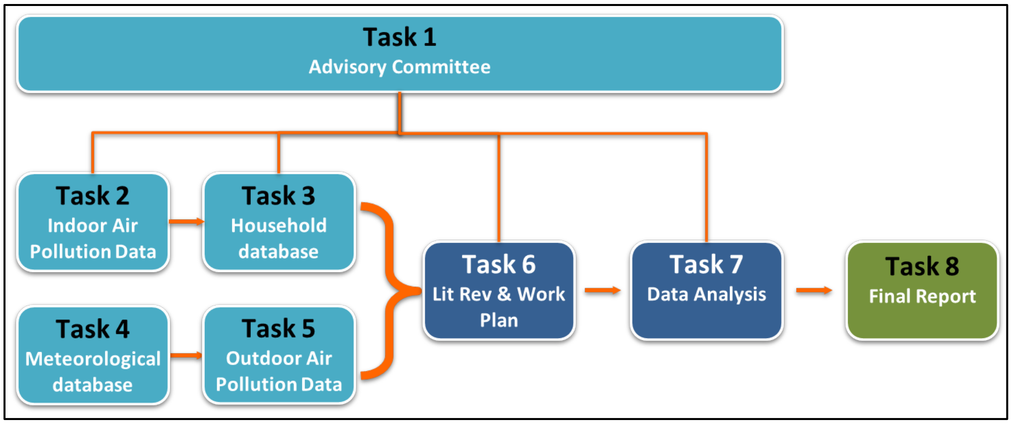

FIGURE 2. PROJECT TASKS FOR THE ANALYSIS OF THE PAIR DATASETS INCLUDING DATA COLLECTION, LITERATURE REVIEW, DATA ANALYSIS AND REPORTING.

Figure 3 provides an overview of the work plan activities undertaken for the data analysis of the collected datasets. This figure outlines the key steps involved in processing and analyzing the data to ensure consistency, accuracy, and statistical validity. Table 1 presents a detailed summary of the data sources, key variables, sampling frequency, and data types used in the PAIR program's data analysis. The datasets included information on indoor air pollution, ambient air pollution, and meteorological parameters, each requiring tailored analytical approaches. The data analysis process followed a structured methodology, incorporating exploratory, descriptive, and inferential statistical techniques. These analyses were applied to the working datasets, which were processed using standardized data collection, quality assurance, and formatting protocols to ensure data integrity. Given the diverse nature of the datasets, the specific analytical steps and statistical methods varied across categories. The subsequent sections provide a detailed description of the data processing and analysis techniques employed for each dataset type.

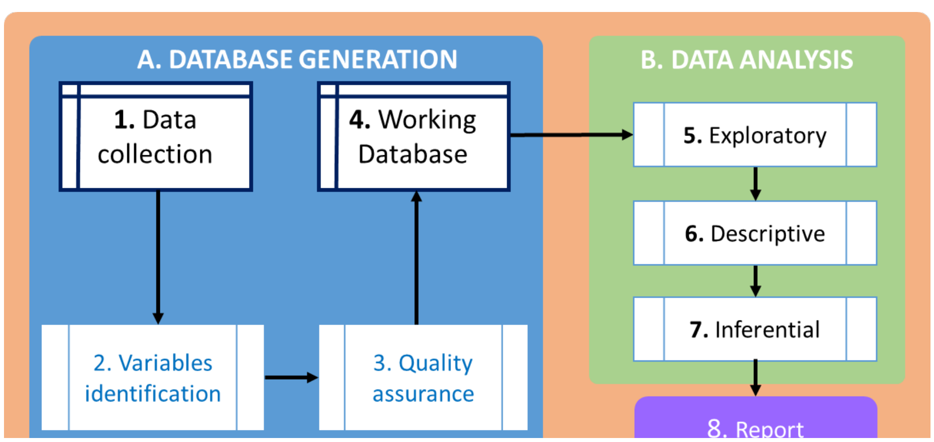

FIGURE 3. DATABASE GENERATION AND DATA ANALYSIS ACTIVITIES THAT FORM THE BASIS FOR THE WORK PLAN.

TABLE 1. DATA SOURCES AND DATA TYPES USED IN THE PAIR PROGRAM.

Database type	Data sources	Variables	Frequency	Data type
Indoor air pollution and meteorology	IQAir AirVisual Pro	• PM_{10}, $PM_{2.5}$, and $PM_{1.0}$, • CO_2 • Relative humidity • Temperature	5 mins, continuous	Interval/Ratio
Household	AirTable	Multiple (e.g., Building type, age, number of residents, square ft, respiratory health conditions, income	NA	Nominal/Ordinal
Ambient air pollution	SDAPCD	• PM_{10}, $PM_{2.5}$ • CO • Nox	1 hour, continuous	Interval/Ratio
Meteorological	SDAPCD	• Wind • Temperature • Pressure • Relative humidity • Precipitation	1 hour, continuous	Interval/Ratio

The database generation process was built on a staggering level approach starting from the raw datasets ("d0" level) containing the data obtained from original data sources, following a quality assurance process ("d1" level), a formatting process ("d2" level), and finally an analysis-ready working dataset ("d3" level). The details of the database generation process, quality assurance, and data management protocols are described in Appendix 1. All statistical analysis were performed for analysis-ready d3 level datasets only after data quality and formatting procedures were finalized. Except for the household information of the participant residents, all the used datasets are publicly available from the listed data sources and were the basis for the quantitative analysis described below. In the case of the household information (AirTable), a qualitative analysis approach was used as described in the section below.

Data Analysis Protocols

Quantitative Analysis

The data analysis process followed a structured approach, incorporating exploratory, descriptive, and inferential statistical methods to evaluate the working datasets. These datasets were prepared through a systematic workflow that included data collection, quality assurance, and formatting protocols to ensure consistency, accuracy, and usability for statistical analysis. Exploratory analysis was conducted to identify patterns, trends, and potential outliers in the data. We obtained descriptive statistics summary measures, including central tendencies and dispersions, to characterize the datasets.

A key focus of the analysis was to assess the impact of the filtration devices by comparing air quality and household conditions before and after their installation. This was achieved by analyzing baseline datasets—collected during the period before the filtration devices were installed—and comparing them with datasets obtained after the devices had been deployed. This comparison allowed for an evaluation of changes in air quality, exposure levels, and other relevant environmental and household parameters. Inferential statistical techniques were then applied to determine whether observed changes between the baseline and post-installation periods were statistically significant, controlling for potential confounding variables.

Data Availability of IQAir Air Quality Monitoring Devices

An important constraint in the data analysis process was the limited availability of indoor air quality data from the air monitoring devices installed in the households of PAIR program participants. The project team collaborated with IQAir staff to retrieve de-identified indoor air quality data from the IQAir Cloud repository. However, after overcoming technical challenges that resulted in delays in accessing these datasets, it was determined that only 47% of the 510 originally installed devices of participating households had successfully uploaded their data to the IQAir Cloud. It is important to note that all devices were actively monitoring air quality, displaying real-time data on the user interface, and storing the information locally in the device's internal memory. However, because the PAIR program was not initially designed as a research initiative, data uploads to the IQAir Cloud were not a priority. The option to share (upload) air quality data to the IQAir cloud platform was presented to participants as a completely voluntary choice during the sensor installation process. Participants were fully informed that enabling data sharing would allow their sensor readings to contribute to a broader, publicly accessible air quality monitoring network. However, they were also assured that choosing not to share their data would in no way affect their participation in the program or their ability to access and use their own air quality information locally. As a result, many participants chose not to enable cloud uploads, limiting the availability of centralized, remotely accessibility of all datasets for analysis.

Furthermore, because the setup process for the IQAir devices was conducted manually, and serial numbers of devices were recorded by hand, during data collection some entries were not legible enough to be accurately matched with the corresponding serial numbers in the IQAir Cloud database. As a result, data from an additional 18 devices could not be retrieved, leading to further data loss. This issue underscores the challenges associated with manual data entry and the importance of implementing standardized digital recording methods to improve data accuracy and traceability. Figure 4 provides a detailed overview of this process, illustrating the manual setup workflow, potential points of error in serial number recording, and the resulting impact on data availability for analysis.

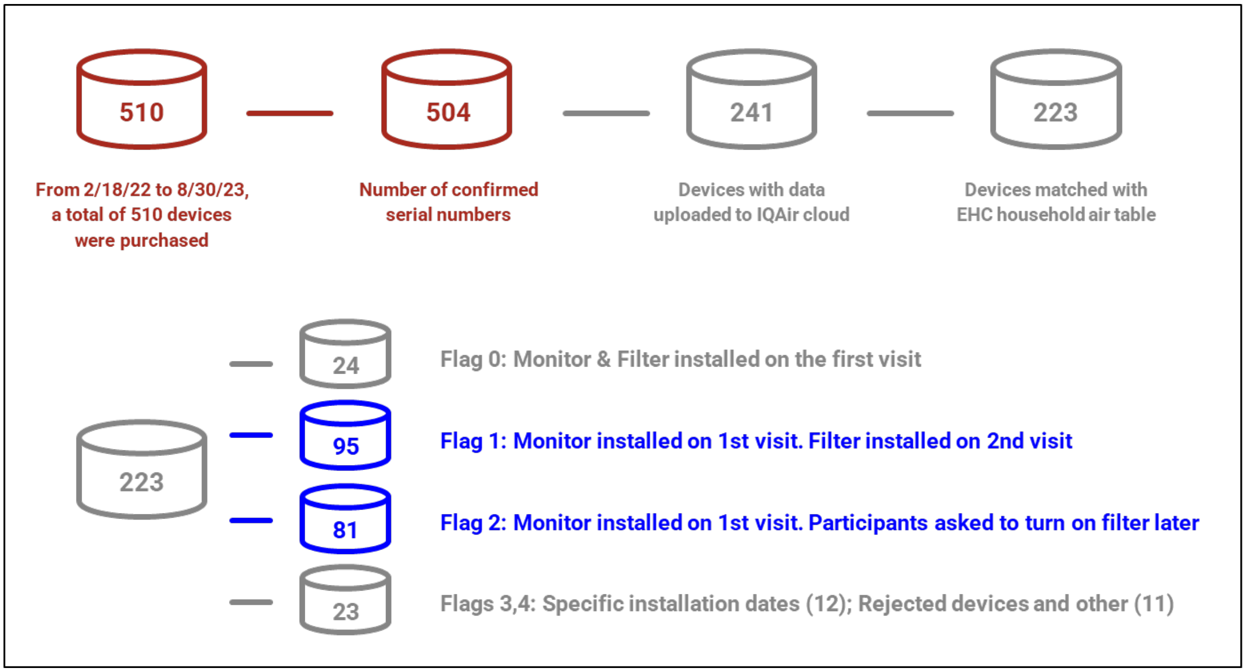

FIGURE 4. DATA AVAILABILITY, FLAGGING PROCESS, AND ANALYSIS OF IQAIR INDOOR MONITORING DATA FOR DEVICES INSTALLED IN HOUSEHOLDS OF PAIR PROGRAM PARTICIPANTS.

Data analysis was conducted on a total of 223 IQAir indoor air quality monitoring devices (see Figure 4). Figure 4 illustrates this categorization process, outlining the criteria used for flagging and the implications for data interpretation. To better assess data reliability and account for potential variability, the devices were categorized, or "flagged," into four distinct groups based on the availability of information regarding the installation dates of the air pollution air purifier devices in each household. This categorization was crucial for identifying potential sources of uncertainty in the dataset and ensuring that comparisons between baseline (pre-installation) and post-installation periods were as accurate as possible. The flagged datasets were later used to refine the interpretation of the statistical hypothesis testing results, providing a more robust discussion of the uncertainty associated with the analysis.

Hypothesis tested

The process for generating hypotheses to evaluate the effectiveness of air purifier devices in participants' households using d3-level datasets is illustrated in Figure 5. This process involved several key steps to ensure a rigorous and systematic approach to data selection and statistical testing. First, the installation dates for both the air quality monitoring device and the air purifier device were identified (see "Monitoring Installed" and "Filter Installed" in Figure 5). These dates served as reference points for defining the sampling periods used in the hypothesis testing. Next, a sampling window of approximately 30 days surrounding the "Filter Installed" date was selected to establish the pre- and post-installation periods. The "Start Date" and "End Date" within this window were determined to create datasets for comparison.

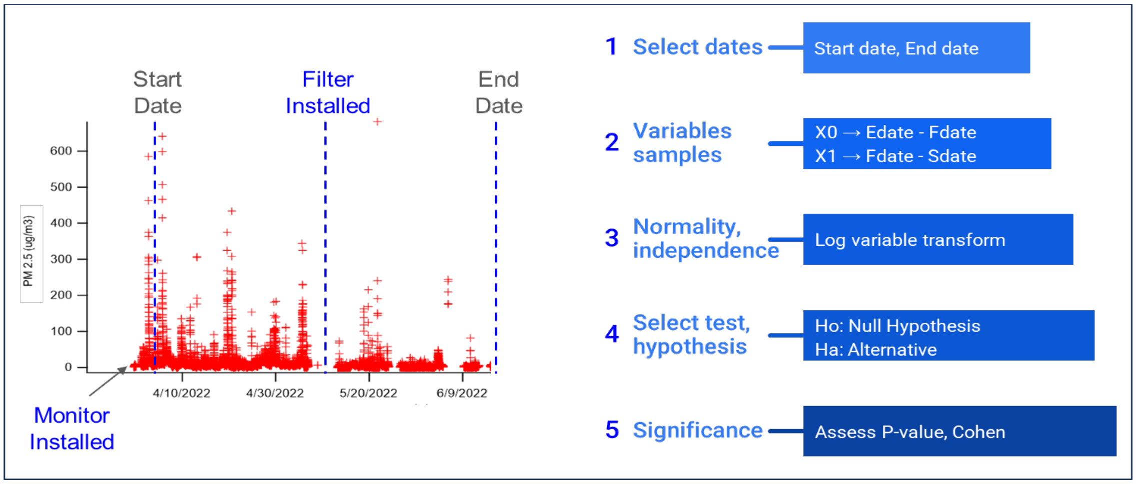

FIGURE 5. DESCRIPTION OF HYPOTHESIS GENERATION AND TESTING PROCEDURE OF THE PAIR DATA.

Once the sampling periods were defined, the datasets were assessed for statistical independence to ensure valid comparisons. If necessary, the data were log-transformed to address skewness or improve normality assumptions required for hypothesis testing. An appropriate statistical test was then selected based on the data characteristics, and the null hypothesis was formulated and tested accordingly. Finally, the significance of the results was evaluated using standard p-values and effect size measures, such as Cohen's d, to assess the magnitude of the observed differences. The specific research questions, hypothesis selected, statistical tests, and associated assumptions for assessing the beneficial impacts of the air purifier devices are described in Table 2. For the purposes of this analysis, the impact of the air purifier devices was calculated in the form of an efficiency estimate as follows:

$$I_{p,i} = \frac{x_{p,i,post} - x_{p,i,pre}}{x_{p,i,pre}} \times 100$$

Where $I_{p,i}$ represents the beneficial impact for the pollutant *p* for the device *i*; $x_{p,i,post}$ and $x_{p,i,pre}$ represent the corresponding statistical measure (e.g. mean or median) of the sample size after and before the air purifier device was installed. The impacts were estimated for each device with available valid data and served as the basis for constructing box plots, calculating quantiles, means, and other statistical measures for assessing the validity of the hypothesis constructed.

TABLE 2. HYPOTHESIS TESTED FOR ASSESSING THE IMPACTS OF AIR PURIFIER DEVICES IN THE PAIR PROGRAM.

Question	Null hypothesis	Statistical test	Assumptions
1. Is there a statistical difference in indoor air pollution levels using the air purifier device?	• There is no significant difference in the mean concentration before and after filtration • There is no significant difference in the median concentration before and after filtration	• Paired t-test • Wilcoxon Signed-Rank Test	• Assumes normality • Nonparametric test for highly skewed data. No normality
2. Is there a statistical difference in the variability of indoor air pollution levels using the air purifier device?	The variance of concentration levels before and after filtration is the same	• F-test for equality of variances • Levene's Test	• Normality • This test is less sensitive to non-normal data distributions
3. Are there differences in filtering efficiency across different time periods?	The filtering efficiency remains consistent throughout the month	• ANOVA or Repeated Measures ANOVA • Friedman Test	• Assumes normality and check for any variations over different time periods (e.g., weekly, weekends, or daily) • Non-parametric (alternative to ANOVA) to detect differences in filtering efficiency across different time periods
4. Is there a statistically significant difference in filtering efficiency at different times of the day?	There is no significant difference in filtering efficiency at different times of the day	• ANOVA • Kruskal-Wallis Test	• Assumes normality • Non-parametric, does not assume normality
5. Is there a statistically significant difference in filtering efficiency and external factors?	The device keeps constant filtering efficiency regardless of external factors (e.g., temperature, pollution levels)	• Multiple regression analysis • Non-parametric multiple regression	• Normality • Does not assume normality

Qualitative Analysis

The qualitative analysis of the household dataset was directed towards a better understanding of the real-world performance of air pollution filtration devices as experienced by PAIR program participants. The dataset used for this analysis was originally collected during the participant selection process by the Environmental Health Coalition (EHC) team, with additional data gathered during the installation of both air quality monitoring and air pollution filtration devices. Data collection was conducted using the AirTable platform, where each participant was assigned a unique identifier to organize and manage the information efficiently. Since all data collection took place prior to the start of this project, one of the initial steps in the research process was to ensure data confidentiality and participant privacy. Before being used for analysis, the dataset underwent a de-identification process, removing all personally identifiable information, including participant names, addresses, phone numbers, and zip codes. This step was conducted following strict data management protocols, as detailed in Appendix 2, to comply with ethical research standards and maintain the integrity of the participants' privacy.

The original AirTable dataset contained approximately 40 parameters, including demographic information such as the number of children living in the home, presence of senior residents, and other household characteristics. For this qualitative analysis, the research team focused on extracting and categorizing information that provided insight into participants' perceived experiences with the air pollution filtration devices. The analysis involved identifying recurring themes and patterns in participant responses, particularly those related to device effectiveness, ease of use, maintenance challenges, and overall satisfaction. By organizing the data into key themes and categories, the research team was able to draw meaningful conclusions regarding the practical benefits and limitations of air filtration devices in real-world household conditions. These insights contribute to a broader understanding of how indoor air quality interventions can be optimized to meet the specific needs of communities affected by air pollution.

In addition to the household dataset, EHC staff conducted a survey among PAIR participants to better understand their perceptions of the beneficial impacts of air purifier devices. A total of 253 participants living within 3 km of the Sherman air monitor were identified using Google Maps. Selection was based on specific flags (1, 2, and 3), which as described above categorized participants according to the timing of their air purifier and air quality monitor installations:

- Flag 1: Participants likely had their monitor installed during EHC's first home visit and the purifier installed on a second visit.
- Flag 2: Participants received both devices during the first visit but were instructed to turn on the purifier after 30 days.
- Flag 3: Participants had specific installation dates unrelated to the home visit schedule.

To randomize the selection process for in-person interviews, a random number function was used, and the first 45 participants were selected for direct outreach. Due to time and resource constraints, EHC aimed to conduct 45 in-person interviews, while the remaining 208 participants were contacted via email or phone. However, due to limited email availability, most participants were reached by phone, with 45 contacted via email—although no responses were received from the email group. In total, the survey yielded 9 in- and person interviews 26 phone surveys. All survey interactions followed a script pre-approved by EHC and were translated into Spanish by an EHC translator. Responses were digitally recorded in Google Forms using laptops or tablets.

Results

Quantitative Results

Descriptive Statistics

Table 3 presents the descriptive statistics for data from indoor air quality monitoring devices involved in the PAIR program, which were included in the analysis. PM data exhibited large variability, as indicated by the standard deviations across the monitoring devices, irrespective of flag classifications. Preliminary graphical comparisons, such as those shown in Figure 6, were used to assess the range of variability in PM levels before and after the installation of air purifiers in participating households. These comparisons suggested that reducing indoor PM levels could have beneficial effects for both low and high pollution levels. This, in turn, indicated that the filtering efficiency of air purifiers may vary depending on indoor pollution levels. To further explore this effect, pollution levels before and after the installation of air purifiers were compared by sorting the sampling data into 1st, 2nd, and 3rd quartiles, with comparisons made within each quartile. The results of the efficiency impacts from individual air purifier devices were further aggregated into box plots to illustrate the overall trends and behavior.

TABLE 3.DESCRIPTIVE STATISTICS OF DATA FROM INDOOR AIR QUALITY MONITORING DEVICES PARTICIPATING IN THE PAIR PROGRAM

Pollutant	Flag	Devices	Sample size	mean	SD	q1	q3
PM1	1	92	5382436	9.18	74.12	1.0	7.0
	2	77	2679560	8.34	35.28	1.0	7.0
	3	11	921093	9.17	34.26	1.0	7.0
PM2.5	1	92	5382436	11.31	81.89	1.0	8.0
	2	77	2679560	10.05	44.16	2.0	8.0
	3	11	921093	12.34	65.87	1.0	8.0
PM10	1	92	5382436	13.17	99.07	2.0	8.0
	2	77	2679560	11.42	57.85	2.0	9.0
	3	11	921093	14.97	96.59	1.0	8.0

Concentrations are in $\mu m/m^3$.

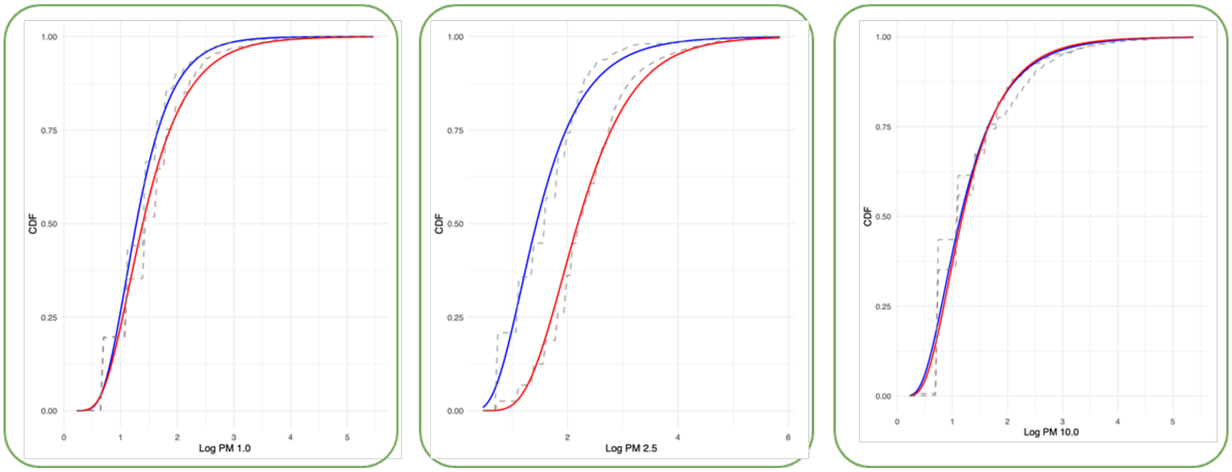

FIGURE 6. EXAMPLES OF VARIABILITY TYPES OBSERVED FOR IMPACTS OF AIR PURIFIER DEVICES USING FITTED (CONTINUOUS LINES) AND EMPIRICAL (DOTTED LINES) CUMULATIVE DISTRIBUTION FUNCTIONS FOR THREE DIFFERENT DEVICES. RED AND BLUE LINES REPRESENT PRE AND POST INSTALLATION

Efficiency of air purifiers

Results of average efficiency impacts from individual air purifier devices were aggregated into box plots as shown in Figure 7 for the case of PM2.5. The results showed that the average efficiency impacts PM1, and PM10 had a variability similar as those for PM2.5. In general, the range of beneficial impacts was smaller for air purifiers marked with flag 1 data types, suggesting some of the observed variability observed for devices marked with flags 2 and 3 is the result of the uncertainty in identifying with precision the installation date of the air purifier in the participants households. As discussed below, the variability in the observed impacts of air purifiers can be attributed to multiple factors, including indoor air pollution sources, external pollution sources such as industrial and traffic emissions, and household-specific conditions. Additional factors, such as the placement and operation of the purifier, airflow patterns, filter type and maintenance practices, household activities, and environmental conditions like humidity and temperature, further contribute to this variability. These complexities underscore the challenges in accurately assessing the sources of variability in air purifier efficiency under real-world operating conditions.

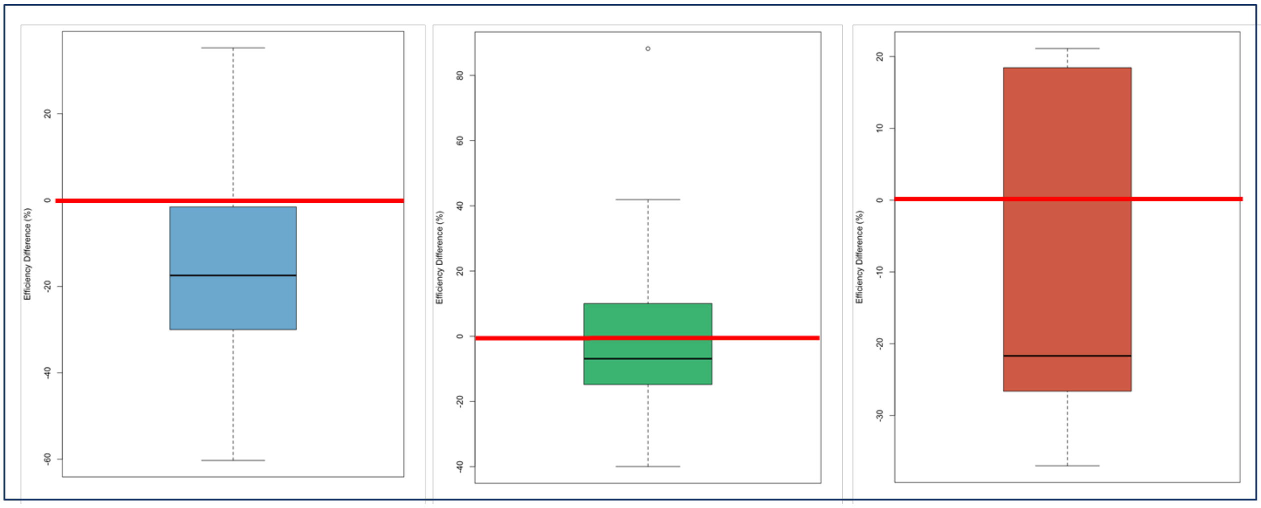

FIGURE 7. BOX PLOTS OF MEAN EFFICIENCY IMPACTS OF AIR PURIFIER DEVICES FOR PM2.5 FOR PAIR PARTICIPANTS HOUSEHOLDS FOR DEVICES WITH FLAGS 1 (LEFT PANEL), 2 (MIDDLE PANEL), AND 3 (RIGHT PANEL).

Table 4 presents the results of the efficiency evaluation of air purifiers under real-world operating conditions for participating households. As shown in Figure 7, efficiency estimates for air purifiers classified with flags 2 and 3 exhibited higher uncertainties compared to those with flag 1. Therefore, to ensure greater reliability, Table 4 includes only the results for air purifiers classified under flag 1—devices for which the installation date is known with high precision and with statistically significant hypothesis tested results. This selection minimizes uncertainty and allows for a more accurate assessment of the purifiers' performance. The results in Table 4 show the percentage reduction in particulate matter concentrations for three pollutant size categories (PM2.5, PM10, and PM1) across different time periods: all data, daytime, and nighttime. The air pollution reductions are reported at three statistical points: the 1st quantile (25th percentile, representing higher efficiency cases), the median (50th percentile, representing typical efficiency), and the 3rd quantile (75th percentile, representing lower efficiency cases).

The results on the overall efficiency across all data indicate that air purifiers effectively reduced PM concentrations, with median reductions of -31.7% for PM2.5, -29.8% for PM10, and -36.8% for PM1. The 1st quantile values show that some households experienced much greater reductions, exceeding -50% in PM concentrations. The 3rd quantile values suggest lower efficiency in some cases, with reductions as modest as -16.7% to -19.2%. Daytime efficiency was consistently higher across all PM sizes, with median reductions of -38.8% for PM2.5, -37.9% for PM10, and -41.1% for PM1. The 1st quantile reductions during the day around -60%, indicating that some households saw substantial air quality improvements. Nighttime efficiency was lower, with median reductions of -25.8% for PM2.5, -20.3% for PM10, and -28.1% for PM1.

TABLE 4. EFFICIENCY [IN %] IMPACTS OF AIR PURIFIERS IN REAL WORLD OPERATING CONDITIONS ESTIMATED FOR THE PARTICIPANT HOUSEHOLDS IN THE PAIR PROGRAM USING DEVICES WITH FLAG 1 CATEGORY.

Pollutant	1st quantile	Median	3rd quantile
All data			
PM2.5	-53.5	-31.2	-17.7
PM10	-48.6	-30.0	-17.1
PM1	-55.6	-33.3	-18.3
Daytime			
PM2.5	-58.1	-35.0	-20.7
PM10	-55.6	-34.4	-18.4
PM1	-55.6	-37.7	-19.3
Nighttime			
PM2.5	-41.1	-23.8	-12.9
PM10	-38.9	-18.2	-11.1
PM1	-50.0	-27.8	-16.1

A notable difference exists between the highest and lowest efficiency cases, emphasizing that household conditions, air circulation, and purifier usage patterns influence performance. The most significant reductions occurred in the 1st quantile during the daytime, suggesting that air purifiers were particularly effective under certain conditions. These results confirm that air purifiers can significantly reduce indoor PM concentrations, especially during daytime hours. However, efficiency varies across households due to factors such as air circulation, pollution sources, and purifier placement. While some households experienced reductions around -60%, others saw more modest improvements, particularly at night.

Indoor air quality and external factors

In this section, we present the results of analyzing the relationships between external factors and indoor air quality following the installation of air purifier devices in participating households. A key consideration when comparing indoor and ambient air quality data is the time resolution of the sampling datasets. Indoor air quality data was collected at a higher frequency (every 5 minutes) compared to ambient data, which is typically recorded at 1-hour intervals. The direct comparison of high-frequency indoor data with lower-frequency ambient data can introduce artifacts or misleading conclusions. To address this, we investigated the impact of varying levels of smoothing applied to the indoor air quality dataset, progressively reducing its frequency, and observed how this adjustment affected the variability and strength of the observed associations. The association between indoor and outdoor parameters was first investigated using Indoor/Outdoor (I/O) ratios and then through linear correlation analysis.

Figure 8 shows the time-series scatter plot of the I/O of PM2.5 for different time resolutions for all devices using post-filtration sampling periods. In the figure, each subplot corresponds to a different time resolution (2 to 24 hours) and a red dashed line serves as a reference threshold for and I/O = 1. The figure shows that the spread of points remains relatively stable across different time resolutions, but some spikes and variations are noticeable at higher resolutions. As discussed below, the fact that some time periods show occasional high I/O ratios, suggests that the filtration from air purifiers may be less effective under certain conditions. A summary of the I/O ratio of PM2.5 across different time resolutions (2 to 24 hours) is shown in box plots of Figure 9. The figure shows that most I/O ratios are below the reference line, with some extreme outliers, suggesting that filtration effectively reduces indoor PM2.5 compared to outdoor levels.

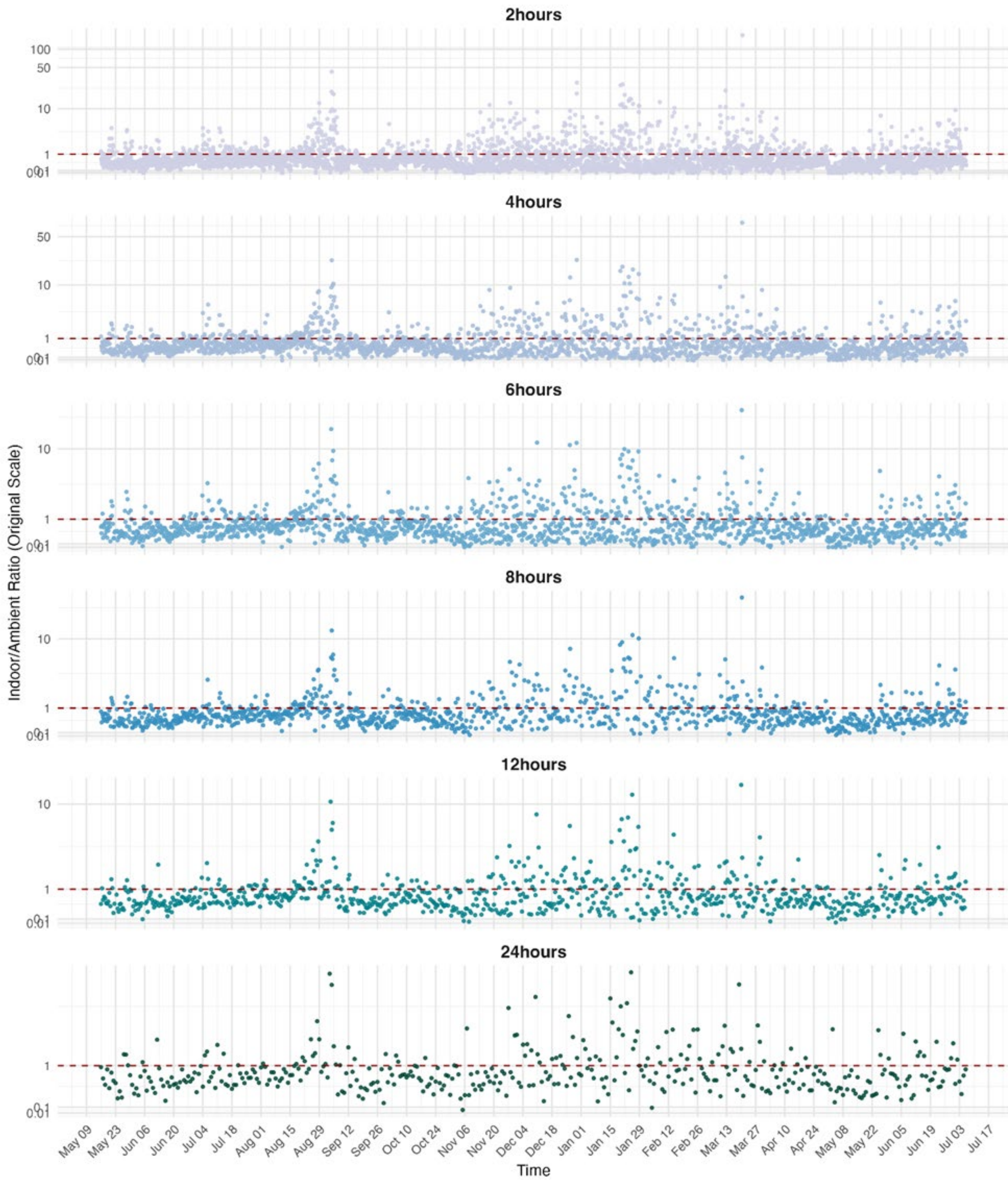

FIGURE 8. PM2.5 INDOOR/AMBIENT RATIO OVER TIME, BY TIME RESOLUTION.

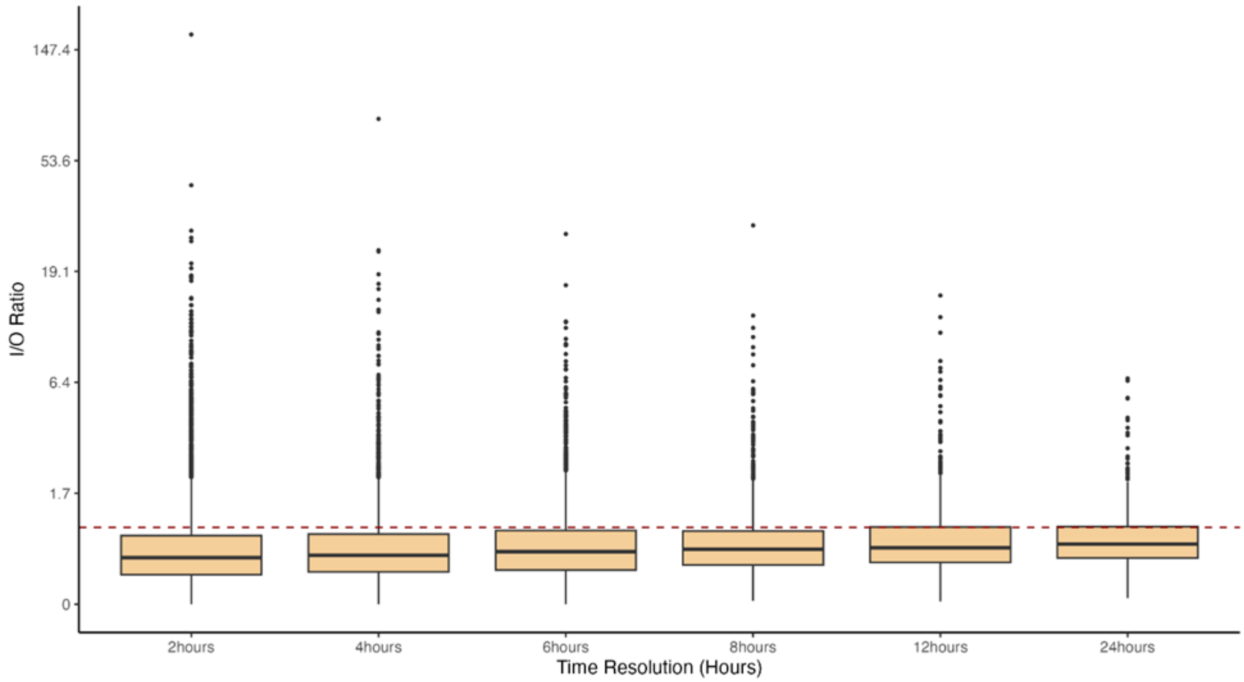

FIGURE 9. SUMMARY OF PM2.5 INDOOR/OUTDOOR (I/O) RATIOS ACROSS DIFFERENT TIME RESOLUTIONS.

The relationship between indoor and outdoor PM2.5 concentrations over different time resolutions during the post-filtration period was also analyzed using linear correlations. Figure 10 shows six scatter plots, each representing a different time resolution (2 hours, 4 hours, 6 hours, 8 hours, 12 hours, and 24 hours) for the normalized median indoor PM2.5 value and the normalized ambient (outdoor) PM2.5 value. Normalization in this regard was performed by dividing each data point by the sampling average. In the figure, each plot illustrates the correlation between indoor and outdoor PM2.5, with a red dashed line indicating a trend or reference line. Overall, the scatter plots indicate a positive relationship between indoor and outdoor PM2.5, indicating that indoor air quality is influenced by outdoor pollution levels. The red dashed line marks a linear trend, but the spread of points at higher values implies variability in how well indoor PM2.5 linearly tracks outdoor levels. At shorter time intervals (2–8 hours), we see greater variability in the data, suggesting more fluctuations in indoor air quality. At longer time resolutions (12–24 hours), the data become more clustered, meaning daily averages smooth out short-term fluctuations.

Figure 11 presents a summary of the linear correlations between indoor PM1, PM2.5, and PM10 and outdoor PM2.5 concentrations at various time resolutions (2 hours to 24 hours). The error bars indicate the variability (standard deviation) of these correlations. The figure shows that the correlation values increase as the time resolution increases, meaning that longer averaging periods lead to stronger correlations between indoor PM levels and outdoor PM2.5. This suggests that outdoor PM2.5 has a more significant influence on indoor PM levels over longer periods, likely due to air exchange and filtration effects. The results show that the linear correlation somewhat increases by increasing the smoothing period, but without necessarily implying linearity in their association by physical processes.

FIGURE 10. NORMALIZED PM2.5 MEDIAN INDOOR VS NORMALIZED AMBIENT OVER TIME, BY TIME RESOLUTION.

Overall, the results show that mean linear correlation values increase with longer averaging periods, indicating that indoor PM concentrations exhibit a stronger relationship with outdoor PM2.5 as the time window expands. Among the different PM types, PM1 consistently exhibits the highest correlation with outdoor PM2.5, followed by PM2.5, while PM10 demonstrates the weakest correlation. These findings suggest that finer particles (PM1 and PM2.5) are more influenced by outdoor air infiltration compared to coarse particles (PM10), which are likely driven more by indoor sources. Error bars in the figures indicate variability in correlation values, with larger uncertainties observed at longer time resolutions. This may reflect the influence of diverse indoor environments, variations in air filtration efficiency, or different ventilation rates among residences.

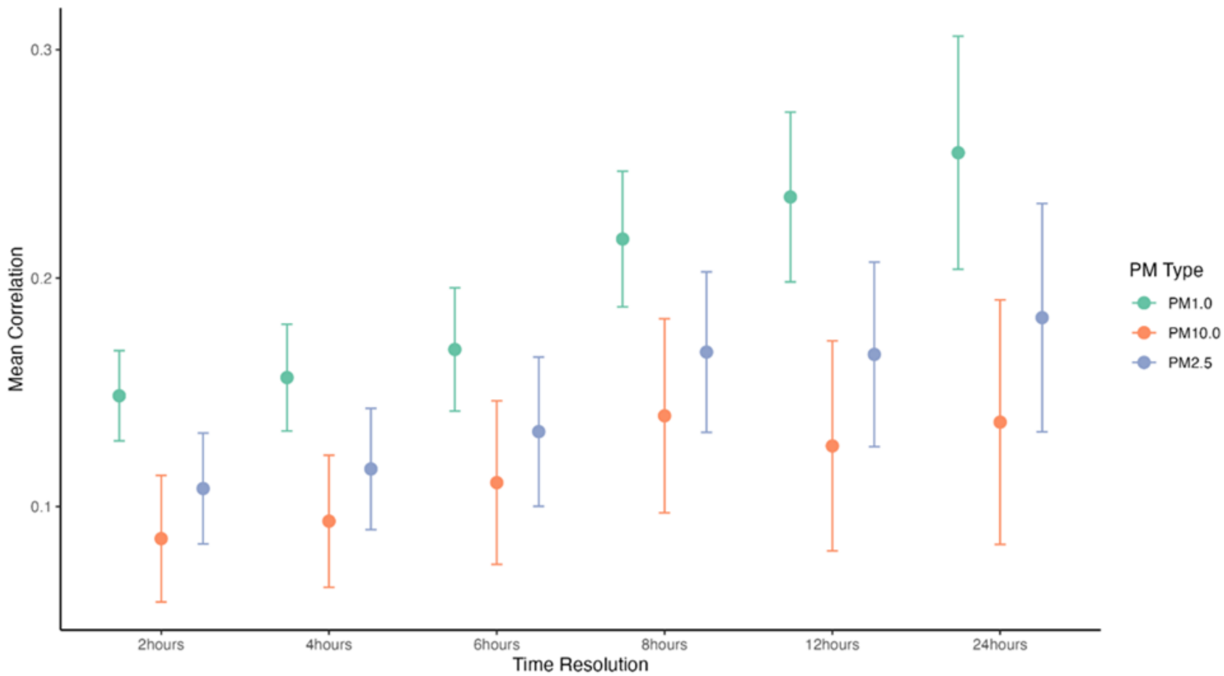

Figure 11. Summary of post-filtration grouped mean correlation for Indoor and Outdoor PM1, PM2.5, and PM10.

The results show that PM1 consistently shows the highest correlation with outdoor PM2.5 across all time resolutions. This may imply that finer particles penetrate indoor spaces more effectively and remain suspended in the air for longer periods, making their concentration more directly linked to outdoor PM2.5 levels. Conversely, PM10 has the lowest correlation values across all time scales. This may indicate that coarser indoor particles are less influenced by outdoor PM2.5 conditions, likely because they settle faster and are more affected by indoor activities rather than infiltration from outdoor air. PM2.5 shows a correlation trend between PM1 and PM10, which aligns with expectations since it is the size fraction being compared to outdoor PM2.5 conditions. The increasing trend over time suggests that outdoor PM2.5 significantly contributes to indoor PM2.5 levels, but with some indoor sources and filtration effects modulating the relationship. The error bars widen as the time resolution increases, indicating that the correlations become less precise over longer periods. As discussed below, the observed variability might be due to differences in indoor environments, filtration effectiveness, or fluctuations in outdoor pollution levels.

Figure 12 presents the mean correlation between wind speed measurements from a single outdoor monitoring site (Sherman Elementary monitoring site) and indoor PM2.5 concentrations across four spatial sectors: Central, North East, North West, and South surrounding the site and different time resolutions (ranging from 2 hours to 24 hours). In the Figure the error bars representing variability in the form of standard deviations.

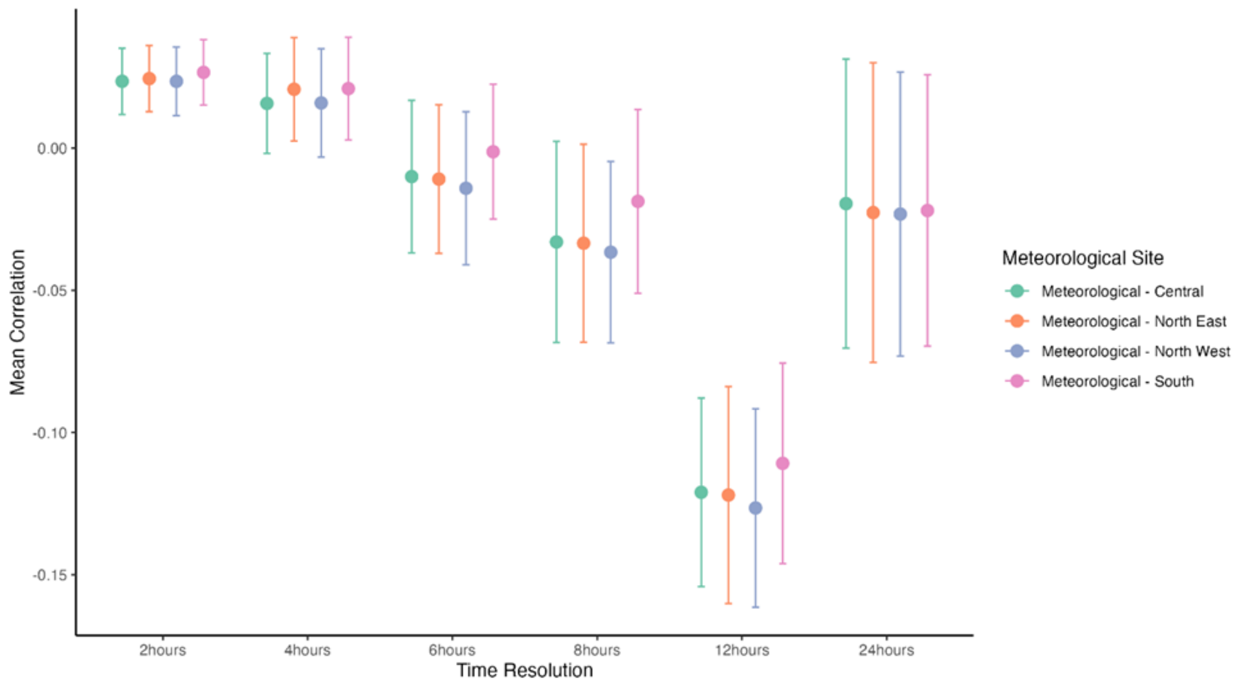

FIGURE 12. SUMMARY OF CORRELATION BETWEEN WIND SPEED MEASUREMENTS FROM THE SHERMAN ELEMENTARY MONITORING SITE AND INDOOR PM2.5 CONCENTRATIONS ACROSS FOUR SPATIAL SECTORS: CENTRAL, NORTHEAST, NORTHWEST, AND SOUTH SURROUNDING THE SITE.

Figure 12 shows that correlations between wind speed and indoor PM 2.5 concentrations tend to decrease as the time resolution increases, moving from near-zero correlations at shorter timescales (2–4 hours) to more negative correlations at longer timescales (8–24 hours). This may suggest that the immediate impact of wind speed on indoor PM2.5 is limited, but over extended periods, wind may contribute to reducing indoor PM2.5 concentrations. The four sectors (Central, Northeast, Northwest, South) correspond to spatially grouping the air purifier devices surrounding the air monitoring site. The results indicate that the four sectors exhibit similar trends, with overlapping confidence intervals and that there is no clear dominant sector where wind speed has a stronger effect, implying a relatively uniform impact of wind on indoor PM2.5 across directions. Nevertheless, the variability of the correlations (indicated by error bars) increases with longer time resolutions, suggesting greater uncertainty in wind's impact on indoor PM2.5 at extended timescales. These results highlight the potential role of wind in modifying indoor PM2.5 concentrations, with the effect becoming more apparent over longer time periods, albeit with increased uncertainty.

Qualitative Results

Device Usage and Satisfaction

The survey results from participants on air purifier and air monitor device usage and satisfaction provide valuable insights into user experiences and habits. The findings highlight regular engagement with the devices, as 80% of respondents reported using the air purifier daily, while another 14.3% used it weekly. Most participants (65.7%) set up the devices themselves, with an equal number of respondents finding the setup process very easy (n=11) or easy (n=11). In terms of perceived air quality improvements, 91.4% of respondents noticed positive changes, primarily reduced humidity and dust levels. Some participants attributed these improvements to alerts from the air monitor, while others described the impact as:

- *"Easier to breathe and sleep at night."*

However, despite the overall positive experiences, one respondent noted a negative change:

- *"The first few days, indoor air quality was really nice. After a few months, it began to give off a foul odor."*

Regarding ventilation practices, 71.4% of respondents relied solely on open windows, while 11.4% used only HVAC systems. Among those using HVAC, the majority (n=6) reported using it in combination with the air purifier during air pollution episodes. The air purifier was most commonly installed in the living room (42.9%) and individual rooms (37.1%), followed by the kitchen (8.6%). In terms of cooking habits, 80% of respondents reported cooking with natural gas, with most preparing meals twice a day (48.6%) or three times a day (25.7%). These findings suggest that the air purifier devices are well-integrated into daily routines and are perceived to improve indoor air quality, particularly in households exposed to cooking-related emissions and external pollutants.

Health Impacts and Perception

Findings on the perceived health impacts of air purifiers provide valuable insights into user experiences. A significant portion of respondents (74.3%) reported noticeable improvements in their health and well-being since using the device. These improvements included fewer allergy symptoms, reduced coughing, fewer illnesses, and easier breathing. Additionally, 74.3% of participants indicated that at least one household member had a pre-existing respiratory condition, such as COPD, asthma, or allergies. Notably, all applicable respondents stated that the air purifier helped alleviate symptoms associated with these conditions and reduced the frequency of respiratory attacks. One participant shared:

- *"I've hardly had any asthma attacks since using the purifier."*

To assess whether the use of air purifiers influenced medication dependence, participants were asked if they or a household member with a respiratory condition had experienced a reduced need for medication. As shown in Figure 13, the majority of applicable respondents (57.1%) reported a decreased reliance on inhalers and allergy medications since incorporating the air purifier into their daily routine. These findings suggest that air purifiers may play a meaningful role in improving respiratory health and reducing the need for medication among individuals with pre-existing conditions.

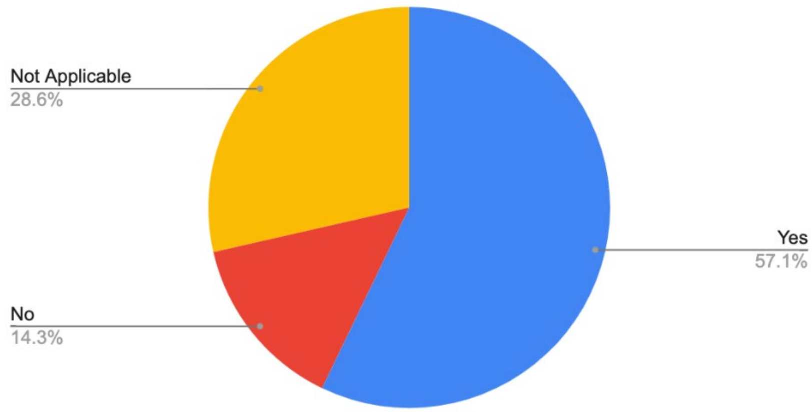

FIGURE 13. PARTICIPANT RESPONSES ON WHETHER AIR PURIFIER USE REDUCED THE NEED FOR MEDICATION IN RESPIRATORY CONDITIONS.

Many respondents (51.4%) reported that other household members, including children, elderly individuals, and even pets, experienced health benefits after the installation of the air purifier. Among those with children, several noted significant improvements. One participant shared:

- *“My son is asthmatic and has allergies, and the purifier has helped alleviate the symptoms.”*

Another respondent described a dramatic improvement in their child's health:

- *“Before the installation of the device, [my daughter] suffered chronically from these issues to the point that I thought she had asthma, but that has all gone away thanks to the air filter.”*

Similarly, a participant with a family member who had undergone lung surgery observed:

- *“They have noticed that it has helped since smoke leads to coughing fits.”*

Also pet owners reported benefits with the use of air purifiers in their households, with some noting that their pets sneezed less frequently after using the purifier. Beyond respiratory improvements, 77.1% of users noticed enhanced sleep quality, increased energy levels, or greater overall comfort. A common sentiment was:

- *“More energy and better sleep.”*

Several respondents also mentioned waking up less frequently at night. Additionally, the purifiers were credited with reducing the effects of mold, cooking fumes, and excess humidity, contributing to a healthier indoor environment. One participant summarized the overall impact by stating:

- *“Sleeping has been made easier, energy levels have increased, and the overall respiratory health of the entire family has improved.”*

Participants widely highlighted a greater sense of well-being, emphasizing better air quality and improved family health since the installation of the air purifiers.

Health Concerns and Recommendations

The results indicated that 65.7% of respondents had concerns regarding indoor air quality affecting their health before receiving the air purifier. As shown in Table 5, common worries included air pollution from nearby sources, such as emissions from NASSCO (National Steel and Shipbuilding Company), as well as indoor pollution from household activities. Additionally, many participants expressed awareness that poor air quality could exacerbate respiratory conditions, further heightening their concerns.

TABLE 5. CLASSIFICATION AND EXAMPLES OF RESPONSES OF PARTICIPANTS REGARDING THEIR CONCERNS ON INDOOR AIR QUALITY AFFECTING THEIR HEALTH.

Category	**Responses**
External Pollution Sources	• *Yes, since air quality is very bad in the area due to NASSCO.* • *Air quality coming from NASSCO.* • *They were next to a freeway and NASSCO, so they were concerned about air quality and health.* • *Yes, because they previously lived near the military base. The house was older which caused health problems to amplify.* • *Area that they reside in is heavily polluted in the air.* • *Live in barrio Logan where the air quality isn't the best so purifier makes air quality so much better.* • *We knew that the air outside of our home was polluted so we knew were breathing the polluted air from outside, inside our apartment, and knew it wasn't any good for us.* • *Particulate matter when you open the door was visible. There is construction 4 months now nearby and it impacted the air quality.* • *At the time I was living in an apartment above commercial space that was being remodeled. Dust particles from construction entered my home as well as Aerosol Spray.*
Indoor Pollution Sources	• *Fumes from cooking and secondhand smoke.* • *Using our gas stove.* • *Was very dusty which in turn made [the participant] wheeze a lot.*
Health Issues	• *She would get bronchitis often* • *Children present in home and other members would get sick frequently raising concerns about air quality* • *Más gripa y alergias antes. / More flu and allergies before.* • *Worse air quality and asthma in community.* • *The air felt very contaminated before, she used to have a light plant very nearby before.* • *Knew that overall environment was impacting her health.* • *Felt that air was polluted and that it would result to more allergies*
Awareness	• *They knew that the air was polluted but weren't aware of how polluted it was inside their homes.*

Nearly all respondents (97.1%) believed that air purifiers provide significant benefits for individuals with respiratory issues, citing improvements in air quality, easier breathing, and reduced allergy symptoms. Many users noted that the purifier helped create a cleaner, healthier indoor environment, alleviating health concerns and enhancing comfort—particularly for children, elderly individuals, and those with pre-existing conditions. Additionally, some participants highlighted the role of air quality monitors in keeping them informed. They appreciated receiving alerts about activities that negatively impact indoor air quality and knowing when it was safe to open windows for ventilation. These findings suggest that air purifiers are widely perceived as effective tools for enhancing indoor air quality and promoting better respiratory health in vulnerable populations.

Challenges and Barriers

Most users reported no significant difficulties with the air purifier or monitor. However, a small percentage (14.3%, n=5) encountered specific challenges, as summarized in Table 6. Among the reported issues, three were related to maintenance, including concerns about the device emitting an unpleasant odor and the inconvenience of replacing filters. Additionally, some users experienced technical difficulties, such as misplacing the monitor cable or encountering problems reconnecting the device to Wi-Fi. Overall, while the majority of participants found the devices easy to use, these challenges highlight areas where improved guidance on maintenance and troubles shooting could enhance user experience.

TABLE 6. CLASSIFICATION AND EXAMPLES OF RESPONSES OF PARTICIPANTS REGARDING CHALLENGES USING THE AIR PURIFIER OR AIR QUALITY MONITOR DEVICES.

Category	Responses
Hardware	• *Lost the cable for monitor, is interested in a new one. Unsure how frequently the filter must be changed.*
Software	• *Had difficulty reconnecting the air monitor after changing their Wi-Fi.*
Maintenance	• *Cost of filters can be a bit pricey, and their house is older resulting in need to replace filters with more frequency.* • *The filter began to give off a bad odor, tried to change the filter and did not help.* • *Changing the filters is a drag.*

Despite the challenges described, the participants indicated that none were severe enough to prevent regular use of the devices. When asked, most participants stated that nothing hindered them from using the devices more frequently. Among the few who reported reduced usage, two participants attributed it to cold winter weather, while another noted that they used the purifier only at night, as they were away from home during the day. A few participants experienced device-related issues, including one who misplaced the monitor cable and another who reported that their purifier emitted an unpleasant odor. Additionally, one participant raised concerns about the device's accuracy and expressed interest in relocating it for better performance. Another user commented that the filter design allowed some dust to pass through, suggesting that it could be improved for better efficiency. Overall, while minor inconveniences were noted, the vast majority of participants found the devices reliable and beneficial for regular use.

Survey results indicate that 91.4% of participants had no concerns or frustrations related to the air quality monitor. Among those who did, all issues were hardware-related. Specifically, one participant misplaced the monitor, another lost the cable, and one reported having "*older cables.*" Additionally, one respondent expressed uncertainty about the accuracy of the data, stating a desire to "…*know where the data is coming from.*" Overall, the most common barriers to more frequent use were related to the cost of replacement filters, cold weather, and occasional maintenance tasks, such as changing filters.

The overall experience with the air purifiers and air quality monitors was highly positive for most respondents, with many expressing satisfaction and gratitude for the program.cA majority (74.3%) of users rated their experience with the air purifier as 'very positive', as shown in Figure 14. Similarly, the air quality monitor received a 'very positive' rating from 57.1% of participants. Notably, no respondents reported an overall 'negative' or 'very negative' experience, further highlighting the perceived benefits and acceptance of the devices.

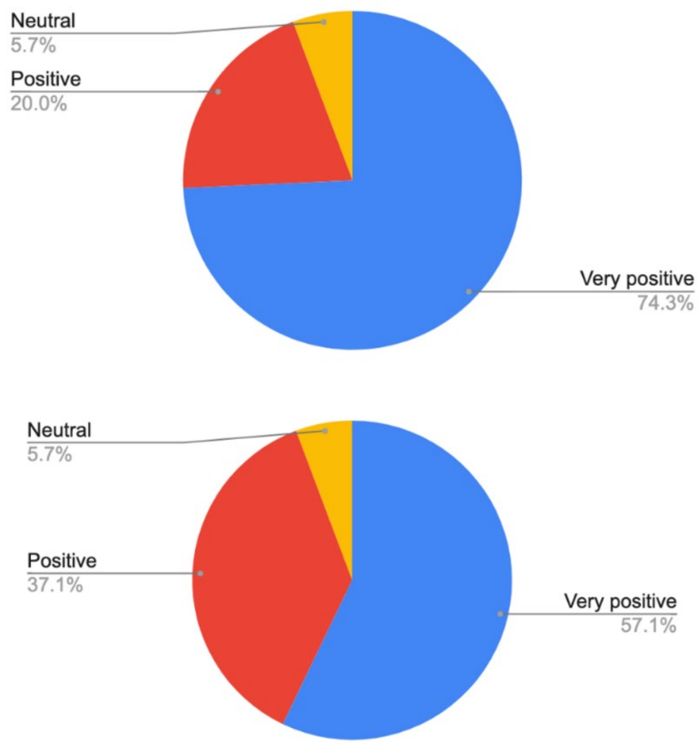

FIGURE 14. RESPONSES RATINGS OF THE PARTICIPANT'S EXPERIENCE USING THE AIR PURIFIER (TOP) AND AIR QUALITY MONITORS (BOTTOM).

All survey respondents reported that they would recommend the devices to others. In response to an open-ended question about their experiences with the devices and the program, participants provided a range of feedback, including positive, neutral, and constructive comments:

- Positive responses (n=19) highlighted satisfaction with the performance of both the air purifier and air monitor, as well as gratitude for the program.

- Neutral feedback included important notes about device usage and requests, such as continuing to use the air purifier after moving, losing the monitor but still valuing the program, and inquiring about how to obtain replacement filters or cables.
- Constructive feedback pointed to areas for improvement:

    - One participant mentioned they were promised a check to help with electricity costs but never received it.
    - Another noted that their air purifier began emitting a bad odor, even after changing the filter.
    - One respondent expressed confusion about a relative with asthma not qualifying for the program.
    - Another participant sought further clarification and resources, stating: *"Would recommend the purifier to others but thinks the monitor could be improved. Tried to reach out to someone but wasn't sure who to contact and wants resources about who to reach out to."*

Discussions

Quantitative Analysis

The findings from this study highlight the effectiveness of air purifiers in reducing indoor PM concentrations under real-world operating conditions. The observed reductions in PM1, PM2.5, and PM10 concentrations confirm the potential benefits of air purifiers in improving indoor air quality for participating households in the PAIR program. However, the variability in efficiency across different households underscores the influence of multiple environmental and operational factors. One key observation is that air purifiers were generally more effective during the daytime than at night. The higher reductions in PM levels observed during daytime hours may be attributed to increased air circulation and household activities that help distribute filtered air more efficiently. Conversely, lower efficiency at night could be associated with reduced air movement and possible resuspension of particles from indoor surfaces. These findings suggest that air circulation patterns play a crucial role in determining the effectiveness of air purification.

Another important aspect is the variability in efficiency among different devices, as indicated by the comparison between air purifiers categorized under flag 1 versus those under flags 2 and 3. The lower uncertainty in efficiency estimates for flag 1 devices reinforces the importance of precise installation date tracking when evaluating air purifier performance. The greater variability observed in flag 2 and 3 devices likely reflects inconsistencies in installation timing and operational differences, which introduce challenges in accurately assessing efficiency.

The results of comparing the efficiency of air purifiers under real-world operating conditions in participating households showed large variability. Understanding these sources of variability is essential for optimizing air purifier effectiveness and improving future implementation strategies. Potential sources of the observed variability can include:

- *Pollution levels*: Variability in indoor air pollution PM levels can influence the effectiveness of air purifiers. Air purifiers may perform differently in households with low versus high pollution levels. We explored the effect by comparing efficiency impacts by quantiles of sampled data and our results indicated higher efficiency of air purifiers under low PM pollution levels.
- *Household size and layout*: Larger households or those with open floor plans may experience different purification efficiency compared to smaller or more compartmentalized homes due to differences in air mixing conditions. Similarly, the natural airflow patterns in the household (e.g., ventilation, windows, doors), can impact its ability to filter air effectively across the entire space. Also, the effectiveness of the purifier may be reduced if the device is unable to circulate air throughout the entire space. Activities such as cooking, smoking, cleaning, or pet ownership also can introduce additional pollutants into the air, affecting the purifier's efficiency. However, since household size and layout data were not available in this study, we were not able to control for these confounding factors.
- *Air purifier device placement*: Placement of the air purifier near sources of pollution (e.g., kitchen, smoking areas) or obstructions (e.g., furniture, walls) can affect its performance. Devices placed in areas with poor air circulation may have reduced efficiency. Unfortunately, this information was not part of the household database of the PAIR program.

- *Filter type and condition*: Different air purifiers use different filter technologies (e.g., HEPA, carbon filters), and the condition of the filters (e.g., clogging or aging) can influence performance. The frequency of filter changes can also impact the efficiency of the device. In the case of this study, these effects are minimized in the results of the comparisons of the efficiencies because all participating households in the PAIR program received the same air purifier model and we used the same criteria for selecting the data periods for analysis.
- *Operating time and user behavior*: Variability in the duration and frequency of the air purifier's operation, including factors like whether it runs continuously or intermittently, can affect how well it reduces pollutants over time. Similarly, variability in how households use the purifier—such as whether they run it at full capacity, adjust settings, or clean the device regularly—can influence the outcomes of efficiency comparisons. The results from the qualitative analysis indicate that most participants of the PAIR program run the air purifier continuously, suggesting that this may be a small source of variability in the observed efficiency comparisons.
- *External factors*: External sources of air pollution, such as traffic from nearby roads, industrial emissions, or regional transport of air pollutants, can impact indoor air quality and thus influence the purifier's performance. High humidity or temperature changes can also impact the performance of certain types of air purifiers, particularly those with filters that are sensitive to moisture or temperature variations. These factors highlight the complexity of assessing air purifier efficiency in real-world settings, where conditions may differ significantly from controlled experimental environments.

With respect to the comparisons between indoor and outdoor ambient PM levels, the observed trends provide key insights into the dynamics of indoor and outdoor particulate matter interactions in post-filtration conditions. The increasing correlation over time suggests that while short-term fluctuations in indoor PM levels may be influenced by transient indoor activities (e.g., cooking, cleaning, or human movement), the longer-term indoor PM concentrations tend to reflect outdoor air conditions more closely. This aligns with prior studies indicating that smaller PM fractions, particularly PM1 and PM2.5, have higher penetration efficiencies and longer residence times indoors.

The higher correlation of PM1 with outdoor PM2.5 likely reflects its ability to infiltrate indoor environments more effectively due to its smaller aerodynamic diameter. In contrast, PM10 demonstrates weaker correlation, which is consistent with the fact that larger particles are more affected by localized indoor sources such as dust resuspension and indoor activities rather than outdoor infiltration.

One key implication of these findings is the effectiveness of air purifier systems in modifying indoor air quality. If filtration were highly effective in removing outdoor-origin PM, we would expect lower indoor-outdoor correlations across all PM fractions. The moderate-to-strong correlation observed for PM1 and PM2.5 suggests that while filtration is likely reducing PM concentrations, it may not be fully preventing outdoor air penetration, particularly for smaller particles. Future work should assess the performance of different filtration systems, ventilation strategies, and air exchange rates to optimize indoor air quality control. Furthermore, the increasing uncertainty at longer time resolutions, as indicated by larger error bars, suggests variability in household behaviors and filtration system efficiency across different residences. This variability may be due to differences in indoor activities, air exchange rates, or the frequency of filter maintenance. Additional investigation into these factors could provide more targeted recommendations for improving indoor air quality management strategies.

Qualitative Analysis

The qualitative results indicate that the air purifiers and air quality monitors were well-integrated into participants' daily routines, with a majority using the devices consistently. The high rate of daily air purifier usage (80%), along with the 91.4% of respondents who perceived improvements in air quality, suggests strong engagement and a clear perceived benefit in reducing indoor air pollution. The most commonly cited improvements, including reduced dust and humidity levels, align with expected outcomes of air filtration systems. Additionally, participants' ventilation practices, with 71.4% relying on open windows and 11.4% using only HVAC systems, highlight differences in air circulation strategies, which may influence the effectiveness of the purifiers in different household environments. These findings underscore the importance of considering household ventilation habits when assessing air purifier performance.

In addition to improved air quality, health-related benefits were widely reported. A large percentage (74.3%) of participants noted improvements in respiratory health, citing fewer allergy symptoms, reduced coughing, and easier breathing. The decreased reliance on medication (57.1%) among those with pre-existing respiratory conditions further suggests a potential positive impact on symptom management. Notably, several participants reported that family members, including children, elderly individuals, and even pets, also experienced benefits, reinforcing the broader household-wide impact of air filtration. These qualitative responses, including statements such as "*I've hardly had any asthma attacks since using the purifier*," add strong anecdotal support to the perceived health benefits. However, controlled studies with objective air quality and health measurements would be necessary to quantify these effects more precisely.

Despite the overwhelmingly positive reception, some challenges and areas for improvement were identified. Maintenance concerns, such as replacing filters and occasional device malfunctions, were among the most common barriers to consistent use. A small percentage of respondents (14.3%) experienced technical issues, including misplaced cables or difficulty reconnecting devices to Wi-Fi. Additionally, one participant raised concerns about the device's accuracy, highlighting the need for greater transparency in data interpretation and usability improvements. While these issues did not significantly hinder overall usage, they suggest potential opportunities to enhance user support and troubleshooting resources. Ensuring that replacement filters remain affordable and accessible could further encourage long-term adoption and consistent use of the devices.

Conclusions

The results from this study strongly indicate that air purifiers can effectively reduce indoor PM concentrations in real-world settings, with median reductions of approximately 30–40% across different PM size fractions. The highest efficiency was observed in the 1st quantile of reductions, where some households experienced PM decreases around 60%, particularly during daytime hours. However, efficiency varied across households due to differences in environmental conditions, air circulation, and operational factors. The findings underscore the importance of proper air purifier placement, regular maintenance, and continuous operation to maximize efficiency. Additionally, the results highlight the need for precise tracking of installation dates and operational conditions to minimize uncertainty in evaluating air purifier impacts. Future studies should consider controlled interventions to better isolate the effects of air purifiers from other influencing factors and explore long-term health benefits associated with improved indoor air quality. Overall, air purifiers represent a valuable tool for mitigating indoor air pollution, particularly in households exposed to high levels of PM, but their effectiveness is contingent upon optimized use and favorable environmental conditions.

The results suggest that air purifiers and air quality monitors were well-received by participants, with high levels of usage and perceived effectiveness. The majority of respondents reported improvements in air quality, respiratory health, and overall well-being, demonstrating a strong alignment between device functionality and user expectations. Additionally, the results highlighted that air pollution concerns, both indoor and outdoor, were prevalent among participants before receiving the devices, reinforcing the need for continued efforts to improve air quality in vulnerable communities. While most users did not encounter major difficulties, some maintenance challenges and technical issues were noted, particularly regarding filter replacements, device connectivity, and data interpretation. Addressing these concerns through better user support, improved device accessibility, and enhanced educational resources could further optimize the program's impact. Overall, the findings indicate that air purifiers and air quality monitors are valuable tools for improving indoor air quality and promoting better respiratory health. Future research could focus on objective measurements of air quality improvements, as well as long-term health outcomes associated with air purifier use, to provide further scientific validation of these perceived benefits.

References

Allen, J. G., MacNaughton, P., Cedeno Laurent, J. G., Flanigan, S., Cao, X., Buonocore, J. J., ... & Spengler, J. D. (2019). Air purifier intervention for reducing indoor fine particulate matter and respiratory symptoms in Beijing, China: A randomized controlled trial. Indoor Air, 29(3), 423-433.

Balakrishnan, K., Parikh, J., Sankar, S., Padmavathi, R., Srividya, K., Venugopal, V., … & Smith, K. R. (2013). Daily average exposures to respirable particulate matter from combustion of biomass fuels in rural households of southern India. Environmental Health Perspectives, 111(9), 1235–1240.

Bruce, N., Perez-Padilla, R., & Albalak, R. (2000). Indoor air pollution in developing countries: A major environmental and public health challenge. Bulletin of the World Health Organization, 78(9), 1078–1092.

California Air Resources Board (CARB). (2023). Portside Environmental Justice Neighborhoods: Community Air Monitoring Plan and Community Emissions Reduction Program. Retrieved from

Chen, C., & Zhao, B. (2020). Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and exposure. Building and Environment, 180, 107021.

Chen, R., Yin, P., Meng, X., Liu, C., Wang, L., Xu, X., & Kan, H. (2022). Associations between long-term exposure to ambient PM2.5 and mortality in China: An analysis of 74 million deaths. The Lancet Planetary Health, 6(5), e386-e395.

Díaz, E., Smith-Sivertsen, T., Pope, D., Lie, R. T., Díaz, A., McCracken, J., … & Bruce, N. (2007). Eye discomfort, headache, and back pain among Mayan Guatemalans exposed to indoor air pollution: A randomized intervention trial. Environmental Health Perspectives, 115(4), 608–612.

Mendell, M. J. (2007). Indoor residential chemical emissions as risk factors for respiratory and allergic effects in children: A review. Indoor Air, 17(4), 259–277.

Naeher, L. P., Brauer, M., Lipsett, M., Zelikoff, J. T., Simpson, C. D., Koenig, J. Q., & Smith, K. R. (2007). Woodsmoke health effects: A review. Inhalation Toxicology, 19(1), 67–106.

Office of Environmental Health Hazard Assessment (OEHHA). (2023). CalEnviroScreen 4.0. California Environmental Protection Agency. Retrieved from https://oehha.ca.gov/calenviroscreen

Pope, C. A., & Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. Journal of the Air & Waste Management Association, 56(6), 709–742.

Puthumana JS, Ngaage LM, Borrelli MR, Rada EM, Caffrey J, Rasko Y: Risk factors for cooking-related burn injuries in children, WHO Global Burn Registry. Bull World Health Organ. 2021 Jun 1;99(6):439-445.

San Diego County Health & Human Services Agency. (2022). Community Health Statistics Unit: Health and Socioeconomic Profiles. Retrieved from https://www.sandiegocounty.gov/content/sdc/hhsa

Weschler, C. J., & Nazaroff, W. W. (2008). Semivolatile organic compounds in indoor environments. Atmospheric Environment, 42(40), 9018–9040.

Vardoulakis S, Giagloglou E, Steinle S, Davis A, Sleeuwenhoek A, Galea KS, Dixon K, Crawford JO. Indoor Exposure to Selected Air Pollutants in the Home Environment: A Systematic Review. Int J Environ Res

Public Health. 2020 Dec 2;17(23):8972. doi: 10.3390/ijerph17238972. PMID: 33276576; PMCID: PMC7729884.

Wang, C., Feng, S., Gao, Y., & Kuraoka, S. (2020). Effectiveness of residential air purifiers in reducing indoor PM2.5: A meta-analysis. Environmental Research, 188, 109759.

Zhang, J., & Smith, K. R. (2003). Indoor air pollution: A global health concern. British Medical Bulletin, 68(1), 209–225.

Zhao, X., Wang, J., Qiu, H., & He, H. (2021). Review on the effectiveness of portable air purifiers in removing airborne pollutants and improving health outcomes. Science of the Total Environment, 777, 145791.

Appendix 1: Project Tasks

The following sections describe each of the tasks undertaken to complete the project. These tasks were designed to ensure a rigorous and systematic approach to data collection, analysis, and reporting, with a strong emphasis on community engagement, data security, and scientific rigor.

Task 1: Establishment of an Advisory Committee

To ensure that the project was conducted in a manner that was both scientifically robust and respectful of community concerns, SDSU collaborated with the Environmental Health Coalition (EHC) and the San Diego Air Pollution Control District (SDAPCD) to establish an Advisory Committee. The committee included representatives from SDSU, SDAPCD, and EHC, each contributing expertise in air quality monitoring, community engagement, and environmental justice.

The primary responsibilities of the Advisory Committee were to:

1. Identify community priority questions – The committee provided critical insights to align the research objectives with the concerns and needs of the Portside Community. This ensured that the study effectively assessed the benefits of indoor air monitors and air purifiers in real-world conditions.
2. Guide the development of data collection and management activities – The committee provided oversight for data handling and security measures during Tasks 2–5, ensuring compliance with best practices for protecting participant confidentiality.
3. Review the literature and develop a work plan – The committee assisted in refining the research methodology (Task 6) to ensure that the study design and analytical methods aligned with scientific and regulatory standards.
4. Oversee data analysis procedures – The committee provided input on statistical methodologies and ensured that findings were presented in a way that respected the privacy of community members (Task 7).
5. Provide guidance on project reporting – The committee played an essential role in reviewing the final report (Task 8) to ensure that key findings were communicated effectively to the community and stakeholders.

A critical aspect of the Advisory Committee's role was overseeing data codification protocols to protect participants' sensitive information. Since the PAIR program involved household-level data collection, strict data anonymization and security protocols were implemented, with EHC supervising data management procedures. Throughout the project, the Advisory Committee held regular meetings to ensure continuous oversight and alignment with community priorities.

Task 2: Development of an Indoor Air Pollution Database

A comprehensive database was established to store and organize indoor air quality data collected through the PAIR program. The database was structured to accommodate the analytical methods identified in Task 6 (Literature Review and Development of a Work Plan) and to facilitate statistical analyses conducted in Task 7 (Data Analysis).

SDSU collaborated with SDAPCD and EHC to gain access to indoor air quality data collected from IQAir AirVisual Pro monitors installed in participant households. The monitored parameters included:

- Particulate Matter (PM10, PM2.5, PM1.0)
- Carbon dioxide (CO_2)
- Temperature and relative humidity

Data was collected at five-minute intervals, 24/7, provided that participants kept the monitors powered on and connected to Wi-Fi. The database also included information on when air purifiers were activated, allowing for comparisons between baseline (pre-filtration) and intervention (post-filtration) periods.

To ensure ethical research practices, SDSU project team members completed training in human subjects' research, information privacy, and data security before handling sensitive data.

Task 3: Development of a Household Information Database

A separate database was established to document household characteristics relevant to the study. SDSU worked closely with EHC to gather information on:

- Housing type (e.g., single-family home, multifamily unit, mobile home)
- Number of residents per household
- Other relevant demographic and housing characteristics

All personally identifiable information (PII) was handled with strict confidentiality measures, and household data was stored in a coded-secure format to protect participant anonymity. Data collection and management procedures adhered to best practices in human subjects' research. Appendix 2 describes the re-codification procedure followed by the project team for the household database.

Task 4: Development of a Meteorological Database

SDSU compiled meteorological data corresponding to the PAIR program study period, ensuring that environmental conditions were accounted for in the analysis. This included the collection of:

- Ambient temperature
- Atmospheric pressure
- Relative humidity
- Precipitation levels
- Wind speed and direction

Meteorological data was retrieved from publicly available air monitoring networks in the Portside region. These data were processed and formatted to align with the analytical framework of Task 6 (Literature Review and Work Plan) and Task 7 (Data Analysis).

Task 5: Development of an Outdoor Air Pollution Database

To assess the broader environmental context, SDSU compiled outdoor air pollution data from regulatory ambient monitoring stations near the Portside area. This database incorporated measurements from locations such as:

- Sherman Elementary School Air Monitoring Station
- Chula Vista Outdoor Monitoring Station

Outdoor air quality data were obtained from the SDAPCD website and were formatted for direct comparison with indoor air quality measurements recorded in Task 2. The database covered at least the same time period as the indoor air monitoring data to facilitate robust statistical analyses.

Task 6: Literature Review and Development of a Work Plan

SDSU conducted a comprehensive literature review on methodologies for assessing indoor air quality using low-cost sensors, particularly the IQAir AirVisual Pro. This review provided the foundation for developing a detailed work plan that outlined:

- Statistical methodologies for analyzing air quality data
- Metrics for assessing air purifier efficiency
- Best practices from previous studies

As part of this task, the Advisory Committee played a key role in refining the research plan by ensuring that it addressed the community's concerns and aligned with the study's objectives.

Task 7: Data Analysis

Using the structured datasets developed in Tasks 2–5, SDSU performed comprehensive data analyses to evaluate the impact of air purifiers on indoor air quality. Key analytical steps included:

- Comparative analysis of indoor air pollution levels before and after purifier installation
- Assessment of meteorological influences on indoor air quality
- Correlation analyses between indoor and outdoor pollution levels

Statistical techniques were employed to quantify the effectiveness of air filtration in reducing indoor particulate matter concentrations. Data analysis procedures were conducted under the continuous guidance of the Advisory Committee to ensure scientific rigor and community relevance.

Task 8: Project Reporting

Throughout the project, SDSU submitted quarterly progress reports to provide updates on key milestones. The final phase of the project involved:

- Preparation of the final report, summarizing research findings and key insights
- Review by the California Air Resources Board (CARB), with an opportunity for feedback and revisions
- Community dissemination efforts, ensuring that results were shared with Portside residents in an accessible and meaningful manner

The Advisory Committee provided final guidance on reporting, ensuring that findings were accurately interpreted and communicated with clarity and transparency.

Appendix 2: Data Management and Recodification process of household data

Data Management

This appendix provides a detailed overview of the guidelines established for the management of databases collected as part of the PAIR project. It outlines the procedures followed for data collection, formatting, quality assurance, and backup to ensure data integrity, consistency, and security. These guidelines were designed to standardize the handling of datasets, facilitate efficient data processing, and maintain accurate records throughout the project lifecycle. The outlined protocols ensure that all collected data undergo rigorous quality control measures before being used for analysis, and that proper backup procedures are in place to prevent data loss. The overall description of the data management process is shown in Figure 15.

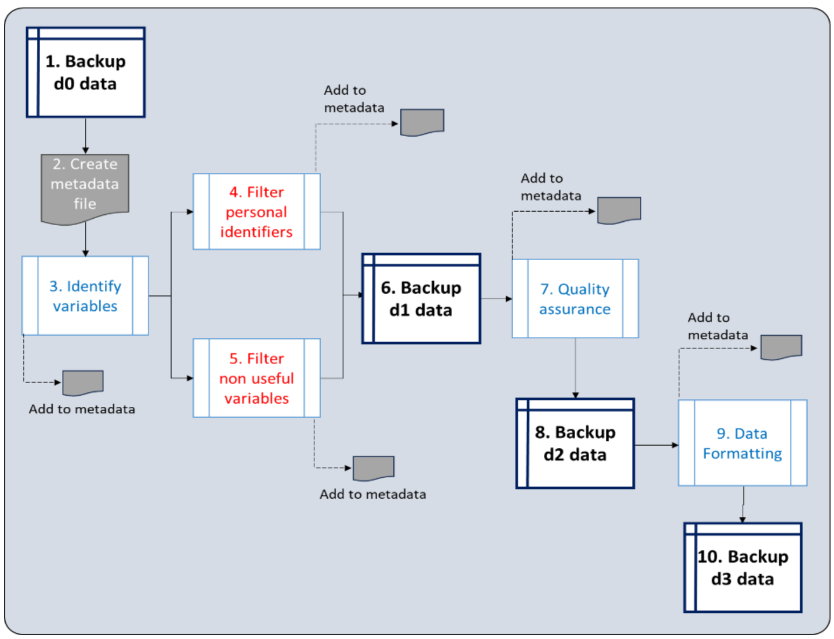

FIGURE 15. STAGGERED PROCESS FOR RECODIFICATION OF HOUSEHOLD DATA.

The protocols outlined in this appendix were implemented for data collected on indoor air quality, outdoor (ambient) air quality, meteorological parameters, and de-identified household information from PAIR program participants. The re-codification process is described in the second section of this appendix.

Database Generation

The database generation process followed a staggered-level approach, beginning with the raw datasets ("d0" level) obtained from the original data sources. The process then proceeded through a quality assurance phase ("d1" level), a formatting phase ("d2" level), and ultimately produced an analysis-ready working dataset (“d3” level). Except for household information from participant residents, all datasets used in the study were publicly available from the listed data sources.

1. Data Collection
The first step involved obtaining datasets required to assess the impact of the air purifier devices implemented in the PAIR project. The PAIR data analysis team identified, collected, and stored household, air pollution, and meteorological data from the sources listed in Table 1.

With the exception of the household dataset, all data were downloaded from their respective sources and stored on a dedicated repository disk. Each collected dataset was preserved in its original form and designated as the first layer of information (“d0” level). These files were renamed as “d0_filename,” where “filename” referred to the original dataset name assigned by the data source. No modifications to formatting were made at this stage. The d0_filename files were stored in a dedicated project directory. Data collection was an ongoing process, repeated as new data became available from each source.

2. Identification of Variables
Once the datasets were collected, the data analysis team identified the key variables from each d0_filename file that were relevant to the analysis. The resulting datasets, incorporating only the selected variables, were stored under the new designation “d1_filename” for each data source.

During this step, all variables in each dataset were characterized based on the following attributes:

- Variable Name (string, no spaces allowed)
- Type (numeric, string, or date)
- Label (string, spaces allowed)
- Values (where applicable)
- Missing Data (flagged where applicable)
- Measurement Scale (ordinal, nominal, or continuous)

With input from the advisory committee, the project team determined which variables would be used to compare air quality conditions before and after the installation of filtration devices. These selected variables were documented in a “meta_filename.txt” file, which was stored alongside the corresponding d1_filename files.

3. Quality Assurance
The next step applied a quality assurance (QA) protocol to the d1_filename datasets to identify outliers, missing data, and errors. The datasets resulting from this QA process were designated as “d2_filename” files. To track all modifications, a metadata file (“meta_filename.txt”) was created within the same working directory as each d1_filename dataset. This file logged all major edits, including:

- Date of modification
- Name of analyst making the changes
- Filename
- Description of edits

The QA process involved descriptive statistical analyses to assess measures of central tendency and dispersion, as well as data visualization techniques to detect outliers, missing values, and anomalies. If necessary, data points were either removed or flagged for correction, with all decisions documented in the meta_filename.txt file. The QA process was repeated as needed whenever new data became available.

4. Working Database
Following QA, the d2_filename files were formatted to be compatible with the statistical software selected for the analysis. Formatting activities included variable renaming, flagging, scaling, and recoding procedures. Missing values were flagged according to the software's requirements.

The final formatted dataset, designated as "d3_filename," was saved in the same directory. A description of all formatting procedures was documented in the meta_filename.txt file. The d3_filename datasets represented the final, analysis-ready datasets used for statistical modeling and interpretation.

Re-codification of household data

Since the raw household data contained general demographic details and limited personal information, EHC established a recodification process to ensure that SDSU researchers worked exclusively with de-identified datasets. This measure safeguarded participants' personal data integrity while enabling a robust analysis of the collected information. The primary objective of the **recodification process** was to transform the collected data into a structured database that:

1. **Excluded personally identifiable information (PII)** to protect participant privacy
2. **Ensured proper formatting** for subsequent statistical analyses
3. **Maintained data integrity and usability** for research purposes

The recodification process followed a ten-step, four-level structured approach, progressing from the original raw dataset (d0 level) to a fully formatted and de-identified dataset (d3 level). The overall recodification procedure was illustrated in Figure 15 and described in detail below. For the purposes of this project, the term "PAIR data analysis team" referred to the collaborative research team members from San Diego State University (SDSU), EHC, and SDAPCD, who were responsible for overseeing the data recodification and analysis procedures.

Recodification Process

1. **Backup of d0 Data**
   The original household data obtained by EHC was designated as "d0", representing the raw dataset in the recodification process. This dataset was collected during the PAIR program recruitment process and stored in AirTable, as provided directly by participants.
   The first step involved creating a backup copy of the d0 dataset onto a designated computer disk, identified by EHC as the sole physical copy of the data. EHC established security conditions to ensure that only EHC-approved personnel had access to this disk. The backup copy was considered the first layer (d0 level) of information and renamed as "d0_filename", where *filename* referred to the original household dataset identity assigned by EHC. This dataset was then saved in a dedicated project working directory.

2. **Creation of a Metadata File**

A metadata file named "meta_filename.txt" was created and stored in the same working directory as the d0_filename dataset. This text document served as a record of all major changes applied to the dataset from d0 to d3 during the staggered recodification process. At each step of the recodification process, the metadata file was updated with the following details:

- Date of the edit
- Name of the analyst
- Filename
- Characteristics of variables
- Description of edits applied

3. **Identification of Variables**
The analyst identified key characteristics for each variable in the d0_filename dataset and recorded them in meta_filename.txt. These included:

    - Name: String {No spaces allowed}
    - Type: Numeric; String; Date
    - Label: String {Spaces allowed}
    - Values: Describe if applicable
    - Missing: Describe if applicable
    - Measurement: Scale; Ordinal; Nominal

4. **Filtering Out Personal Identifiers**
In collaboration with EHC**,** the designated analyst identified variables containing direct or indirect personal identifiers**.**
    - Direct personal identifiers included names, phone numbers, and email addresses.
    - Indirect personal identifiers consisted of data points that could be used to infer a participant's identity, such as home addresses, birth dates, and racial/ethnic information.

    EHC determined which variables contained direct and indirect identifiers, and these were filtered out from the d0_filename dataset**.** The removed variables were documented in meta_filename.txt**.**

5. **Filtering Out Unnecessary Variables**
The PAIR data analysis team, in collaboration with EHC, determined which variables from the d0_filename dataset were not relevant for statistical analysis. These non-essential variables were identified, removed, and documented in meta_filename.txt**.**

6. **Backup of d1 Data**
After filtering out both personal identifiers (Step 4) and irrelevant variables (Step 5), the resulting dataset was saved as "d1_filename" and stored in the project working directory.

7. **Quality Assurance of Data**
A quality assurance (QA) process was conducted on the d1_filename dataset to identify:

- Outliers
- Missing data
- Potential errors

The PAIR data analysis team assisted in conducting basic QA procedures, including:

- Descriptive statistics (measures of central tendency and dispersion)
- Data visualization (plotting variables to detect inconsistencies)

The team determined whether data points should be removed or flagged for correction. All modifications were recorded in meta_filename.txt**.**

8. **Backup of d2 Data**
   The dataset resulting from the QA process (Step 7) was saved as "d2_filename" and stored in the same directory. The d2_filename dataset contained the cleaned and validated variables required for data analysis in the PAIR program**.**

9. **Data Formatting**
   The d2_filename dataset was formatted according to the specifications required by the statistical software selected for analysis. The formatting procedures were documented in meta_filename.txt**.**

10. **Backup of d3 Data**
    The final recodified and formatted dataset was saved as "d3_filename" and stored in the working directory. The d3_filename dataset was now ready for statistical analysis in the PAIR program.