Thesis Proposal

Low-Cost Fine Particulate Monitors and their Applications

Michael D. Taylor

Tuesday, November 10th, 2015

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Illah R. Nourbakhsh, Aaron Steinfeld, Albert Presto, James Longhurst

Copyright © 2015 Michael D. Taylor
**Keywords:** Low-cost sensing, PM2.5, calibration, air quality, fine particulate matter, sensor characterization, citizen science, building management, building automation
Abstract

Air quality has long been a major health concern for citizens around the world, and increased levels of exposure to fine particulate matter ($PM_{2.5}$) has been definitively linked to serious health effects such as cardiovascular disease, respiratory illness, and increased mortality. $PM_{2.5}$ is one of six attainment criteria pollutants used by the EPA, and is similarly regulated by many other governments worldwide. Unfortunately, the high cost and complexity of most current instruments results in a lack of detailed spatial and temporal resolution, which means that concerned individuals have little insight into their personal exposure levels. This is especially true regarding hyper-local variations and short-term pollution events associated with industrial activity, heavy fossil fuel use, or indoor activity such as cooking.

Advances in sensor miniaturization, decreased fabrication costs, and rapidly expanding data connectivity have encouraged the development of small, inexpensive devices capable of estimating $PM_{2.5}$ concentrations. This new class of sensors opens up new possibilities for personal exposure monitoring and building instrumentation. It also creates new challenges related to calibrating and characterizing inexpensively manufactured sensors to provide the level of precision and accuracy needed to yield actionable information without significantly increasing device cost. Additionally, we must develop new methods for visualizing and presenting data in an interactive fashion such that the wealth of data presented by many spatially distributed sensors continues to empower individuals and communities to better understand their personal exposure.

This proposed thesis seeks to pursue the following three questions:

1. Can an inexpensive air quality monitor based on mass-manufactured dust sensors be calibrated efficiently in order to achieve inter-device agreement in addition to agreement with professional and federally-endorsed particle monitors?
2. Can an inexpensive air quality monitor increase the confidence and capacity of individuals to understand and control their indoor air quality?
3. Can networks of inexpensive air quality monitors be used in tandem with existing building monitoring systems to characterize and control air quality in large multi-occupant spaces such as offices and university buildings?
In the proposed experiments, we will utilize the Speck fine particulate monitor, developed over the course of four years. The Speck processes data from a low-cost dust sensor using an asymmetric low-pass filtering algorithm. We have optimized the parameters for the algorithm through short-term co-location tests with professional HHPC-6 particle counters, and verified typical correlations between the Speck and HHPC-6 units of $r^2 > 0.90$. To account for variations in sensitivity, we have developed a calibration procedure whereby fine particles are aerosolized from a container resting on a 6-inch speaker cone. This allows us to produce Specks for commercial distribution as well as the experiments presented herein. Over the course of the thesis, we will continue to refine this process to increase accuracy and precision as well as automation and throughput.

Drawing from previous pilot studies, we will distribute low-cost monitors through local library systems and community groups. We will use pre-deployment and post-deployment surveys to characterize user perception of personal exposure and the effect of a low-cost fine particulate monitor on empowerment.

We will also deploy monitors in academic and industry campuses in order to explore the potential for small internet-connected particulate monitors to provide higher spatiotemporal resolution for air quality data in building management and automation applications.
Contents

1 Introduction 1

2 Problem Statement 3

3 Related Work 5

3.1 Health Effects of $PM_{2.5}$ 5

3.1.1 Harvard study 5

3.2 Survey of existing $PM_{2.5}$ monitoring technologies 6

3.2.1 Federal Reference Method 6

3.2.2 GRIMM Laser Aerosol Spectrometer (LAS) 6

3.2.3 Beta Attenuation Monitor (BAM) 7

3.2.4 Tapered-Element Oscillating Microbalance (TEOM) 8

3.2.5 Handheld Optical Particle Counter (OPC) 9

3.3 Existing Low Cost Air Quality Monitors 9

3.4 Building Management Systems 10

3.5 Studies With Low Cost or Portable Air Quality Monitors 11

3.5.1 Backpack study 11

3.5.2 Clarkson HEPAiRx Study 12

4 Prior Work 15

4.1 Speck Design 15

4.2 Speck Pilot Study 17

4.2.1 Experimental methods 17

4.2.2 Results and Discussion 18

4.3 Comparison with handheld particle counters 22

4.3.1 Calibration Pre-Test 22

4.3.2 Cooking Test 23

4.3.3 Incense Test 24

4.3.4 Results and Discussion 26

4.4 Rapid Autonomous Group Exposure (RAGE) Calibration Method 28

4.5 Comparison with Federal Equivalence Methods 32

4.5.1 Outdoor tests with South Coast Air Quality Management District 32

4.6 ESDR 33

4.7 Building Management System Integration 36
# List of Figures

3.2 Schematic of the laser chamber in the GRIMM Model 1.109 LAS[4] .......................... 7  
3.3 Diagram of BAM operation[24] ............................................ 8  
3.4 Diagram of TEOM operation[23] ............................................ 9  

4.1 The Speck air quality monitor .............................................. 15  
4.2 Calibration setup showing five Specks, the HHPC-6+ (left), and the HHPC-6 (right). Displayed values are scaled down from recorded ppl counts ............................................ 22  
4.3 Plot of calibration data over time. Note the large discrepancy between the HHPC-6 channels ............................................ 23  
4.4 Plot of cooking data over time ............................................. 24  
4.5 Comparison of additional HHPC size channels for the cooking test, showing ppl vs time (minutes) ............................................. 25  
4.6 Table of cooking $r^2$ values. Shaded cells indicate $r^2 > .9$ ............................................. 26  
4.7 Incense setup. Notice the close spacing of the instruments, which may have disrupted readings (see Results and Discussion). ............................................. 26  
4.8 Comparison of additional HHPC size channels for the incense test, showing ppl vs time (minutes) ............................................. 27  
4.9 Table of incense $r^2$ values. Shaded cells indicate $r^2 > .9$ ............................................. 28  
4.10 Current calibration setup ................................................. 29  
4.11 Future calibration chamber ................................................. 31  
4.12 One week of time-series AQMD data at default device sampling rates ............................................. 32  
4.13 One week of time-series AQMD data at hourly averaged sampling rates ............................................. 33  
4.14 One week of time-series AQMD with 24hr averages, reported hourly ............................................. 34  
4.15 Early concept of interactive building air quality visualization using simulated data ............................................. 37  

5.1 An Acu-Rite 01057 weather station used to measure windspeed and direction as well as temperature, humidity, and rainfall[1] ............................................. 42
List of Tables

4.1 Table of pairwise $r^2$ values for all Specks for the calibration test . . . . . 23
4.2 Table of scalar calibration values . . . . . . . . . . . . . . . . . . . . . . 24
4.3 Table of pairwise RMS error between devices at hourly sampling averages . 33
4.4 Table of pairwise $r^2$ values between devices at hourly sampling averages . 34
4.5 Table of pairwise $r^2$ values between devices at 24hr sampling averages, re- 34
ported hourly . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
4.6 Comparison of features afforded by ESDR vs. other available platforms . . 36
Chapter 1

Introduction

Fine airborne particles that are smaller than 2.5 microns, collectively designated as $PM_{2.5}$, pose a serious health risk to the public. These airborne particulates have been linked to significantly increased risk of cardiopulmonary and respiratory illnesses, particularly in the elderly and other sensitive populations.[16][17][20]

$PM_{2.5}$ can be composed of any number of airborne particulate substances, and is a particularly useful metric in predicting adverse health effects.[8] Outdoor fine particulates commonly include products from combustion such as diesel exhaust and coal particulates[18] but not larger particles such as pollen. $PM_{2.5}$ is often present indoors as well, and includes fine dust and particulates created from certain kinds of cooking, such as frying.

Inside the home, $PM_{2.5}$ is composed of particles both from indoor sources as well as outdoor sources. Open windows and doors as well as insulation can affect how much particle interchange occurs. Because air quality dynamics are complex, there are times when it is beneficial to open a window if the outside air is cleaner than the indoor air, and vice versa. Different cooking oils have varying smoke points, which can affect the number of particulates generated. Kitchen smoke hoods should ideally vent outside, but it is common for these hoods to simply pass cooking emissions through a coarse filter and then exhaust back into the kitchen. A HEPA air purifier can dramatically lower indoor particle counts if located in an appropriately sized room. Unfortunately, air quality is largely an unknown until adverse symptoms present themselves and prompt the need for personal indoor air quality monitoring.

Because $PM_{2.5}$ is a measurement that can indicate the presence of a diverse range of pollutants,[15] it is an ideal single metric for personal air quality monitors. Existing particle counters are almost solely available in the scientific and industrial markets, and as such are much more expensive than is practical for a home budget. As a result, $PM_{2.5}$ measurements are sparse and information on personal exposure is often unavailable or difficult to understand. A new class of low-cost air quality monitors is emerging, however. This thesis will examine the potential for this technology to increase awareness and effect changes in behavior as an individual and also at a community level. We will also examine the effects of inexpensive air quality monitors spatially distributed in mass-occupied spaces like open office buildings and university campuses.

In addressing these questions, we will also pursue new methods of calibration, char-
acterization, and visualization necessary for this new class of sensors to present precise, accurate, and actionable information to users.
Chapter 2

Problem Statement

Air quality measurement and analysis with high temporal and spatial resolution has yet to be achieved due to prior limitations on instrument cost, size, and technical complexity. Rather than precision optics and complex control of temperature, flow, and humidity of incoming air, low cost air quality monitors may utilize an inexpensive mass-manufactured dust sensor and on-board signal processing to estimate particle count and $PM_{2.5}$. This thesis will seek to answer the following two questions:

1. Can an inexpensive air quality monitor based on mass-manufactured dust sensors be calibrated efficiently in order to achieve inter-device agreement in addition to agreement with professional and federally-endorsed particle monitors?
2. Can an inexpensive air quality monitor increase the confidence and capacity of individuals to understand and control their indoor air quality?
3. Can networks of inexpensive air quality monitors be used in tandem with existing building monitoring systems to characterize and control air quality in large multi-occupant spaces such as offices and university buildings?

Calibration of inexpensive sensors is generally challenging because of the low precision of manufacturing, large numbers of sensors to be calibrated, and the goal to keep the calibration inexpensive relative to the cost of the monitor. This challenge is made more difficult for air quality, where it can be hard to generate repeatable levels of air pollution. We have explored a method using large low-frequency speakers to aerosolize fine dust and spread it across a bench of sensors to be calibrated. In this thesis, we seek to refine this procedure through the development of a calibration enclosure, controlled distribution of source particle sizes, and additional instrumentation using an optical $PM_{2.5}$ federal equivalence monitor.

We will deploy low-cost monitors through local pilots in the Pittsburgh region in order to study the effect of these monitors on individual empowerment. One vector for distribution will be through the public library system. Through surveys conducted at the time of check-out and at check-in, we will evaluate changes in user confidence and understanding of their indoor air quality while in possession of the monitor. We will also use the surveys to gauge whether people were able to use their collected and observed data to make in-home changes to improve their air quality by reducing $PM_{2.5}$. We will conduct additional
in-person interviews when possible to gather anecdotal evidence to support the results of the surveys.

In evaluating the use of inexpensive air quality monitors in tandem with building management systems, we will deploy monitors both in university buildings such as Newell Simon Hall as well as in office buildings in the United States and India. Deployment of the current Speck models will allow us to begin collecting data while gaining insight which will inform the co-design of a monitor specifically for integration into building management systems. The hardware co-design will focus on discovering a form factor and communication protocol that allows for integration into the largest percentage of existing systems. The collected data will allow us to design separate visualizations specific to three user groups: individual office occupants, building managers, and public relations specialists. Analysis of the data will identify the appropriate device density for segregated offices and for large open-office spaces. We will conduct surveys and interviews to answer the following questions:

1. What is the effect of access to air quality measurements on the comfort and behavior of individual office occupants?

2. Does increased access to spatio-temporal air quality data enable building managers to improve occupant personal exposure through HVAC control and building modification?

3. Can visualizations be effectively used by public relations specialists to promote building innovations leading to cleaner air quality technology?
Chapter 3

Related Work

3.1 Health Effects of $PM_{2.5}$

Exposure to elevated levels of particulate matter has been linked to numerous health impacts, including but not limited to respiratory illness, cardiovascular problems, low birth weight, and reduced productivity. Particulate matter is one of the EPA’s six criteria air pollutants, alongside ozone, carbon monoxide, nitrogen oxides, sulfur dioxide, and lead. While particles less than 10 microns are recognized to pose a significant threat to human health, the EPA is especially concerned with particles smaller than 2.5 microns, collectively referred to as fine particulate matter or $PM_{2.5}$. Particles within this size range are small enough to enter the lungs, and the smallest of these can additionally enter the bloodstream. Populations at increased risk of negative health effects include children and older adults, and premature death can result in people with heart or lung disease.

$PM_{2.5}$ is also of concern to employers and workers whose jobs regularly place them in environments with higher exposure to fine particulate matter. Increased levels of pollution, including fine particulate matter, has been linked with lower productivity in the workplace. This in turn results in higher employment costs, which often constitute the greatest fraction of operational expenses for many businesses.

3.1.1 Harvard study

The Harvard Six Cities Study\textsuperscript{11} followed 8111 participants in six cities around the United States in order to determine the effects of air pollution on mortality and cause of death over a period of 14 to 16 years. The study found that fine inhalable particulate matter and sulfate particles had the greatest effect on mortality rates, more so than total particulate pollution, aerosol acidity, sulfur dioxide, or nitrogen dioxide. Fine particulates were especially significant in association with cardiopulmonary illness, lung cancer, and other respiratory illnesses. The study acknowledges potential confounding factors including cumulative lifetime exposure to particulate matter, though controlling for as many factors as possible yields similar results. Smoking was also shown to have an extreme effect on mortality due to lung cancer, as additionally supported by most other relevant studies.
3.2 Survey of existing $PM_{2.5}$ monitoring technologies

3.2.1 Federal Reference Method

The federal reference methods (FRM) for measuring $PM_{2.5}$ and $PM_{10}$ involve circulating air through carefully pre-weighed filters over a period of time\([19]\). The change in mass over time, given a constant and known flow rate, can be used to determine the particulate concentration of a volume of air. This method is very labor intensive, however, and outdated by more recent methods. It has the advantage of being the only widely-accepted means of measuring mass directly (rather than through resonance, beta attenuation, etc.). Its downsides include non-continuous sampling over large intervals, and the large potential for human error. Additionally, volatile substances may be initially captured as fine particulates, but they have the ability to sublimate off of the filters, and are thus only intermittently recorded.

3.2.2 GRIMM Laser Aerosol Spectrometer (LAS)

The GRIMM EDM 180 is a federal equivalence monitor that calculates mass at multiple size cutoffs by counting and sizing particles optically\([21]\). The GRIMM is capable of discriminating between 32 distinct size ranges of particles. Because the measured particles are categorized by size, no cyclones or impactors are required; only very coarse filtering to prevent bugs and singly-visible particles from entering. Because the instrument is based on optical measurement, however, the device must account for or reduce hydroscopic growth due to humidity. The GRIMM accomplishes this by transporting newly sampled air down through a long nafion tube. Dry, non-sample air flows upward in a concentric pipe around the nafion tube, extracting moisture from the intake air in the process. An additional concentric pipe recirculates the air back into the base unit for expulsion. The pressure needed to circulate this air necessitates a large and relatively noisy pump, but this system only receives power at an observed humidity of around 70%.
3.2.3 Beta Attenuation Monitor (BAM)

Beta attenuation monitors (BAMs) operate on the principle that beta radiation from any particular source is attenuated by obstructing solid matter. Typically, a continuous ribbon cycles past a window of controlled airflow. The ribbon is made of fine filter paper capable of capturing 2.5 matter. Two detectors on either side of the window measure the beta attenuation level before and after that section of ribbon is exposed to the airflow, and the difference in attenuation can be used to calculate the increase in mass of the ribbon. Given a known flow rate and a known exposure time, airborne particle concentration is easily determined.

Some shortcomings of this method include the necessity to pre-filter the air to the appropriate size particle distribution (usually $PM_{2.5}$ or $PM_{10}$) using a cyclone assembly.
The minimum sampling rate is typically once per hour, whereas optical or resonant methods can report samples every minute (though shorter sampling times lead to greater error, especially at low concentrations). Additionally, while the GRIMM is capable of delivering estimates for $PM_{2.5}$ and $PM_{10}$ simultaneously, this is not possible with the BAM. The BAM also suffers from the volatile sublimation issue present in the FRM.

![Figure 3.3: Diagram of BAM operation](image)

3.2.4 Tapered-Element Oscillating Microbalance (TEOM)

TEOMs operate by directing particles from an airstream onto or through a small filter or plate attached to a thin oscillating beam. The total mass on the plate can be determined by calculating the resonant frequency of the beam and plate. Like the BAM, particle concentration is calculated using the difference in accumulated mass across the sample interval. These intervals can be short (e.g. one minute) but shorter intervals can sometimes result in imprecise readings. This is especially true if a large number of the particles are volatile and evaporate from the plate or filter, as this can cause a net loss of mass across a small time interval, which is then used to calculate an impossible negative particle concentration. Also, similar to the BAM, cyclones and impactors must be used to separate out particles that are larger than the cutoff size of interest.
3.2.5 Handheld Optical Particle Counter (OPC)

Handheld optical particle counters such as the Hach Met One HHPC-6+ are significantly less expensive and more portable than the previous devices, but are not considered federal reference methods. The principle of operation behind these devices is similar to that of the GRIMM, where a laser is used to count and size particles drawn through the sensing chamber using a small pump. Typically, the sample volume is smaller and little or no preconditioning is performed on the sample air. Temperature and relative humidity are frequently measured in these units at the point of particle detection. Devices like the HHPC-6+ frequently incorporate a basic user interface, on-board storage of varying capacities, and a rechargeable battery. These devices can be very useful for point sampling, but are not designed for continuous operation. Additionally, while some devices like the TSI DustTrak will estimate mass, the HHPC-6+ delivers only particle counts across different size bins, making comparison with mass-only methods like the BAM difficult. If the particle source is well characterized, it is possible to estimate mass with an OPC that possesses at least micron-level size resolution[7].

3.3 Existing Low Cost Air Quality Monitors

Related works[14] have established that inexpensive sensors designed to detect larger dust particles can also be used to detect particles in the 2µm range. Typically, these experiments use a moving average of raw sensor readings, with a linear scale factor used to match these
values to particle counts or 2µm values from a professional monitor.

There are a small number of personal air quality monitors available on the market. The Air Quality Egg,[10] available from Wicked Device, uses a similar dust sensor to those in [14] in its particulate sensor add-on. The Egg and daughterboard, at a combined cost of $243, upload data to a publicly-available map currently populated with 1094 units. The Dylos,[3] priced from $199 to $425, is widely used in experiments and publications where inexpensive and often mobile particulate monitoring is desired. The basic Dylos units use custom laser optics and report particle counts larger than 2.5µm and larger than .5µm on two separate channels. Advanced models feature additional functions such as additional size channels, datalogging capability, and battery power. In China, where air quality is a major concern within cities, several low-cost particulate monitors are available, though these have not been as extensively characterized or used in air quality studies in comparison with the Dylos. These include the BPEER[13] and Air.Air![5] devices, both of which rely on ambient airflow through a central passage.

We have observed that poor manufacturing tolerances, a lack of calibration, and a lack of sensitivity can cause these sensors to deliver varying qualities of information. We believe that the quality control, calibration, and more sophisticated algorithmic implementation can produce much higher data quality.

A few qualitative methods are sometimes used by civilians and citizen scientists to determine relative amounts of pollution. These often include either basic grab-samples (capturing air in a non-reactive jar or bag) or sticky traps made of white paper and double-sided tape. The sticky traps can be investigated with a microscope to begin to determine the composition of larger particles. Higher powered microscopes are needed to examine fine particles.

Although personal air quality monitors are becoming more available and affordable right now, it is also important to develop appropriate knowledge and support infrastructure for users. This type of scaffolding can help guide everyday citizens as they explore ways to improve their air quality and overall home health. A goal for this study is to gain a better understanding of what this type of holistic air quality tool kit could look like.

### 3.4 Building Management Systems

Building management is another rich area for potential enrichment through Speck sensors and data visualization. Because of the high cost of air quality monitors, important health factors like carbon dioxide, PM$_{2.5}$, and other criteria pollutants are often only measured in a small number of locations within a building, if any. Particulate matter sensors are very uncommon even in more advanced and health-conscious buildings.

Elevated levels of fine particulate matter have been linked to reduced worker productivity.[9] Because lower productivity leads to higher employment costs, and given that employment costs constitute the largest fraction of operating expenses for many businesses, building managers should be greatly concerned with monitoring and controlling particulate matter levels in the workplace. With the increased ability of building management systems to track comfort factors like temperature and humidity, building managers are now able to
optimize thermal comfort and energy expenditures. This is in large part due to the quantification of the impact of discomfort on employee productivity and the resulting increase in personnel costs. As particulate matter sensors are only now becoming miniaturized and affordable, however, these systems have yet to implement spatially-distributed $PM_{2.5}$ monitoring, which would provide another variable for assessing employee comfort. Some recently launched products such as the Airassure $PM_{2.5}$ monitor from TSI are designed to connect to networked building management systems, though at this time the Airassure $PM_{2.5}$ monitor is not yet available in the USA or Europe.

Automated Logic currently offers one of the most advanced commercially-available building automation systems. A full building solution allows building managers to monitor environmental conditions as well as the status of energy and climate related building infrastructure. Additionally, control of these systems can be automated using schedules and instrumentation feedback rules. WebCTRL, one of Automated Logic’s flagship products, integrates most of these features into a single interface which can be accessed from any device with a network-connected web browser. Automation rules can be created using a largely graphical, object-oriented interface. Set points for these rules are easily created and modified as necessary by the building manager. Visualization of current, projected, and past data is presented either as floor-based heatmaps, trend graphs, or detailed device or subsystem animations. WebCTRL EnergyReports summarize data over specified periods of time for easy review of large campuses by management staff. WebCTRL Time-lapse allows for building managers to review past data as a replay over periods of up to 24 hours. This functionality allows reviewers to identify performance or unexpected changes in energy use or climate after the event has already passed. Automated Logic’s software is largely customizable, which allows it to be used in a variety of buildings or campuses.

3.5 Studies With Low Cost or Portable Air Quality Monitors

3.5.1 Backpack study

Spira-Cohen et al. ([22]) measured personal exposure to fine particulates in asthmatic schoolchildren in their oft-cited South Bronx backpack study. This study found a significant correlation between respiratory symptoms and fine particulate matter, especially diesel exhaust. These symptoms included increased coughing, wheezing, allergic inflammation, decreased lung function, and other asthma symptoms. The data for this study was obtained through qualitative and quantitative self-assessment of symptom presence and severity along with quantitative data from mobile and stationary air quality monitors.

Stationary monitors included a TEOM 1400a $PM_{2.5}$ sampler, AE-21 black carbon aethalometer, and an ACCU filter sampler as well as monitors for $O_3$, $CO$, $SO_2$, and $NO_x$, all located at the school. Additionally, background $PM_{2.5}$ data was available from a rooftop monitor operated by the New York State Department of Environmental Conservation. The personal monitors carried by the children were placed in rolling backpacks, which the children were instructed to keep with them 24/7. The backpacks contained filter-based
24-hour integrated weekday $PM_{2.5}$ samples as well as motion sensors to determine if the backpacks were being carried at all times.

The wide assortment of sensor data gathered in this study allowed the research team to investigate many independent correlations, and represented a significant amount of research effort as each of the 40 students required regular intervention to change, collect, and analyze filter samples in addition to gathering symptom data. One shortcoming of this study is that because of the 24-hour integrated samples, the lag effect between exposure and the presentation of symptoms could only be resolved as same-day or in one-day increments. Continuous sampling as is made possible by optical nephelometer-based devices could provide significant time resolution in similar experiments. These devices also do not require changing filters or regular maintenance over a period of only a few months, and with automatic wireless data uploads, the interventions of the research team could be decreased to symptom collection. Additionally, smaller and more lightweight optical monitors could reduce the impact and intrusion of the equipment on the participants’ behavior. With new technology, these monitors may become wearable, thus increasing the accuracy of personal exposure monitoring. The scope of the experiments can also be extended to a larger population if the cost of the equipment is reduced along with the personnel expense associated with tending to filter-based samplers.

3.5.2 Clarkson HEPAiRx Study

In a study conducted by Xu et al.\cite{xu2017}, air quality monitors were placed in the homes of 30 asthmatic children for 18 weeks along with a HEPAiRx air cleaning and ventilation unit. The unit was kept on for 12 of the 18 weeks for each household, with half of the group turning the unit on during the first 12 weeks and the other half turning it on for the last 12 weeks. During this period, the AirAdvice air quality monitor was used to track the effects of the HEPAiRx operating status on $PM_{10}$, $CO$, $CO_2$, TVOC, temperature and humidity. Every sixth day, measurements would be taken from participants of their exhaled breath concentrate (EBC), pH, and peak expiratory flow (PEF). The study found that the air purifier reduced $PM_{10}$ by 72%, VOC by 59%, $CO_2$ by 19%, and $CO$ by 30%. Additionally, these reductions were shown to have a statistically significant effect in improving EBC nitrate concentrations, pH, and PEF. These results were made possible by the relatively small and unobtrusive AirAdvice unit, which can rest on a table or countertop and requires only one household power outlet. Data is relayed wirelessly on a minute by minute basis, thus simplifying measurement collection. The typical business model of AirAdvice is for contractors to place the monitor in the home for a short time, typically less than one hour. This data is automatically transmitted and analyzed by the AirAdvice servers. The monitor itself is not available to consumers, potentially because of cost and usability constraints. A lower-cost monitor with greater ease of use could potentially expand the scope of similar experiments to a greater number of houses. The AirAdvice platform solves many of the shortcomings associated with using other technology, such as the difficulty of manually collecting data or the requirement for multiple and/or large monitoring devices in the home. Clarkson University was able to confirm the value of the HEPAiRx unit by utilizing the AirAdvice monitor over a long period of time. With a commercially available
and affordable air quality monitor, individuals can verify the effectiveness of their own air purifier, including whether it has sufficient airflow to properly clean any specific room or floor.
Chapter 4

Prior Work

4.1 Speck Design

The Speck utilizes an inexpensive DSM501a dust sensor[2] rather than custom optics, but also employs a small fan to increase airflow. The Speck contains on-board signal processing and storage in addition to a color LCD touchscreen for the user interface. Power is supplied via USB, and data can be downloaded directly to any computer. The interface allows users to view the current estimate of 2µm particle concentration as well as a scaled estimate of $PM_{2.5}$ in $\mu g/m^3$. The interface can also graph the past hour or past 12 hours of data on-screen, allowing for quick access to historical data.

![The Speck air quality monitor](image)

Figure 4.1: The Speck air quality monitor

The output of the DSM501a dust sensor is a digital pin which is pulled low when particles are detected in the optical chamber. According to the datasheet, the duty cycle is approximately proportional to the number of detected particles. The period of the sensor varies greatly, however, especially at low particle concentrations. While the duration of a
low pulse (indicating detected particles) rarely exceeds 100ms, the duration between pulses can last from under one second to more than one minute. We observe that single-cycle readings are too noisy to be used directly. Instead, our algorithm samples the sensor 10,000 times per second, and uses the number of low samples each second as an input to an asymmetric filtering function. This input is herein referred to as the raw sensor value.

\[
est_{t+1} = \begin{cases} 
\frac{(A \times \text{raw}_t - \text{est}_t)}{B} + \text{est}_t & : \text{raw}_t > 0 \\
(1 - D) \times \text{est}_t & : \text{raw}_t = 0
\end{cases}
\]  

The piecewise function given in eqn 4.1 describes the second-by-second processing of the raw sensor values, where \( \text{est}_t \) is the Speck’s 2\( \mu \)m particle count estimate at time \( t \), \( \text{raw}_t \) is the raw sensor value at time \( t \), and \( A, B, \) and \( D \) are constants. We observe that the individual raw sensor values are frequently zero in all but visibly smoky environments, though the non-zero values tend to increase linearly with particle concentration. Because of this, we design our filtering algorithm to give more weight to non-zero raw values. Each second, if the raw value is non-zero, we increment or decrement our current estimate at a rate proportional to the difference between the estimate and the raw value scaled by a constant. If the raw value is zero, the estimate exponentially decays toward zero at a lower rate. The resulting behavior is that the estimate quickly responds to non-zero raw values, but decays toward zero slowly due to the potential for long pauses between pulses.

The constants in eqn 4.1 were empirically determined in previous experiments through post-processing optimization of raw data from a single prototype Speck. The cost function selects values for the best fit of the estimate against the 2\( \mu \)m channel of the HHPC-6.

Figure 1: The Speck air quality monitor (top); home screen display with current air quality reading (left); on-screen view of 12-hour historical data (right). Note: both images depict the version of Speck developed at the time of this study (summer 2014).

The Speck interface allows users to view a scaled estimate of \( PM_{2.5} \) levels in micrograms per cubic meter (\( \mu g/m^3 \)). In addition, the screen clearly rates the level of air quality from a scale of Good to Hazardous, to facilitate data interpretation. Historical data is also available on the Speck screen. Users are able to toggle between the current fine particle level, as well as graphs of readings from the past hour and past 12 hours (Figure 1). Finally, at the time of this study, the color coding and scale used for the Speck interface were aligned with the U.S. Air Quality Index for easier comparison between indoor Speck readings and outdoor federal monitor data. It should be noted that the Specks design has changed significantly since this study took place. The newer version displays current estimates of 2\( \mu m \) particle concentration as well as a scaled estimate of \( PM_{2.5} \) levels in \( \mu g/m^3 \), and uses a different color palette and scale. These modifications were partially informed by findings of this study.
4.2 Speck Pilot Study

4.2.1 Experimental methods

Overview

This was a proof of concept study to explore whether and how an environmental sensor can empower a community of users to improve the air they breathe. To this end, we provided a sample of Pittsburgh residents with a Speck air quality monitor along with supporting material, and monitored how using this device affected their perspectives and behaviors. We hypothesized that, given a sufficient support structure, everyday citizens can gain the knowledge and confidence they need to investigate, discover and take action to mitigate air quality problems in their homes.

A secondary objective of this study was to discover the usability and utility of the Speck and associated support material, from the perspective of end users. Research participants were asked to rate the device and supporting material twice during the study period, so as to capture their feedback at varying stages of use.

Supporting Material

Based on feedback received during prior prototype testing, it was clear that people who use the Speck also require a support structure to help guide their process of discovery and action. The primary questions asked regarded what to do when residents discover high levels of particulate matter in their homes. While there is no single answer to this question, there are a few steps people can take as they investigate sources of and solutions to their air quality issues. Through this study, we wanted to understand which modes (or collection of modes) were most effective for conveying this type of information to the community of Speck users. We designed print material in the form of a Speck setup guide as well as a Speck website (http://specksensor.org) to test with research participants. These media contained instructions for how to use the Speck, along with guidance on recognizing pollution sources and experimenting to uncover ways to resolve air quality problems.

In addition, it was important to design a platform through which individual Speck users could connect with others in their community (as well as researchers) to discuss air quality problems and brainstorm solutions collectively. To leverage the popularity of existing social media frameworks, we created a private Facebook page where study participants could interact with one another about their experiences with the Speck. Joining the Facebook group was an optional aspect of participation in the study, and we ensured that only participants and researchers had access to this page.

Community Selection

The target audience for this study was people in the Pittsburgh region who had some interest in learning more about their air quality. A total of 47 participants were recruited through the research groups network of local community partners. Email solicitations were
sent out to individuals and community groups, who in turn passed on the call for research participants to their contacts via email and social media.

Interested individuals met in person with researchers to receive an overview of the study, provide consent to participate (if willing), and pick up their Speck. The majority of consenting participants were white females, who were approximately 25 to 50 years old. This study was approved by Carnegie Mellon University’s Institutional Review Board, and accordingly, we followed appropriate protocol during the course of the study.

Data Collection and Analysis

Each participant received a Speck to use in their home for 2 to 3 weeks, and was asked to complete a total of 3 surveys. The first survey was administered before participants received the Speck to gauge their baseline level of knowledge about indoor air quality. About a week after picking up their Speck, participants received a second survey to inquire about their experiences to-date. When it was time to return their Speck, participants completed a final survey to provide their overall feedback about the Speck and supporting material, describe what they learned from the study, and share any relevant personal stories.

In addition to survey data, we also captured discussions and posts that occurred on the participant Facebook page. These anecdotes provided valuable insight into how participants used their Speck, the discoveries they made, actions they took, and air quality conversations that were sparked by the Speck. This social media platform was also useful for participants to pose questions about the Speck and its applications, which could be answered by researchers as well as other participants. Thus, the Facebook group served as a virtual community meeting place where participants could discuss their air quality findings and concerns, as well as share knowledge and ideas with each other.

Findings from this study were primarily qualitative and anecdotal. As such, survey results and Facebook posts were qualitatively examined to detect data trends and common themes in participant responses. In addition, participant quotes were compiled to create a storyboard of user experiences and feedback.

4.2.2 Results and Discussion

Speck Assessment

All participants found the Speck easy to install and use, and most (77%) agreed that the monitor met their expectations. When asked to describe their overall experience with the monitor, the majority indicated that it was informative and/or useful (Table 1). The chief complaint was the level of noise caused by the fan in the device; this is reflected in the Other category.

Table 1: Participant perspectives on their overall experience using the Speck Air Quality Monitor Overall User Experience Rating Number of Responses* % of Responses** Enjoyable 20 43% Useful 29 62% Informative 39 83% Challenging 0 0% Confusing 5 11% Frustrating 5 11% Other 11 23% * Participants could select more than one rating ** This percentage is based on the total number of survey respondents (47) The most popular
locations for in-home Speck placement were the kitchen, bedrooms and living room, and on average, most participants left the monitor in each location for 1 to 6 days. Participants made note of their Speck reading multiple times a day, with the majority (62%) looking at the screen each time they walked into the room the Speck was in.

In terms of the hardware, participants were mostly happy with the Speck, but would have liked to see an even smaller and more portable design (including a battery pack), and longer USB cable to connect to the power supply. Some respondents also requested a better option for securing the Speck to a surface (with added weight or foot grips) so it wouldn't slide off as easily.

Apart from including a quieter fan and more robust micro-USB port, participants were primarily interested in Speck modifications that would allow them to:

• Receive alerts about significant changes in air quality; the majority wanted the alert to be in the form of a blinking light (43%) or text message (26%)
• Personalize Speck settings to be able to choose when they get alerts, alter the types of historical views that are displayed, change the color coding, etc.
• Connect to their device through a mobile phone app
• Download their data and run independent analyses on it

Overall, the Speck design and functionality were satisfactory to participants, who found the device to be a useful addition to their homes. Note that the data download feature was already implemented. However, given that participants would only have the Speck for a short time, we did not include training on how to utilize this feature as part of this small-scale study.

Perception and Behavioral Impact

Based on pre and post data from participants, we found that 53% had a different perception of their in-home air quality after using the Speck. This suggests that the use of Speck can promote greater awareness about a persons indoor environment; making the invisible particles more visible to users. Additionally, 55% of participants reported an increased level of confidence in their ability to act when the air quality is poor in their homes, hinting at a sense of empowerment among this group of users. At the end of the study, when asked if the Speck helped them feel empowered to understand and manage their air quality, 72% of participants responded positively. This provided further evidence of the Specks potential to motivate and support citizens to become more aware of their home environment and take action to improve their air quality. While the Speck was in their possession, several participants reported that they:

• Became more aware of how cooking affected air quality, and/or altered cooking habits (e.g. turning on stove vents, frying less, etc.) to mitigate these effects
• Gained a better understanding of how cleaning affected air quality, experimented with different cleaning methods, and in general cleaned more often or thoroughly
• Experimented with opening and closing windows and discovered more about how those actions impacts air quality differently
• Changed or cleaned air filters
• Added air purifiers to their homes

Thus, participants were able to gain useful insight from using the Speck, and were thereby motivated to take explicit steps to improve their air quality.

**Effectiveness of Supporting Material**

A notable aspect of this study was its pairing with supporting material and a social media platform for Speck users. Most participants (87%) indicated that the setup guide was useful. On the other hand, 39% of participants never visited the Speck website to obtain further tips and information. Findings highlighted that the Speck website may need to be reconfigured and supplemented (e.g. through workshops, instructional videos, step-by-step instructions, etc.) in order to improve user engagement and overall effectiveness. The Facebook group, on the other hand, proved to be very actively utilized. 77% of participants joined the group, and 72% reported that they found it useful to connect with other users through this social media outlet. These results indicated that a community forum can enhance an individuals experience, and produce a broader body of shared knowledge and ideas. A key component missing from these community interactions was input from air quality, health and other relevant experts who could potentially further enrich the impact and utility of such forums. In particular, if policy-level decision makers are included in the conversation, these community interactions could have longer-term impact on achieving better air quality for all citizens.

**Participant Stories**

Overall, results from this pilot study were encouraging and support the notion that the Speck could serve as a tool to empower a community of users to rally around common issues related to air quality. A few noteworthy examples of participant comments that further bolster this idea are given below:

• My husband and son have asthma and the Speck has been very informative about why their symptoms are worse at night. The readings in the bedrooms are much higher vs the front of the house. Appreciate the tips to improve the air quality and on the lookout for more.

• Vacuuming my hardwood floor made it spike from mid 20s to an unhealthy 67 in seconds. Guess I should use a mask while cleaning, seriously.

• Very excited to be testing the Speck. First day with it we turned on the broiler. Dangerous levels! Woah.

• I really like the speck... I am putting in it my bedroom tonight to see what the readings are. My daughter has been keeping detailed notes in the notebook :)

• I absolutely love being able to measure the air quality of what I’m breathing in- it’s a very empowering tool.
• Its been fun watching the reading [on the Speck] go up and down in response to predictable stimuli. For example, the reading shot up very quickly when a guest lit a cigarette.
• I think these [Specks] will be super useful to many, many people in the future (like asthmatics like myself)...
• I never thought about the invisible particles floating around before. It was interesting to watch the monitor spike and then drop. After the first time, I was more confident that the spike would drop again. I totally knew the cause.
• I’ve found it [the Speck] very beneficial as someone who is very concerned about my health.
• I have learned a surprising amount about my home and work environment and have found solutions to problems. I’m breathing easier!
• Mostly the Speck has been fun to have and has certainly been the topic of conversation whenever someone sees it for the first time - so it has been good at initiating conversations about air quality - which is a good thing in my opinion :-)
• I think this [the Speck] is going to a very popular and empowering tool (just imagine this being used worldwide!)
• I really liked the speck. My whole family participated in following it.
• I found myself checking the display each time I walked past. It was a useful and interesting data point about our lives that had not been revealed before.
• Overall, it was a good experience. Not hard to set up, easy to monitor and it gave me a sense of relief to know that the air in my home was basically good as I live near a coke plant that exceeds particulate emission standards many days a year!
• The biggest surprises were how different household activities affected the readings. I had thought opening the windows would freshen the air - not so!... The [Speck] study opened our eyes to all of the ways that we could improve and control the air quality in our home. Fantastic information!

Limitations

While findings from this study did offer useful insight to support our proof of concept, certain aspects of the study design limited the extent to which results could be generalized. A primary drawback was that participants opted into the study and thereby were not a representative sample of the larger Pittsburgh population. In particular, people of color and individuals from lower socio-economic backgrounds were under-represented in or absent from our sample. It would be critical to cast a wider net for the next phase of this research, so as to capture more people from communities who are potentially more likely to experience and be vulnerable to impacts of poor air quality in their homes.

Additionally, the timing of participation was not consistent across all individuals, given their varying schedules and availability. As such, participation in the social media conversation occurred in phases, with the majority of contributions coming from the first and
largest group of participants. Therefore, latter groups had a less impactful experience given lower participation rates during their phase of the study. Coupled with this was the short term nature of the project, which did not allow much time for participants to explore and discuss air quality topics in a deeper and more impactful manner. Longer-term studies would be needed to examine this type of interaction.

Finally, limited data collection methods to surveys alone, did not allow for more in-depth conversations with participants. Such approaches could have enriched our data and provided further insight into how such technology should be deployed to empower communities of users.

4.3 Comparison with handheld particle counters

4.3.1 Calibration Pre-Test

The two primary experiments in this section were preceded by a basic calibration test used to select and adjust three Specks. We began with five Speck units running in parallel with one HHPC-6+ particle counter and one HHPC-6 particle counter. The HHPC units log particle counts within six size ranges. The HHPC-6 unit measures $0.3\mu m$, $0.5\mu m$, $0.7\mu m$, $1\mu m$, $2\mu m$, and $5\mu m$ sizes, while the newer HHPC-6+ unit measures $0.3\mu m$, $0.5\mu m$, $1\mu m$, $2\mu m$, $5\mu m$, and $10\mu m$ sizes. We exposed these seven instruments to one half-hour cooking event in order to gather an appropriate calibration dataset with high dynamic range. The cooking event involved frying papadum (Indian lentil crackers) in a medium-sized kitchen with windows closed and the stove hood off. After approximately 15 minutes, we ceased cooking and allowed the air to clear naturally.

Our basic calibration involves scaling each of the Speck instrument outputs in order to minimize the mean percent error between the Speck signals and the $2\mu m$ channel of the HHPC-6+. This minimization is given in eqn 4.2, where $C$ is the calibration constant and $H_t$ is the HHPC-6+ $2\mu m$ measurement at time $t$. Of the five Specks, we select the
Figure 4.3: Plot of calibration data over time. Note the large discrepancy between the HHPC-6 channels.

Table 4.1: Table of pairwise $r^2$ values for all Specks for the calibration test

<table>
<thead>
<tr>
<th>Speck</th>
<th>Speck 1</th>
<th>Speck 2</th>
<th>Speck 3</th>
<th>Speck 4</th>
<th>Speck 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speck 1</td>
<td>1</td>
<td>0.935</td>
<td>0.991</td>
<td>0.991</td>
<td>0.993</td>
</tr>
<tr>
<td>Speck 2</td>
<td>0.935</td>
<td>1</td>
<td>0.902</td>
<td>0.955</td>
<td>0.922</td>
</tr>
<tr>
<td>Speck 3</td>
<td>0.991</td>
<td>0.902</td>
<td>1</td>
<td>0.979</td>
<td>0.985</td>
</tr>
<tr>
<td>Speck 4</td>
<td>0.991</td>
<td>0.955</td>
<td>0.979</td>
<td>1</td>
<td>0.985</td>
</tr>
<tr>
<td>Speck 5</td>
<td>0.993</td>
<td>0.922</td>
<td>0.985</td>
<td>0.985</td>
<td>1</td>
</tr>
</tbody>
</table>

three with the highest inter-Speck $r^2$ values (given in table 4.1) for use in subsequent tests, specifically Specks 1, 4, and 5 from this calibration set.

$$\min_c \frac{|C \times est_t - H_t|}{H_t}$$  \hspace{1cm} (4.2)

4.3.2 Cooking Test

In the cooking test, the three most consistent Specks and the two HHPC particle counters tests were exposed to a second cooking test, with environmental conditions similar to those used in the calibration pre-test. The outputs of the Specks were scaled by the calibration values calculated in the pre-test. Figure 4.4 shows that the calibrated Specks performed similarly in magnitude and shape to the HHPC-6+ $2\mu m$ channel. Additionally, table 4.6 shows that the Speck units correlate strongly with the $2\mu m$ channels for both HHPC instruments. While the channels of the two HHPC units correlate strongly to one another,
Table 4.2: Table of scalar calibration values

<table>
<thead>
<tr>
<th></th>
<th>Calibration Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speck 1</td>
<td>4.09</td>
</tr>
<tr>
<td>Speck 2</td>
<td>1.19</td>
</tr>
<tr>
<td>Speck 3</td>
<td>2.92</td>
</tr>
<tr>
<td>Speck 4</td>
<td>2.78</td>
</tr>
<tr>
<td>Speck 5</td>
<td>3.21</td>
</tr>
</tbody>
</table>

there is a significant discrepancy between the magnitudes of the HHPC-6 and the HHPC-6+ measurements, shown in figure 4.5

Figure 4.4: Plot of cooking data over time

4.3.3 Incense Test

The same three Specks and two HHPC particle counters were then placed in a small, closed room, where incense was burned for approximately five minutes, followed by a 45 minute rest period, and another five minute burning period. Ten minutes after the second burning, a HEPA air purifier was turned on until particle counts approached the initial baseline. In this experiment, the Specks underestimated the 2µm particle concentration with respect to the HHPC-6+, as shown in 4.8. The correlation values between all five instruments in table 4.10 were notably less than those from the cooking experiment. Despite the weaker correlation between HHPC instruments, the magnitude of the errors between these
Figure 4.5: Comparison of additional HHPC size channels for the cooking test, showing ppl vs time (minutes)

instruments is similar to those in the cooking experiment.

Figure 4.8: Plot of incense data over time
4.3.4 Results and Discussion

In both the calibration pre-test and the cooking test, we observed that the pairwise $r^2$ value between every Speck unit exceeds .90, and exceeds .98 for all but one unit. This demonstrates that the Speck design is internally consistent in its ability to detect pollutants from cooking. Furthermore, the $r^2$ values between the three test Specks exhibit a similarly strong correlation with the 2$\mu$m channels from the HHPC-6 and HHPC-6+ instruments, as well as the 1$\mu$m channel from the HHPC-6. This suggests that the Specks are most capable of detecting particles in the 2$\mu$m range.

The two HHPC monitors show strong correlation values between their corresponding size channels during the cooking test, which demonstrates the instruments’ capability to detect the same particle signals. We also note the strong correlation between the HHPC-6 1$\mu$m and 2$\mu$m channels, which may indicate a possible overlap in detection for these two sizes on this instrument. Though not as strong, the correlation is also high between these two channels on the HHPC-6+ instrument, which could also indicate that the ratio between these two particle sizes remained relatively constant, or that many particles fell...
Figure 4.9: Comparison of additional HHPC size channels for the incense test, showing ppl vs time (minutes)

within the 1µm to 2µm range.

The incense test was less conclusive, in that each of the five instruments demonstrated much weaker correlation values. The strongest correlations in this experiment were between the 1µm, 2µm, and 5µm channels of the HHPC-6. The similar magnitude in error between the HHPC-6 instruments across both experiments strongly suggests that the instruments differ in calibration. The magnitude of the error between the calibrated test Specks and the HHPC-6+ 2µm channel is notably less than the error between the two HHPC units in both experiments. We conclude that the Speck can be calibrated to closely match the performance of handheld optical particle counters intended for the professional market.

Because of the correlation discrepancy between the HHPC instruments in the incense test, we conclude that factors such as airflow and obstructions may have interfered with the even dispersal of particles between the instruments. In future work, this experiment should be repeated in a more controlled environment, with the instruments closer to the pollutant source and away from obstructions.

Both experiments exposed the instruments to very high particle concentrations. The shape of the .3µm data from both HHPC instruments and both experiments suggest that the instruments may saturate at levels just under $5 \times 10^5$ ppl. Another future experiment should include tests performed at ambient household particle concentrations in order to evaluate the performance of the Speck and corresponding calibration values at lower pollution levels. Ideally, calibration can also be performed at these lower concentrations for ease of scalability.

Ultimately, we plan to have the Speck measure particle counts as well as mass concen-
Figure 4.10: Table of incense $r^2$ values. Shaded cells indicate $r^2 > .9$

tration in $\mu g/m^3$, as this is the most common reporting method for $PM_{2.5}$ among state and federal monitoring stations. This would allow users to compare their indoor air quality with the outdoor air quality of their local region. This raises several challenges, however, as our sensor is optical rather than mass-based, and typical $PM_{2.5}$ measurements are the product of averaging over one to 24 hours. Additionally, the relationship between particle counts and mass will vary with particle composition. At present, we estimate mass readings using a linear scale factor generated from fitting Speck particle count data to that of a co-located tapered element oscillating microbalance (TEOM) monitor owned by the Allegheny County Health Department. Future works will incorporate onboard humidity and temperature measurements to refine both the particle count accuracy as well as the mass estimate.

4.4 Rapid Autonomous Group Exposure (RAGE) Calibration Method

The aforementioned kitchen calibration process demonstrated the potential for individual units to be calibrated to a handheld particle counter when exposed to varying levels of particulate matter. Mass production and distribution, however, requires a calibration method that meets the following key criteria:

- **Repeatability**: We require a means to generate repeatable and controlled $PM_{2.5}$ particulate emissions so that each calibration instance is nearly identical.
- **Scalability**: The system must be replicable so that multiple manufacturing sites can calibrate prior to shipping.
- **Consumable cost**: We require a calibration substance which can be readily sourced, reclaimed, and controlled at a low enough cost so as to not affect the price of the monitor.
- **Efficiency**: The calibration method must maximize throughput without sacrificing calibration quality or placing large demands on time or space resources.
• Automation: As much of the calibration and quality control process as possible should be automated to eliminate both the demand on human resources and the need for specialized training and experience.

The first challenge is to determine an inexpensive and repeatable means of generating controlled concentrations of particulate pollution. We experimented with several initial methods, including off-the-shelf essential oil vaporizers, hand-held vacuum cleaners with removed filters, and robotic vacuum cleaners with dust bins removed. Our unique method that has proven effective involves a mid-size subwoofer underneath a container of fine polydisperse particulates. Our initial substance was local construction dirt, primarily silicate, with a mixture of fine and large particulates. We discovered that a frequency of 100Hz was sufficient to vibrate the smaller particles into the air, which we could then propel in a controlled direction using a small fan. Using handheld particle counters, we determined that the airborne particles were largely at or smaller than 2 microns. Later, we replaced the construction dust with food grade diatomaceous earth, which exhibits a similar prevalence of small particles when aerosolized by the subwoofer at 100Hz.

At present, the particle generator is located approximately four feet away from the calibration bench. The bench holds one HHPC-6+ handheld particle counter closely surrounded by 10 Speck units. Using a second particle counter, we have determined an effective radius of about one foot around the HHPC-6+ within which the particle concentrations are largely consistent. This dictates our batch size of 10 Specks, given that each Speck must have ample room to intake and exhaust air. This number of Specks also allows us to analyze the pairwise $r^2$ correlation between each device to identify outliers, which are reconditioned (usually by replacing the sensor) prior to recalibration.
The calibration constants are generated by the same means as described in the kitchen calibration tests, producing individual scaling constants that minimizes the percent error between each Speck and the HHPC-6+. Unlike the kitchen tests, however, the constants are generated in real-time by a host server connected via Wi-Fi to each Speck and via Ethernet to the HHPC-6+. Upon the completion of the calibration batch, calibration files are automatically written to the host server, where the Specks automatically download their specific calibration file and begin displaying the new calibrated values.

The operator inspects each Speck visually to confirm functioning touch screens, fans, and Wi-Fi. In addition, the operator ensures that no Speck’s mean $r^2$ coefficient deviates significantly from the batch, or that the graphed output of the sensor deviates from that of the other Specks or the HHPC-6+. Outliers are immediately removed for reconditioning.

A typical calibration batch takes approximately 20 minutes, yielding 10 Specks per batch, or 30 per hour, or two minutes per Speck. Within a 20 minute batch, approximately five minutes are required for setup, and 15 minutes for particle exposure. The first five minutes of data are unused for calibration as the Specks acclimate to the initial particle concentration level of 500 $2\mu m$ particles per liter. After 10 minutes, the particle concentration is increased to 1000 particles per liter. At minute 15, all Speck calibration files are written and transmitted.

We plan to further automate this procedure. Presently, the operator manually controls the particle concentration by increasing or decreasing the amplitude of the 100Hz signal. The computer prompts the operator to increase particle counts at the appropriate time. We intend for this to be computer controlled in the next calibration procedure iteration. The time delay in this feedback loop will be minimized by another planned alteration, where instead of calibrating in an open, ventilated room, calibration will occur inside a sealed metal cabinet. This smaller space, equipped with circulating fans, should hopefully yield a uniform particle concentration within the chamber. Finally, we are presently seeking to create quantitative rules for singling out Specks for reconditioning based on $r^2$ values and absolute error.

Some operator oversight will remain necessary. For example, confirming the functionality of the fan and touchscreen will remain necessary. Additionally, all Specks are now labeled with key information such as calibration batch and date, serial number, and MAC address. These labels are automatically generated by plugging the Speck into a second computer, however this currently must occur one at a time, and the operator affixes the labels. These labels are typically generated for a batch of newly calibrated Specks while the next batch is running.

Theoretically, parallel calibration chambers may be run at once, however this will require a means of isolating each batch so that any given Speck only connects to the wireless router associated with its chamber. We intend to test whether we can accomplish this simply by constructing the chamber as a faraday cage.
Figure 4.12: Future calibration chamber.
4.5 Comparison with Federal Equivalence Methods

4.5.1 Outdoor tests with South Coast Air Quality Management District

In April and May of 2015, the South Coast Air Quality Management District (AQMD) performed one month of colocation tests using two calibrated Specks, one beta attenuation monitor (BAM) and one GRIMM federal equivalence monitor. Due to technical difficulties with each of the monitors, only one week of data is available where all four monitors are functioning normally. Despite this, the data provides several insights into the performance of the Speck against federal equivalence monitors, in addition to insight regarding the performance of the BAM and GRIMM federal equivalence monitors side-by-side. In fig. 4.13, we see the unfiltered data as reported by each of the four monitors across a one-week period. The Specks and GRIMM collect samples once per minute, whereas the BAM collects hourly averages. In order to better examine error and correlation, we convert the Speck and GRIMM data to one-hour moving averages with timestamps identical to those recorded by the BAM, as seen in fig. 4.14.

Some qualitative observations can be made at first glance from fig. 4.14. The hourly BAM data is somewhat erratic, frequently fluctuating by more than 10 $\mu g/m^3$ between readings. By contrast, this signal frequency is largely absent from the GRIMM and Speck data. The GRIMM also consistently reads lower particle concentrations than either the BAM or the two Specks. The median error between the GRIMM and the BAM is approximately a factor of 1.7. This makes assessing the absolute error of the Specks difficult, as the discrepancy between the federal equivalence monitors leaves uncertainty as to the actual
Figure 4.14: One week of time-series AQMD data at hourly averaged sampling rates

Table 4.3: Table of pairwise RMS error between devices at hourly sampling averages

<table>
<thead>
<tr>
<th></th>
<th>BAM</th>
<th>GRIMM</th>
<th>Speck 1</th>
<th>Speck 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAM</td>
<td>0</td>
<td>8.3023</td>
<td>6.7416</td>
<td>6.0384</td>
</tr>
<tr>
<td>GRIMM</td>
<td>8.3023</td>
<td>0</td>
<td>12.0914</td>
<td>10.5030</td>
</tr>
<tr>
<td>Speck 1</td>
<td>6.7416</td>
<td>12.0914</td>
<td>0</td>
<td>2.1349</td>
</tr>
<tr>
<td>Speck 2</td>
<td>6.0384</td>
<td>10.5030</td>
<td>2.1349</td>
<td>0</td>
</tr>
</tbody>
</table>

particle concentrations present at the test site. In light of this, we present our analysis as pairwise comparisons between each of the four monitors in terms of median absolute error and $r^2$ correlation values, as seen in tables 4.3 and 4.4 respectively.

Larger averaging windows will increase the correlation coefficients between each pair. As an example, table 4.5 demonstrates very high correlation coefficients when values from each instrument report hourly averages of the past 24 hours. Daily averages are used to determine EPA AQI values, though as can be seen in fig. 4.15, this removes many of the potentially important signal dynamics.

4.6 ESDR

We have established the importance of dense spatiotemporal sensing and visualization. We predict, however, that many spatially-located sensors collecting samples once per minute will quickly exceed the number of data points that Excel or even Google Maps can easily visualize. Thus, we require a new data storage resource that enables the collection of a
Table 4.4: Table of pairwise $r^2$ values between devices at hourly sampling averages

<table>
<thead>
<tr>
<th></th>
<th>BAM</th>
<th>GRIMM</th>
<th>Speck 1</th>
<th>Speck 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAM</td>
<td>1</td>
<td>0.5299</td>
<td>0.2791</td>
<td>0.1641</td>
</tr>
<tr>
<td>GRIMM</td>
<td>0.5299</td>
<td>1</td>
<td>0.5603</td>
<td>0.3679</td>
</tr>
<tr>
<td>Speck 1</td>
<td>0.2791</td>
<td>0.5603</td>
<td>1</td>
<td>0.8785</td>
</tr>
<tr>
<td>Speck 2</td>
<td>0.1641</td>
<td>0.3679</td>
<td>0.8785</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.5: Table of pairwise $r^2$ values between devices at 24hr sampling averages, reported hourly

<table>
<thead>
<tr>
<th></th>
<th>BAM</th>
<th>GRIMM</th>
<th>Speck 1</th>
<th>Speck 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAM</td>
<td>1</td>
<td>0.9111</td>
<td>0.7720</td>
<td>0.7063</td>
</tr>
<tr>
<td>GRIMM</td>
<td>0.9111</td>
<td>1</td>
<td>0.8715</td>
<td>0.8030</td>
</tr>
<tr>
<td>Speck 1</td>
<td>0.7720</td>
<td>0.8715</td>
<td>1</td>
<td>0.9836</td>
</tr>
<tr>
<td>Speck 2</td>
<td>0.7063</td>
<td>0.8030</td>
<td>0.9836</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 4.15: One week of time-series AQMD with 24hr averages, reported hourly
large amount of data. Additionally, this system must provide a method for interactive
data exploration by both scientists and the public. For data sources that vary greatly
within short time intervals but also demonstrate longer-term trends, subsampling and data
summaries are important for the data to remain easily decipherable when viewed over long
intervals such as months or years. Normally, this would require client-side computation.

Precomputing various summary levels on the server side both eliminates the computa-
tional load on the client while providing a bandwidth-efficient means to transmit data
at lower zoom levels (across longer time intervals). Additionally, as with any typical data	
pling system, the storage cost increase is only twice that of the raw data.
The ability for users to annotate the data (providing relevant metadata) is an additional
key feature in encouraging dialogue and shared learning. The combination of context and
data allows for rich storytelling that carries much more rhetorical power than data or
anecdotal evidence alone.

We have conducted a large-scale survey of existing data repository systems, both open-
source and for-profit, and have established that no extant repository system is able to
provide the features required for a community-based air quality reporting system with
authentication, privacy, annotation, and the ability to handle massive spatial-temporal
resolution. Thus, as part of the proposed work, we will develop an open-source Environ-
mental Sensor Data Repository (ESDR) customized for these particular requirements.

We have examined existing data repositories and visualization tools, and as shown in
Table 1, no current tools have both the features we need for data upload flexibility and
exploration. See the data plan for more information on system architecture and data
security.

There are several use cases for ESDR, each of which requires a diverse set of visualization
techniques. Our system will allow for the user to create and tailor views that meet their
inquiries. For example, if a user is interested in personal exposure inside his or her own
home, a timeline view of the sensor or sensors installed in that location may make the most
sense. This timeline will be zoomable along both the time axis and the vertical value axis.
Annotation capabilities will allow the user to mark when specific events such as cooking
occur.

On a community level, multiple individuals may be interested in comparing ambient
data side by side. Stacked timelines may be appropriate for this, which we have demon-
strated and implemented in BodyTrack (http://bodytrack.org/info.html) and Fluxtream
(https://fluxtream.org/) projects. These too will have adjustable scales and resolutions
for better comparison.

In other cases, spatial information is important, such as any data collected using the
bicycle monitoring kits from GASP. In this case, heat maps (Wilkinson and Friendly, 2009)
with variable time ranges and zoomable scales will be essential. We have experience with
this type of visualization, as evidenced by our visualization of 599 AirNow stations from
across the United States (http://explorables.cmucreatelab.org/explorables/airnow.html).
Multiple sources can be used, including any indoor sensors with geolocation whose owners
are open to sharing their data.

Privacy is a major concern, and so any users who are not comfortable with openly
sharing their data will have the ability to abstain or to selectively share. It is our hope,
4.6 Comparison of features afforded by ESDR vs. other available platforms

<table>
<thead>
<tr>
<th>Tools/Features</th>
<th>Upload Data</th>
<th>Custom Data Sources</th>
<th>Exportable Spatiotemporal Interaction</th>
<th>Text Annotation</th>
<th>Photo Annotation</th>
<th>Spatiotemporal Sharing</th>
<th>Selective Sharing</th>
<th>Embed Capability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESDR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Habitat Map</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluxstream</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Weather Underground</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phant</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>AirCasting</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xively</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netatmo</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SafeCasting</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>airNow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

however, that users will see the value in aggregating as much data as possible to address community issues by removing spatial resolution so that individual sensor positions are dissociated from publicly viewable data. Cross-authentication services allow each participant to choose to share their data at geospatially low-fidelity with the community as a whole.

The versatility of ESDR will allow us to additionally import data from other sources as necessary or appropriate, such as wind and weather data, and data from AirNow and other air quality services. CREATE has already imported all data streams from AirNow, delivering a first time-explorable interactive for public consumption showing all federal air quality monitors nationally.

For each of these scenarios, it is important to capture qualitative data from the community. This collection of metadata involves being able to annotate the data with comments such as times when users notice strange smells in their homes or outdoors, or when they notice traffic to be particularly congested. This feature will facilitate the sharing of stories that contain both qualitative experience and quantitative data.

4.7 Building Management System Integration

Building management is another rich area for potential enrichment through Speck sensors and data visualization. Because of the high cost of air quality monitors, important health factors like carbon dioxide, $PM_{2.5}$, and other criteria pollutants are often only measured in
a small number of locations within a building, if any. Particulate matter sensors are very uncommon even in more advanced and health-conscious buildings.

### 4.7.1 Demonstration of Integration Potential

Using pre-existing example data from a whole-room test of 144 specks in a 12x12 grid, we have prototyped an initial concept for a building management system. This early prototype shows a cutaway view of an office complex, with each room colored according to the (simulated) measurement of air quality within this room at specific times. The time slider can be used to either explore different times or automatically animate the air quality dynamics over time. Clicking on any given room generates a line plot of the air quality in that area that aligns with the time slider. Multiple plots can be generated and overlaid to demonstrate linked air quality between nearby rooms.

In this particular example, the data is arranged such that an apparent pollution event from outside infiltrates the building through stairwells first before dispersing through individual offices. Some areas remain clean while others experience highly increased particle
concentrations.

The usefulness of such a visualization for real buildings with real-time air quality data is that it can enable building managers to identify effective strategies for ventilating the building to optimize both energy usage and air quality. Built-in temperature and humidity sensors can help building managers to estimate occupant comfort levels. Furthermore, the future addition of carbon dioxide sensors gives a simultaneous indication of air freshness in addition to occupancy levels. If carbon dioxide levels or particle concentration levels are high, an articulate HVAC system can circulate new, filtered air into those areas. Additionally, external building sensors can be used to control how much air is recirculated versus how much new air to draw in from outside.

Such a system also has Human Resources benefits as well. By identifying the offices with the cleanest air, managers can proactively offer these spaces to employees with known breathing conditions, those who may be pregnant, and other sensitive populations. This may help to simultaneously increase productivity and employee comfort while reducing the rate of workplace illnesses.

4.7.2 Newell Simon Trial Deployment

The low cost and small size of the Speck, in addition to its wireless connectivity, will allow building managers to densely install air quality sensors in areas of interest such as offices, open workspaces, and factory floors. These sensors can be easily moved and reconfigured as necessary and do not require planned or pre-existing infrastructure beyond wireless connectivity and power. We predict that the Speck will not require additional calibration or maintenance beyond simply cleaning the optics with canned air, so the cost of this increased capability does not greatly exceed the cost of initial installation. In preparation for our first major deployment of Speck sensors on the Infosys campus, we installed five Speck sensors on the third and fourth floors of Newell Simon Hall on Carnegie Mellon University’s campus. These sensors were spatially distributed in a combination of open and closed spaces. One sensor each on the third and fourth floors are located in open areas. While this initial deployment was very sparse because of the currently high demand and low supply of Specks, the data collected from these five units were intended to give us example data from the following locations:

- Unoccupied double-occupant offices
- Regularly occupied double-occupant offices
- Intermittently occupied multi-occupant offices
- Copy and printer rooms
- Regularly occupied offices with constant open hallway airflow

The only campus wireless SSID that the Specks can connect to is unsecured, but requires all MAC addresses to be pre-approved. More importantly, however, is the fact that this network seems to have a hidden SSID and broadcasts weakly and intermittently. As a result, two of five Specks are typically behind on ESDR. This raises an important question for future deployments, however, regarding how real-time visualizations should deal with stale or absent data. We will address this in the thesis work for Infosys deployments.
Chapter 5

Proposed Work

5.1 Calibration Refinement

We have established the necessity of individual calibration for low-cost air quality sensors in previous chapters. Manufacturing differences inherent in complex but low-cost devices ensures that no single set of calibration values will work for all sensors. Additionally, the importance of quality control procedures is much higher than in many other consumer devices because users will be basing important health decisions around the readings shown by their device, often in comparison with local federal monitors, co-located Specks, or Specks owned by nearby neighbors. As described previously, in order to keep cost low and reliability high, our calibration system must be judged based on the following criteria:

- **Repeatability**: We require a means to generate repeatable and controlled $PM_{2.5}$ particulate emissions so that each calibration instance is nearly identical.
- **Scalability**: The system must be replicable so that multiple manufacturing sites can calibrate prior to shipping.
- **Consumable cost**: We require a calibration substance which can be readily sourced, reclaimed, and controlled at a low enough cost so as to not affect the price of the monitor.
- **Efficiency**: The calibration method must maximize throughput without sacrificing calibration quality or placing large demands on time or space resources.
- **Automation**: As much of the calibration and quality control process as possible should be automated to eliminate both the demand on human resources and the need for specialized training and experience.

We already have a largely automated system with the potential to meet all of these goals, however, further improvements can be made to enhance our adherence to each of these criteria.

Automatic regulation of particle concentrations can be done via a feedback control loop between the HHPC-6+ reference instrument and the amplitude of the 100Hz audio signal. This will increase repeatability as well as automation, since it removes the cognitive load of the operator significantly. It also reduces the observed tendency for operators to
overestimate the degree and rate to which the amplitude should be increased due to the significant time taken for additional particles to become airborne, cover the devices, and achieve equilibrium. The calibration particle signal will thus be more consistent and stable.

Encasing the entire calibration system within a smaller enclosure (as opposed to an open room) will reduce this time lag, thus making automatic control simpler. It also increases repeatability by eliminating external factors such as air currents in the ventilated open room. Scalability is improved because these calibration chambers can be produced identically and in quantity, allowing for other calibration locations to have identical equipment and environments. If multiple calibration chambers can be used in the same location via Wi-Fi isolation, efficiency is also increased by improving parallelism and throughput. Additionally, whereas it is currently not possible to reclaim used particles from an open room, some particles may be reclaimed from a closed chamber by emptying an underlying hopper collection system. This reduces consumable cost.

We also plan to make the particle source smaller, such that a smaller amount of calibration substance (likely to remain food-grade diatomaceous earth) can be used. This also reduces consumable cost, and the source can be replenished after every calibration batch using a fixed quantity of particles to improve repeatability.

Additional automation improvements may be possible through the use of automatically printed labels and on-screen messages identifying which labels should be affixed to which Specks. These on-screen messages can also be used to identify Specks for reconditioning based on new quantified rules for Specks whose dynamics deviate from simultaneously calibrated units, or whose calibration data does not meet a maximum error threshold in comparison to the reference instrument.

Finally, if we are able to improve the repeatability and quality of the calibration process significantly, then the HHPC-6+ can be replaced with a single trusted and well-calibrated Speck. These master Specks will be calibrated and individually assessed by a calibration specialist against the HHPC-6+ to ensure that they will be an appropriate substitute for the reference instrument. Thus, calibration for consumer Specks can be achieved for lower cost using a process with one degree of separation from the HHPC-6+ monitor, eliminating the need for multiples of these expensive monitors in one location. Recalibration and regular cleaning of the reference Specks can be performed at regular intervals as determined through our initial testing.

5.2 Large-scale community deployment

A section of this thesis will focus on the impact of Specks in individuals and communities. Specifically, we will seek to qualitatively and quantitatively understand and improve the effect of Specks on the empowerment of individuals to understand and control their air quality. The first step toward this goal is to collectively determine the definition and dimensions of empowerment in conjunction with Carnegie Mellon’s Environment and Public Policy program. We will achieve this in collaboration with faculty members and graduate students who are familiar with the codification of empowerment in other domains such as social safety and awareness. The Speck will be used in student self-guided capstone
projects in the Fall semester of 2015, led by Professor Baruch Fischhoff. These projects will likely provide significant insight and data regarding the use and impact of Specks on individuals. Throughout this process, we will serve as technical advisors to the students and share the insight we have already gained through our pilot deployment of Specks and feedback from our existing users.

The Metro 21 project is currently placing Specks in classroom settings, which will allow for us to understand how children and educators can use low-cost air quality monitors to better understand their environment. We are particularly interested in the ways in which data visualization and storytelling can be merged to develop compelling rhetoric in children.

Additionally, Specks are gradually being deployed through Pittsburgh’s Carnegie Library system. As part of Metro 21, we will administer surveys and interviews to gauge empowerment in users who check out Specks for short periods of time. These surveys will be administered on a voluntary basis using Chromebooks placed in local libraries. We have observed that there is significant interest in placing air quality monitors in outdoor areas, specifically where local polluters are of community concern. We are working closely with Allegheny County Clean Air Now (ACCAN) to deploy Specks around the Shenango Coke-works plant, also in collaboration with the Allegheny County Health Department (ACHD), which has installed a PM$_{2.5}$ BAM and weather station along with other sensors to monitor the air quality in the nearby Avalon community. The Specks deployed in these areas are installed on or near volunteer individuals’ homes, allowing them to gain insight into their hyperlocal outdoor air quality. In addition to codifying the impact of Specks on this community, we must seek to address several technical challenges that outdoor installations of Specks produce. The Specks are designed for indoor use, and as such, are not inherently equipped to deal with extreme temperatures, exceedingly high humidity, rain, or wind. The Specks are placed in covered enclosures to prevent immediate damage from natural elements. They are also placed in multiples of two to identify potential malfunctions by observing deviations in data and behavior. While the functioning colocated Specks tend to match each other, there are sometimes significant deviations between their readings and the ACHD BAM, even between the closest locations. We will seek to understand and remedy this discrepancy.

We know that high humidity affects optical particle behavior through hydroscopic growth. We will seek to employ basic humidity correction using built-in humidity and temperature sensors, in addition to identifying times when condensation is likely to occur. Additionally, we have recently discovered that rapid increases in Speck particle counts are highly correlated with high gusts of wind. This connection is only made possible through our observation of minutely wind data, as these events occur quickly and are not evident in hourly averages. Our current hypothesis is that gusts of wind lift larger particles into the air which the Specks then detect. While the Speck is less sensitive to larger particles, as previously established, the sensitivity is non-zero. By contrast, the BAM and other federal reference methods use mechanical cyclones and high air flow to separate out large particles at a 2.5 micron sharp cut. We believe this may be the source of the discrepancy, along with the fact that the Specks are located much closer to the ground than the nearby BAM unit. We will attempt to confirm this hypothesis by adding pre-conditioning cyclones to at least one Speck enclosure to see if this eliminates the temporary spikes along with the effects of
wind and larger particles. This is a temporary and non-scalable solution, however, as the cyclones are expensive, and even appropriately sized filters will remove smaller particles along with those larger than $PM_{2.5}$ particles. If we can confirm that larger particles are the source of the discrepancy between Specks and the BAM, we will immediately recommend the installation of inexpensive anemometers along with outdoor enclosures, and we will seek to design inexpensive inlets along with the faculty in Carnegie Mellon’s air quality research labs which can be used to mitigate if not eliminate this effect.

Finally, we may seek to deploy our own GRIMM optical particle counter and gravimetric monitor next to a local Speck installation. The GRIMM is capable of measuring particles within 31 size bins ranging from .25 to greater than 32 microns. This will easily let us see what fraction of the Speck signal is due to large particles, and whether this fraction increases significantly as we expect during short wind gusts.

5.3 Building Management System Integration

5.3.1 Office Buildings: Infosys Deployments

In September 2015, we began our first major deployment of Speck sensors within the Infosys campuses in Bangalore and Chennai. These buildings are now outfitted with 26 and 14 sensors respectively, with at least one sensor per floor with the remainder located on a single floor in an open floorplan. In order to facilitate secure and efficient transmission of the data between India and the ESDR server in Pittsburgh, the Specks connect first to a
local Raspberry Pi based server, which then makes the single secure outgoing connection to ESDR. The Raspberry Pi also generates and hosts formatted Speck data which is accessible by campus BACnet systems already in place for building management.

For this initial deployment, we will continue using our existing personal monitor technology as a proof of concept. While the system is functional, it is not suited for widespread deployment because of the difficulty of adapting the home monitors to a secure office campus. During installation, we gained key experience that we will use to co-design air quality monitoring technology specifically for building management. The following is a list of key hardware changes to be made:

- Change the form factor to a wall or ceiling mount design to avoid protruding into walking spaces
- Eliminate or reduce on-device screen and emphasize use of networked visualizations
- Make the device permanently mountable to prevent theft or tampering
- Upgrade WiFi chip to be capable of secure WPA2 Enterprise encryption
- Potentially integrate reliable CO2, temperature, and humidity capability
- Potentially enable any device to be hardwired and act as a BACnet gateway
- Potentially upgrade particle sensor or use multiple sensors in each unit for redundancy

With a device capable of measuring particulate matter, CO2, temperature, and humidity, we will have created a monitor that can measure the key parameters for occupant comfort as well as air quality factors that are related to health and productivity. For large campuses, many building sensors are used directly in the loop for ventilation, heating, and cooling automation. We will work to integrate our technology with existing building management systems, for example those from Automated Logic or Johnson Controls. Some of these software products are designed to allow building managers to add a variety of sensors and devices as long as they communicate using common protocols such as BACnet. Time and resources permitting, we will conduct a second deployment to demonstrate the use of building sensor prototypes and software integration for control of building features. First, we will simulate building automation by comparing the prescribed behavior from installed sensors with the automated behavior from existing air quality monitors. When it is evident that sensor accuracy and reliability are high enough to control real building parameters without adversely affecting occupant comfort or productivity, we will begin testing actual automation.

At the end of this deployment, we will perform a cost-benefit analysis of our system for a variety of building sizes and configurations. We will develop recommendations for sensor placement density, installation guidelines, and estimated return-on-investment rates based on predicted changes in occupant productivity.

5.4 Visualization concepts for building management

The utility of a building management system is largely dependent on the way in which it presents data to stakeholders. We have identified the importance of spatiotemporal
data visualization, and have developed several key technologies that may be of use in building management scenarios. While most existing building management platforms are primarily concerned with presenting data to those in charge of operating and monitoring the building’s systems, we identify three classes of stakeholders with different sets of needs, each of which can be designed for separately.

5.4.1 Visualization for occupants

Occupants may wish to know their immediate environmental conditions. Most frequently, they are likely to be interested in temperature. In manufacturing environments, air quality may also be of natural interest. With the ability to view temperature, humidity, \( CO_2 \) levels, and particulate matter measurements in real time, however, we expect individuals to gain a better understanding of how these factors influence their comfort and productivity. One key aspect of designing for occupants will be that these readings should be presented on-demand rather than spontaneously in order to avoid inducing placebo effects. While occupants may be interested in historical data as well, we predict that their primary interest will be in current local conditions, so their visualization portal may resemble a dashboard or dock on their computer or phone. Additionally, not all occupants may have access to data from the entire campus, and may be restricted both for security and simplicity reasons. While these are our initial assumptions, we will conduct user studies to determine what is needed and appropriate for this group.

5.4.2 Visualization for management

This class of visualizations will draw largely from existing systems, augmented with our own visualization technology. If possible, we will directly integrate into an existing building management software package such as WebCTRL in order to maximize impact and familiarity among existing users. We will conduct user studies here as well in order to present building managers with the most useful information. It is likely that these users will need to view the data on adjustable spatiotemporal scales, either at a whole-campus, floor by floor, or partition level. WebCTRL currently allows for regional viewing as well as subsystem viewing for heating, cooling, and ventilation. Timeline viewing of historical data will be important for identifying trends, as well as a method of viewing cyclical patterns on a daily, weekly, monthly, or yearly basis. Potentially, quantities such as a comfort index can also be calculated and visualized using a combination of sensor data, such as temperature and humidity.

5.4.3 Visualization for building evaluation

Upper management and marketing may also be interested in visualizing climate and air quality data. In these cases, users may desire easily-explorable summaries, similar to those for building management teams but with less granularity and more generalization. These summaries may be used for promotion of environmental technologies, broad evaluation of
working conditions, and potential marketing to clients who may lease the space in the case of independent building owners.
Chapter 6

Schedule of Work

The three thrusts of this thesis, being the Device Focus, Community Focus, and Building Management Focus, will be undertaken largely in parallel due to each of these becoming activated in the latter half of 2015.

Most of the Device Focus work will occur early in 2016, as we have need of a new device design due to the discontinuation of several key components. We will ideally complete this design in the first third of 2016, after which we will characterize this sensor extensively alongside federal monitors as well as other personal monitors that are rapidly becoming available.

The Metro21 library deployments are currently in progress, and we will be collecting data from library users over the coming months. We will use this along with customer feedback and projects conducted by the Fall 2015 EPP class using Speck to improve educational resources on our website to better help people understand and use their Speck data on fine particulate matter. In parallel, we will test the device (ideally prototypes of the second version) in more robust outdoor enclosures to attempt to discover if interferences from wind and external conditions can be eliminated or minimized.

We have currently developed a first prototype of a building management visualization based on the Bangalore deployment. This visualization is not integrated with any existing systems, however, and will undergo significant refinement in the first third of 2016. Work to design an integrated building monitor will commence in the Spring, and shortly thereafter, we will begin work on building automation integration using the current deployment. If a second deployment can be made, we will transition to using those monitors for automation as well.

The Fall of 2016 will be used to finalize documentation and write the thesis manuscript, to be presented in December.
<table>
<thead>
<tr>
<th>Thrust</th>
<th>Device</th>
<th>Community</th>
<th>Building Management</th>
<th>Thesis manuscript</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration chamber dev</td>
<td>Interactive air quality visualization</td>
<td>Metro 21 Data Collection</td>
<td>NSH/India deployment</td>
</tr>
<tr>
<td>Nov-15</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dec-15</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jan-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Feb-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mar-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Apr-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>May-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jun-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jul-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Aug-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sep-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Oct-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Nov-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dec-16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Chapter 7

Contributions

7.1 Calibration and Characterization

As previously mentioned, efficient and effective calibration is necessary for low-cost sensors to become an important part of a monitoring ecosystem. The calibration techniques described in this thesis along with the quality control methods employed are generalizable to other sensors as well. These generalizations can be used by others seeking to develop low-cost but accurate sensors or monitors by enabling multi-sensor calibration and comparative quality control.

Additionally, the characterization presented in this thesis can be used as a basis for characterizing other sensors as well by highlighting the limitations of traditional comparative techniques like correlation coefficients and cumulative absolute error. The value of appropriate filters and real-world evaluation conditions will be emphasized, along with the qualitative importance of a sensor’s ability to detect events of significance and interest.

7.2 Community Resources

Proper integration of personal low-cost monitors into society requires that we provide a means for individuals to obtain proper data fluency and an understanding of the parameters they seek to measure, such as fine particulate matter. In this thesis, we will present educational resources and interactive visualizations that promote intuition of spatiotemporal data and an exploratory mindset for determining environmental causes and effects, specifically focused on air quality and health.

We will also present methods for dispersing low-cost monitors throughout communities using libraries, citizen science groups, and educational outreach. We will demonstrate the importance of community in understanding data in larger regional contexts.
7.3 Building Management

This thesis will explore the means by which air quality data can become a staple of building management. We will demonstrate the added value to building managers of measuring parameters such as fine particulates and $CO_2$ in the context of comfort indices and occupant productivity. The prototype building monitors will demonstrate the potential of integrated temperature, humidity, $CO_2$ and fine particulate sensors into a single unit which can easily be installed in existing buildings because of wireless communication.

The building visualizations developed in this thesis will show that environmental data in the workplace can be of common interest to building occupants and managers alike. Finally, we will explore and present new methods of using this spatially distributed data to identify and model environmental patterns, and subsequently create automation rules to regulate comfort and clean air.
Appendix A

Pilot Study Surveys
Hello,

Thank you again for participating in our Speck user study!

Before you begin using the Speck, we would like to understand how much you know about air quality in general, and how aware you are of the air quality in your home. This initial survey should only take about 5 minutes to complete. Please be open and honest in your responses – there are no wrong answers!

Sincerely,
The CREATE Lab Speck Team

**1. Speck Number (can be found on the box or on the bottom of the Speck)**

**2. How knowledgeable would you say you are about air quality in general?**

- Very knowledgeable
- Quite knowledgeable
- Somewhat knowledgeable
- Not knowledgeable

**3. Do you think air quality can cause or exacerbate the following health issues? Please check all that apply.**

- Asthma and other respiratory illnesses
- Heart disease
- Diabetes
- Lung Cancer
- Stroke
- Epilepsy
- Allergic responses

**4. What kind of a residence do you currently live in?**

- Rented house
- Owned house
- Rented apartment or condo
- Owned apartment or condo
- Other (please specify)

**5. On average, how would you rate the air quality in your home?**

- Very good
- Good
- Fair
- Poor
- Very poor
6. Are you confident that you will know what actions to take, if you learned that your indoor air quality was poor?
- Very confident
- Quite confident
- Somewhat confident
- Not confident

7. What do you think are some of the sources of pollution inside your home? Please check all that apply.
- Cooking
- Vacuuming
- Smoking
- Microwave oven
- Gas heating
- Fireplace
- Open windows
- Insulation
- Refrigerator
- Other (please specify)

8. Are you aware of what kind of air filtration system is installed in your home?
- Yes
- No
- I don't have one

9. Do you have a hood vent or ventilation fan over your stove?
- Yes
- No
- I don't know
10. In the past, have you taken any of the following steps to improve the air quality in your home? *Check all that apply.*

- [ ] Changed the filter
- [ ] Cleaned the filter
- [ ] Installed a stove hood vent
- [ ] Installed an air purifier
- [ ] I have not taken any steps to improve the air quality in my home
- [ ] Other (please specify)

11. Please use this space to share any other comments or thoughts you may have about the Speck or this user study.

[Blank space for comments]

*
Thank you again for participating in our Speck user study!

This second survey asks a few questions about your experiences with the Speck so far and should only take 5 – 10 minutes to complete. We truly appreciate your open and honest feedback!

Sincerely,
The CREATE Lab Speck Team

1. Speck Number (can be found on the box or on the bottom of the Speck)

2. Have you been able to successfully set up the Speck in your home?
   - [ ] Yes
   - [ ] No
   - [ ] Have not tried it yet

3. How would you describe your experience installing the Speck?
   - [ ] Very easy
   - [ ] Easy
   - [ ] Moderate
   - [ ] Difficult
   - [ ] Very difficult

4. If you had difficulty installing your Speck, please describe your experience. Please check all that apply.
   - [ ] Confusing
   - [ ] Frustrating
   - [ ] Overwhelming
   - [ ] I did not experience any difficulty setting up the Speck
   - [ ] Other (please specify)

5. Based on your experiences so far, how would you rate the following resources?

<table>
<thead>
<tr>
<th>Resource</th>
<th>Essential</th>
<th>Very Useful</th>
<th>Somewhat Useful</th>
<th>Not Useful</th>
<th>N/A (I have not used this)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speck Setup Guide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speck Website (<a href="http://specksensor.org">http://specksensor.org</a>)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speck Users Facebook Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. If you saw a spike in the Speck reading in the past week, how did that make you feel?

Please check all that apply.

☐ Nervous
☐ Stressed
☐ Uncertain about what to do
☐ Curious to find the source
☐ Motivated to fix the issue
☐ Skeptical about the reading
☐ I did not see a spike

7. If you saw a drop in the Speck reading in the past week, how did that make you feel?

Please check all that apply.

☐ Satisfied
☐ Relieved
☐ Confused
☐ Motivated to identify the cause
☐ Skeptical about the reading
☐ I did not see a drop

8. Have you become aware of or learned anything new about the air quality in your home after using the Speck? Please explain briefly.

☐

9. Based on readings you’ve seen on the Speck during the past week, have you made any changes in your home to improve the air quality? Please briefly explain your response.

Yes, I have: 

☐ Not yet, but I plan to: 

☐ No, I have not and don’t plan to:

10. Please use this space to comment on the experiences you’ve had so far with the Speck and supporting material (setup guide, website and Facebook group).

☐
Hello,

Thank you again for participating in our Speck user study!

Now that you have been using the Speck for about two weeks, we would love to hear your overall thoughts about the Speck and supporting material. This final survey is more detailed and should take 20 – 30 minutes to complete. Please be open and honest - we truly appreciate your feedback!

Sincerely,
The CREATE Lab Speck Team
**Baseline Information**

*1. Speck Number (can be found on the box or on the bottom of the Speck)*

*2. On average, how would you rate the air quality in your home?*

- Very good
- Good
- Fair
- Poor
- Very poor

*3. What do you think are some of the sources of pollution inside your home? Please check all that apply.*

- Cooking
- Vacuuming
- Smoking
- Microwave oven
- Gas heating
- Fireplace
- Open windows
- Insulation
- Refrigerator
- Pets
- Other (please specify)

*4. Are you aware of what kind of air filtration system is installed in your home?*

- Yes
- No
- I don't have one

*5. Do you have a hood vent or ventilation fan over your stove?*

- Yes
- No
- I don't know

*6. Are you confident that you will know what actions to take, if you learned that your indoor air quality was poor?*

- Very confident
- Quite confident
- Somewhat confident
- Not confident
7. How knowledgeable would you say you are about air quality in general?

- [ ] Very knowledgeable
- [ ] Quite knowledgeable
- [ ] Somewhat knowledgeable
- [ ] Not knowledgeable
Experience with the Speck

8. How would you describe your overall experience using the Speck? Please check all that apply.

- Enjoyable
- Useful
- Informative
- Challenging
- Confusing
- Frustrating
- Other (please specify)

9. Would you say that the Speck met your expectations as an air quality monitor?

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Disagree

10. In which rooms/locations of your home did you place the Speck? Please list.

11. On average, how long did you keep the Speck in each room?

- 1 to 3 days
- 4 to 6 days
- 7 to 9 days
- 10 to 14 days
- Other (please specify)
12. How often did you look at the Speck reading?
- Each time I walked into the room it was in
- Multiple times a day
- Once a day
- Less than once a day
- Only when I thought there was a problem
- Other (please specify)

13. Were you surprised by Speck readings in any specific rooms/locations? *Please explain briefly.*

14. What action, if any, did you take to improve your air quality based on what you learned from using the Speck? *Please explain briefly.*

15. How would you like to be alerted about significant changes in your air quality, if at all?
- Blinking light
- Sound
- Text message
- Email
- Voicemail
- I do not want alerts
- Other (please specify)

16. Would you want the capability to personalize the Speck settings so that you can, for example, choose when you get alerts, alter the types of historical views that are displayed, change the color coding, etc.? *Please explain briefly.*
17. The next version of the Speck will include a quieter fan and a more robust micro USB port. Apart from these modifications, what other changes would you suggest for the Speck (e.g. shape, size, color, screen, displayed information, portability, etc.)? Please explain briefly.
**18. Based on your experience in this study, how would you rate the following resources?**

<table>
<thead>
<tr>
<th></th>
<th>Essential</th>
<th>Very Useful</th>
<th>Somewhat Useful</th>
<th>Not Useful</th>
<th>N/A (I did not use this)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speck Setup Guide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speck Website (<a href="http://specksensor.org">http://specksensor.org</a>)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speck Users Facebook Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

19. Is there anything you would change in the Speck Setup Guide to make it more useful? *Please explain briefly.*

20. What other information (if any) would you have liked to see on the Speck website (http://specksensor.org)? *Please explain briefly.*

*21. Do you think it is useful to have a discussion platform (on Facebook or elsewhere) for Speck users?*

- [ ] Essential
- [ ] Very Useful
- [ ] Somewhat Useful
- [ ] Not Useful
- [ ] Other (please specify)
**22. What is your opinion about the following statement?**
"The Speck helped me feel empowered to understand and manage my air quality."

- [ ] Strongly Agree  
- [ ] Agree  
- [ ] Neutral  
- [ ] Disagree  
- [ ] Strongly Disagree

**23. Has the Speck been useful for your home?**

- [ ] Strongly Agree  
- [ ] Agree  
- [ ] Neutral  
- [ ] Disagree  
- [ ] Strongly Disagree

**24. Are you likely to purchase a Speck in the near future?**

- [ ] Strongly Agree  
- [ ] Agree  
- [ ] Neutral  
- [ ] Disagree  
- [ ] Strongly Disagree

**25. How much would you be willing to spend on a Speck?**

- [ ] Between $50 and $100  
- [ ] Between $100 and $150  
- [ ] Between $150 and $200

26. Please use the space below to share any anecdotes or notable experiences you had during this study. *We'd love to hear your stories!*  

27. If you have any other thoughts, questions or comments about your overall experiences using the Speck and associated material, please list them here. *We truly appreciate your open and honest feedback.*
Bibliography


[12] David Gobeli, Herbert Schloesser, and Thomas Pottberg. Met one instruments bam-


