Chapter 1

Introduction

Earth, our home planet, is the only place in the known universe affirmed to host life. The life on earth is characterized by three components, air, land, and water. Each element has its own special property and is required in its proper proportions to maintain the healthy life of all living beings [1]. The quality of life relies upon the quality of the environment. However, industrial development and other manmade activities create a disturbance in the natural environment. The process of making environment unsuitable and unsafe for the living is called pollution.

Pollution changes the quality of the atmosphere and is transboundary, as they travel thousands of miles [1]. The presence of Persistent Organic Pollutants (POPs) [2] in the environment make adverse impact on human health and sorrounding as it resist degradation by microorganisms [3]. Many types of pollution contribute to global warming and climate change which are the major issues tackled by environmental scientist these days. Out of different types of pollution, air pollution plays a major role in global warming.

Air contamination alludes to release of pollutants into the environment that is unfavorable to human wellbeing and the planet in general. Air pollution is stated as a complex mixture of gases and particles whose sources and composition vary over space and time [4]. The boom in the development of industries and technology have created an alarming situation that is, degrading the quality of air. Contamination of air is a matter of serious concern and society is quite unaware of the impact that it causes to human health as well as to the surrounding. The World Health Organization (WHO) reported that death rate estimates are around 7 million every year as 9 out of 10 people breathe polluted air [5] [6]. This has made many motivated individuals like researchers and communities to work towards creating awareness among the people. Various studies are underway on air pollution due to its increasing effect on the environment day by day. This brings to the reason on choosing my research topic as I always wanted to work towards a social concern. Having the background of electronics and combining with advanced computer skills seems to be a perfect combination for problem solving. In the following section, I will discuss the background to introduce my research contribution.

1.1 Air Pollutants and Measurement Metrics

There are various pollutants that contribute to the contamination of the environment. These pollutants differ from region to region depending on the human activities. For example, an industrial area which manufactures products from raw material, such as production of iron from its ore or production of gasoline from crude oil, releases inorganic carbon compounds into the atmosphere [7]. Urban areas are the major sources of particulate matter and carbon compounds produced from the burning of fuels in the vehicle. Based on the severity of health impact and the kind of human activities, the National Ambient Air Quality Standards (NAAQS) has set six common criteria pollutants that harm human health, environment or even property damage. The NAAQS was established by the United States Environmental Protection Agency (EPA) under the Clean Air Act which specifies the pollutants that are harmful for the public health and environment [8] [9]. The pollutants specified by NAAQS are Particulate Matters (PM), Ozone (O_3), Nitrogen Dioxide (NO_2), Carbon Monoxide (CO), Sulphur Dioxide (SO_2), and Lead (Pb).

Different countries measure different set of pollutants, for example, India measures 8 major pollutants such as (PM), (O_3), (NO_2), (CO), (SO_2), Ammonia (NH_3), and Benzene

(C_6H_6) (in some places (Pb) instead) [10]. Most other countries measures a subset of these criteria pollutant for example, Canada measures PM, O_3, NO_2, SO_2 and CO [10]. Particulate matters are measured at two levels; 2.5 microns size particles ($PM_{2.5}$) and 10 microns size (PM_{10}), and they are measured in micro-grams per cubic meters ($\mu g/m^3$). CO is measured in parts per million (ppm) and other gases are measured in parts per billion (ppb). These individual measurements are not provided in layman terms and therefore are not helpful in understanding the cumulative impact of the air quality. Taking this into account the government agencies of each country around the globe developed their own indices corresponding to the NAAQS for representing the quality of air. The different indices are Air Quality Health Index (AQHI), Air Quality Index (AQI), Air Pollution Index (API), Pollution Standard Index (PSI), Comprehensive Air Quality Index (CAI), Daily Air Quality Index, Common Air Quality index (CAQI) are few used in different countries. Out of all these indexes the most common are AQI and AQHI which are proposed and used by different countries [10].

India, USA, UK, and many other countries use AQI and Canada, Hong Kong uses AQHI. These metrics are designed by carefully examining those pollutants which are harmful to human health and environment. AQI is a piecewise linear function of the pollutant concentration [11] and is measured using the following formula.

$$AQI = Max\{I_i | i = 1, ..., 8\} \tag{1.1}$$

where I_i is an air quality subindex corresponding each pollutant and it is computed as

$$I_i = \lceil (\frac{I_{high} - I_{low}}{C_{high} - C_{low}}) \rceil \times (C - C_{low}) + I_{low} \tag{1.2}$$

where C is concentration of the i^{th} pollutant. C_{low} and C_{high} are lower and upper concentration breakpoints of C respectively. I_{low} and I_{high}, respectively, are index breakpoints corresponds to C_{low} and C_{high}. The value of AQI varies from 0 to 400+ as shown in the Figure 1.1 and each colour code shows the quality of air in the atmosphere.

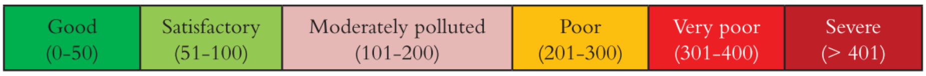


Figure 1.1: Air Quality Index (AQI) [12]

In Canada, the Environmental agency developed AQHI to make the common public aware about the quality of air that surrounds them. It was based on five major pollutants $PM_{2.5}$, O_3, NO_2, SO_2, and CO initially and later the last two pollutants were dropped from the calculation as they were identified to contribute very less in predicting health effects. AQHI is computed using the following formula.

$$AQHI = \lceil (\frac{1000}{10.4}) \times [e^A - 1] + [e^B - 1] + [e^C - 1] \rceil \tag{1.3}$$

where $A = 0.000537\times$ concentration of O_3, $B = 0.000871\times$ concentration of NO_2 and $C = 0.000487\times$ concentration of $PM_{2.5}$. The value of AQHI varies from 1 to 10+ as shown in Figure 1.2

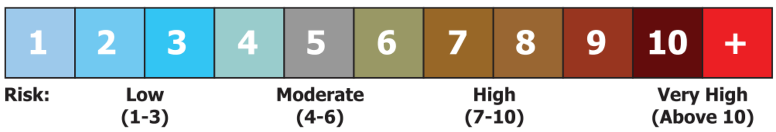


Figure 1.2: Air Quality Health Index (AQHI) [13]

Now the question is which is a better measure, AQI or AQHI? Health index was created with a different objective in mind that is to provide the effect of pollutants to the human health. AQI is based on a single pollutant and hence it is hard to see its direct relationship with health [10]. On the other hand AQHI is based on the exposure-response relationship between the major air pollutant mix and human health [10].

1.2 Impact of Air Pollution

Air pollution has significantly increased after industrialization and urbanization have taken place. The burning of fossil fuels, exhaust from factories and industries, and mining operations are major contributors towards air pollution. Exposure to air pollutants causes premature deaths, cardiovascular disease, stroke, and other respiratory diseases. The state of global air 2017 has discussed the effects of long-term exposure to harmful air pollutants such as particulate matter which contributes to over 4 million premature deaths and is estimated to double by 2050, if the issue remains unattended [4]. Air pollution accounts for highest death rates annually.

Particles with a diameter of less than 10 microns (PM_{10}), and less than 2.5 microns ($PM_{2.5}$) causes the greatest threat to health, as they are capable of penetrating the lungs which leads to cardiovascular and Chronic Obstructive Pulmonary Diseases (COPD) [5] [14]. Exposure to Carbon Monoxide (CO) which is a colorless and odorless gas, results in absorption of the gas into bloodstream and reduces the ability of lungs to transfer oxygen which inturn affects the functionality of vital organs such as brain and the heart [15] [16]. Respiratory issues such as decrease in responsiveness of airways, inflammation in airways and lung infectivity occurs due to exposure of high concentration of ozone ($O3$) [17]. Another pollutant is nitrogen dioxide ($NO2$) which causes illness such as wheezing, coughing, bronchitis and increases severity of flu symptoms.

The exposure of mixed air pollutants leads to health effects like birth defects, development delays in children, skin irritation or even cancer [18]. Apart from the health issues, there is a large effect in environment such as rise in sea-level, global warming that contributes to increase in overall temperature and melting of ice caps, acid rains which leads to land and crop destruction, ozone layer depletion, wildlife destruction and eutrophication which affects on aquatic life [19]. Having discussed the effect of pollution the next section will brief about the history of air pollution and how pollution is measured.

1.3 Background

The quality of the air has declined since the industrialization has taken place in the mid 19th century. Ever since it has affected not just the environment but also human health. Development of heavy industries across Europe and North America used coal as their major source of energy which contributed to black smoke pollution [20] [21]. Coal was not just used in industries but also in houses for heating in winter which made the pollution even worse [22]. These emissions resulted in serious health impacts on residents in urban areas that increased the mortality rate during 19th century.

One such important event in the history of pollution is the great smog of London which killed as many as 12,000 people mostly infants. This was caused due to the combination of cold weather with smoke and lasted for several days [23]. There was a string of similar events reported in New York and England around the same time. All these incidents led to the development of Environmental Protection Agency (EPA) who enforced the need for installing the monitoring stations. The government along with these environmental agencies established acts like clean air act, motor vehicle air pollution act, air pollution control acts for a better quality of air [24]. Apart from that, they took initiative to monitor the air pollution by installing systems which could measure the concentration of pollutants and could give warnings to public as well as industries regarding how polluted the atmosphere is.

1.3.1 Existing Environmental Monitoring station

The government and environmental agencies are making an effort to install monitoring stations for improving the air quality. The EPA monitors the criteria pollutants along with any special pollutant that is dominant in that area. These monitoring stations are fixed in a location and are operated by environmental agencies. The stations are equipped with instruments not just monitoring criteria pollutants but also analyzes other parameters like wind speed, humidity, precipitation. These analytical instruments work by the principle of sampling of the air collected from the atmosphere. There are two main methods for pollutant sampling: 1) passive sampling, and 2) active sampling [25] [26]. These sampling techniques are considered as one of the most significant development for air quality measurement and used widely for monitoring purposes.

In passive sampling the pollutants are collected by physical process such as diffusion through a static air layer or membrane. These pollutants in the air are diffused to adsorb on the sampling media due to difference in the chemical composition of pollutants. The analysis of the pollutant on the sampling media gives the time-averaged contaminant concentration [27]. On the other hand, active sampling works with an air sampling pump which actively pulls the air through a collection device like filter and weighted concentration is calculated. However these static instruments have a major drawback of temporal resolution as they are large in size and needs high maintenance. These instruments are expensive and are financially impractical to expand to more than one station. Table 1.1 gives the average estimated cost for purchasing air quality monitoring equipment produced by US Environment Protection Agency (EPA).

Pollutant/Parameter	Estimated cost
NOx	10,4440 USD
$SO2$	35,000 USD
COx	28,000 USD
Ox	6,600 USD
PM	37,700 USD
FTIR Analyzer	100,000 USD

*FTIR: Fourier Transform Infrared spectroscopy measures multiple gases

Table 1.1: Estimated cost of Reference Instruments [28]

As a result of high cost, only a few monitoring stations are installed for an area. Thus, we can claim that spatial resolution is limited by these conventional monitoring equipments. This made the researchers and scientists to work more on portable sensor networks to understand air pollution in a better way.

1.3.2 Sensor Network for Air Quality Monitoring

There is an emerging trend of using air sensors to understand the quality of air because of its low cost, small size and low power consumption [29]. These sensors are cheaper than reference station and can be replaced easily in case of any damage. The advancement in the Micro-Electro-Mechanical System (MEMS) made it possible to integrate all the functions into a single electronic circuit which makes it more compact. These sensors do not incur deployment cost, infrastructure cost or does not require frequent maintenance which makes the overall expense even less.

Sensors are used in variety of application such as military monitoring purpose, surveillance and tracking, traffic density calculation and road condition and environmental applications [30]. Sensor units can be accommodated on an automobile, building tops or even could be in a wearable form in order to track or measure. This enables high spatial resolution without establishing huge reference station [31]. Sensors measuring different compounds are integrated together to make a sensor node system. Extensive research works have been done with the available sensors in the market.

1.4 Motivation

One of the important components in solving the issue of air pollution is to increase the awareness among public about the current situation and its impact so that they can act on it. The conventional method of monitoring the air quality with the help of a few heavyweight expensive stationary monitoring systems typically installed by the state may not be effective enough for this task. To achieve the goal effectively and without further delay, pollution monitoring must become part of daily activity for everyone. For that, the devices to monitor pollution must be small, portable, inexpensive, and part of a global system. With the technological advancement of low-cost computing, communication, and sensing devices, and the revolution of open source software [32], we believe it is possible to build pervasive air pollution monitoring system with the available sensors from the market and building an open source software. Now the question is how to design such pollution monitoring devices faster and make them accessible to as many people as possible.

Achieving the above stated goal requires a suitable system framework that can help to accelerate the process of the design and implementation of an air pollution monitoring system using the off-the-shelf hardware and building an open source tool for representing the collected data from sensor. Some recent attempts to build a low cost air pollution monitoring system have been done, however, none of them are simple and easily re-producable. This thesis is an attempt to fill that gap by first proposing a simple and comprehensive framework and then demonstrating its feasibility and use by creating our own low cost and easy to use pollution monitoring system that is operational in our lab. We have also added a step of calibration by implementing a web-based tool to ensure that we measure high quality data. Our contribution is a step towards inspiring and motivating not only the public to use the device, but also many amateur electronic hobbyists to buy the hardware locally and download the associated software to build their own pollution monitoring device.

1.5 Research Problem

This thesis focuses on design and development of an air pollution monitoring system using off-the-shelf hardware and build an open source software that focuses on three different types of users with the following objectives in mind:

1. To give an idea on how to integrate the hardware components to a processor and also make an independent software, which can be accessed world-wide.

2. To influence the behaviour of people by including air quality health index which provides a measure of adverse affects on health by pollutants.

3. To encourage public citizen contribute to solve the issue of air pollution and give more understanding to the impact it causes to human health and environment.

4. Improving the quality of data obtained from the system by integrating a calibration tool along with the software.

1.6 Thesis Contribution

There are three major contribution from thesis:

1. Air pollution monitoring system: The system itself which measures the pollutants from the atmosphere. This include the sensors with the processor and the data trasferring module. We believe that this could be a way to show that low cost system could be used for data collection.

2. Air pollution visualization software: Next major contribution is the the software that could be used for data visualization. The main idea is to make the collected data user sensitive and hence we came up with the idea of building the complete tool from the scratch as the already available tools in the market are paid service.

3. Integration of Calibration tool : Macro Analysis Tool (MAT) developed by Environmental Protection Agency (EPA) which was an excel based tool was modified into a web-based tool which provides an access for any user to understand the quality of data obtained.

1.7 Research Question

Some important research questions to be addressed related to the issue of air quality are:

1. What kind pollutants affect the human health most? How to measure them?

2. As the pollution is in the atmosphere, should it not be measured everywhere all the time? If so, what kind of infrastructure is needed to facilitate such a ubiquitous measurement?

3. What all factors should be considered while selecting sensors for measurement of pollutants?

4. Which processor to be used for processing the data collected from sensors?

5. How should the hardware component i.e the sensors and the processor, should be packaged?

6. What is the difference between different metrics which is used to show the effect of pollution?

7. How can the seriousness of pollution be made aware?

8. How can the representation of Air quality metrics be done?

1.8 Research Challenges

The major challenges experienced on creating a complete system are:

1. The integration of different sensors with the processor.

   As each sensor has its own property, there can be issues when connecting all of them together in a single platform. Most of the sensors used for measurement are heat sensor and needs to be powered atleast 12 hours before operating. There is a particular operating temperature range for each sensor connected. As each of the sensor needs different environment for working, all the factors should be incorporated for a complete working environment.

2. Integration of communication module which is WiFi with processor as it requires a different voltage level to drive.

   The data transferring module used here is a WiFi module which needs a separate voltage level of 3.3 Volt when compared to the sensors which are driven with a voltage of

5 Volt each. A voltage more than 3, will result in burning the entire module and hence the WiFi module it needs a separate voltage line.

3. Calibration of each sensors based on the data sheet.

   As I am selecting commodity sensors which says that it is already calibrated but usually it should be done for each environment. The data collected changes for each environment and this should be taken care of. This can be a tough task as for each of the sensors it changes based on the slope of internal resistance and the internal voltage values given in data sheet.

4. Packaging of hardware component.

   To make the whole system portable it should be packed carefully. For certain sensor like dust sensor, it should be exposed to the environment so as to estimate the particulate matter. This packing should be done in such a way that the it does not hinder with the observed value.

5. Transferring the data to a platform or a database using the WiFi module.

6. Checking whether the collected data is accurate.

1.9 Structure of the Thesis

Chapter 2

Literature Survey

One of the matured innovation in the field of electronics was in the development of Wireless Sensor Network (WSN) which paved the way for various research works. The miniaturization of the components has allowed the user to experiment in various field of science such as health care, military applications, traffic control, monitoring, data collection [33] [34]. Out of these urban air quality monitoring has gained a lot of attention as it is one of the major issues faced by the society. With the help of low cost and energy efficient sensor network there are various research work done. This chapter explains the different methods of understanding and detecting air pollution which is classified based on the platform which is used for the measurement [35] [36].

1. Vehicle based sensor network
2. Wearable sensor network
3. Community sensor network
4. Static sensor network

The work done in these different categories will be explained in detail further in this chapter.

2.1 Vehicle based sensor network (VSN)

The compact and low cost sensors are installed and attached to a vehicle like bus or cars in order to achieve a spatially resolved data. There have been an increasing number of mobile vehicles in the urban areas. This was taken as a medium for obtaining air pollution data from the environment in cities.

Vehicular wireless sensor network [37] proposed a system to study *micro-climate monitoring* which means the changes in concentration of a single pollutant measurement ($CO2$ in this case) by mounting the sensor node onto a vehicle. The system is equipped with $CO2$ sensor, Global System for Mobile (GSM) module, GPS receiver and an intra-vehicular network by ZigBee module. The collected data was transferred through GSM short messages to the server and is displayed on google maps for results. Later this same work was further improved in [38] by creating a simulation model and making changes to the existing prototype of the system. The working prototype is now divided into two which consists of a central unit which is inside the vehicle and an external unit to collect the environmental data. The external units sends the data to central unit which reports the data. The work also addressed two network related problem of reducing reporting overhead and reducing communication overhead on the cellular network as sensing short messages incurred charge.

In Mobile Air Quality Monitoring Network (MAQUMON) [39] which is supported by microsoft proposed a prototype in which the sensor node measuring $O3$, CO, and $NO2$ concentration are mounted on cars for monitoring a large area. The sensor node is connected to a microcontroller and is powered by a Li-ion battery whose life is limited for few hours. It is equipped with a GPS module for determining the location and the collected data is transferred to a laptop through a bluetooth module. When the system is in coverage of a Wi-Fi network the collected data is transferred to web server and visualized in sensor map web application. The issue with this system is that it cannot transfer the data immediately for visualization. Another approach for Vehicular based mobile architecture is represented in [40], which is a fine grained approach was provided to get spatially dense data of pollutants. It propose two data framing models one for a public transportation infrastructure and other for a personal sensing device. In public transportation a mobile sensing box (MSB) was installed in buses which contained the microcontroller, sensors for measuring *PM* AND *CO*, a GPS module and a cellular modem. The module was powered through bus batteries and the collected data were transferred to server and visualized using google fusion tables interfaces. In the second framing, the system was installed in a car and connected over a bluetooth to a smartphone. This gave the user aware about the air quality while driving through a specific area.

The wireless sensor deployment in bus service [41] proposed the idea of placing the sensor nodes in public transport buses in order to obtain real time data of pollutants. The WSN here is divided into sensor node and sink node in which the former collects the data from environment and latter aggregates the data from the sensor node and transmits to a server via a long range radio bands. These sink nodes are either installed in a T-junctions or an X-junction where most of the buses cross. Another paper that discusses an effective method to acquire air quality data in urban area is [42] in which the proposed system has the sensing unit kept on a public vehicle which keeps on moving in and around the city. The system has made use of garbage trucks and garbage patrol vehicles. The main focus is on pollutants like *PM*, *CO* and *SO*2 and is also aided with GPS module. To remotely monitor the sensing conditions and to check for maintenance, a control center tool has also been developed. It consists of a map, graph which tracks the route of the vehicles and the sensor data acquired by each vehicle. There was a monitor which was developed along with the system to send the users the collected data. It estimates the amount of pollution inhaled by the user, by the acquiring the user's location from the mobile application and mapping it to the location-sensors value which has been already computed. It also considers the respiratory volume of the user to estimate the pollution inhalation.

A further method of transportation which was taken for understanding the air quality was public bicycle system. The bicycle borne sensor [43, 44] deploys an exhaust gas sensor and *PM* sensor, GPS module and a microprocessor. The system collects data along with the location from the GPS module and is stored in micro SD card. The system is powered with the help pf Lithium polymer batteries. When the subscriber returns the bicycle to dock station the data is transferred to data centre via bluetooth module and is then visualized using 'baidu' heat map.

In [45] to understand air quality they used Unmanned Aerial Vehicle (UAV) as a medium. In this the sensor board which measures six different pollutants that contribute to AQI index is mounted on the flight. These flights are carried out in 30 days with 20 minutes of monitoring. Along with the sensor board, a smart phone is attached along which collects data from air sensor by establishing a bluetooth connection. The collected data is transferred to air quality analysis software that will display the real time monitoring value along with AQI value.

In a nutshell, the possibility of collecting redundant data is high in VSN as the chances of vehicle to get stuck in traffic congestion or for a car to be in a parking lot will collect execcisve data that will reduce the data quality. The efforts to calculate AQI or AQHI is not taken as a priority excpet in [45] which uses UAV. The representation of these indexes will make people aware about the air quality but most of the work in this category have failed to bridge the data gap.

2.2 Wearable sensor network

The individual effect of pollutant on health depends on the extend to which a person is exposed to the polluted environment. The understanding of the health effects could be achieved by observing the exposure-response relationship [46] that will give an idea about the amount of pollutant inhaled by an individual. This could be achieved by using wearable sensor which helps to understand the effects of individual impact when exposed to polluted air. The works done in this section mainly focuses on improving the understanding of the personal health and exposure to air pollution [47]. One such development is Mypart [14] from the university of california, a wrist worn particle sensor that measures particulate matter of 10 microns or less. The design of MyPart is based on traditional laser based photo diode system with improvement in airflow to remove light leakage, integration of structural design and circuitry for ambient visualization, BLE transceiver for low power networking and also mobile application for visualization. Two main issues tackled by MyPart is accuracy and calibration of sensor which no existing consumer sensor has addressed. A mobile application was also developed in addition to the hardware in order to demonstrate the volume of pollutant.

Another related work is 'Eco-mini' [48] is a wearable stand alone device for clinical use that measures Volatile Organic compounds (VOCs), sound level, air quality, temperature and humidity values. This system is based on a low power microcontroller (Atmel Xmega 128k) and consists of a GPS module for position identifying and a bluetooth module for data transfer. They modified the web server which was developed on a simple java script application. Another wearable work is 'citisense' [49] which is attached to a bag stripe is a system which measures air pollutants like $NO2$, CO and $O3$ along with environmental parameters such as temperature, humidity and barometric pressure. The collected data from the sensor is processed by the microcontroller (ATMEGA1284p) and transfers the data using a bluetooth module to a smartphone which does the data storing, analysis and data aggregation. The collected data is then transferred to back end to a web server where the user can get a personalized view of their data. They also developed a citisense android application that runs on the smartphone.

In [50], the research group in US developed an expressive T-shirt called 'WearAir' which indicates the measured volatile organic compound (VOC) through expressive patterns. The T-shirt is designed with a metaphor of a car emitting gases with four vertical arrays of LEDs which shows different frequencies depending on the concentration of VOC gas in the surrounding. When the wearer is exposed to dense VOCs the LED will blink rapidly. The authors of [51] developed a novel system consisting of several sensors that gives a real-time feedback of an individuals exposure dose. This consist of arm sensors, chest sensor or even wrist sensors which measures various pollutant concentration (CO in this case) and transfers to an android or ios application via bluetooth. They also calculates the inhaled dose of pollutant by calculating the volume of air inhaled into a person's lung per minute through an algorithm developed in [52]. The inhaled dose of pollutant was calculated and compared during various activities like jogging, bicycling and driving.

There are also wearable sensor projects initiated in vancover in association with university of british columbia like TZOA [53] that can be clipped to the clothing and measures the air quality. It mainly measures PM values and display in an application. These devices decreases the gap between individual and their understanding about the polluted air. In Newyork a striking project named 'Aircasting' [54] which provides the health and environment data using android 'Aircasting app'. The 'Aircasting' [55] platform includes a palm-sized air quality monitor which measures $PM2.5$, relative humidity and temperature. The outside air is drawn through a sensing chamber and the particles are measured through light scattering method. It also includes a LED wearable apparel named 'Air casting luminisence' [56] that illuminate LEDS according to the obtained sesnor measurement; variying from red for high intensity, then orange, then yellow and finally green for low intensity.

The emergence of such wearables make individuals to be more cautious about their health and encouraged them to stay fit. At the same time the public are not aware about these devices and do not prefer wearing T-shirts or carry a device for understanding the impact of pollution. Another issue to be focused is about the cost and stability of connection due to which the measured data values wont be able to visualize.

2.3 Community sensor sensor

The development of portable sensor devices have paved way for a novel paradigm for monitoring the pollution known as crowdsourcing or participatory sensing. This gives an opportunity for any citizen to collect data and transfer it to a common platform like a webinterface. The collected data from the participants gives a spatiotemporal view of the effect of pollution [57]. In sydney a low cost participatory system is deployed named 'Haze-watch' [58] for monitoring pollution in urban areas. In this, mobile sensor units were attached to vehicles and collected data were transferred using bluetooth to a mobile application which tags its location with date and time information. This data is then send to a cloud-based server which stores data and applies interpolation models [59] to generate spatio-temporal estimates. Then using a web application the geo-referenced data is depicted as a contour map. Intel has developed a prototype named 'Common Sense' [60] which is based on mobile participatory sensing [61] that enables citizen to collect relevant data and involves in the decison making process. The system includes a handheld device that measures a couple of pollutants and uploads the value for visualization over the web using bluetooth or GPRS radios. This work was further tested by deploying on muncipal fleet of street sweepers in the city of San Francisco [62]. Another community-driven sensing developed is 'OpenSense' [63] which focuses towards the utility of data by giving an idea about how the data collected from sensors needs to consumed. They have provided two use-cases first one is smart healthcare which by giving alerts on identifying the pollution induced diseases (like asthma, particle allergies, etc.) and next is urban planning by identifying polluted areas and identifying alternative routes. The system is deployed on mobile vehicles and stationary stations in Swizerland and the collected data is pipelined to a Global Sensor Network (GSN) from where the streamed data is processed and represented. In [64] an outdoor participatory monitor was introduced called 'GasMobile' by connecting low-cost ozone sensor to an android smartphone. The collected data from the sensor is transffered to the phone and from which it is visualized using application as well as webserver.They have also implemented calibration procedures to the low cost sensors and the work claims to have high accuracy when compared to static measurement. The above mentioned research work in 'OpenSense' and 'GasMobile' have made opening for a further participatory sensing research in Swizerland supported by samsung called 'Exposuresense' [65] that monitors user activities like walking, running, etc., from smartphones and understanding their exposure from obtaining datas from the already installed 'OpenSense' and 'GasMobile'. Their main idea here is to make use of the available smartphone for next generation healthcare. The growth of Micro-Electro-Mechanical System (MEMS) and Wireless Sensor Network (WSN) have made difference in the way how data is collected and understood from the physical world. 'G-Sense' [66], for Gloabl-Sense, is a work initated from the university of Florida in which they combine features of sensing platform applications like Location-Based Services (LBS) for tracking and location identification, Participatory Sensing (PS) for determining pollution index and other environment data, and Human-Centric Sensing (HCS) for heath related datas for specific group of users. The sensors collects the data and sents to a first-level integrator where all the data gets collected and from there using a data transport network it is transferred to the server that stores and performs data processing. It is from the server where the data visualization takes place which reports the data. Later there was another work which is considered to be the subset of G-Sense and named as 'P-Sense' [67] or Pollution-Sense. The architecture of this system is based on G-Sense in which external sensors are integrated into an Arduino development board. In this the data collection is based on Participatory Sensing (PS) and the goal here is to provide government official, doctors, and community developers with data so as to get a deeper understanding. They have also pointed out the research oriented challenges that needs to be addressed when building a community networked system like security, privacy, data visualization and working towards to achieve them.

The work discussed in this category seems to be more promising but at the same time the quality of data obtained, getting public involvement for data collection and privacy issues [35] are few challenges faced under this and researchers are trying to work towards it. The cost of maintenance in such community network in case of any damage is a crucial factor.

2.4 Static sensor network

In this, the system is kept at a fixed location like traffic lights, street lights or any planned areas [36] which collects the pollutant values and transfers it to a visualization platform where the users can view it instantly. These systems are inexpensive and can be easily replicated or replaced. The system can be used for measuring either indoor or outdoor pollutants. Pollution in urban areas is increasing rapidly and due to which the number of people suffering chronic illness, permanent disability or even death are also increasing [68]. The already existing station based environmental monitoring system are complex and costly hence there is a need to develop portable and low cost environmental monitoring system. The Integrated Environmental Monitoring System (IEMS) [68] integrates different environmental detection sensors in a single system and data from this system is used for processing and visualization. IEMS consist of Integrated Environmental Monitoring Device (IEMD) which consists of Microcontroller units, sensors, wireless communication module. They developed Handheld Remote Control Panel(RCP) for the system which is an android application that acts as an interface for the device control and handles the data exchange between IEMD and Web Server. Finally the web server that provides a real time data visualization and data analysis. These systems were placed on bus stops, bridges and even in the construction sites. In mauritius a research team developed a Wireless Sensor Network Air Pollution Monitoring System (WAPMS) [69] that designed a data regression algorithm named Recursive Converging Quartiles (RCQ) to remove duplicate data and then calculate AQI values. The array of sensor node collects the pollutant data and transfers it to cluster heads where the RCQ is applied to improve the efficiency and alleviate congestion problem. From the cluster head the data is send to server and represented using line graphs for each areas.

Another related work is 'AirSense' [70] which is an excellent approach to asses the indoor air quality. The author tries to introduce the idea of indoor air quality to the society by proposing a system which measures indoor pollutants. The system works through electronic sensors which are coupled to an Arduino (processing unit). The system not only extracts the data, but also provides its users with very effective visualization and analysis of the data. The researchers of the paper have done an excellent work on developing this robust system that would sense the pollution and provide education and awareness among the users. This system has made use of some machine learning algorithms to predict the pollution sources and forecast their behavior so as to provide intelligence to the system. The system has also got a smart phone application which gives the users a very effective interface for visualizations and understanding the data. In [34] a different group of researchers from China developed the system 'Air-Sense' to monitor and predict the quality of air based on ZigBee network. The system uses 4 different type of sensor, respectievly are humidity, temperature, $PM2.5$, TVOC (includes the general organic gases) and a ZigBee transreciever for communication with network nodes. The prototype is tested for different areas in the house. It collects data in real time and using Bayesian mathematical statistics it predicts how accurate is the collected value to standard value predicted by WHO.

A static work which focuses on indoor air quality is demonstarted in [71] in which the main focus is to understand the pollution in office environment where the pollution is triggered from the electronic devices and machines. In this the primary pollutant measured is ozone which is mainly emmited from a photocopier machines. The system is designed into different nodes where the sensing node contains the Arduino which process which collects the data from the ozone sensor. The measured data is transferred through a bluetooth link to a gateway node from where the data is formated as IP packets and forwarded over the Ethernet network to processing node. The data is saved in database and using 2D graph the concentration gets visualized. A research group from Harvard university developed a wireless networking testbed called as 'CitySense' [72] in which multiple environmnetal sensors are attached to street lights. These sensors were deployed in Cambridge and data was uploaded to server using mesh networking like RoofNet [73], TFA [74] and CUWin [75]. Using a Web-based interface the data can be pulled from the server and made available to end users. The main feature of this research work is that the sensor nodes are powered from street lights and there is no constraint about long battery life. Liu J.H. et al [76] developed a micro-scaled air quality monitoring system for understanding the CO emmision from vehicles by integrating the sensor nodes with a gateway. The collected data from sensor is transferred to gateway using ZigBee communication link and from here meteorogical data and collected sensor data is forwarded to a central system through short message service via GSM. This centralised control system is supervised by the LabVIEW [77] programming which helps in storing the data into MySQL database. They deployed the system on the main roads of Taipei city and obtained accurate values of pollutant concentration.

Our research work falls under static sensor network in which we have tried to integrate a system that measures all the major pollutant in the city of Prince George and providing AQHI and AQI values. Unlike the other system mentioned in the literature our main focus is to give a user specific data for better understanding of pollution.

2.5 Summary

In this chapter the work done in sensor network for understanding the air quality is categorized into four. We went through the system which is attched to a vehicle for understanding the pollutant concentration and gets categorized as vehicular sensor network. This category provides great mobility but at the same time accumalation of redundant data is high. Next category we mentioned is the sensors which could be worn or attached to a person. This gives a better understanding of individual health effects of pollutant and also the amount of pollutant inhaled. The work under this category has not gained a lot of attention as it demands the individual to carry the device. Participatory sensing is the next category in which the citizens performs the collection of data and it gets transferred to a common platform. Although the work done in this is more promising it faces challenges like privacy, data quality and maintenance cost. The final category is in which the system is placed at a planned area and called as static sensor network. Our research work falls under this section and we focus not just on collecting pollutant value but also making an effective visualization to reduce the data gap for users.

Chapter 3

Design and Implementation of Air pollution Monitoring System

The progress in technology has made it easy for researchers to explore more in different areas. One such development was is in the area of sensor technology which completely changed the outlook of different application level problems. In this chapter, I share my hands-on experience in the development, design, integration, and operation of the air pollution system using commodity sensor. Earlier, the approach for understanding air pollution used complex and stationary equipment which collects data and used these data for analyzing, but things have changed after the low cost, easy to use, portable sensors came in markets [78].

3.1 Design Goals

There are many factors which need to be considered for the development of a simple yet reliable system. In this section, we have mentioned the factors which should be considered for an effective air pollution monitoring system.

1. Sensor Identification

The very first task is to figure out which all sensors need to be included for the completion of the system. There are sensors available in the market for the measurement of almost all types of gases in the atmosphere. It should be very clear that which all gases need to be measured and this definitely changes from region to region as in certain places the concentration of a particular gas is more. Having said that, there will be a certain set of gases which must be included for measurement regardless of the region.

2. Communication Module

   As the system is completely based on wireless sensors the selection of data transmission is another crucial factor. The communication between the server and the sensors should be taken into consideration. The collected data from the sensors should be transferred over a database or to the sever. For that the type of communication module can be either Wi-Fi or bluetooth module.

3. Reliability

   The success of the system depends upon how much accurate the data is. The value which we obtain from the sensor should make sense to the audience. There will be a lot of noise coming with the collection of data, the sensor should have the ability to remove the noise data or it should allow the programmer to make changes or apply certain algorithm so that the data sets will be refined.

4. Easy Integration

   The integration of sensors with the processor is one important factor that needs to be kept in mind. Some sensors can be easily integrated with any processor but others needs driver codes to be written in order to work with the processor.

5. Printed Circuit Board

   The final system should be build on a printed circuit board as it is more dependable. Circuit build on basic breadboard might even come out as it is not permanently fixed and this will cause frequent breakdown. Its always easy to work on breadboard but that will be useful only for the initial set up. The system should be transformed to PCB.

6. Maintenance

   In case of any sensor damage it should be easily replaceable which means the complete system should be a plug and play type model. On building up such a model like that will help in debugging the problems caused by sensors if any. It should also be considered that the sensors selected for the system should be easily available in market so that it can be replaced if needed.

7. Easy Replication The idea behind creating such a system is that it can be replicated by anyone without even knowing the dept knowledge. The system should be designed in such a way that it should use the most available sensors and processors in the market. The programming part of the sensors to processor will be easy if the selection of processor is simple. This could definitely bring down a lot of work done at the hardware level.

8. Low Cost

   Within the available sensors in the market one could find sensors ranging from a very low price to costliest of all. There was a budget set for the the complete system and finding the right sensors with the affordable cost is one crucial factor.

3.2 Targeted Pollutants

Our surrounding is filled with various gases, these gases will become harmful if the concentration of it increases to an undesired level. On the development of a air pollution system measurement of all the gases in the atmosphere is not necessary as the collected data from all the sensor will make no sense to the public. Our main idea here is to make the general people aware about the dominant gases and the extend of health hazard caused by these gases. This can be identified through different indexes know as Air Quality Health Index(AQHI) which is a scale from one to ten developed by health and environmental professionals [79] and Air Quality index (AQI)which gives the level of air quality status in an area [80].

The development of such indexes by the scientists will give the general public more idea of the pollution. The main gases to be included for the measurement for the indexes are $PM_{2.5}, O_3, NO_2$, and CO along with temperature and humidity sensor for awareness. These gases are mainly caused due to industrialization, urbanization and motorization [81]. Industrial and vehicles release greenhouse emissions which are largely responsible for air pollution [82]. The sensors thus can be limited to five which will also make the system compact.

3.3 System Architecture

In this section we describe the design for air pollution monitoring system which can be categorized into system hardware architecture and system software architecture. The system is designed in such a way that it collects the data through the sensors, performs specific mathematical equation on the collected data and calculate the indexes which then transfers these data to the software where it is visualized.

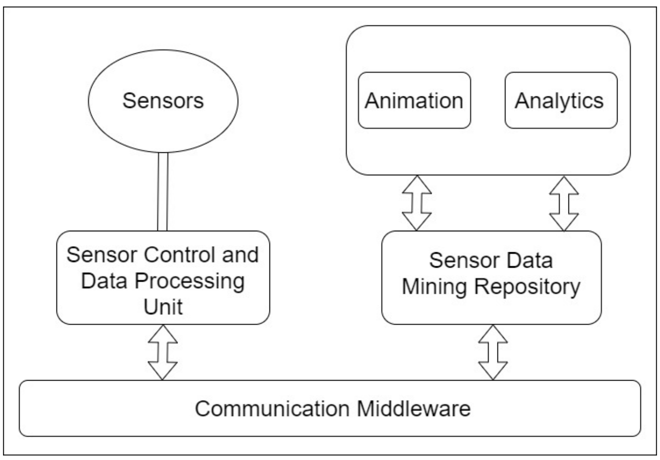


Figure 3.1: System Architecture

The hardware section includes multiple sensors, processor and also on the wireless communication module for transmitting and receiving signals. Second, we will discuss the air pollution software which mainly focuses on the visualization part. The overview of the system is as shown in the figure 3.1 and each part along with the sensor specification, implementation, design will be discussed further in the section.

3.4 Hardware Architecture

3.4.1 Sensor Control and Data Processing Unit

This unit is the core for the pollution monitoring as it is where all the other modules are connected including the communication middleware. This module does the following function:

1. Control the sensors in collecting data.

2. It filters and processes the collected data and forward to the software.

3. Provide the necessary voltage for all the hardware connected to it.

For simplicity and ease of programming we have selected one of the most popular processor in market, Arduino Ethernet board which has ATmega328 microcontroller as shown in Fig.3.2.

Arduino is an open source physical computing platform that is divided into two parts, one is the hardware which is the board itself in which the external components are added to and other is the software which is the development environment for the processing language. It is very easy and simple to use the board with any external devices such as sensors or actuators and is widely used by researchers. There are different features which makes arduino popular and can be listed as [83]:

1. Arduino is multi-platform and can be used with windows,mac and linux.

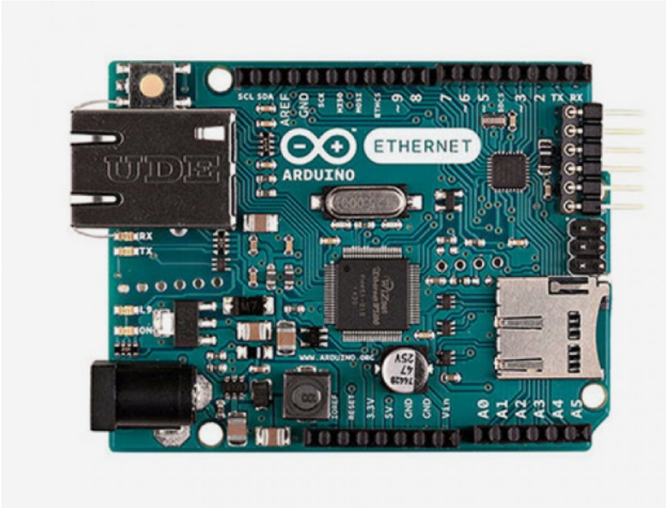


Figure 3.2: Arduino Ethernet Board

2. The programming is done via USB cable and not serial port and is useful as modern computers don't have serial port.

3. As it is open source hardware and software all the necessity for an external hardware to be worked on like circuit diagram, the code can be easily downloaded.

4. There is an Integrated Development Environment (IDE) which can be used as an interface for talking with the hardware and is very simple.

5. There is an active arduino forum in which many researchers or developers who are working on projects contribute their ideas and will help in trouble shooting.

6. The cost of the hardware is very cheap and will come under 60 CAD and is easily affordable.

The board can be powered either by using a FTDI cable/USB serial connector, external power supply or using an optical Power Over ethernet module (PoE). The ATmega328 has 2 KB of SRAM and 1 KB of EEPROM. The detailed specification [84] of the board is given below in the table 3.1.

Having all these specification gives the freedom for researchers to explore with different electronic devices easily. The board is programmed with the help of arduino programming

Description	specification
Microcontroller	ATmega328P
Operating Voltage	5V
Input Voltage Plug (recommended)	7-12V
Input Voltage Plug (limits)	6-20V
Input Voltage PoE (limits)	36-57V
Digital I/O Pins	14 (of which 4 provide PWM output)
Analog Input Pins	6
Flash Memory	32 KB (ATmega328P) of which 0.5 KB used by bootloader
SRAM	2 KB (ATmega328P)
EEPROM	1 KB (ATmega328P)
Clock Speed	16 MHz

Table 3.1: Technical specification

language which is very close to embedded C language and it is done on Arduino development environment.

3.4.2 Sensor

Sensor networks are instruments useful to detect the conditions in remote places in the physical world in environmental monitoring applications such as pollution monitoring, transportation management, intrusion detection and many more [85]. With constant development in electronics industry, it is possible to collect data remotely and collected data can be transferred to the required platform at a short period of time. There are different sensors available in market which can measure the pollutants and display the value, but the idea here is to select the one which is of low cost and also gives the most accurate values.

Having said that there is a variety of options available for sensors based on the way it measures the pollutants. One such category is Metal Oxide Semiconductor (MOS) gas sensor also known as semiconductor gas sensor, which is used to detect the concentration of any hazardous gases in the atmosphere by changing its resistance. The most popular series available in market for this category is MQ-XX sensors which is popular for its wide detecting scope, long life, stability, high sensitivity, fast response and also simple drive circuit [86].The sensing material is made up either from Aluminum Oxide (Al_2O_3) or Tungsten trioxide ($WO3$)based ceramic and has a coating of Tin Oxide SnO_2 that acts as the sensing material for the desired gas. The sensing element is heated through Platinum wires which is connected to leads made up of Nickel-Chromium, well known conductive alloy. The gas to be detected has a specific temperature at which it gets ionized and the task of the sensor is to work at that temperature. Once the gas gets ionized it gets absorbed by the sensing material which changes the resistance and inturn changes the voltage across the sensor and can be read by the microcontroller [87]. The voltage value along with reference voltage and other resistor's resistance is used to find the resistance of sensor. Once the resistance of the sensor is known then by using the sensitivity curve the concentration could be found out.

Another popularly used MOS sensor which we explored was MICS which are MEMS based whose mode of operation is similar to the above said sensor as both of them are metal oxide. Here, oxidizing gas or the pollutant gas add to the insulative oxygen species causing the resistance to increase [88]. Other than MOS sensors, we also took a look at optical sensor which are spectroscopic devices which uses light scattering principle to find the concentration of pollutants. These sensors are known for its detecting capability of particulate matter of different sizes and is one of the recent development in the field of air quality monitoring. Highly responsive, reliable and long life are the main highlights of this sensor.

Selected Sensor

After understanding the wide range of sensor options in the market the sensors were selected based on their performance, availability, ease of integration, and cost. The selected sensors were:

1. MQ-2 Sensor: This is a a semiconductor gas sensor which has an electrochemical sensor which detects multiple gases such as Carbon monoxide, LPG, methane, and other combustible steam. The sensor is connected in series with a variable resistor to form a voltage divider circuit and the variable resistor is used to change the sensitivity. The sensitivity curve for the sensor is shown for different types of gases.

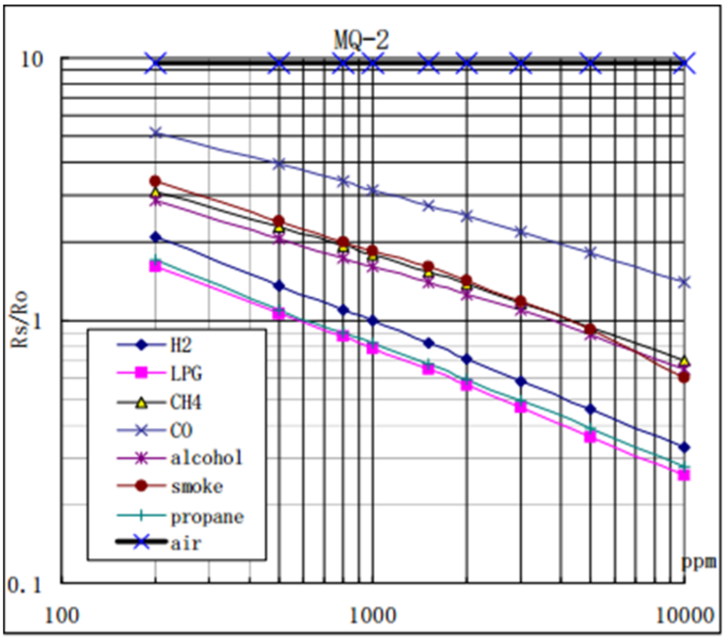


Figure 3.3: Sensitivity characteristics curve [86]

where R_o is the sensor resistance in clean air and R_s is the sensor resistance when exposed to gas.

From the curve in Fig.3.3, the voltage across the sensor is found out depending on which gas one wants to detect and there after using this voltage value the concentration of pollutant is calculated. The range of detection of gas is from 100ppm to 10,000 ppm and has a high sensitivity and fast response time. The sensor is small and portable and provides integration in famous MCU platforms like arduino, raspberry pi etc.

2. MQ 131: Another MQ-XX series sensor that we used for the system is for ozone. The working of the sensor is similar to MQ-2 sensor. It decreases the resistance when exposed to ozone and becomes more conductive when exposed to large concentration of the gas. This can be used to measure the concentration of ozone in air. The detecting concentration scope is from 10ppb to 2 ppm of ozone and also has a fast response time and long life.

3. MICS-2714: This is a robust MEMS sensor used for the detection of Nitrogen dioxide (NO_2). The detection range is from 0.05 to 10 ppm and has a response time of 10 seconds. The sensor is comparatively small and of low cost.

4. PPD42NS Particulate sensor: The sensor detects the particulate matter through light scattering mechanism and consists of infrared LED positioned in forward angle to a photo diode. As soon as there is a variation in light density, the photo diode detects this and changes the current from the diode [89]. The circuit generates a measurable signal known as Low Pulse Occupancy (LPO) which is proportional to PM concentration [90]. This sensor can measure both PM2.5 and PM10 concentration.

5. DHT11: This a very low cost sensor available in market for temperature and humidity measurement and has a calibrated digital output. The measurement range for temperature is from 0 degree to 50 degree Celsius. The device can be integrated to almost all platforms of Microcontroller and is considered to be the best choice for many applications.

3.4.3 Communication Middleware

The collected data needs to be transferred to the software so that user can understand and interpret the data from the sensor. Choosing the communication middleware was a tough task as we need a stable system that will constantly send data over the network. In the beginning we used an ESP module for this purpose. The WiFi module, ESP8266, which is a highly integrated SOC that meets the requirement of user demand of efficient power usage, compact design and reliable performance [91]. The ESP module can be connected to a processor or it can also be flashed on its own as the module itself is a MCU unit. Unlike the other sensor modules connected to the processor which needs a 5V power supply for its working, this module needs a 3.3V for its power up. The ESP module comes along with installed firmware from AI-thinker and it can be communicated with AT commands. On typing the AT command in the serial monitor the output would come as ’OK’ if there is a successful connection.

This firmware can be replaced with user’s own code which gives the power of flexibility of connecting with any IOT platform. The arduino platform supports the programming of

ESP module which makes the integration much easier. The code can be written in the Arduino IDE and can be flashed into the ESP board by connecting it separately. However this module did not provide a reliable connection to the network and after few tests we had to replace it to another module. We choose a WEMOS[INSERT ABOUT WEMOS]

3.4.4 System Overview

This section will show about the integration of the above mentioned sensors with the processor. After selecting the sensors and processor for the air pollution data measurement the next task was to put together so as to make it as a working system. This step was done one at a time so that troubleshooting will be easy. The complete circuit diagram of how the components are integrated to the Microcontroller is as shown is the figure3.4. Each of the

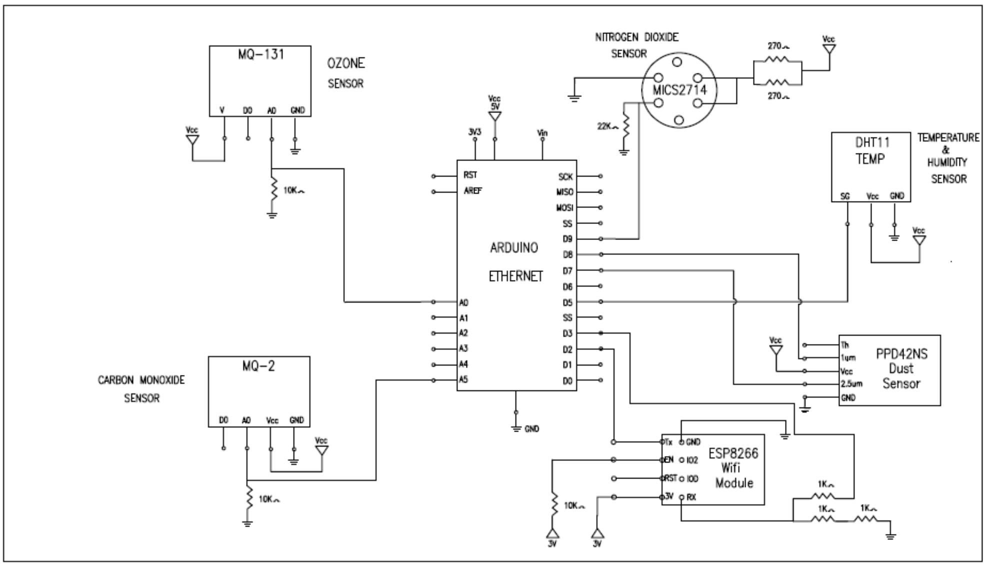


Figure 3.4: Circuit daigram of air pollution system

sensors were carefully arranged so that it reduces the spatial interferences. As majority of the sensors are heat sensors and discharges heat it affects the working of the sensor placed near to each other. The figure 3.5 shows the initial circuit implementation that was done in the breadboard with wires. Building the circuit in breadboard gave the flexibility of working with different arrangements.

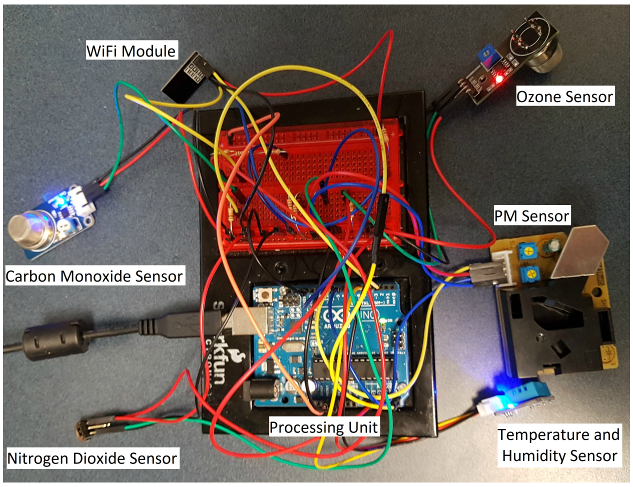


Figure 3.5: Prototype Implementation

After building the circuit in breadboard, it was migrated to a Printed Circuit Board (PCB) developed by Flemings Solution Ltd, India. The design of the board is as shown in the figure 3.6. The PCB was designed in such a way that each sensors and the microcontroller itself was plug-in model to board.

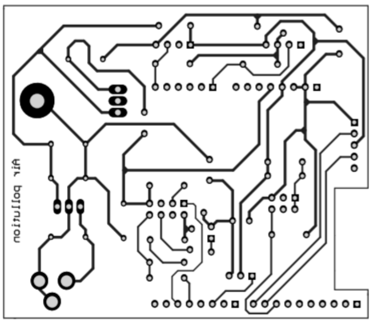


Figure 3.6: Printed Circuit Board

This was done so that in case of any malfunctioning of any of the sensors or processor, it could be replaced easily. The main advantage of using a PCB design is that all the components are fixed and there is no wiring at all. This produces a zero error circuit and there wont be any short circuiting.

3.5 Software Architecture : Customizable Layered Visualization (CLV)

The hardware is the one responsible for the data collection but making these data available to user is done by an effective visualizing software. The main objective kept when developing the system was the data collected should be accessed from anywhere and thats why IoT platforms were selected. We developed a customizable layered visualization of data that potentially involve different stakeholders. We believe an effective visualization of concerned data is critically important for effectively combating major issues such as mentioned above. The software is driven by the following four key elements:

- Stakeholder specific visualization.
- Author-guided and user-driven interactiveness.
- Hierarchical approach to visualization.
- Providing alternate options.

3.5.1 Framework

"Seeing is believing because seeing is seduction" [92].

The pollution monitoring stations collect huge amount of data of pollutants and there are various ways to make people aware about these data. There can be different groups of users who will have access to these data which can be categorized as common people, educationist or the researchers who will be analyzing on the data, and the policy or law makers who will take necessary actions.

Accurately identifying all the stakeholders of air pollution is challenging. In [93], stakeholders in environmental risk decisions are identified into four categories of stakeholders as risk losers, risk gainers, risk perpetrators, risk managers. Risk losers in [93] are those who may be adversely affected by environmental risk decision, in terms of health or economic values and risk gainers are those who may be favorably affected by an environmental risk decision, typically through economic gains. The people who create the risk are the perpetrators and those who takes the action against the risk are the risk managers. In a similar way, here we have put forward an approach of visualization based on the different stakeholders, in which we have classified into three categories: risk losers, risk analyzers and risk managers.

Risk losers are those who are prone to pollution which include the common people or the society as a whole. Risk analyzers are those who identify the cause and increase in pollution and suggests pollution control methods. The next category is the one who implements laws and regulation by looking at the trends of pollution for a certain period of time, which are categorized as risk managers. We believe this categorization applies to air pollution too, and therefore based on this categorization we offer our data visualization strategy.

The visualization at first go should be very simple and be able to drive the user to which details they have to go through. It should not be representing all the data in a single screen which will be annoying to user and will end up loosing the interest of different stake holders. There should be options given to users so that the detailed data will be displayed. The user should be able to choose what data they need to view. If this approach is included in visualization, the data that different stakeholders are looking for will be passed onto them in a very efficient manner.

3.5.2 Implementation of CLV of Air Pollution Data

The complete software implementation of visualization was done with the combination of Node.js and python.The server used for storing the data from the system is Thingspeak [94]. The collected data from the system is stored in the server of Thingspeak server [94] which is a visualization paid service that was used for initial testing. According to the classification of stakeholders, each groups are looking for different kind of data. The data can be represented in the form of huge spreadsheets or documents. There are different methods for visualization like heat-maps, bubble plots, box plots, bar graphs or a plane graphs. The most common method of displaying the data is graphical method, which can show the data corresponding to a certain time or day. We believe that instead of visualizing the entire data through graphs or plots, it would be much easier for the public to interpret the data if it is a single instantaneous value. The representation implemented for the first stakeholder which is the layman is as in figure 3.7, which shows a single value of the pollutant which is all the data that as a citizen will be looking for.

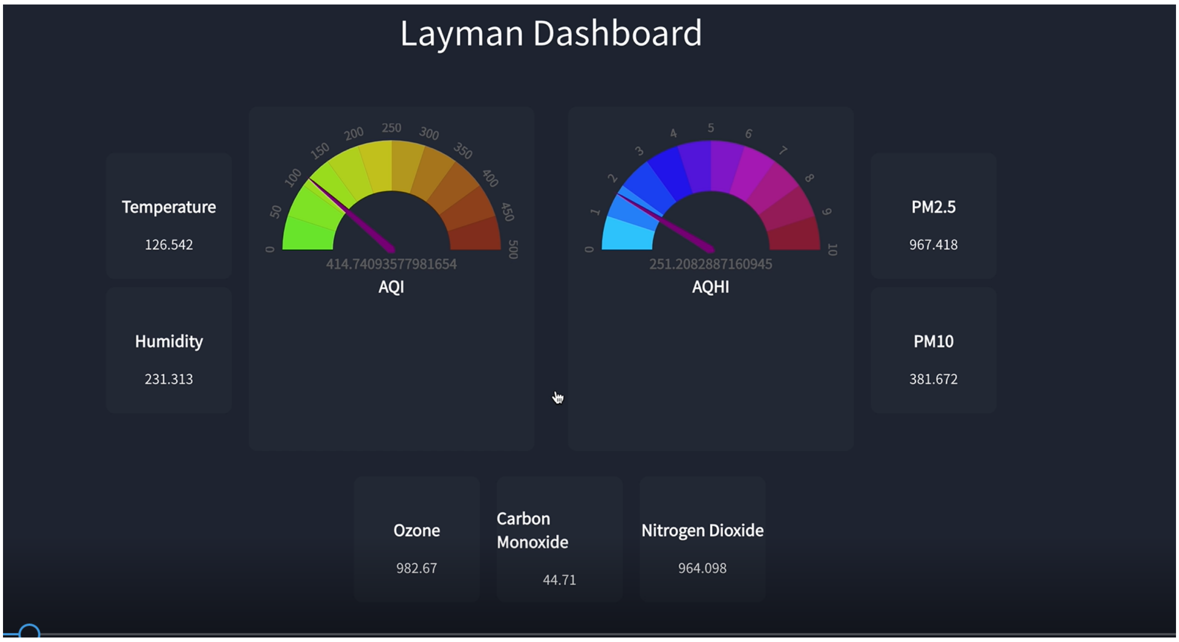


Figure 3.7: Lay man view

The implemented representation is very easy to understand the data and it shows the concentration that each pollutant adds to atmosphere. The idea of this kind of visualization is that the user in the first category that is the common lay men, simply needs to read the data from display. The next implentation is done for the datascientist or researchers in figure 3.8 who need to get a detailed view of the collected data and time. The ideal way to represent detailed information is through a line graphs as it shows the value obtained for a given time.

The value of pollutant could be seen when the pointer is hovered over the graph. These graphs gets updated everytime when the data gets collected at the system as it is important for researchers to go through every collected data. The last categorization of the user is the

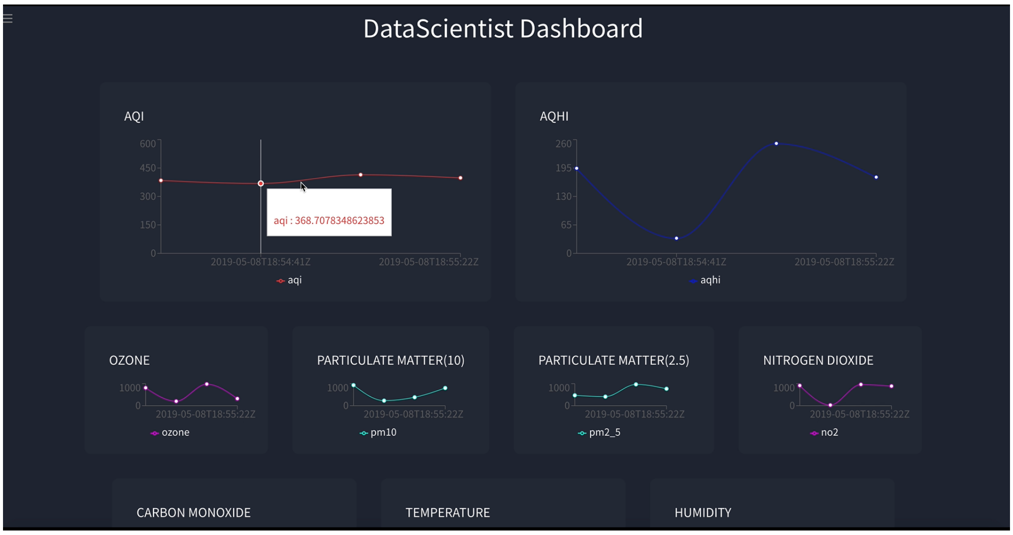


Figure 3.8: Data Scientist view

policy makers who implements action against change in the concentration of pollutant. They are the once who enforce laws by looking at the pollutant data and hence they are the ones who need to see a combined view of all the pollutant as shown in figure 3.9 to understand the trend.

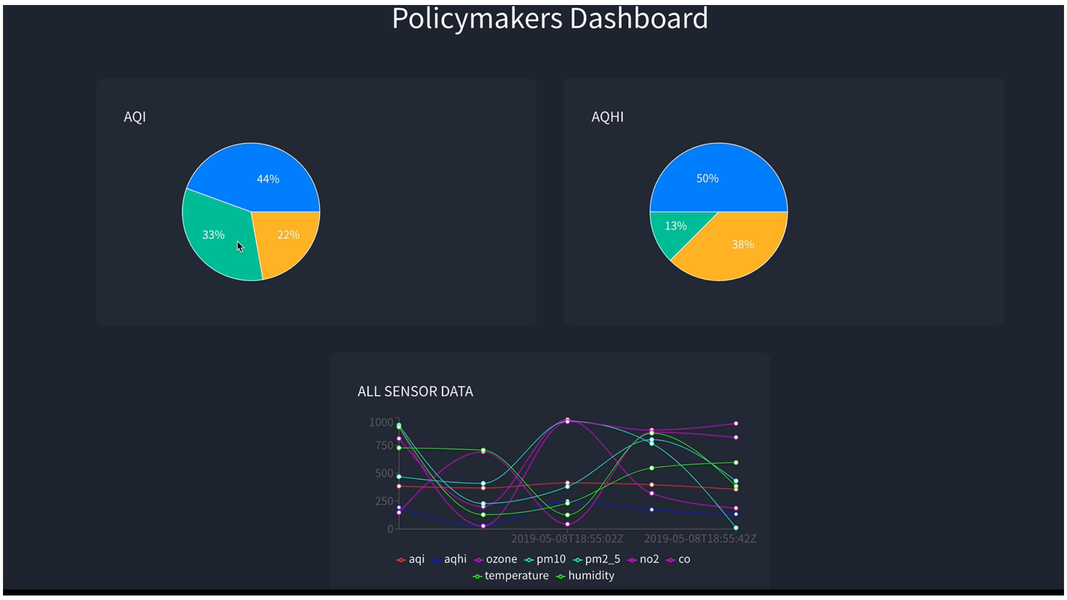


Figure 3.9: Officials view

This can be viewed through a pie-chart which shows the percentage of each pollutant that contributes to the indexes. There is also a combined line graph that shows the concentration of different pollutant data. By looking at the pie-charts and the combined graphs it would be easy to understand the trends and also could find the pollutant which contributes more to the developed indexes.

The idea of such a hierarchial level of visualization is that it changes according to the people who are focusing on these data. Albert Cairo, a renowned information designer claims that the fundamental goal of data visualization should be to make people better informed [92] [95], similarly we also believe that a well communicated data is important for visualization.

Chapter 4

Calibration Of Low Cost Sensor

With the development of sensor technology for air pollution it has attracted a majority of researchers as well as common people to explore and understand more about the pollutants and its effects. This has given freedom for one to set up their own monitoring system at residences, office or even at schools. The problem with this is to identify how accurate the data collected from these commodity sensors to the reference monitoring system. The reference system are the pollution monitoring devices that are developed to a clearly defined standard for a specific criteria pollutant [s] and has completed a rigorous testing and analysis protocol [96]. If the system is giving values which is way too different from the reference value then it brings down the advantages of this technology. This issue can be dealt through calibration which will reduce the uncertainty in data and makes the output more accurate.

Calibration can be defined as the act of evaluating and adjusting the precision and accuracy of measurement equipment [97]. If the measured output from the sensor is not equals to the actual output then it shows that there is a need for calibration. Usually all the electronic instruments are calibrated according to a particular conditions in the laboratory and acquires a certification of calibration before its sold out in the markets. Even then the measured value does not reach accuracy as the condition or the environment where it was calibrated changes that leaves the user with some raw values that gives no information. This issue was taken up and explored by a few researchers in the US Environmental Protection Agency(EPA) and suggested with three 'Straw-Man Approach' to improve the usability of such data and presented in the Air sensor Workshop [98].

The first approach was by a signal-based calibration technique which requires the data from the remote stations, which is the reference station, to be broadcasted to the local station where the sensor is located and will receive this data and performs a single point calibration of the response. This approach would have been easy if the sensor was already equipped with the data collection and would process automatic calibration.

The next option for calibrating is called the direct sensor calibration that involves placing the sensor in a chamber in which a known concentration of pollutant is set and response is observed. As the concentration of the pollutant is already known the output curve can be compared with it and calibrated accordingly. This is the most common method used for calibration as is often called as laboratory evaluation. Another way of approaching direct calibration is by inspecting the pre-defined response given by the manufacturer and checking how accurate the sensor is to the given concentration. In either case the calibration requires equipment and skills to give accurate concentration value.

The last option is by secondary data normalization in which the concentration values of the pollutant from the low cost sensor is normalized in accordance with the federal reference method (FRM) or federal equivalent method (FEM) analyzers. This approach is cost effective when compared to the other techniques and is less complex. This could be achieved by the use of a linear mathematical equation model that will convert the non-calibrated data into data of an acceptable form. The linear relationship equation $y = mx + c$ where m is the slope and c is the intercept of the sensor raw data is compared with analyzer data and a relationship pattern is found out from this. The drawback of this is that the sensor data does not always give a linear response and thus not applicable for all the curves.

Even though these 'Straw-Man Approach' was defined well it was still hard to implement it practically. This eventually led to the development of the 'Air Sensor Toolbox' by the Environmental Protection Agency (EPA) [99] which introduced a tool called 'Macro Analysis Tool' (MAT) for performing comparisons of air sensor data with reference data and interpreting the results [100].

4.1 Calibration Procedure

The Air Sensor Toolbox provide guidelines for researchers, citizen scientists and developers with working of low cost sensor and its calibration to give more insight about air quality [101]. The Air Sensor guidebook [100] provides a three step procedure for calibrating a sensor:

1. Comparing the data from the low cost sensor with a reference instrument.

   This means the data collected from the sensor system should be compared with the data from an already calibrated system which is placed by the local authority. This type of comparison is called as 'collocation' and in order to collocate, first find out where the reference system is placed and get access to the data from these system. After getting access to these system place the sensor system at the same level as the standard reference. Once this is done the data collected from both the system can be downloaded in the desired format.

2. Creating a calibration curve with the help of Macro Analysis Tool (MAT).

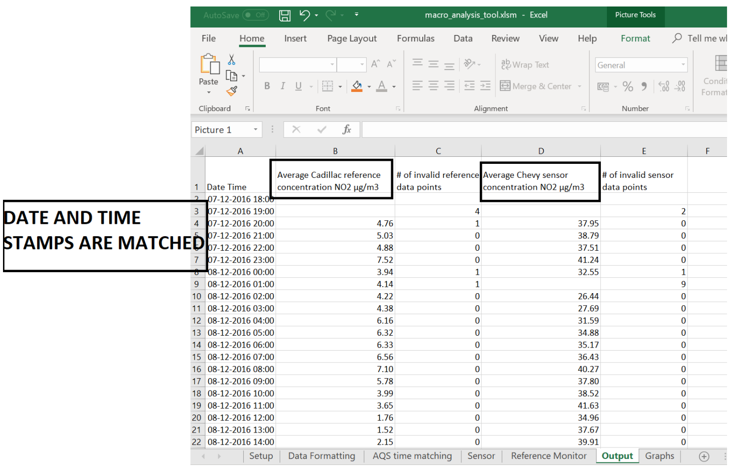
   The relationship between the response of an analytical instrument to the concentration or amount of an analyte introduced into a known instrument is referred to as the "calibration curve" [102]. The MAT tool by EPA is used for drawing the calibration curve and will give an error function as output. This error function is an equation which can be used for calibrating the sensor.

3. Repeating the calibration periodically.

   This procedure of calibration should be done periodically as the performance of instruments changes often. There will be changes in the equation and graph each time and this should be noted so as to get accurate value.

4.2 Macro Analysis Tool -MAT

The development of MAT tool gives a huge reliefs for the researchers in calibrating sensor data. This is an Excel based user-friendly macro tool that compares sensor data with reference data [103] even if measurements weren't recorded at precisely the same time, or were collected at different time intervals, such as 1-minute versus 5-minute intervals [104]. The user needs to insert the data from the sensor into the sensor page of the tool and the values from the reference station is inserted into the reference page of the tool. After these values are inserted the details of pollutant, time interval, measurement units and data completeness (amount of usable data obtained) needs to be added in the set up page of the MAT. After all the required fields are filled the output page as shown in the figure 4.1 shows the two sets of data being compared as per the control panel setting made in the tool [103]. This output page basically gives the statistical comparison of two different data sets. From this page it could be seen that the date and time stamps are averaged to a single value for both the system. There is another column in the sheet which is the invalid data points that gives the total number of non validated data points during the observed time. This non validated data points occurs due to either the fault by instrument or it could be the fault in reading by MAT or it could be even due to unacceptable range specified by the user.



	A	B	C	D	E
1	Date Time	Average Cadillac reference concentration NO2 µg/m3	# of invalid reference data points	Average Chevy sensor concentration NO2 µg/m3	# of invalid sensor data points
2	07-12-2016 18:00				
3	07-12-2016 19:00		4		2
4	07-12-2016 20:00	4.76	1	37.95	0
5	07-12-2016 21:00	5.03	0	38.79	0
6	07-12-2016 22:00	4.88	0	37.51	0
7	07-12-2016 23:00	7.52	0	41.24	0
8	08-12-2016 00:00	3.94	1	32.55	1
9	08-12-2016 01:00	4.14	1		9
10	08-12-2016 02:00	4.22	0	26.44	0
11	08-12-2016 03:00	4.38	0	27.69	0
12	08-12-2016 04:00	6.16	0	31.59	0
13	08-12-2016 05:00	6.32	0	34.88	0
14	08-12-2016 06:00	6.33	0	35.17	0
15	08-12-2016 07:00	6.56	0	36.43	0
16	08-12-2016 08:00	7.10	0	40.27	0
17	08-12-2016 09:00	5.78	0	37.80	0
18	08-12-2016 10:00	3.99	0	38.52	0
19	08-12-2016 11:00	3.65	0	41.63	0
20	08-12-2016 12:00	1.76	0	34.96	0
21	08-12-2016 13:00	1.52	0	37.67	0
22	08-12-2016 14:00	2.15	0	39.91	0

Figure 4.1: MAT Output

Another output page provided by the tool is correlation graph which plots a graph between the reference monitor and the sensor data as in the figure 4.2. The graph drawn will be a scatter plot and a line called as 'slope-intercept' will be drawn through the data points and is represented simply by the equation $y = mx + c$. This equation represents the average behavior of the sensor data (vertical axis, represented by y) compared with the reference data (horizontal axis, represented by x) [103]. The slope of the graph shows the similarity or difference in sensor measurements when compared with the reference instrument, on average.

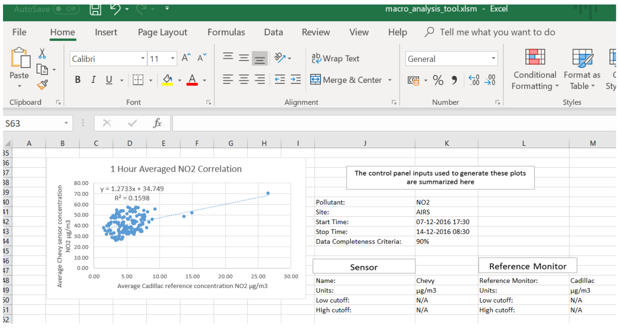


Figure 4.2: Correlation Graph

The tool also generates a coefficient of determination (also known as correlation coefficient in statistics) or R-Squared (R^2) shows how close the value is to the slope-intercept line. The value ranges from 0 to 1 and the closer R^2 is to 1, the stronger the agreement between the sensor and the reference data [105]. Along with these output the tool also generates a time series output graph which will show the concentration of pollutant for both the system as in the figure 4.3

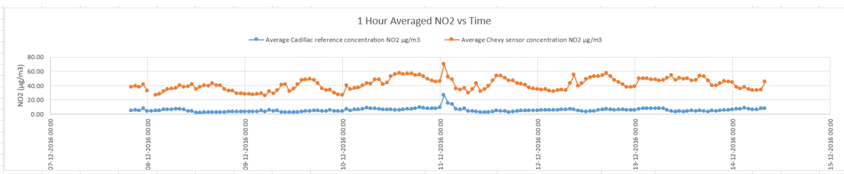


Figure 4.3: Time-Series Graph

From the Time-Series graph it could be seen how the closely value from two different system is related to. Having included all these functionality in the tool gives a strong solution for the calibration. The development of the tool was a year-long project of EPA with other community group(Clean Air Carolina) and one tribal nation(Eastern Band of Cherokee Indians). The next section will expand in detail about the background concept about the working of tool which is linear regression.

4.3 Linear Regression

Calibration can be done through different statistical methods and the most popular approach is Linear Regression method. In statistics, the term regression is used to describe a group of methods that summarize the degree of association between one variable (or set of variables) and another variable (or set of variables) [106]. Linear Regression establishes relationship between two variables one is the independent variable which will be on the x axis and the dependent variable in the y axis by drawing a straight line. If there are many observations and is plotted as a scatter plot then a line called as regression line as in the figure 4.4 could be drawn through it by Least Square method.

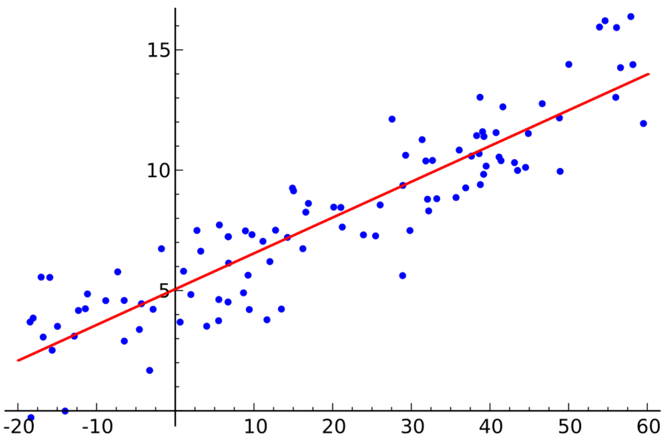


Figure 4.4: Linear Regression Model [107]

The method of Least Squares is a procedure to determine the best fit line to data [108] by minimizing the errors between the observed value and the actual value. This method considers that, for each value of x, there is a sub-population of y values normally distributed, that the means of all the sub-populations of y lie on the same straight line and all the sub-populations of y values have equal variances [109] [110]. The line drawn will have a slope m and is given by the formula

$$m = \frac{\sum_i[(x_i - \bar{x})(y_i - \bar{y})]}{[\sum_i (x_i - \bar{x})^2]} \text{ [111] [112] [106]}$$

where m is the slope, x_i and y_i are the reference and observed data values, $\bar{x}$ and $\bar{y}$ are the mean values of the data. The mean values will act as the centroid of the regression line and it has to pass through these points. The intercept equation c is

$$C = \bar{y} - m\bar{x} \text{ [111] [113]}$$

On getting the slope and intercept value the equation of line can be found. Once the Regression line is drawn, to find out how much the data is scattered and to what extend, correlation coefficient (R^2) is used. In other words R^2 gives the measure of degree to which the values of x and y are linearly correlated [111]. The equation for finding the regression value is

$$R^2 = \frac{\sum_i[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{[\sum_i (x_i - \bar{x})^2 [\sum_i (y_i - \bar{y})^2]]}} \text{ [111]}$$

The range of value for R^2 is between 0 to 1 and closer the value is to 1, the stronger the correlation between the two data. All these equation can be easily calculated with the help of Excel and thus all these equations are integrated to the MAT tool.

Chapter 5

Experimental Evaluations

Bibliography

[1] Wikipedia. Natural Environment. https://en.wikipedia.org/wiki/Natural_environment, 2017. [Online; accessed 26-January-2018].

[2] J Ritter, L., Solomon, K.R. & Forget. Persistent organic pollutants: An assessment report on DDT, aldrin, dieldrin, endrin, chlordane, heptachlor, hexachlorobenzene, mirex, toxaphene, PCBs, dioxins and furans. Report for the International Programme on Chemical Safety (IPCS) within the framewo. page 43 pp., 2005.

[3] Wikipedia. Persistent organic pollutant. https://en.wikipedia.org/wiki/Persistent_organic_pollutant, 2013. [Online; accessed 10-February-2019].

[4] Health Effects Institute. State of Global Air 2017. *Health Effects Institute*, page 15, 2017.

[5] World Health Organization. How air pollution is destroying our health. https://www.who.int/air-pollution/news-and-events/how-air-pollution-is-destroying-our-health, 2019. [Online; accessed 04-January-2019].

[6] WHO. Preventing disease through healthy environments exposure to air pollution: A major public health concern. 2010.

[7] Daniel Vallero. Tool Steel - an overview (pdf) | ScienceDirect Topics. *Fundamentals of Air Pollution (Fifth Edition)*, pages 1–10, 2014.

[8] OAR US EPA. NAAQS Table.

[9] Definition of National Ambient Air Quality Standards (NAAQS). *ohioepa.custhelp.com*.

[10] H Chen and R Copes. *Review of Air Quality Index and Air Quality Health Index*. 2013.

[11] Nikhil Soni and Neetishree Soni. Benefits of pedestrianization and warrants to pedestrianize an area. *Land Use Policy*, 57:139–150, 2016.

[12] Government of India and Forests & Climate Change Ministry of Environment. National Air Quality Index. page 58, 2014.

[13] Frequently Asked Questions Air Quality Health Index. The air quality health index. *Air Quality health index FAQ*.

[14] Rundong Tian, Christine Dierk, Christopher Myers, and Eric Paulos. MyPart: Personal, Portable, Accurate Airborne Particle Counting. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1338–1348, 2016.

[15] M Artha P Atricia Sierra-vargas and L U I S M Teran. INVITED REVIEW SERIES : AIR POLLUTION AND LUNG HEALTH Air pollution : Impact and prevention. (January), 2012.

[16] Farideh Golbabaei, Hamid Hassani, Asghar Ghahri, and Mirghani Seyed Someah. Risk Assessment of Exposure to Gases Released by Welding Processes in Iranian Natural Gas Transmission Pipelines Industry. 4(2):6–9, 2012.

[17] Morton Lippmann. Health Effects Of Ozone A Critical Review. *Journal of the Air Pollution Control Association*, 39(5):672–695, 1989.

[18] Elias Castanas Marilena Kampa. Human health effects of air pollution. 151:362–367, 2007.

[19] Rinkesh. What is Air pollution? `https://www.conserve-energy-future.com/causes-effects-solutions-of-air-pollution.php`, 2013. [Online; accessed 2-June-2018].

[20] K Heidorn. A Chronology of Important Events in the History of Air Pollution Meteorology to 1970. *Bulletin of the American Meteorological Society*, 59(12):1589–1597, 1978.

[21] Timothy Hatton. Quantifying the impact of smoke pollution on health in the 19th century. `https://voxeu.org/article/smoke-pollution-and-health-19th-century`, 2017. [Online; accessed 26-January-2019].

[22] Viktor P Ál. The Environmental Consequences of Industrialization in Western European Core Countries and the Borsod Basin of Hungary , 1850 – 1945 : A Comparative Outline. 3:75–84, 2016.

[23] Wikipedia. Great Smog of London. `https://en.wikipedia.org/wiki/Great_Smog_of_London`, 2019. [Online; accessed 26-January-2019].

[24] Wikipedia. Air Pollution Control Act. `https://en.wikipedia.org/wiki/Air_Pollution_Control_Act`, 2018. [Online; accessed 26-January-2019].

[25] K Balakrishnan, M Brauer, G Chen, and J Chow. *To Humans Outdoor Air Pollution*, volume 109. 2015.

[26] Eddie Salter. Active versus Passive Air Sampling . `https://www.envirotech-online.com/article/health-and-safety/10/skc/active-versus-passive-air-sampling-eddie-salter/923`, 2011. [Online; accessed 26-January-2019].

[27] Ministry of Environment. 4 How to Monitor? – Monitoring Methods | Ministry for the Environment, 2009.

[28] Daniel C Mussatti, Margaret Groeber, and Dan Maloney. Generic Equipment and Devices for continous emission minitoring system. *Evironment Protection Agency*, pages 4–1 – 4–40, 2000.

[29] Li Sun, Ka Chun Wong, Peng Wei, Sheng Ye, Hao Huang, Fenhuan Yang, Dane Westerdahl, Peter K.K. Louie, Connie W.Y. Luk, and Zhi Ning. Development and application of a next generation air sensor network for the Hong Kong marathon 2015 air quality monitoring. *Sensors (Switzerland)*, 16(2), 2016.

[30] Abdullah Kadri, Elias Yaacoub, Mohammed Mushtaha, and Adnan Abu-Dayya. Wireless sensor network for real-time air pollution monitoring. *2013 1st International Conference on Communications, Signal Processing and Their Applications, ICCSPA 2013*, pages 1–5, 2013.

[31] David Hasenfratz and Enabling. *Enabling Large-Scale Urban Air Quality*. 2015.

[32] Gary Anthes. Open source software no longer optional. *Communications of the ACM*, 59(8):15–17, 2016.

[33] Kavi Kumar Khedo and Vishwakarma Chikhooreeah. Low-Cost Energy-Efficient Air Quality Monitoring System Using Wireless Sensor Network. *Wireless Sensor Networks - Insights and Innovations*, (October), 2017.

[34] Yang Huang Liu, Liang Hu, Disheng Yang, and Hengchang. Air-Sense: indoor environment monitoring evaluation system based on ZigBee network. *IOP Conference Series: Earth and Environmental Science*, 81(1):12208, 2017.

[35] Wei Yi, Kin Lo, Terrence Mak, Kwong Leung, Yee Leung, and Mei Meng. *A Survey of Wireless Sensor Network Based Air Pollution Monitoring Systems*, volume 15. 2015.

[36] Movva Pavani and P. Trinatha Rao. Urban air pollution monitoring using wireless sensor networks: A comprehensive review. *International Journal of Communication Networks and Information Security*, 9(3):439–449, 2017.

[37] Shu Chiung Hu, You Chiun Wang, Chiuan Yu Huang, and Yu Chee Tseng. A vehicular wireless sensor network for CO2monitoring. *Proceedings of IEEE Sensors*, pages 1498–1501, 2009.

[38] Shu Chiung Hu, You Chiun Wang, Chiuan Yu Huang, and Yu Chee Tseng. Measuring air quality in city areas by vehicular wireless sensor networks. *Journal of Systems and Software*, 84(11):2005–2012, 2011.

[39] Péter Völgyesi, András Nádas, Xenofon Koutsoukos, and Ákos Lédeczi. Air quality monitoring with SensorMap. *Proceedings - 2008 International Conference on Information Processing in Sensor Networks, IPSN 2008*, pages 529–530, 2008.

[40] Srinivas Devarakonda, Parveen Sevusu, Hongzhang Liu, Ruilin Liu, Liviu Iftode, and Badri Nath. Real-time Air Quality Monitoring Through Mobile Sensing in Metropolitan Areas. 2013.

[41] Debanshee Saha. IoT based Air Quality Monitoring System using Wireless Sensors deployed in Public Bus Services. 2017.

[42] Yoshinari Shirai, Yasue Kishino, Futoshi Naya, and Yutaka Yanagisawa. Toward On-Demand Urban Air Quality Monitoring using Public Vehicles. *IWSC@Middleware*, pages 1:1–1:6, 2016.

[43] Chaosheng Xiang, Xiaofeng Liu, Aimin Jiang, Bin Yan, and Jing Xia. Poster: A Bicycle-borne Sensor Network for Monitoring Urban Air Quality. *Mobisys'16: Companion Companion Publication of the 14Th Annual International Conference on Mobile Systems, Applications, and Services*, (March):88, 2016.

[44] Xiaofeng Liu, Bin Li, Aimin Jiang, Shixin Qi, Chaosheng Xiang, and Ning Xu. A bicycle-borne sensor for monitoring air pollution near roadways. *2015 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-TW 2015*, (January 2017):166–167, 2015.

[45] S. D. Zhi, Y. B. Wei, and Z. H. Yu. Air Quality Monitoring Platform based on Remote Unmanned Aerial Vehicle with Wireless Communication. *Proceedings of the International Conference on Future Networks and Distributed Systems - ICFNDS '17*, (July):1–7, 2017.

[46] Evi Dons, Michelle Laeremans, Juan Pablo Orjuela, Ione Avila-Palencia, Glòria Carrasco-Turigas, Tom Cole-Hunter, Esther Anaya-Boig, Arnout Standaert, Patrick De Boever, Tim Nawrot, Thomas Götschi, Audrey De Nazelle, Mark Nieuwenhuijsen, and Luc Int Panis. Wearable Sensors for Personal Monitoring and Estimation of Inhaled Traffic-Related Air Pollution: Evaluation of Methods. *Environmental Science and Technology*, 51(3):1859–1867, 2017.

[47] Ke Hu, Timothy Davison, Ashfaqur Rahman, and Vijay Sivaraman. Air Pollution Exposure Estimation and Finding Association with Human Activity using Wearable Sensor Network. pages 48–55, 2015.

[48] Alyssa I. Robinson Fletcher, Richard Ribón Nicolas M. Oreskovic, M.D. Design and Clinical Feasibility of Personal Wearable Monitor for Measurement of Activity and Environmental Exposure. pages 874–877, 2015.

[49] Piero Zappi, Elizabeth Bales, Jing Hong Park, William Griswold, and Tajana Šimunić Rosing. The CitiSense Air Quality Monitoring Mobile Sensor Node. *Proceedings of the 11th ACM/IEEE Conference on Information Processing in Sensor Networks, Beijing, China*, (January), 2012.

[50] Sunyoung Kim, Eric Paulos, and Mark D Gross. WearAir: Expressive T-shirts for Air Quality Sensing. *Proc. TEI '10*, pages 295–296, 2010.

[51] Ke Hu, Yan Wang, Ashfaqur Rahman, and Vijay Sivaraman. Personalising pollution exposure estimates using wearable activity sensors. *IEEE ISSNIP 2014 - 2014 IEEE 9th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Conference Proceedings*, (April):21–24, 2014.

[52] Gabriele Valli, Mattia Internullo, Alessandro M. Ferrazza, Paolo Onorati, Annalisa Cogo, and Paolo Palange. Minute ventilation and heart rate relationship for estimation of the ventilatory compensation point at high altitude: A pilot study. *Extreme Physiology and Medicine*, 2(1):1, 2013.

[53] brian p rubin. Tzoa's Wearable Enviro-Tracker Wants To Clear The Air. `https://readwrite.com/2015/05/20/tzoa-wearable-air-quality-sensor-crowdfunding-indiegogo/`, 2015. [Online; accessed 06-March-2019].

[54] Chris C. Lim Michael Heimbinder, Raymond Yap. The aircasting platform. `http://aircasting.org/`, 2006. [Online; accessed 01-January-2019].

[55] Dan Han. Air Quality Monitoring , Fushun-Kokkola. pages 1–2, 2010.

[56] Aircasting Luminescence, New York Hall, The Aircasting Luminescence, Iem Heng, Raymond Yap, City Tech, Andy Zhang, The Aircasting, and Luminescent Vest. Make Your Own AirCasting Luminescent Apparel.

[57] Salil S. Kanhere. Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7753 LNCS:19–26, 2013.

[58] Vijay Sivaraman, James Carrapetta, Ke Hu, and Blanca Gallego Luxan. HazeWatch: A participatory sensor system for monitoring air pollution in Sydney. *Proceedings - Conference on Local Computer Networks, LCN*, pages 56–64, 2013.

[59] Duanping Liao, Donna J. Peuquet, Yinkang Duan, Eric A. Whitsel, Jianwei Dou, Richard L. Smith, Hung Mo Lin, Jiu Chiuan Chen, and Gerardo Heiss. GIS approaches for the estimation of residential-level ambient PM concentrations. *Environmental Health Perspectives*, 114(9):1374–1380, 2006.

[60] P Dutta, Paul M Aoki, N Kumar, and A Mainwaring. Common Sense: Participatory urban sensing using a network of handheld air quality monitors. *SenSys '09*, pages 349–350, 2009.

[61] Tarek Abdelzaher, Yaw Anokwa, Péter Boda, Jeff Burke, Deborah Estrin, Leonidas Guibas, Aman Kansal, Samuel Madden, and Jim Reich. Mobiscopes for human spaces. *IEEE Pervasive Computing*, 6(2):20–29, 2007.

[62] P M Aoki, R J Honicky, A Mainwaring, C Myers, E Paulos, S Subramanian, and A Woodruff. Common sense: Mobile environmental sensing platforms to support community action and citizen science. *Adjunct Proceedings Ubicomp 2008*, (October):Paper 201 OR – Human Computer Interaction Institu, 2008.

[63] Karl Aberer, Saket Sathe, Dipanjan Chakraborty, Alcherio Martinoli, Guillermo Barrenetxea, Boi Faltings, and Lothar Thiele. OpenSense. (January):39–42, 2010.

[64] David Hasenfratz, Olga Saukh, Silvan Sturzenegger, and Lothar Thiele. Participatory air pollution monitoring using smartphones. *Proc. 1st Int'l Workshop on Mobile Sensing: From Smartphones and Wearables to Big Data*, pages 1–5, 2012.

[65] Bratislav Predic, Zhixian Yan, Julien Eberle, Dragan Stojanovic, and Karl Aberer. ExposureSense: Integrating daily activities with air quality using mobile participatory sensing. *2013 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2013*, (3):303–305, 2013.

[66] Alfredo J. Perez, Miguel A. Labrador, and Sean J. Barbeau. G-Sense: A scalable architecture for global sensing and monitoring. *IEEE Network*, 24(4):57–64, 2010.

[67] Diego Mendez, Alfredo J. Perez, Miguel A. Labrador, and Juan Jose Marron. P-Sense: A participatory sensing system for air pollution monitoring and control. *2011 IEEE International Conference on Pervasive Computing and Communications Workshops, PERCOM Workshops 2011*, pages 344–347, 2011.

[68] Man Wong, Tsan Yip, and Esmond Mok. Development of a Personal Integrated Environmental Monitoring System. *Sensors*, 14(11):22065–22081, 2014.

[69] Kavi K. Khedo, Rajiv Perseedoss, and Avinash Mungur. A Wireless Sensor Network Air Pollution Monitoring System. *International Journal of Wireless & Mobile Networks*, 2(2):31–45, 2010.

[70] B. Fang, Q. Xu, T. Park, and M. Zhang. AirSense: An intelligent home-based sensing system for indoor air quality analytics. *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 109–119, 2016.

[71] M. F.M. Firdhous, B. H. Sudantha, and P. M. Karunaratne. IoT enabled proactive indoor air quality monitoring system for sustainable health management. *Proceedings of the 2017 2nd International Conference on Computing and Communications Technologies, ICCCT 2017*, (February):216–221, 2017.

[72] Rohan Narayana Murty, Geoffrey Mainland, Ian Rose, Atanu Roy Chowdhury, Abhimanyu Gosain, Josh Bers, and Matt Welsh. CitySense: An urban-scale wireless sensor network and testbed. *2008 IEEE International Conference on Technologies for Homeland Security, HST'08*, pages 583–588, 2008.

[73] John Bicket, Daniel Aguayo, Sanjit Biswas, and Robert Morris. Architecture and evaluation of an unplanned 802.11b mesh network. page 31, 2005.

[74] Joseph Camp, Joshua Robinson, Christopher Steger, and Edward Knightly. Measurement driven deployment of a two-tier urban mesh access network. page 96, 2006.

[75] CUWin Foundation. open source community hosting that makes a difference. `https://www.cuwireless.net/`, 2006. [Online; accessed 10-April-2019].

[76] Jen Hao Liu, Yu Fan Chen, Tzu Shiang Lin, Da Wei Lai, Tzai Hung Wen, Chih Hong Sun, Jehn Yih Juang, and Joe Air Jiang. Developed urban air quality monitoring system based on wireless sensor networks. *Proceedings of the International Conference on Sensing Technology, ICST*, pages 549–554, 2011.

[77] NATIONAL INSTRUMENTS. LabVIEW. (June):1–89, 2013.

[78] Emily G. Snyder, Timothy H. Watkins, Paul A. Solomon, Eben D. Thoma, Ronald W. Williams, Gayle S. W. Hagler, David Shelow, David A. Hindin, Vasu J. Kilaru, and Peter W. Preuss. The Changing Paradigm of Air Pollution Monitoring. *Environmental Science & Technology*, 47(20):11369–11377, 2013.

[79] Frequently Asked Questions. The air quality health index. *Air Quality health index FAQ*.

[80] J S Asha and P Sindhu. Website : www.ijirset.com Assessment of Air Quality in Two Cities of Kerala based on AQI by USEPA Method and. (2009):9284–9292, 2017.

[81] Debanshee Saha. IoT based Air Quality Monitoring System using Wireless Sensors deployed in Public Bus Services. 1952.

[82] Various Contributors. Essay on Air pollution: Introduction, Causes, Sources, Impact and Control Measures. 2013.

[83] Massimo Banzi and First Edition. *Getting started with Arduino*. :Books, an imprint of Maker Media, a division of O'Reilly Media, Inc., october 20 edition, 2008.

[84] Manuel Ruiz Guti. Arduino + Ethernet Shield. pages 1–42, 2017.

[85] Young Jin Jung, Yang Koo Lee, Dong Gyu Lee, Yongmi Lee, Silvia Nittel, Kate Beard, Kwang Woo Nam, and Keun Ho Ryu. Design of sensor data processing steps in an air pollution monitoring system. *Sensors*, 11(12):11235–11250, 2011.

[86] Technical Data. MQ-2 Semiconductor Sensor for Alcohol. *Structure*, pages 3–5, 2012.

[87] Vaibhav Jain. Insight - Learn the Working of a Gas Sensor. `https://www.engineersgarage.com/insight/how-gas-sensor-works`, 2013. [Online; accessed 10-November-2018].

[88] SGX Sensortech. SGX Sensortech Data Sheet MiCS-2714 MOS Sensor for Nitrogen Dioxide. pages 1–5.

[89] Tracy Allen. De-construction of the Shinyei PPD42NS dust sensor. (510):0–3, 2002.

[90] Joel Kuula, Timo Mäkelä, Risto Hillamo, and Hilkka Timonen. Response characterization of an inexpensive aerosol sensor. *Sensors (Switzerland)*, 17(12), 2017.

[91] Espressif Systems. ESP8266EX. pages 1–22, 2018.

[92] Katherine Hepworth. Big Data Visualization : Promises & Pitfalls. pages 7–19.

[93] Mary R English. Who Are the Stakeholders in Environmental Risk Decisions? ... How Should They Be Involved? *Risk: Health, Safety & Environment*, 11(243):352–362, 2000.

[94] Thingsspeak. Things speak- IOT platform. `https://thingspeak.com/`, 2017. [Online; accessed 26-January-2018].

[95] Alberto Cairo. Ethical Infographics. *Investigative Reporters and Editors, Inc. The IRE Journal*, 37(2):25–27, 2014.

[96] Eric S Hall, Surender M Kaushik, Robert W Vanderpool, Rachelle M Duvall, Melinda R Beaver, Russell W Long, and Paul A Solomon. Federal Reference Method (FRM), Federal Equivalent Method (FEM), National Ambient Air Quality Standards (NAAQS), Sensors; Federal Reference Method (FRM), Federal Equivalent Method (FEM), National Ambient Air Quality Standards (NAAQS), Sensors. *American Journal of Environmental Engineering*, 4(6):147–154, 2014.

[97] Davis) Kejuruteraan, Fakulti Tekn, NIST Engineering Statistics Handbook Jenny Wu (University of California and 3. Instrument Calibration. pages 1–7, 2018.

[98] Ron Williams, Tim Watkins, and Russell Long. Low Cost Sensor Calibration Options The Straw-Man Approach. 2013.

[99] Ron Williams, Vasu J Kilaru, Emily G Snyder, Amanda Kaufman, Timothy Dye, Andrew Rutter, Ashley Russell, and Hilary Hafner. Air Sensor Guidebook. *Epa/600/R-14/159*, (1):1–5, 2014.

[100] U.S. Environmental Protection Agency. Macro Analysis Tool -MAT. `https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=340520&Lab=NERL`, 2018. [Online; accessed 04-December-2018].

[101] U.S. Environmental Protection Agency. Air Sensor Toolbox for Citizen Scientists, Researchers and Developers. `https://www.epa.gov/air-sensor-toolbox`, 2018. [Online; accessed 04-December-2018].

[102] Epa and Ord. Calibration Curves: Program Use/Needs Final. (October), 2010.

[103] National Exposure Research Laboratory Office of Research and Development. How to Evaluate Low-Cost Sensors by Collocation with Federal Reference Method Monitors. pages 1–45, 2017.

[104] Instruction guide and macro analysis tool for community-led air monitoring. `https://www.epa.gov/air-research/instruction-guide-and-macro-analysis-tool-community-led-air-monitoring`, 2016. Accessed: 2018-12-06.

[105] Ron Williams. Emerging Sensor Technologies. 2018.

[106] Shaun Burke and High Wycombe. Regression and Calibration. pages 13–18.

[107] The free Encyclopedia Wikipedia. Linear regression. `https://en.wikipedia.org/wiki/Linear_regression`, 2018. [Online; accessed 04-December-2018].

[108] S.Miller. The Method of Least Squares and Signal Analysis. pages 1–7, 1992.

[109] A. M. Almeida, M. M. Castel-Branco, and A. C. Falcão. Linear regression for calibration lines revisited: Weighting schemes for bioanalytical methods. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 774(2):215–222, 2002.

[110] Wayne W Daniel and Chad Lee Cross. *Biostatistics: a foundation for analysis in the health sciences*. Wiley, 2018.

[111] D Stone and J Ellis. Calibration and Linear Regression Analysis: A Self-Guided Tutorial (Part 2). pages 1–7, 2001.

[112] Diana Arsova, Sofia Babanova, and Petko Mandjukov. Linear Calibration – Is It so Simple ? pages 35–42, 2009.

[113] David Harvey. 5.4: Linear Regression and Calibration Curves - Chemistry LibreTexts. *LibreTexts*, pages 1–12, 2016.