# IoT Network for PM Monitoring: Development and Deployment

Thesis submitted in partial fulfillment
of the requirements for the degree of

*Master of Science*
*in*
***Electronics and Communication Engineering***
*by Research*

by

Ayu Parmar
2021702003
ayu.parmar@research.iiit.ac.in

International Institute of Information Technology
Hyderabad - 500032, INDIA
August 2023

International Institute of Information Technology
Hyderabad, India

## CERTIFICATE

It is certified that the work contained in this thesis, titled ”IoT Network for PM Monitoring: Development and Deployment” by Ayu Parmar, has been carried out under my supervision and is not submitted elsewhere for a degree.

| | |
|---|---|
| Date | Advisor: Dr. Sachin Chaudhari |

To

My Family and Friends

# Acknowledgments

I would first like to express my deepest gratitude to my advisor Prof. Sachin Chaudhari for guiding me through the right paths throughout my master's course. His expertise, experience, and innovative ideas helped me explore new technical subjects. His constructive remarks and supervision motivated me continuously to complete the journey of my master's degree with joy. He persistently pushed me to get better every day, helped me to enhance my technical expertise, and learn many new things throughout the writing of this dissertation.

I would also like to thank Prof. K. S. Rajan for his unwavering support and guidance throughout my research journey. He has been a guiding light since the early stages of my research, offering invaluable domain expertise that has allowed me to incorporate cutting-edge ideas and methodologies in my work. I am indebted to him for broadening my perspective on problem viewing and analysis. I would also like to acknowledge the guidance I received from Dr. Kavita Vemuri in the research collaborations we had during my time here research. I would like to extend my sincere appreciation to Mrs. Anuradha Vattem for recognizing my potential and appointing me as a research engineer at the Smart City Research Center. It is from this platform that I have embarked on a remarkable journey toward pursuing my Masters degree.

I would also like to thank my project partners, Spanddhana Sara, Ayush Diwedi, Ishan Patwardhan, Nitin Nilesh and Rajashekhar along with my colleagues and friends, for providing me assistance continuously at several instances. I want to extend my deepest thanks to my friends Usha and Jayati for supporting, encouraging and for being with me through the tough times. I would also like to thank my labmates Ruchi, Souradeep, Sasanka, Ihita, Viswanadh, Rishi, Nilesh and everyone else. Special thanks to Ritik Yelekar and Vivek Gupta for helping me with deployment and maintaining the big IoT network. Finally, I would like to express my deepest gratitude to my mother, especially my late father, who supported me throughout this journey. I know he would be immensely proud to see me complete my master's degree. It would have been impossible for me to complete my study without their tremendous understanding and encouragement over the past two years.

I would also like to thank Smart City Living Lab for supporting me financially for this dissertation. This exclusive fellowship made me more confident in my technical skills and research. I would also like to thank the Society for Cyberabad Security Council (SCSC) for giving permission and helping us in deploying devices on roads and junctions.

# Abstract

Monitoring and analyzing air quality is a challenging task without understanding influencing factors. Conventional air pollution monitoring is limited to few locations and accessing data is difficult. Internet of Things (IoT) provides a solution by enabling cost-efficient networks of particulate matter (PM) monitoring devices that are easily connected to the internet. PM monitoring portable sensors combined with IoT offer a more efficient and widespread PM monitoring solution, overcoming issues with bulky and expensive traditional systems. This thesis focuses on development and deployment of PM monitoring device and the establishment of a dense PM monitoring network.

This thesis focuses mainly on four aspects. In the first aspect, the development of an end-to-end low-cost IoT system for a densely deployed PM monitoring network carried out an extensive field study in a region of the Indian metropolitan city of Hyderabad. A total of 49 devices were deployed in an area of $4\,\text{km}^2$, with 43 of them developed specifically for this study and the remaining six obtained from external sources, a density never realized in any metropolitan city worldwide and mostly in developing countries like India in the past. A total of 15 devices were connected to Wi-Fi, primarily within the campus area and wherever Wi-Fi connectivity was available, whereas 34 devices were equipped with 2G eSIM. The low-cost sensors were carefully calibrated to account for seasonal variations by utilizing a highly precise reference sensor. Also, a robust device was made that can cache data to avoid loss due to communication outages.

The second aspect of this thesis focuses on gathering a significant quantity of data, nearly 20.7 million. This unique dataset offers great opportunities for future research and is examined to evaluate whether a concentrated deployment of PM monitoring devices was necessary. Third, various analytical methods, such as mean, variance, spatial interpolation, and correlation, were utilized to produce informative findings about the fluctuations of PM over time and across seasons. In order to understand the impact of firecracker detonations during Diwali festival evenings, a spatio-temporal analysis of PM values was conducted. As part of our analysis, we also tried to answer an important question concerning decision-makers about the optimum density required for effectively monitoring street-level pollution. For the considered scenario, we demonstrated that PM monitoring devices should be deployed at most $350\,\text{m}$ apart to accurately capture the spatial variability of PM.

Finally, this thesis examines various challenges encountered during the pre-deployment and post-deployment of PM monitoring devices. It addresses issues such as the design of low-cost devices, pre-deployment calibration, and seasonal calibration. Also explores challenges related to power supply, theft, environmental factors affecting sensor performance, and hardware failures. Additionally, the study investigates hardware reset and corrupt or redundant data. This thesis examines the challenges that arise during the deployment of PM monitoring devices and to propose feasible solutions to mitigate these issues, thereby offering valuable insights into the effective implementation of PM monitoring technology. Overall, the thesis highlights the importance of dense deployment and addresses the challenges to ensure reliable and accurate data collection.

# Contents

# List of Figures

# List of Tables

*Chapter 1*

# Introduction

## 1.1 Motivation

Internet of Things (IoT) refers to the network of physical devices, vehicles, buildings, and other items embedded with electronics, software, sensors, and connectivity, enabling these objects to connect and exchange data. IoT allows for the seamless integration of the physical and digital worlds, creating smart environments in which devices can communicate and interact with one another. This creates new opportunities for businesses and individuals to improve efficiency, reduce costs, and gain insights from data. IoT has greatly increased the amount of data generated by connected devices [1]. IoT devices collect and transmit vast amounts of real-time data, enabling organizations to gain valuable insights and make data-driven decisions in various fields such as manufacturing, healthcare, transportation, and smart cities. Overall, IoT has greatly expanded the potential for data generation and insights. The combination of sensors and the IoT allows for the creation of smart environments in which monitoring Air pollution is one particular field.

Air pollution has been an issue of grave concern across the world for decades [2]. PM occurring from local activities is a significant contributor to air pollution, which causes serious health implications. In India, the average PM concentration is 55.8 $\mu g/m^3$, 11 times higher than the WHO guideline [3]. According to an estimate, in 2019, around 6.7 million premature deaths globally were associated with air pollution [4]. Studying and continuously monitoring the various patterns related to air pollution is essential to address the challenge comprehensively. Many countries have established elaborate structures for air-quality monitoring based on beta attenuation monitor (BAM) and tapered element oscillating microbalance (TEOM) often deployed by pollution control boards and other governmental agencies to monitor air quality [5]. Although the PM data from these stations is very accurate, this approach has the limitation of scalability. Often the centralized AQMS are expensive, bulky and large in size [6]. Thus, they cannot be densely deployed. For example, in a big metropolitan city like Hyderabad, with a population of over 6.7 million [7] and area over $650\,\mathrm{km}^2$, the Central Pollution Control Board (CPCB) and the Telangana State Pollution Control Board (TSPCB) have deployed only 12 AQMS [8]. All of this points out the issue of the expensive and time-consuming setup of the air quality monitoring apparatus and a significant mismatch between the requirement of PM data and its availability. Consequently, the resolution of available air quality data is limited as very few stations are typically responsible for an entire city region. Low resolution is not enough for a deeper understanding of PM as the pollutant levels can vary drastically even within smaller blocks in a city [9].

Citywide deployment of such devices has been studied to increase pollution data's spatial-temporal resolution [10, 11, 12, 13], but there is a lack of literature on densely deployed low-cost PM monitoring IoT network. This is the primary motivation for this work. To overcome the limitation of scalability, this thesis focuses on the development and deployment of low-cost sensor-based monitoring nodes with the fine spatiotemporal resolution, localized monitoring, and real-time analysis of outdoor PM. A thorough analysis of data collected for seven months has been presented to establish the need for dense deployment of PM monitoring devices.

## 1.2 Summary of Contributions

The main contributions from this thesis are presented in the chapters mentioned as follows -

- **Chapter 4**
  - For the high spatial resolution of outdoor PM, 49 IoT-based PM monitoring devices have been developed, calibrated and deployed at various outdoor locations.
  - The developed device is designed to be robust against the issue of data loss due to connection and power outages. The device maintains an offline cache in the event of an outage. The stored data is offloaded in bulk once the power and communication are restored.
  - All PM sensors were calibrated for seasonal variations by co-locating with a reference sensor. Also, each device was calibrated individually.
  - The devices were deployed at 49 outdoor locations covering a $4\,\text{km}^2$ area in Gachibowli, Hyderabad, India. The field locations were selected to include urban, semi-urban, and green regions. Few devices were deployed at busy traffic junctions and roadsides. The data were recorded at a frequency of every 30 seconds (sec) spanning over all the seasons for seven months, thus aggregating 20.7 million data points.
  - Different analyses were carried out by observing seasonal mean and variance, spatial interpolation, event-driven variation and correlation. Results show the optimal deployment across a varied landscape and can be a key factor in identifying the release of high concentration of PM in real-time.

- **Chapter 5**
  - We highlighted the challenges and issues encountered during and after deployment of PM monitoring devices, which may be helpful for on-field researchers and readers.
  - A novel approach was proposed that can estimate the Air Quality Index (AQI) without using any PM sensors, by utilizing machine learning and traffic data, which can help to avoid the hassle associated with sensor usage.

  **Note:** This particular setup dataset has been used for training the algorithm that can estimate the AQI by using machine learning and traffic data and without using PM sensors. Its training and result is not discussed in this dissertation. Credits belong to Nitin Nilesh.

## 1.3 Structure of Thesis

The rest of this thesis is organized as follows-

- **Chapter 2** briefly introduces IoT, its applications, and the challenges involved.
- **Chapter 3** gives an overview of the literature survey on low-cost air pollution monitoring and the current state of IoT monitoring networks for air pollution.
- **Chapter 4** describes the dense air pollution monitoring IoT network developed and deployed in $4\,\mathrm{km}^2$ area in Gachibowli Hyderabad. The description includes the hardware development & deployment of the sensor nodes, calibration of the PM sensor, Indian dataset collection for more than one year as well as the spatiotemporal analysis of the dataset.
- **Chapter 5** presents the challenges faced during deployment and maintenance.
- **Chapter 6** conclude this thesis.

*Chapter 2*

# An Overview on IoT

This chapter provides an overview of IoT. It is followed by a summary of IoT components. Following that, IoT applications and use cases are addressed. Finally, the key issues confronting IoT are discussed. This chapter only provides a brief overview of IoT; interested readers may learn more about IoT in numerous books, including as [1, 14, 15, 16, 17].

## 2.1 Definition of IoT

IoT is a network of physical devices, vehicles, buildings, and other objects that are embedded with sensors, software, and connectivity which enables these objects to collect and exchange data. Two definitions of IoT are mentioned as follows-

- Gartner Research [18] defines it as *the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment.*

- United Nations International Telecommunication Union [19] defines it as *a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.*

The IoT allows for the seamless communication and integration of these devices, resulting in the ability to monitor, control, and automate various systems and processes. IoT devices are connected to the Internet, and this allows for real-time data collection, monitoring, and analysis. This also enables remote access and control of the device. IoT is used in many industries like smart homes, healthcare, agriculture, automotive, manufacturing, and many more.

IoT can be represented as a network of connected devices. Fig. 2.1 shows an illustration of devices connected representing an IoT network. These networks are now commonly found in smart homes, factories, and other automation-based industries. IoT allows us to link computational, mechanical, and virtual devices by allowing them to share data over the internet or another form of communication.

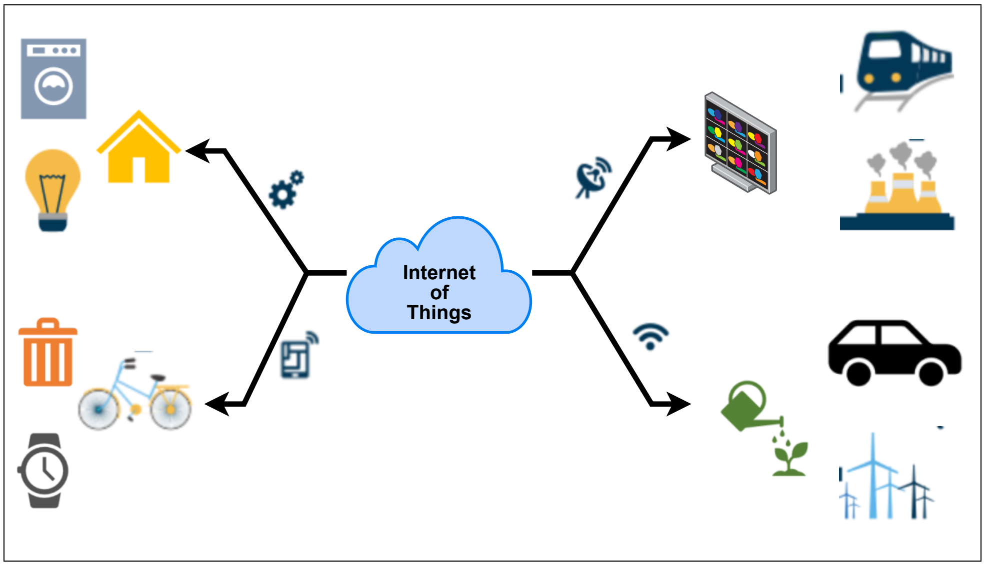



Figure 2.1: Illustration of connected devices to the internet

## 2.2 Architecture

The components of IoT architecture are critical for the functioning of IoT systems. Few of the IoT models are mentioned in [20]. There are a few models that consider five layers, and there are several others that include four or three stages for an IoT project. Fig. 2.2 shows the architecture of the IoT system having four layers [21]. They include sensors and actuators that collect data from the environment, a gateway that facilitates communication between devices and the internet, and a cloud/data server that stores and processes large amounts of data. Lastly, the user interface allows users to interact with the system and access the data. These components work together to create a seamless and integrated IoT system that can collect, transmit, and process data for smart systems and applications. The components of IoT architecture are briefly mentioned below:

### 2.2.1 Sensing and Actuator

Sensors and actuators are crucial components of the IoT architecture as they are responsible for collecting data from the environment. These sensors such as temperature, humidity, light, or motion sensors, can be equipped with various technologies and they can collect data in real-time. The data collected by these sensors is then transmitted to other components of the IoT system, such as the gateway, cloud server, or application layer, for further processing and analysis. The data collected by these sensors play a key role in enabling the creation of smart systems and applications as it provides the raw data that is used to make decisions and automate processes [22] [23]. Apart from having the sensors, nodes may have a feedback system and contain actuators that react based on the input received by the device from the sensors. Sensors and actuators come in various shapes and sizes, ranging from small wearable

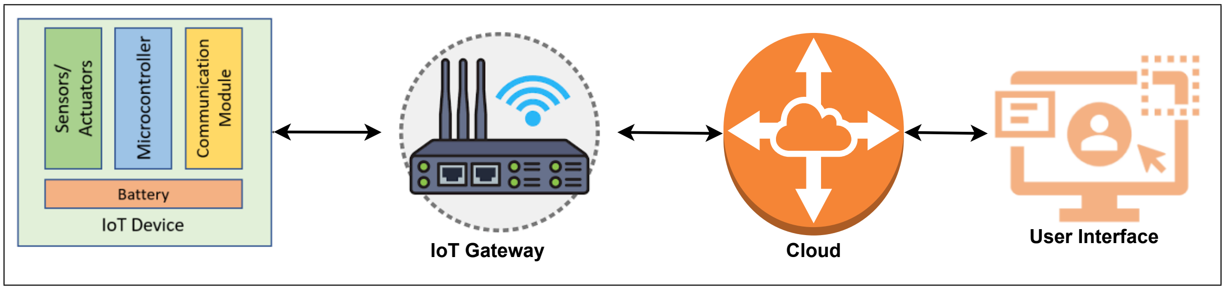



Figure 2.2: IoT architecture

devices to large industrial sensors. They can be connected to the internet via Wi-Fi, Ethernet, or cellular networks, allowing for real-time data transmission. They are also designed to be low-power and energy-efficient, allowing for longer battery life and less maintenance. Sensor and actuator-based IoT systems may either be standalone (as exemplified) or integrated by a pipeline following an automation cycle of sense $\rightarrow$ analyze $\rightarrow$ actuate.

### 2.2.2 Gateway

The gateway is a critical component of the IoT architecture that facilitates communication between devices and the Internet. It acts as a bridge between the devices and the cloud, allowing for data transmission and ensuring seamless communication between the different components of the IoT system. The gateway collects data from the sensors and devices and transmits it to the cloud for storage and analysis.

The gateway can be a standalone device or integrated into other devices, such as smart home hubs or industrial control systems. It can be connected to the internet via Wi-Fi, Ethernet, or cellular networks, allowing for real-time data transmission [23]. It also serves as a security layer, providing encryption and authentication to ensure secure communication between devices and the internet. In addition to its role as a communication facilitator, the gateway can also perform local data processing and analysis. This enables the IoT system to make real-time decisions and take action, reducing the need for data transmission to the cloud and reducing the overall system latency.

### 2.2.3 Cloud/Data Server

The cloud or data server is a critical component of the IoT architecture as it stores and processes large amounts of data. It acts as the central repository for all data collected by the sensors and devices, allowing for data analysis and decision-making. The cloud server can be a public or private cloud, or it can be a dedicated data center [24] [25] [26].

The cloud server is responsible for storing and managing the vast amounts of data that IoT devices and systems generate. This data can be used for various purposes such as real-time monitoring, predictive maintenance, or customer behavior analysis. With the growth of IoT, the demand for cloud storage

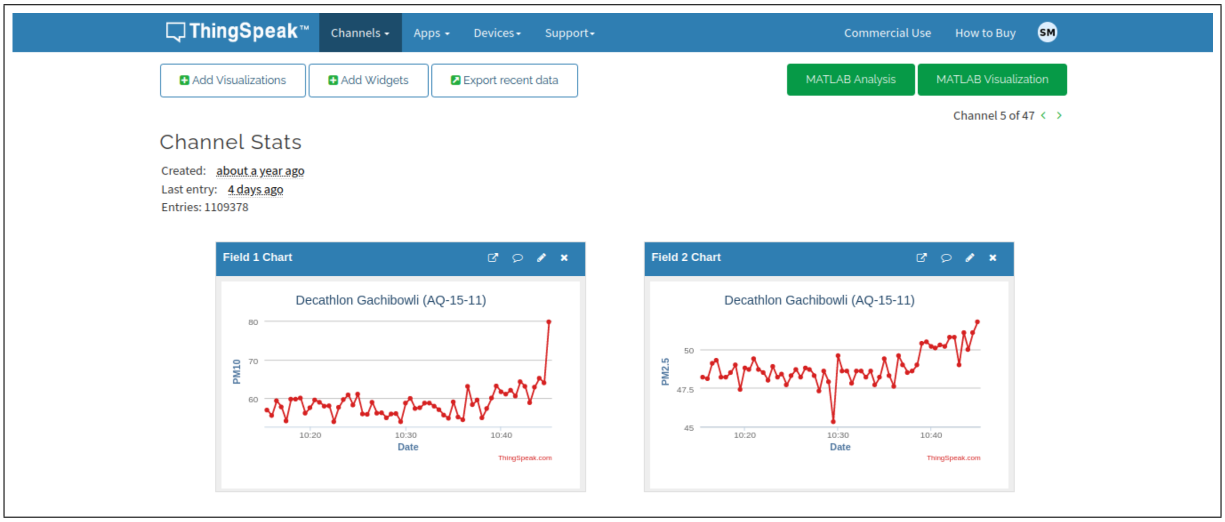



Figure 2.3: ThingSpeak User Interface

and processing capacity has increased dramatically, leading to the development of highly scalable and secure cloud infrastructure. In addition to data storage, the cloud server performs complex data analysis and processing. This can include data mining, machine learning, or artificial intelligence algorithms. These algorithms can be used to identify patterns in data, make predictions, or automate processes.

### 2.2.4 User Interface

The user interface is an important component of the IoT architecture, providing a way for users to interact with the system and access data. The user interface can be web- or mobile-based, allowing users to access and interact with IoT data from anywhere, at any time. It provides a graphical and user-friendly interface for accessing and interacting with IoT data, making it easy for users to understand and use. It can include dashboards, visualizations, and reporting tools that allow users to view and analyze data in a variety of formats, such as charts, graphs, and tables.

In addition to data visualization, the user interface provides functionality for controlling IoT devices. This can include controlling device settings, scheduling actions, and receiving notifications and alerts. Throughout this dissertation, ThingSpeak and AirIoT was used as one of the interfaces. As shown in the screenshot in Fig. 2.3, Thingspeak offers an easy way to track devices in real-time, with the ability to monitor data patterns and perform mathematical operations on the time-series data produced by IoT devices.

The user interface must be designed to be intuitive and user-friendly, with clear and concise information presented in an easy-to-understand format. It should also be responsive and adaptable to different devices and screen sizes, allowing users to access and interact with IoT data from a variety of devices.

## 2.3 Applications

The introduction of IoT has expanded the potential of multiple industries, representing a significant technological advancement due to its ability to store vast amounts of sensor data and perform edge computing - a feature absent in previous products and devices. Various innovative developments have emerged in global industries with examples including healthcare, wearables, automotive, and more, summarized briefly below with detailed information available in the provided sources for interested readers [27, 28, 29, 30, 31, 32, 33, 34, 35, 36].

- **Smart cities:** The main idea behind smart cities is to improve the lifestyle and address the problems related to pollution, traffic, logistics, education, public transport, etc., for sustainable development. The term "smart" refers to an intelligent system that can make its own decision up to a certain extent without any manual intervention. For example, a traffic light adjusts the duration of red and green signals in real-time according to the vehicle density at the junctions. The smart Heating Ventilation and Air Conditioning (HVAC) systems in smart buildings ensure good ventilation. In 2015, the Government of India launched the Smart City Mission to improve the existing infrastructure and enhance the efficiency of systems for the public's benefit by leveraging technology. The mission aimed to develop one hundred cities across India as smart cities. In line with this, the IIIT (International Institute of Information Technology) Campus in Hyderabad adopted the smart city approach and created a Living Lab within its premises to promote sustainable development across three value domains - social, economic, and environmental. This initiative aims to establish a livable and renewable urban center in Hyderabad through the implementation of various smart technologies, including air pollution monitoring, smart street lamps, water monitoring, smart public transport and more.

- **Smart homes:** Automating and controlling lighting, heating, cooling, and other household appliances with smart devices. Smart homes use IoT technology to automate and control household appliances like lighting, heating, and cooling. This is achieved through the use of smart devices that are connected to the internet and can be controlled remotely. The goal of a smart home is to make household tasks more convenient and efficient, reducing energy consumption and increasing comfort for the occupants. Overall, smart homes represent a step forward in home automation and technology integration in daily life.

- **Healthcare:** Monitoring and tracking vital signs, medicine intake, and other health-related data to improve patient care and health outcomes. The healthcare industry is one of the areas that has greatly benefited from the application of IoT. IoT-based solutions in healthcare allow for the monitoring and tracking of vital signs, medicine intake, and other health-related data. This helps healthcare professionals to better understand the health status of their patients and improve patient care. IoT devices such as wearable fitness trackers, smart pills, and remote patient monitoring systems have made it easier for healthcare providers to access critical information about their patients, leading to better health outcomes. With IoT, patients can also be better managed and cared for in their own homes, reducing the need for hospital visits and lowering healthcare costs. IoT has the potential to revolutionize the healthcare industry and change the way we approach healthcare delivery.

- **Agriculture:** Using sensors and devices to monitor soil moisture, temperature, and other environmental conditions to optimize crop yields. Agriculture is another industry that has seen the benefits of IoT implementation. IoT devices and sensors are used to monitor soil moisture, temperature, and other environmental conditions that are critical to crop growth. By collecting this data, farmers can make informed decisions about when to plant and harvest crops, how much water to use, and what fertilizer to apply. This leads to improved crop yields, reduced waste, and increased efficiency in agriculture. IoT also allows for precision agriculture, where specific areas of a field can be targeted for different treatments based on the data collected by the sensors. In addition, IoT devices can also monitor animal health, reducing the spread of disease and improving the overall health of livestock. The use of IoT in agriculture is expected to have a significant impact on food production and global food security.

- **Manufacturing:** Manufacturing is another industry that is embracing IoT. IoT technology is being used to improve production processes by enabling real-time monitoring and control of machinery, tools, and other industrial equipment. This allows manufacturers to quickly identify and resolve any problems, reducing downtime and improving efficiency. In addition, IoT-enabled machinery can collect and transmit data on usage and performance, enabling predictive maintenance and reducing the need for costly repairs. Furthermore, IoT can also provide real-time monitoring of production lines, enabling manufacturers to track inventory and make more informed decisions about production. By implementing IoT, manufacturers can improve the overall efficiency of their operations and remain competitive in a rapidly evolving industry.

- **Energy management:** Energy management is an important application of IoT. IoT technology is used to monitor and control energy usage and distribution, aiming to optimize energy usage and reduce waste. By integrating sensors and devices into the energy grid, energy providers can gain real-time insight into energy usage patterns and adjust energy distribution accordingly. This enables providers to more effectively balance supply and demand more effectively, reducing the need for energy generation and ultimately lowering costs. Additionally, IoT-enabled energy management systems can also automate energy-saving processes, such as turning off lights and appliances when they are not in use, reducing energy waste. Through these applications, IoT is helping to make energy management more efficient and sustainable.

- **Environmental monitoring:** Environmental monitoring involves using IoT devices to gather data on various environmental factors like air quality, water quality, etc. This helps in tracking the changes in the environment and taking necessary measures to manage it effectively. The collected data can also be used for research and analysis to develop new strategies for environmental conservation. The goal is to understand the environment better and work towards preserving it for future generations.

## 2.4 Challenges

We have discussed several applications of IoT in the previous section. But, IoT is a relatively new research area and suffers from a few challenges that must be addressed to utilize the resources best. A few of such challenges are discussed below -

- **Reliability of sensors:** Reliability is a key challenge in the implementation of IoT. As discussed in previous sections, IoT systems that rely on small, low-cost devices are not ideal substitutes for high-quality instruments. Low-cost, small IoT sensors have limitations that make it difficult to achieve high data density. The low-cost devices suffer from certain inaccuracies that need to be considered before deployment and they need frequent calibration to maintain accuracy due to sensitivity to environmental conditions. This frequent calibration requirement poses a major challenge to the reliability of the collected data. Additionally, the high failure rate of low-cost sensors further exacerbates the challenges.

- **Development Cost:** The cost of implementing an IoT solution can be a major challenge as it requires investment in devices, infrastructure, and technology. The cost of devices such as sensors, actuators, and other hardware components can add up, especially when deploying a large-scale solution. The cost of setting up and maintaining the infrastructure to support these devices, such as servers, cloud storage, and networking equipment, also adds to the overall cost. Companies may have to find cost-effective alternatives or opt for a gradual roll-out of IoT devices to achieve cost-effectiveness.

- **Power management:** Power management is a significant challenge in IoT as the devices need to function for extended periods without requiring frequent battery replacements or power source changes. Long battery life is crucial for IoT devices, often deployed in remote or hard-to-reach locations. Power efficiency is also important as it helps to minimize energy consumption and reduce costs associated with frequent battery replacements or maintenance. IoT devices must be designed to minimize power consumption through the use of low-power microcontrollers, low-power communication protocols, and efficient power management strategies.

- **Interoperability:** Interoperability is a crucial aspect of IoT as it involves the seamless integration of devices from different manufacturers. A lack of interoperability can result in a fragmented system that fails to provide a unified solution. Communication protocols and data formats must be standardized to achieve interoperability. This is essential for ensuring compatibility between devices and enabling them to work together smoothly. Interoperability is also important for ensuring that users can add new devices to their existing systems without facing compatibility issues. In short, interoperability is vital for the smooth functioning and growth of the IoT ecosystem.

- **Scalability:** Scalability is a major challenge in the growth of the IoT ecosystem. As the number of IoT devices increases, the system must be able to handle the rising number of connected devices and the growing volume of data they produce. The architecture of the system must be able to support a large number of devices and the processing of vast amounts of data, while also ensuring that performance is not compromised. To achieve scalability, the system must have the ability to distribute processing power, storage capacity, and network bandwidth as needed, and to integrate new devices and services as they come online seamlessly. This requires a flexible, modular, and scalable infrastructure that can be easily adapted to meet changing demands and challenges.

- **Security:** In earlier sections of Chapter 2, it was noted that standard protocols like Wi-Fi and BLE are integrated into small, low-end MCUs. However, these chips may not implement the protocols with all their features, making the data sensed by IoT nodes vulnerable. Outdated encryption methods used by low-end chips can pose a privacy issue at the device level, as they may be less secure and easier to hack. Manufacturers prioritize cost and efficiency over advanced encryption methods, which can result in compromised security. There are also security concerns regarding transferring data from the gateway to the cloud securely. While cloud providers like Amazon, Google, and Microsoft offer secure cloud services, low-end devices still have a long way to go in terms of addressing privacy and security concerns.

- **Privacy:** Privacy is a major concern for many users in the context of IoT. The vast amounts of data generated by IoT devices can include sensitive personal information, such as location data, financial transactions, and health records. To protect privacy, it is important to ensure that personal data is collected, stored, and used ethically and in compliance with privacy laws and regulations. This requires secure data storage, encrypted communication, and strict access controls to ensure that only authorized individuals can access personal data. Privacy is a complex issue that requires a multi-faceted approach, involving technical, legal, and policy-related solutions. Ensuring privacy in the context of IoT requires a commitment from all stakeholders to prioritize privacy and ensure that it is protected in a way that is consistent with users' expectations and rights.

- **Integration with legacy systems:** Integrating IoT solutions with legacy systems can pose a challenge, as the integration process often involves significant effort and resources. The compatibility of the new technology with the older systems needs to be thoroughly checked and tested. It can also be challenging to find the right personnel with the required technical expertise to manage the integration process. In order to ensure seamless integration, a well-planned strategy is crucial.

*Chapter 3*

# Overview of Air Pollution Monitoring Networks

This chapter briefly outlines the reason for working on dense air pollution monitoring. A complete literature study of previous approaches and traditional monitoring sensor networks, low-cost sensors for air pollution monitoring, and a thorough survey of various current IoT air pollution monitoring networks worldwide are briefly covered.

## 3.1 Motivation

The importance of addressing air pollution cannot be overstated, as it poses a serious threat to the health and well-being of individuals of all ages, as cited in [4]. PM is a particularly dangerous air pollutant, capable of causing respiratory and heart related illnesses and even premature death, especially for those exposed to it over long periods, as stated in [37]. Rapid and unplanned urbanization has led to increased air pollution, particularly in developing nations with high population and traffic density, where vegetation is being degraded, and metropolitan areas lack proper ventilation, putting residents at risk, as discussed in [38]. To combat this issue, monitoring air quality and taking appropriate action are critical. The present thesis aims at dense PM monitoring IoT networks and evaluates the effectiveness of dense deployment of low-cost PM sensors.

## 3.2 World Initiatives for Air Pollution Monitoring

There are several global initiatives aimed at improving air pollution monitoring and control. Some of these initiatives include:

- **National Clean Air Programme:** In India, the Central Government launched National Clean Air Programme (NCAP) [39] as a long-term, time-bound, national-level strategy to tackle the air pollution problem across the country in a comprehensive manner with targets to achieve a 20% to 30% reduction in PM concentrations by 2024. The air quality of cities is monitored by State Pollution Control Boards, which publish their results from time to time.

- **World Health Organization (WHO) Air Quality Guidelines:** The WHO provides guidelines on air quality standards and health effects, and promotes the development of air quality monitoring networks to ensure that these standards are met [40].

- **United Nations Environment Programme (UNEP):** UNEP works with governments and other organizations to promote sustainable practices and to mitigate the impacts of air pollution. This includes the development of air pollution monitoring networks and the use of technology, such as IoT devices, to improve air quality monitoring and control [41].

- **European Environment Agency (EEA):** The EEA is a European Union agency that provides information and support on environmental issues, including air pollution. The EEA operates an air quality monitoring network across Europe and provides regular reports on air quality and health impacts [42].

- **US Environmental Protection Agency (EPA):** The EPA is the lead agency in the United States for the protection of human health and the environment. The EPA operates an air quality monitoring network across the country and provides regular reports on air quality and health impacts [43].

- **Asian Development Bank (ADB):** The ADB works to improve the quality of life in Asia and the Pacific region by reducing poverty and promoting sustainable economic growth. One of the ADB's priorities is to mitigate the impacts of air pollution, including through the development of air pollution monitoring networks and the use of technology, such as IoT devices, to improve air quality monitoring and control [44].

  These initiatives demonstrate the global commitment to improving air quality and mitigating the impacts of air pollution and highlight the importance of air pollution monitoring and control in achieving these goals.

## 3.3 Related Work

Many initiatives have been made across the world to detect air pollution. To evaluate air quality, the studies include static, mobile, and current image processing-based solutions. The subsections that follow address a few pertinent techniques.

### 3.3.1 Stationary Networks

Many countries have established elaborate structures for air-quality monitoring based on BAM and TEOM often deployed by pollution control boards and other governmental agencies to monitor air quality [5]. Although the PM data from these stations is very accurate, this approach has the limitation of scalability.

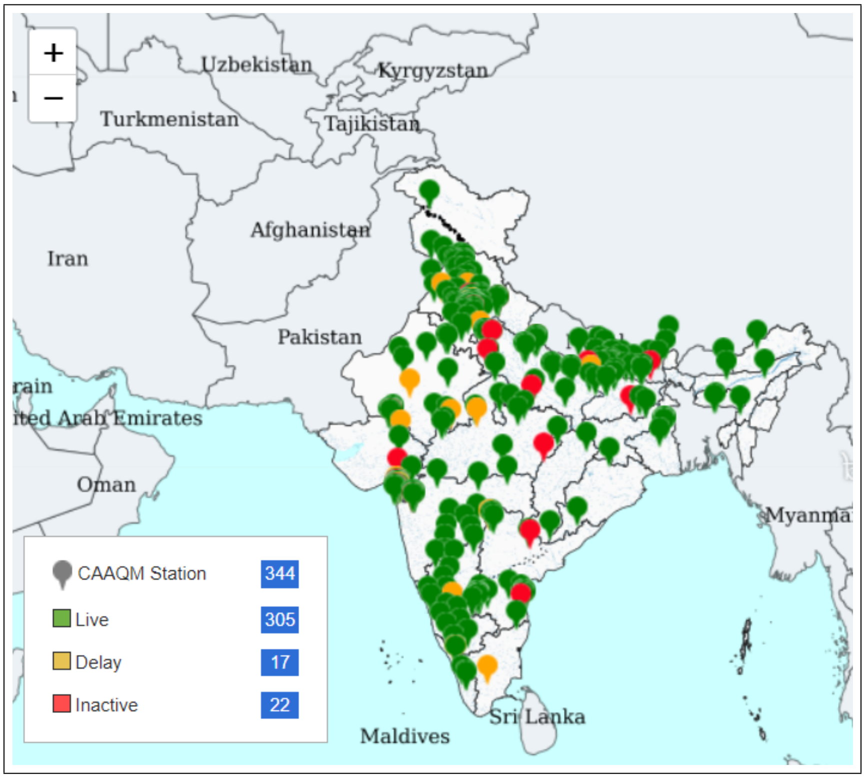



Figure 3.1: CPCB monitoring stations under NAMP

The high spatial variance of PM concentration has been demonstrated in previous studies, indicating that it can vary within a few meters [45]. Various low-cost IoT-based solutions have emerged to mitigate this issue, which enhances the spatial resolution of PM data by utilizing hardware devices that can be installed on poles, walls, traffic lights, etc. Few of the network are CPCB, London Air Quality Network (LAQN) [46], CitySense [47], AirBox [48].

In India, CPCB in India is responsible for regulating air and water pollution. To carry out this role, they have established the National Air Monitoring Programme (NAMP), which focuses on regularly monitoring three specific pollutants: sulfur dioxide, nitrogen dioxide, and PM. The CPCB also considers meteorological factors such as temperature, humidity, wind speed and direction, in addition to pollutants, in their monitoring efforts. They work in collaboration with other entities such as State Pollution Control Boards, Pollution Control Committees, and the National Environmental Engineering Research Institute in Nagpur to ensure consistent and uniform data on air quality. CPCB provides technical and financial support to operate the 804 monitoring stations located in 344 cities/towns across 28 states and 6 union territories as shown in Fig. 3.1.

In LAQN [46], the network comprises 33 monitoring units located throughout London City. In addition, data from a few more nodes deployed by local authorities are also included in the network. This information is publicly accessible and can be viewed in real-time on the website.

The citysense [47] project aims to build a wireless sensor network that covers the urban area of Cambridge. The network includes air quality monitoring sensors and is made up of a total of 100 Linux-based PCs that are placed at various locations such as streetlamp posts and poles. The nodes are equipped with radios that function as a mesh network. The data collected by the sensors is continuously transmitted to servers and made available to the public through a web application.

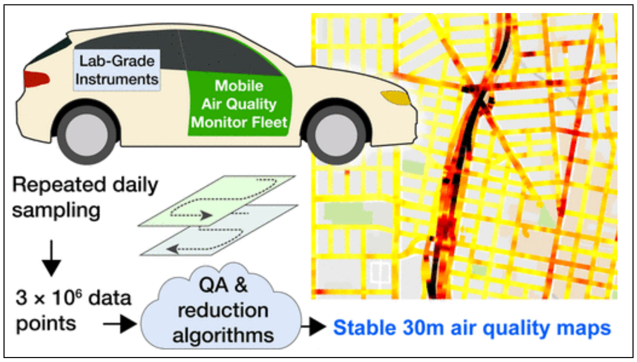
In Taiwan [48], local communities have volunteered to install 438 AirBox low-cost PM2.5 sensors of various types over most of the highly urbanized parts of the study area. Similarly, in Los Angeles, [49] 361 devices were placed at various traffic and industrial locations. They developed a machine learning model integrating spatially dense measurements from a low-cost air sensor network, highly resolved traffic data, and a suite of spatiotemporal variables to estimate hourly intra-urban PM2.5 distribution patterns in Los Angeles.

### 3.3.2 Mobile Sensing Systems

In mobile sensing system/networks, the hardware devices are small and portable. They can be easily carried from one place to another and have wireless interfaces such as Wi-Fi, BLE, and 4G. The data collected from these devices can be transferred to the cloud through a smartphone. There are two types of these networks: community-based monitoring networks as shown in Fig. 3.2, in which the devices are given to the public, researchers, and professionals to carry around, and Vehicle-based Sensing (VSN) systems, in which the sensors are carried on public transportation vehicles. The devices used in these networks can be easily carried and have a form factor like a smartphone. They are equipped with wireless interfaces and protocols to connect to a smartphone. The data collected is then transferred to a central server. The benefits of these systems include accurate and reliable data, high mobility, and low cost. However, there are also challenges such as uncontrolled mobility, redundant sampling, and a trade-off between spatial and temporal resolution. Examples of these networks include City Scanner, Google Street View vehicles, and MegaSense.

City Scanner [50] is a cost-effective way to gather a large amount of information about various aspects of a city using a modular sensing system that is mounted on top of garbage trucks. This system operates within a centralized IoT framework, allowing for near real-time visualization of collected data. During a real-world trial in Cambridge, Massachusetts, data was collected by a drive-by approach for eight months and then sent to the cloud for processing and analysis. This application demonstrates the potential of using various non-dedicated vehicles to optimize data collection in urban areas, both in terms of the amount and quality of data gathered over time and space.

Google [51] has fitted its Street View vehicles with a fast-acting air pollution measurement system and has taken readings on every street in a 30-square-kilometer area of Oakland, California. This has resulted in the largest collection of urban air quality data of its kind. The maps produced show the annual levels of NO, NO2, and black carbon at a 30-meter resolution and reveal consistent, long-lasting pollution patterns with striking small-scale variations due to local sources. These variations can be up to five to eight times higher within individual city blocks. Given that local differences in air quality




(a) Google car (VSN)

(b) Megasense (Mobile)

Figure 3.2: Mobile sensing device.

significantly impact public health and environmental justice, these findings have important implications for how air pollution is measured and managed.

The MegaSense [52] gadget is a component of the Helsinki Open Source Environmental Project (HOPE) aimed at monitoring air quality with public participation. It is a small, battery-powered device equipped with sensors, which individuals can carry to the places they visit daily. An Android app allows real-time air quality monitoring, and the data collected is sent to servers for further analysis, resulting in the creation of a district-level map of pollution data based on contributions from multiple users.

All these air pollution monitoring systems claim to have a better spatial and temporal resolution than conventional systems, but there have not been any comparisons made among them regarding real-time performance, spatial and temporal resolution, and quality of service. In the Indian context, there is a lack of this type of dense deployment and data to show better spatial and temporal resolution than conventional systems. This is an area for future research.

*Chapter 4*

# Development of End-to-End Low-Cost IoT System for Densely Deployed PM Monitoring Network

This chapter involved the development and deployment of 49 PM monitoring devices in a specific area of Hyderabad, a metropolitan city in India. The collected data was analyzed over a period of seven months, providing evidence for the necessity of dense PM monitoring device deployment. Various techniques, including mean and variance calculation, spatial interpolation, and correlation, were utilized to gain insights into temporal and seasonal variations of PM. Furthermore, an event-driven spatio-temporal analysis was conducted to examine the effects of firecracker bursting during the Diwali festival evening on PM levels.

## 4.1 Introduction

A dense IoT system with low-cost portable sensors is required to monitor outdoor PM in real time. Currently, due to the limited deployment of devices, there is limited coverage in a large metropolitan city like Hyderabad. Only 12 devices have been deployed by CPCB, covering a $650\,\mathrm{km}^2$ area. The sparse deployment also results in a lack of pollution data in places of personal interest to the public, such as residential areas, offices, and schools. To address this issue, we have deployed low-cost IoT networks for monitoring air pollution to understand local pollution more deeply.

The specific contributions of this thesis are:

1. For the high spatial resolution of outdoor PM, 49 IoT-based PM monitoring devices were developed, calibrated, and deployed at various outdoor locations.

2. The developed device is designed to be robust against the issue of data loss due to connection and power outages. The device maintains an offline cache in the event of an outage. The stored data is offloaded in bulk once the power and communication are restored.

3. All PM sensors were calibrated for seasonal variations by co-locating with a reference sensor. Also, each device was calibrated individually.

4. The devices were deployed at 49 outdoor locations covering a $4\,\text{km}^2$ area in Gachibowli, Hyderabad, India. The field locations were selected to include urban, semi-urban, and green regions. Few devices were deployed at busy traffic junctions and roadsides. The data were recorded at a frequency of every 30 seconds (sec) spanning over all the seasons for six months, thus aggregating 20.7 million data points.

5. A web-based dashboard was developed and deployed to visualize the data in real time.

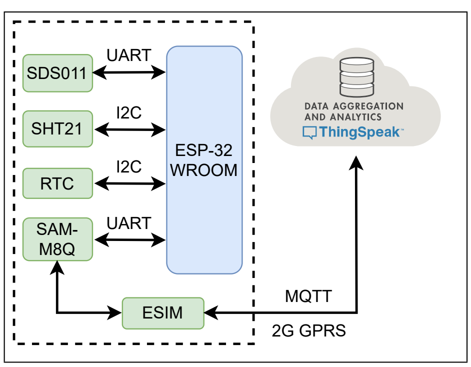
6. Different analyses were carried out by observing seasonal mean and variance, spatial interpolation, event-driven variation, and correlation.

   Results show the optimal deployment across a varied landscape and can be a key factor in identifying the release of high concentration in real-time.

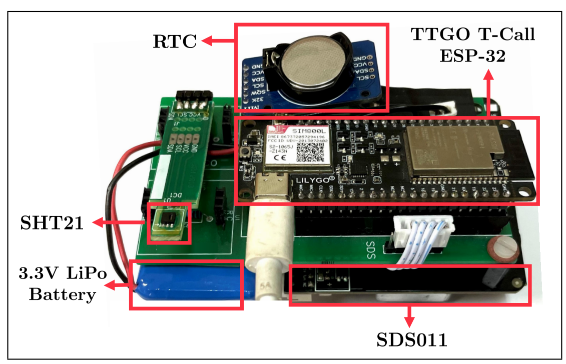
## 4.2 Hardware Implementation

### 4.2.1 Hardware Specification

For dense deployment, 49 devices were developed. Fig. 4.1 shows the hardware architecture and circuit board for the developed PM monitoring device in this work. The basic architecture consists of sensors (PM sensor SDS011 and temperature humidity sensor SHT21), a communication module (SIM800L and eSIM), a real-time clock (RTC), and a lithium polymer (LiPo) battery.




(a) Block Architecture




(b) PM monitoring circuit board

Figure 4.1: Block Architecture and the circuit board of the deployed PM monitoring device.

All these components are connected to the microcontroller TTGO T-Call ESP32. The controller reads data from all the sensors periodically every 30 sec and offloads it to ThingSpeak, a cloud-based server employing message queuing telemetry transport secured (MQTTS) over a 2G or 4G or Wi-Fi network and the data packet size is 28 bytes. The device is powered with an AC-DC power adapter and a 1000 mAh battery and enclosed in an IP65 box made of ABS filament, the enclosure offers complete protection against dust and a good level of protection against water. The form dimensions are: width =

Table 4.1: Specifications of the components used in the developed PM monitoring device.

| Component | Specification | Value |
|---|---|---|
| SDS011 [53] | Operating Voltage | 4.7 V to 5.3 V |
| | Operating Temperature | $-20\,°C$ to $+50\,°C$ |
| | Operating Rel. Humidity | 0 % to 75 % |
| | Measurement Parameters | PM2.5 & PM10 |
| | Measurement Particle Size | 0.3 to 10 $\mu$m |
| | Measuring Range | 0.0 to 999.9 $\mu g\, m^{-3}$ |
| | Serial Data Output Freq. | 1 s |
| | Maximum Current | 100 mA |
| | Signal Output | UART, PWM |
| SHT21 [54] | Operating Voltage | 2.1 V to 3.6 V |
| | Operating Temperature | $-40\,°C$ to $+125\,°C$ |
| | Operating Rel. Humidity | 0 % to 100 % |
| | Temperature Resolution | 0.01 °C |
| | Humidity Resolution | 0.04 %RH |
| | Temperature Accuracy | $\pm 0.3\,°C$ |
| | Humidity Accuracy | $\pm 2.0$ %RH |
| | Response Time | 8 s to 30 s |
| | Signal Output | I2C |
| TTGO T-Call ESP32 [55] | Operating Voltage | 3.3 V |
| | Operating Temperature | $-40\,°C$ to $+85\,°C$ |
| | Max Operating Frequency | 240 MHz |
| | RAM | 540 KB |
| | Wi-Fi | IEEE 802.11 b/g/n |
| | SIM Module | SIM800L |
| eSIM [56] | Operating Voltage | 1.62 V to 5 V |
| | Operating Temperature | $-40\,°C$ to $+105\,°C$ |
| | Available Memory | 128 KB or more |
| | Technology | 2G GPRS |
| | Bandwidth | 25 MHz |

125 mm, depth = 125 mm and height = 125 mm. The SDS011 and SIM800L modules are connected to the controller through the UART protocol, while the SHT21 and RTC are connected through the I2C protocol. The overall cost of the device after adding the cost of individual hardware components is 7000 INR (approximately 85 USD). The specifications of the individual hardware components are listed in Table 1.

### 4.2.2 Price Comparison of Consumer-Grade PM Monitoring:

Table 4.2 lists the price for constructing the AirIoT PM monitoring device and compares it with other low-cost consumer-grade PM monitoring devices. The monitors were selected considering available online devices for the measurements of outdoor PM quality.

Table 4.2: Low-cost consumer-grade monitors and the associated price.

| **Device** | **Retail Price (INR)** |
|---|---|
| AirIoT | 7000 |
| Airveda | 35000 |
| Atmotube PRO | 15000 |
| Prana CAAQMS | 64900 |

### 4.2.3 Working Mechanism of the Device

Fig. 4.2 illustrates the flowchart of the sensing algorithm developed to avoid data loss in the event of a connection outage. The microcontroller first reads the sensed data every 30 sec; however, the time to offload the sensed data depends upon the network. Next, the controller checks the network connectivity for pushing the data to the server. If the network is available, the data is transmitted instantaneously. However, if the network is unavailable, the data is stored locally in a part of the microcontroller RAM until the device reconnects to the network. Note that the size of microcontroller RAM is 540 KB and part of it is used for code and header files (created while pushing data), while the part of the remaining memory can be used to store data. We define $S = 20000$ as the maximum number of data points that can be stored. Every stored data point contains the value of sensed parameters and the time of sensing. Once the connection restores, the stored data is uploaded to the cloud server in a bulk transmission and subsequently cleared from the device. In the case device memory is filled with back-logged data in the event of a long connection outage, i.e., if the number of stored data points ($s$) is equal to the $S$, the data is cleared in first-in-first-out (FIFO) format to make space for the new incoming data.

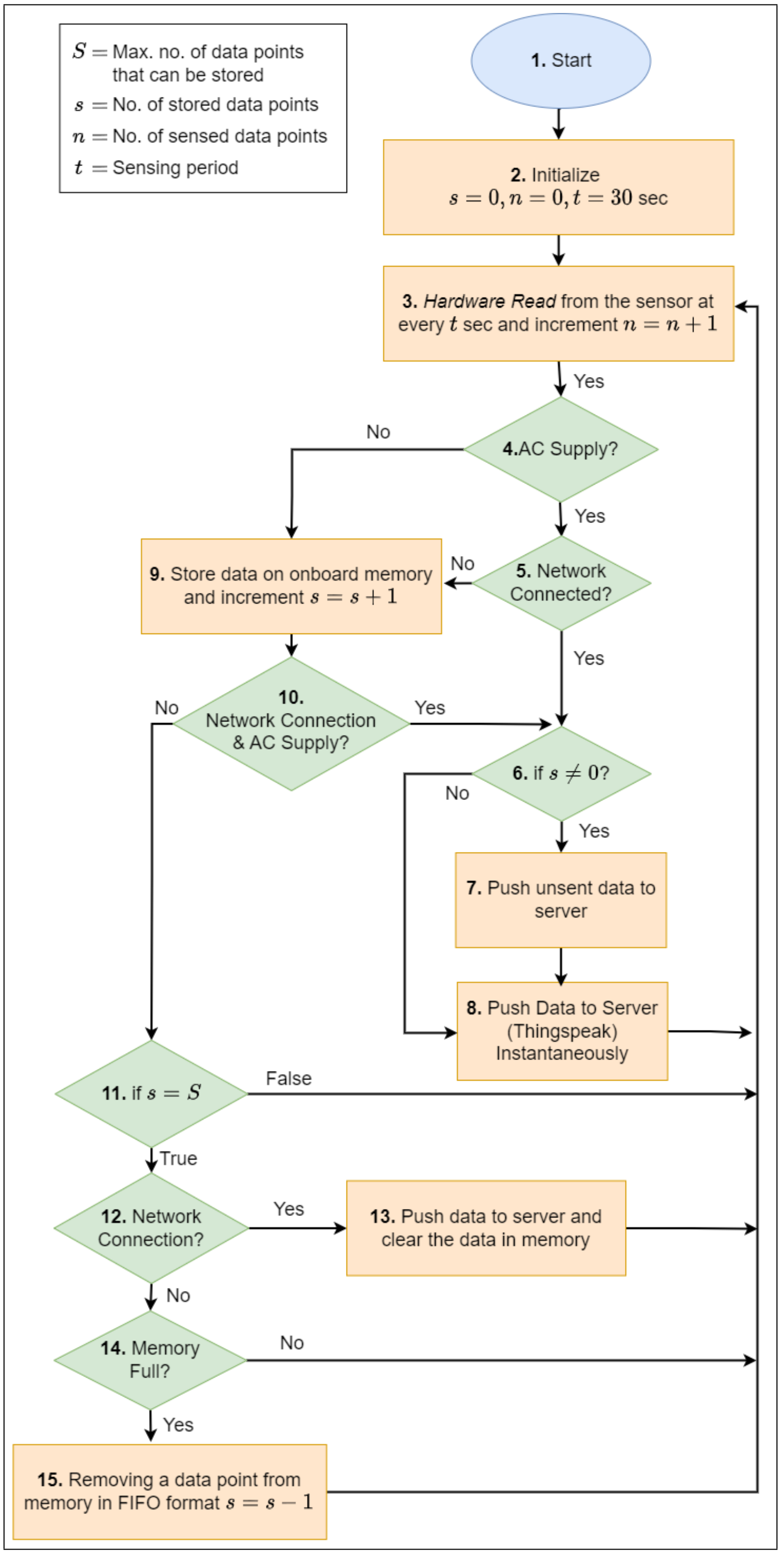



Figure 4.2: Working mechanism of the device.

Table 4.3: Deployment Setup

| **Device** | **No. of Location** | **Location Type** | **Network Type** |
|---|---|---|---|
| IIITH (AQ) | 43 | L1 (07 devices)<br>L2 (05 devices)<br>L3 (15 devices)<br>L4 (16 devices) | Wi-Fi (2 devices)<br>2G eSIM (32 devices)<br>4G Jio-Fi (9 devices) |
| Airveda (AV) | 6 | L1 (04 devices)<br>L2 (01 device)<br>L3 (01 device) | Wi-Fi (4 devices)<br>2G eSIM (2 devices) |

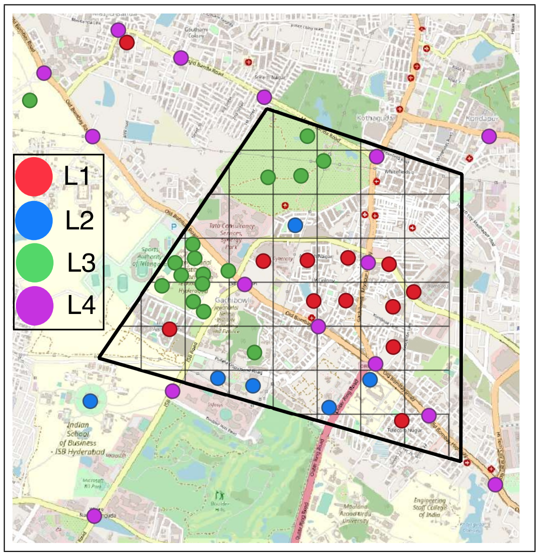
## 4.3 Network Deployment Strategy

The deployment was done in the Gachibowli region of Hyderabad, the capital city of the Telangana state and the fourth largest populated city in India [7]. Fig. 4.3a shows the plan for the field deployment of devices, while Fig. 4.3b shows an example of the deployed device at one of the locations. A total of 49 devices were deployed in a region of approximately $4\,\mathrm{km}^2$ to understand the variation of PM across different environments and areas.

Based on the landscape pattern of Gachibowli [57], [58], the entire region is divided into three categories: urban, semi-urban, and green. A few devices have also been deployed at busy traffic junctions and roadsides. Fig. 4.3a shows the deployment plan of all the devices with exact locations of the following location types:

- Location type L1: Urban region
- Location type L2: Semi-urban region
- Location type L3: Green region
- Location type L4: Traffic junctions and roadsides poles

The $4\,\mathrm{km}^2$ area has been divided into approximately 42 boxes. Every square box in Fig. 4.3a represents an area of $400 \times 400$ m$^2$. An attempt has been made to deploy 1 device in each box depending on the availability of power, network and consent availability. However, in some boxes, more than 1 device has been deployed, as shown in Fig. 4.3a. The field deployment of devices was completed in July 2021 and the data collection started in Aug. 2021. As part of experimentation, along with the devices developed at IIITH, a few devices from an Indian manufacturer, Airveda, were also deployed [59]. Table 4.3 summarizes all the deployed devices with their network configuration.




(a) Deployment plan


(b) Example field deployment

Figure 4.3: Deployment plan covering urban, semi-urban, green region, junctions, and roadsides poles.

## 4.4 Data Collection, Preprocessing and Calibration:

Fig. 4.4 shows a flow diagram depicting different steps in collecting usable data analysis data. This involves data collection, creating the dataset, removing outliers and interpolating missing data, followed by calibration. Each of the steps involved is explained in detail.

### 4.4.1 Data Collection

To create the data set, the air quality was sensed at a frequency of $t = 30\,\mathrm{sec}$ for 43 IIITH devices. For 6 Airveda devices, $t = 1\,\mathrm{sec}$, averaged over 30 sec. All the devices were deployed for almost one year and are still deployed. However, usable data were collected for seven months (Aug. 2021, Nov. 2021, Dec. 2021, Jan. 2022, Apr. 2022, May 2022, and June 2022). The loss in the data is because the devices had to be brought back to the lab due to the frequent failure of low-cost sensors requiring regular repair and maintenance. Additionally, the devices were brought for seasonal calibration at regular intervals and to make a major upgrade in the use of ThingSpeak from MQTT to MQTTS (in Mar. 2022). A total of 20.70 million usable data points have been collected. As shown in Fig. 4.4a, the collected dataset has PM2.5, PM10, temperature and RH parameters. Hereafter, all the concentration values of PM10 and PM2.5 are mentioned in $\mu\mathrm{g\,m^{-3}}$. The temperature and RH values are mentioned in $^{\circ}\mathrm{C}$ and $\%$, respectively. Corresponding to every device, a vector of data points sent from the device is stored on the cloud server for each sensing instance having the following elements:

- **created_at**: Timestamp at which the sensor value is read. This timestamp is recorded utilizing the RTC module of the device.

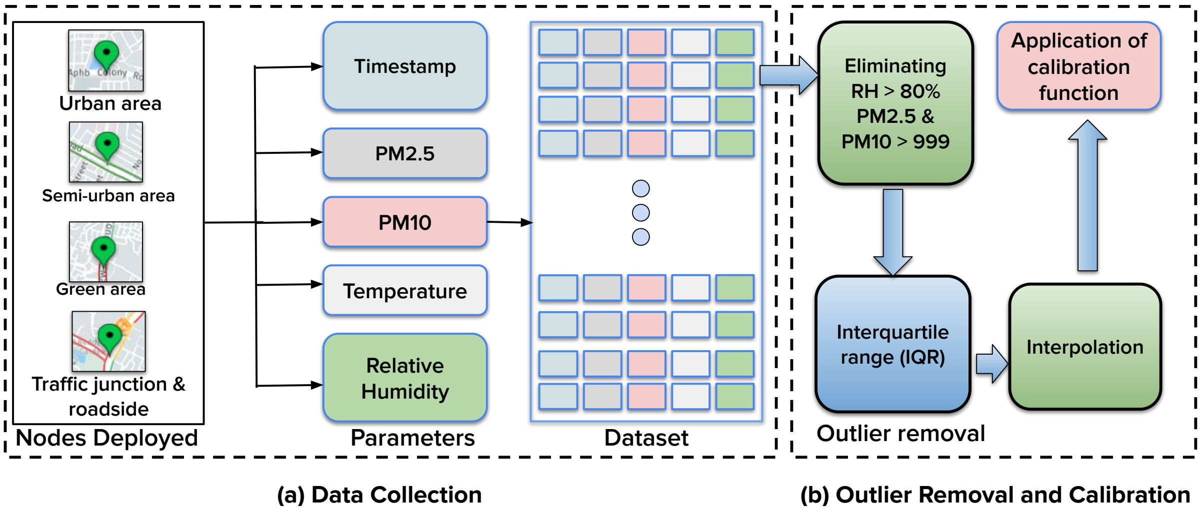



Figure 4.4: Data collection, preprocessing, and calibration.

- **PM10**: Raw concentration of PM10 read by SDS011.
- **PM2.5**: Raw concentration of PM2.5 read by SDS011.
- **RH**: Raw RH value read by SHT21
- **Temp**: Raw temperature value read by SHT21 sensor.

The size of this payload (or sensor data sent from the device) for each sensing instance is 24 bytes. In addition to this, the following static information is stored in the cloud server

- **Device_id**: ID for device identification like IIITH device as AQ-XX and Airveda device as AV-XX, where XX denotes the device sequence number.
- **Location**: Latitude and longitude according to the deployment location.

### 4.4.2 Data Preprocessing

The following methods have been employed for preprocessing the raw data received from the PM monitoring device:

#### 4.4.2.1 Outlier Removal

Environmental conditions like RH and temperature, sensor behavior and anthropogenic activities occasionally result in outliers in the sensed data. Hence the raw data received from the devices need to be preprocessed to make it statistically significant, as shown in Fig. 4.4b. PM values are unreliable at higher RH levels (RH>80%). Apart from this, errors may cause raw values to be out of the PM sensor range (0-999). These unreliable points are thus removed. In the dataset, nearly 0.5% values have been found unreliable.

Further, to identify and remove outliers, the interquartile range (IQR) method [60] is used. In this method, the data is separated into four equal parts and sorted in ascending order using three quartiles ($Q_1$, $Q_2$ (median), $Q_3$). Let the difference between the first ($Q_1$) and the third quartile ($Q_3$) be represented by the $I_{QR}$, which is a measure of dispersion. A decision range is set to detect outliers with this approach, and every data point that falls outside this range has deemed an outliers. The lower and upper values in the range are given by

$$L_r = Q_1 - 1.5\, I_{QR}, \tag{4.1}$$

$$U_r = Q_3 + 1.5\, I_{QR}. \tag{4.2}$$

Any data point less than the $L_r$ or more than the $U_r$ is called an outlier. In the collected dataset, nearly 1.4% values have been found as an outlier.

#### 4.4.2.2 Interpolation

Interpolation is a technique to estimate the missing (or removed) data point between two existing data points. In the data set, only 1.9% data is an outlier which is very less and easy to interpolate. In this work, simple linear interpolation was used for this purpose.

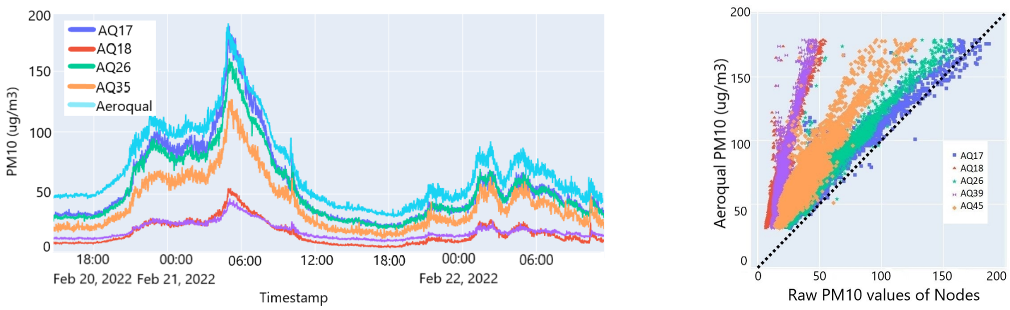
### 4.4.3 Calibration

For calibration, the low-cost PM sensors were co-located with a reference sensor (Aeroqual S500 [61, 62]) in a ventilated room for a week. Data points were collected at a frequency of 30 sec. A raw dataset of approximately 20,160 data points for each sensor was collected to perform the calibration. Fig. 4.5a shows the time series plot of PM10 averaged hourly for a few devices before deployment in the field. It can be observed that all the sensors follow the reference sensor in trend but differ with an offset in absolute value. Therefore, there is a need for calibration. It can also be observed that the offsets for each sensor are different. Although not shown to maintain brevity, the same is also valid for raw PM2.5 values.

This thesis uses simple linear regression to compensate for the difference between the values of the low-cost sensor and the reference sensor. Although many complex algorithms have been used in the past for calibration, linear regression has been chosen since it can compensate for the offset well while preserving the trend in the data, as shown in our previous work for SDS011 in [63]. The calibrated data $y(i)$ corresponding to the $i^{\text{th}}$ data point can be written as

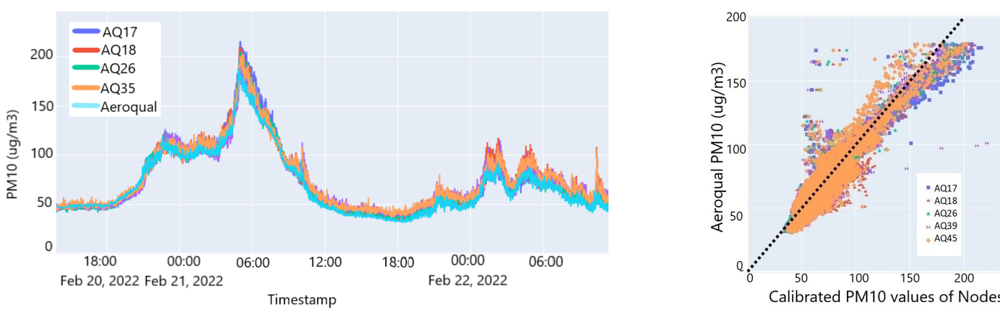
$$y(i) = m\, x(i) + c, \tag{4.3}$$

where $x(i)$ is the $i^{\text{th}}$ raw data point and $m, c$ are the learned parameters. Each sensor will have a different value of $m$ and $c$.




(a) Timeseries Plot (b) Scatter Plot

Figure 4.5: Time series and scatter plot of raw PM10 data (1-hour average).




(a) Timeseries Plot (b) Scatter Plot

Figure 4.6: Time series and scatter plot of calibrated PM10 data (1-hour average).

Fig. 4.6a shows the calibrated data of PM10 for a few devices. It can be observed that the low-cost sensors match well with the reference sensor after calibration. Similar results are obtained by training separate functions for PM2.5 as well. It can also be observed from Fig. 4.5b and Fig. 4.6b that every low-cost sensor differs uniquely from the reference sensor. The same is also concluded in [63] for 3 sensors. Therefore, PM10 and PM2.5 of each low-cost sensor have to be calibrated using unique functions for each sensor. Moreover, it is observed that the sensor behaves differently in different seasons. Hence, separate calibration functions have been calculated for different seasons by repeating the process at the season's onset.

## 4.5 Results and Analysis

This section presents mean and variance analysis results and the spatial interpolation for PM values in different seasons. Further, the event-driven variation analysis is done for the data collected during the festival of Diwali. This is followed by correlation analysis to understand the range, after which the correlation between the two points is insignificant. Note that similar observations have been made for PM2.5 as well.

### 4.5.1 Mean and Variance

Fig. 4.7 and Fig. 4.8 show the mean and variance of PM10 in monsoon (Aug. 2021), winter (Dec. 2021) and summer (May 2022). It can be observed that the mean and variance values are highest in winter and lowest in monsoon. This is expected as the surface temperature inversion (cold air near the ground and warm air on top) in winter trap PM near the ground. On the other hand, frequent rains during monsoons settle the PM, reducing their concentration in the air. It can also be observed that there is a lot of variation in the mean and the variance of the PM values among the various devices in the same geographical region, demonstrating the need for dense deployment to understand street-level pollution.

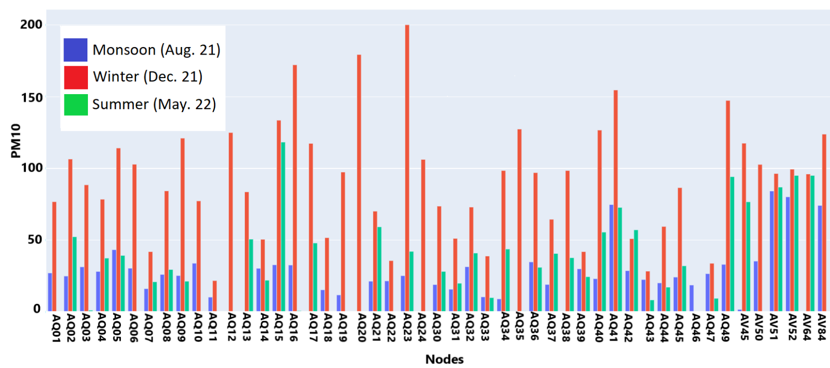



Figure 4.7: Mean and of PM10 concentration at the different locations in different seasons.

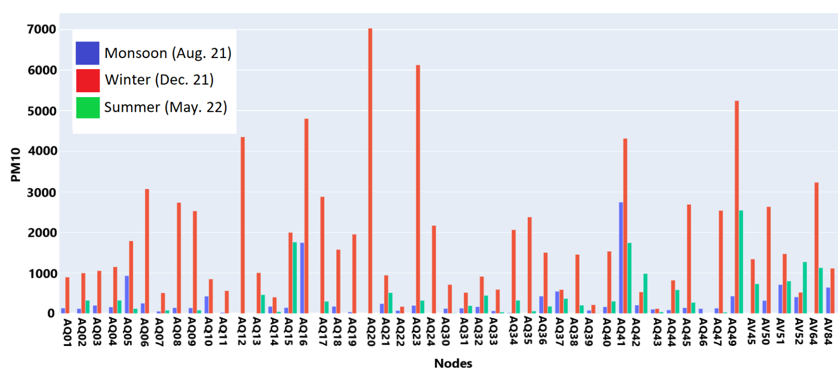



Figure 4.8: Variance of PM10 concentration at the different locations in different seasons.

In Fig. 4.7, the three devices with the highest mean PM10 values among the 49 devices are AQ23 (Traffic junction), AQ20 (Green region), and AQ16 (Roadside), while the three devices with the lowest mean PM10 values are AQ11 (Residential area), AQ43 (Roadside), and AQ22 (Roadside). Devices like AQ23 near the traffic junctions have high PM10 exposure due to heavy traffic. Similarly, devices like AQ16 near traffic lights have sluggish traffic flow leading to high mean PM10 concentrations. AQ20 is placed in high vegetation area but still shows a high mean due to ongoing construction activities in the region. Among the ones with low mean values, AQ11 is placed in a residential area with fewer anthropogenic activities. Similarly, AQ43 and AQ22 are otherwise placed on the roadside but still experience low mean PM10 due to the free flow of traffic and less anthropogenic activities.

### 4.5.2 Correlation Analysis

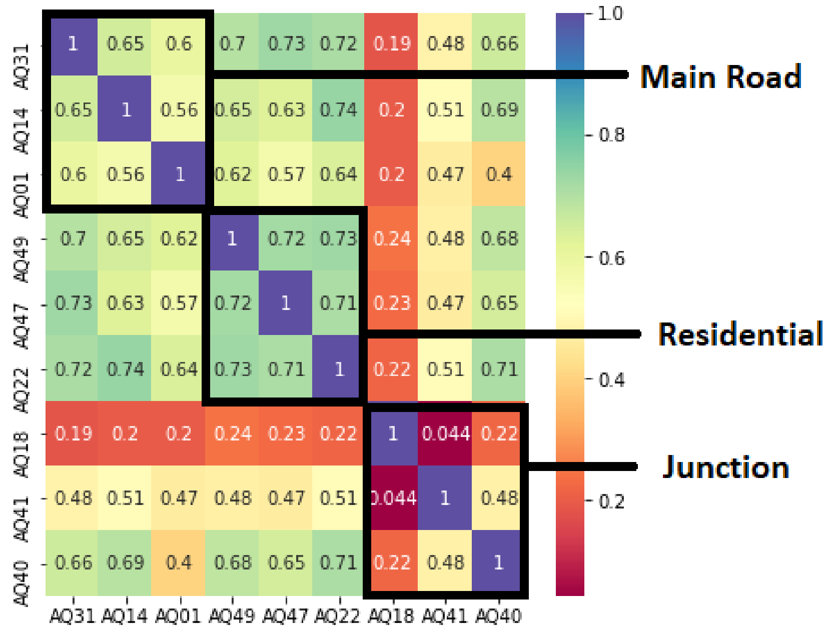



Figure 4.9: Correlation analysis of nodes at different locations

Correlation is a type of bivariate analysis that evaluates the direction and strength of an association between two variables. A statistical instrument called the correlation coefficient is used to assess how closely two variables are related when compared to one another. Different correlation coefficients, such as Pearson and Kendall, exist. One of the most widely used correlation coefficients is Pearson's, although it makes a number of assumptions about the data such as normally distributed variables, linearly related variables, the complete absence of outliers and homoscedasticity (homogeneity of variance). On the other hand, Kendall's tau is more appropriate for this study's work as it does not require the presumptions indicated above. The correlation coefficient's value ranges from +1 to -1 depending on the strength of the association. Kendall's correlation coefficients $\tau$ between the 49 sensor nodes have been calculated using hourly averaged PM10 and PM2.5 samples. The values of Kendall's correlation coefficients are shown in Fig.4.9 for a few of the nodes deployed at different location types. The coefficient varies from a value of 0.044 to 0.74 for PM10 samples. The significant variation between correlation values highlights the spatial variability between the PM values at different nodes. The maximum amount of correlation has been observed between AQ14-AQ22 for PM10. The most minor correlation is observed between nodes AQ18 and AQ41 and mainly with all other nodes for PM10. Interestingly, although AQ01 and AQ40 are merely 350m apart, they correlate less. It can be attributed to the fact that AQ01 is approximately 5m away from the main road, whereas AQ40 is deployed right at the traffic junction. This decrease in correlation shows how local activities affect the PM values and the necessity of densely deploying PM monitoring devices for better understanding.

### 4.5.3 Spatial Interpolation

Inverse distance weighting (IDW), one of the most popular spatial interpolation techniques, is used for spatial interpolation in this thesis. IDW follows the principle that closer devices will have more impact than farther devices [64]. A linearly weighted combination of the measured values at the devices is used to estimate the parameters at the nearest location. The weights are a function of the inverse distance between the device's location and the estimate's location.

Figs. 4.10, 4.11 and 4.12 are the IDW-based interpolation maps for PM10 in monsoon (Aug 21), winter (Dec 21) and summer (May 22), respectively. For all three seasons, the interpolation results are shown at three different times of the day, 1100 hrs, 1400 hrs and 2100 hrs, based on hourly averaged PM values. Similar to the observations from Figs. 4.7 and 4.8 it can also be observed in these figures that the PM concentrations are lowest in monsoon and highest in winter.

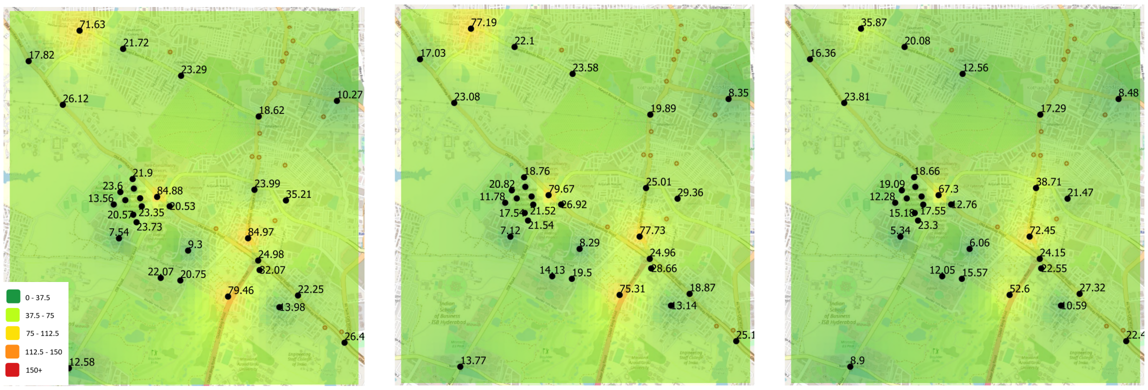
It can be observed from Figs. 4.10, 4.11 and 4.12 that PM concentration was high at 1100 hrs and 2100 hrs and low at 1400 hrs. At 1100 hrs, PM concentration was high, primarily due to heavy traffic. As the day progresses, the density of traffic decreases and the PM concentration decreases at 1400 hrs. However, with the onset of night, PM concentrations can be seen as increased at 2100 hrs, falling in peak traffic hours.

### 4.5.4 Event Driven Variation Analysis

Diwali, also known as the festival of lights, is celebrated during the start of the winter. As part of this five-day festival, people burst large numbers of firecrackers in the late evening of the third day of Diwali (4 Nov. 2021). The bursting of firecrackers leads to a significant increase in PM values during those times. Fig. 4.13 shows a time series plot of hourly averaged PM10 values for a few devices over a few days around Diwali. A few critical observations can be made from this figure. First, there was a sudden drop in the PM10 values on the afternoon of $4^{\text{th}}$ Nov. because of rain. The same has been observed on $5^{\text{th}}$ and $6^{\text{th}}$ Nov. afternoons.

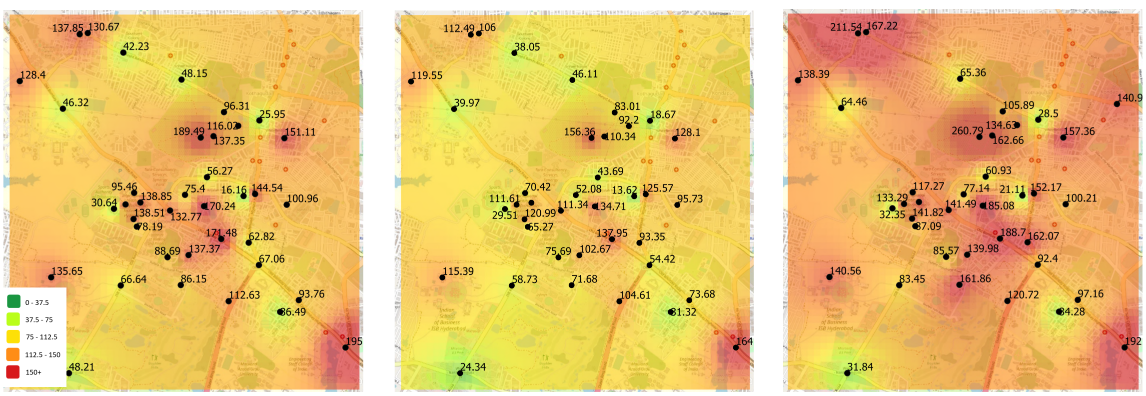
Second, a clear peak is observed for all the devices during the late evening on 4 Nov. 2021, roughly after 2000 hrs. For example, the PM10 values in AV64 increased from 40 to 307 before and after bursting crackers. This peak can be attributed to the widespread bursting of firecrackers during the festive celebrations. Third, it can be observed that the PM10 concentrations decrease sharply after a few hours, indicating that the rise was temporary and activity driven.

Further, we see the effect of sparse deployment on the event-driven analysis. Figs. 4.15, 4.16 and 4.17 show the IDW-based interpolation maps for PM10 using all 49 devices and sparse deployment of 12 and 4 devices, respectively, at different time instances on 4 Nov. 2021. For sparse deployment, 4



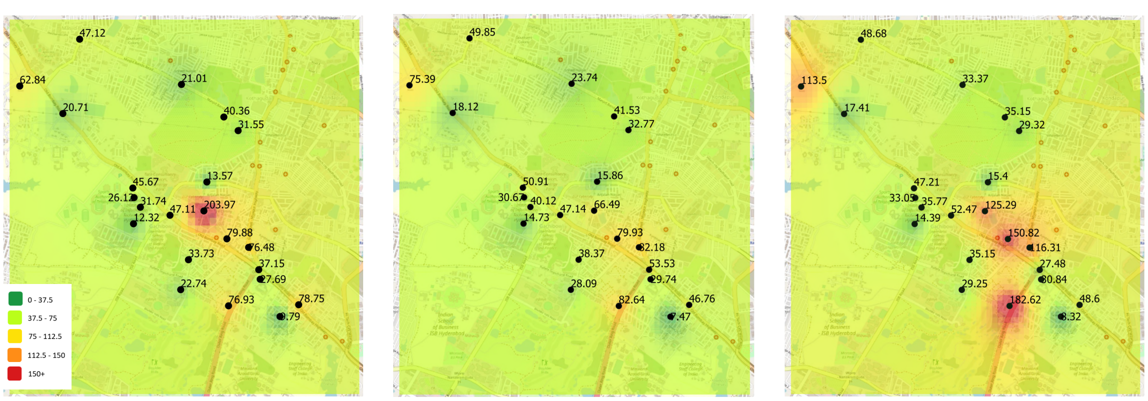
(a) At 1100 hrs (b) At 1400 hrs (c) At 2100 hrs

Figure 4.10: Spatial interpolation of PM10 values in Monsoon (Aug. 2021) using IDW.



(a) At 1100 hrs (b) At 1400 hrs (c) At 2100 hrs

Figure 4.11: Spatial interpolation of PM10 values in Winter (Dec. 2021) using IDW.



(a) At 1100 hrs (b) At 1400 hrs (c) At 2100 hrs

Figure 4.12: Spatial interpolation of PM10 values in Summer (May 2022) using IDW.

and 12 devices were chosen randomly with the constraint that they should not form a cluster and should also cover different types of regions. It can be seen from Fig. 4.15 that the interpolation plot with all 49 devices can identify the event as well as the local hotspots of pollution. Although the interpolation in Fig. 4.16 (with 12 devices within $4\,\mathrm{km}^2$) can identify the event but is not able to identify the hotspots. On the other hand, the interpolation (with 4 devices within $4\,\mathrm{km}^2$) misses the event entirely and can be misleading. The root mean square error (RMSE) for different number of deployed devices (12 and 4 devices) is calculated. The RMSE for sparse deployments like 4 devices (59.29) is significantly more compared to 12 devices (32.47).

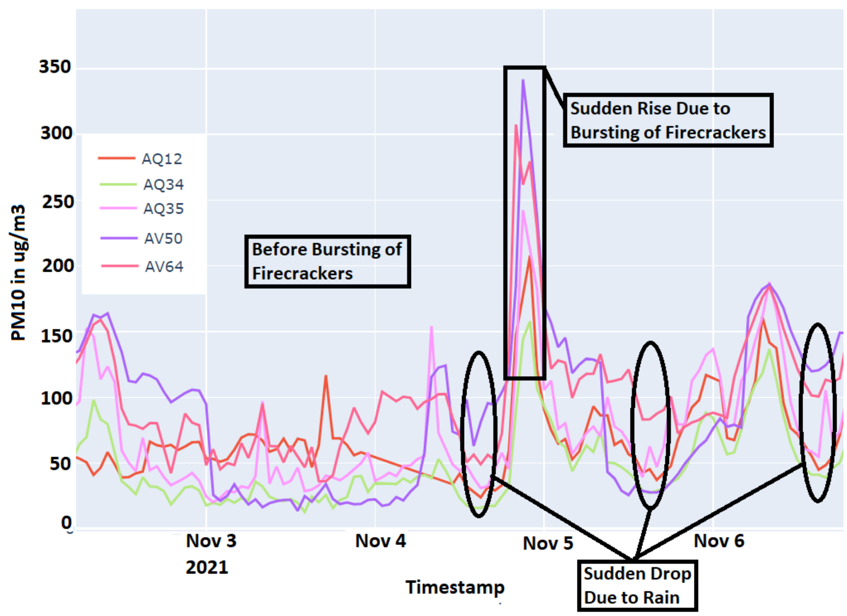



Figure 4.13: Time Series of PM10 (1-Hourly Average) showing the rise in PM10 due to bursting of firecrackers during Diwali.

### 4.5.5 Correlation Analysis

Correlation is a type of bivariate analysis that evaluates the direction and strength of an association between two variables [65]. Kendall's tau method is used as it does not require any presumptions on the data and suits the work in this study. The correlation coefficient's value ranges from -1 to +1 depending on the strength of the association. Kendall's correlation coefficients $\tau$ between the 49 sensor devices have been calculated using hourly averaged PM10 samples.

A two-term exponential fit is obtained on the correlation values when plotted against the distance between the devices. The fitted model can be written as

$$f(x) = a\,e^{b\,x} + c\,e^{d\,x}, \tag{4.4}$$

where $a = 0.4801, b = -0.0124, c = 0.7380$ and $d = -0.0001$ are the coefficients of the best fit for PM10. Fig. 4.14 shows the correlation of PM10 plotted against distance. It can be observed that

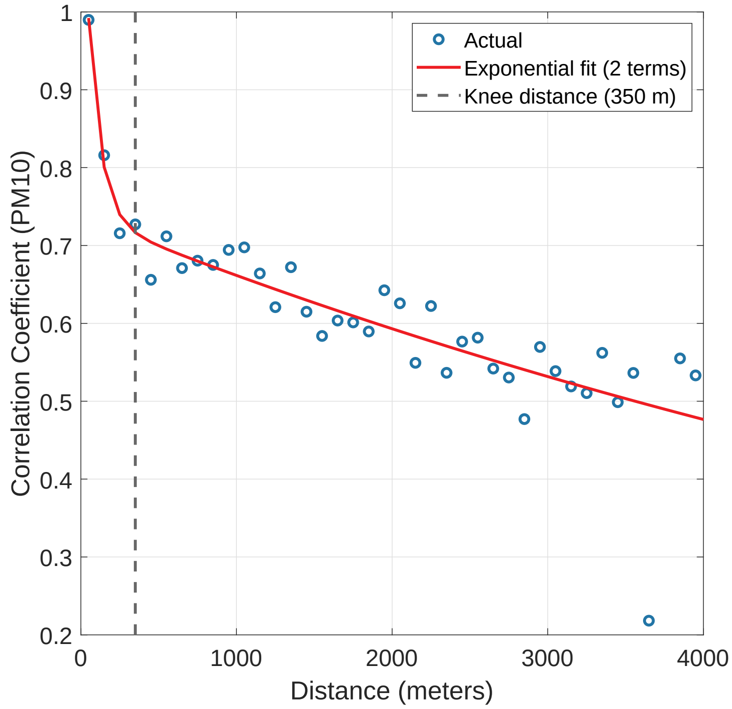
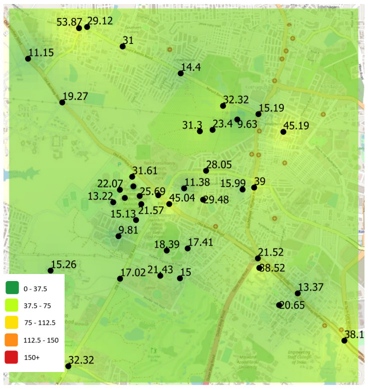
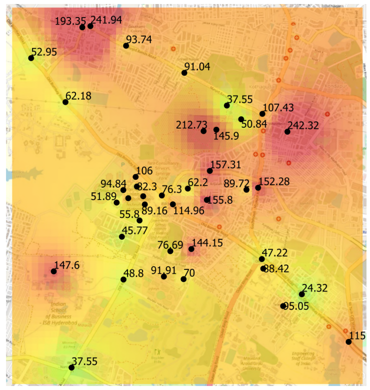



Figure 4.14: Correlation coefficient of PM10 *w.r.t* distance between the deployed devices to find the optimum distance between two deployment locations.

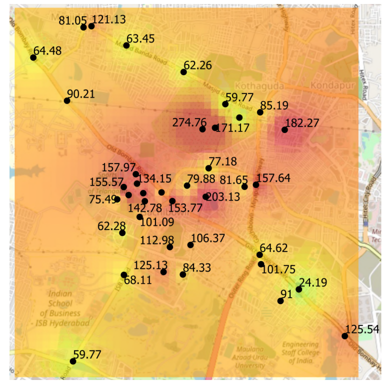
the change in the correlation coefficient under 350 meters is significantly large, after which the decline is gradual. The $\tau$ change rate between 0 to 350 meters is very fast compared to distances above 350 meters. Similar results were obtained for PM2.5 as well. It indicates that the PM monitoring devices shall be deployed at most 350 meters apart to accurately capture the spatial variability of PM.



(a) At 1700 hrs

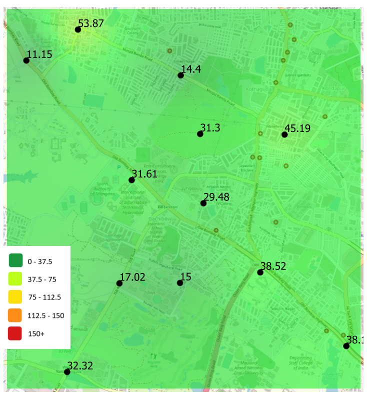


(b) At 2100 hrs

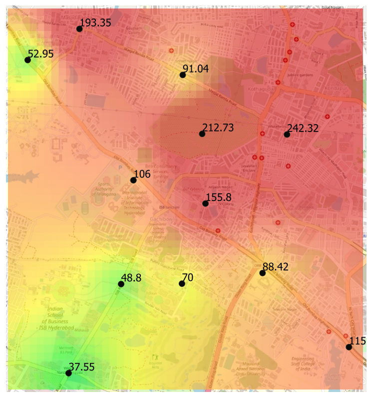


(c) At 2300 hrs

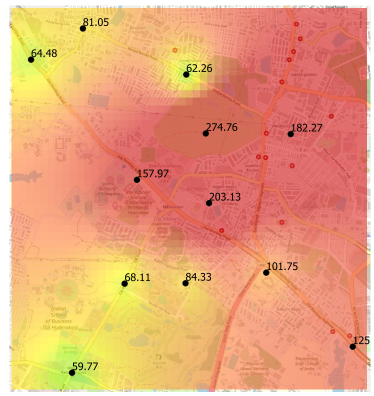
Figure 4.15: Spatial interpolation of PM10 values from densely deployed devices during Diwali 2021 using IDW.



(a) At 1700 hrs

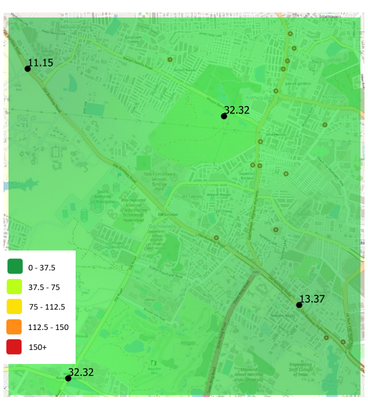


(b) At 2100 hrs



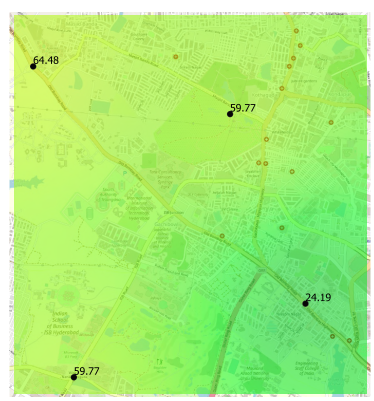
(c) At 2300 hrs

Figure 4.16: Spatial interpolation of PM10 values from 12 sparsely deployed devices during Diwali 2021 using IDW.



(a) At 1700 hrs

(b) At 2100 hrs



(c) At 2300 hrs

Figure 4.17: Spatial interpolation of PM10 values from 4 sparsely deployed devices during Diwali 2021 using IDW.

*Chapter 5*

# Challenges

The field deployment of the devices was started in April 2021 and completed in July 2021. All devices were deployed for almost one year and are still deployed. AirIoT was tested thoroughly regarding embedded hardware and firmware credibility before taking it to the field. Still, plenty of challenges were faced during the deployment that resulted in corrupt/ redundant data or, sometimes, hardware failure. Most of the significant issues were faced during the monsoon season. Also, after addressing this problem we faced a few minor challenges. A few of such challenges and issues are listed in two subsections, i.e., during deployment and after deployment from personal experience, which might be insightful to this thesis's on-field researchers and readers

### 5.0.1 Pre-Deployment Challenges

- As discussed in Chapter 4 the device in each $400 \times 400\,\mathrm{m}^2$ box was a big task and a careful selection of location was necessary.

- The process of designing and creating a small and compact device was time-consuming, primarily to ensure that the device would be difficult to tamper with and would have adequate ventilation.

- Creating a low-cost device that can provide accurate and dependable data was a challenging task, as the device needed to be both affordable and reliable

- Post-deployment calibration is challenging, particularly for a large number of devices (49 devices). A proper close pack room is required in which at a certain interval expose to different ways to generate PM like incense sticks, deodorant, room freshener, etc.

- The power issue in the botanical garden is one of the biggest challenges. We deployed a device inside the jungle where there was no power supply and laid down 40 meters of wire to provide power connectivity to the devices. The Figs. 5.1 and 5.2 show the power cable laying inside the jungle.

- While deploying the device, one of the biggest challenges was explaining its use to people, as they had many questions, which made the process a little complex. Some of the questions are mentioned below:

  - Why are we deploying?

  - What does this device do?

  - How much electricity is that device going to consume?

  - If it shows our place polluted and will you take any steps?

  - Can you show our area/place green on your dashboard?

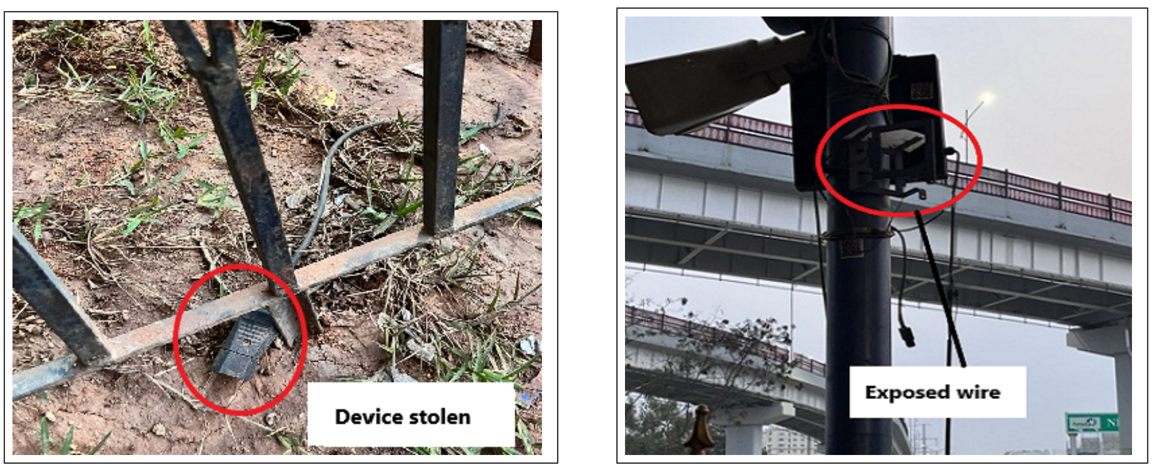
Figure 5.1: Providing power connection in Botanical garden

Figure 5.2: Providing power connection in Botanical garden

### 5.0.2 Post-Deployment

- **Faulty SDS011 sensor:** After deployment, the first problem was encountered with the SDS011 sensor. It was determined that the cause of the failure was a batch of faulty sensors. As a result, all of the devices had to be returned. This unexpected setback was frustrating for the team, as it delayed their progress on the project. However, they quickly identified the issue and took action to resolve it, which helped to minimize any long-term impacts.

- **Theft:** As a result of individuals mishandling the device, there have been several attempts at theft, including instances of leaving the 240V electrical wire exposed, which poses a shock hazard to the public, as well as turning off devices deployed on main roads and intersections. The accompanying images, as shown in the figure, illustrate this issue.




(a) Device Stolen

(b) After stealing the device wire is exposed which can be dangerous

Figure 5.3: Theft

- **Environment factor:** Working and testing the device in a controlled environment was relatively straightforward. However, when environmental factors were introduced, particularly during the monsoon season, challenges arose. The rain and high relative humidity caused dust particles to form a sticky, hard substance that clogged the inlet of the dust sensor. This resulted in the device not functioning as intended and required regular maintenance. Despite these difficulties, the team persevered and found ways to mitigate the effects of the environment on the device's performance by use of one filter, which will only allow PM2.5 and PM10 particles in side the sensor. The clogging of the inlet of the sensor due to a nest of ants and insects was the second challenge faced in deployed devices.

- **Power supply:** The devices were directly connected to the AC supply and had a battery backup of 6 hours, however, due to the unmaintained CCTV poles by the authority and frequent power out-

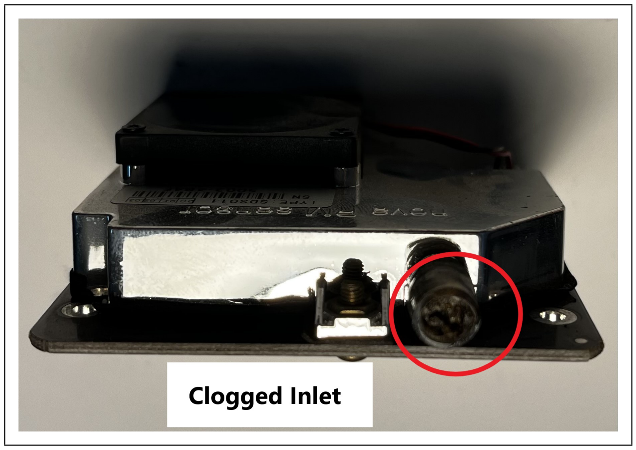



Figure 5.4: Clogged inlet of dust sensor due to rain and humidity

ages, the team faced challenges and was often dependent on third parties for maintaining CCTV poles.

- **Hardware reset:** This issue was faced only in a few nodes(2-3 nodes), especially are deployed on roads and junction. The device fails to read sensor data and starts sending NAN values. To resume the device hardware rest is required manually.

- **Seasonal maintenance and calibration:** The low-cost sensors had a limited life span and required proper maintenance. As a result, the team had to bring back all the devices before every season, including summer, monsoon, just after monsoon, and winter, to clean and perform necessary maintenance tasks. This included cleaning all 50 devices, conducting calibration, and other required maintenance procedures. The process was time-consuming but essential to ensure the proper functioning of the devices. Once maintenance was completed, the devices were redeployed to their respective locations.

Figure 5.5: Ensuring accuracy through seasonal maintenance of dust sensors

- **Firmware update:** One of the biggest challenges faced during the deployment was the upgrade of the cloud server from MQTT to MQTTS, where the devices were sending data. The task of upgrading the nodes was difficult as it needed to be done quickly to avoid significant data loss, especially during winter, which is the peak pollution time. The team worked efficiently to minimize downtime and ensure seamless data transmission from the devices to the cloud server. The upgrade was a critical step towards improving the overall performance of the system and ensuring the security of the data collected.

- **Reliability of sensors:** The low-cost IoT devices may have inherent inaccuracies that need to be considered before deployment. While techniques like intelligent sensing and calibration can address some of these inaccuracies, they are not a complete solution and may not be effective in the long term. It is important to understand the limitations of these devices and develop appropriate strategies to mitigate their inaccuracies in order to ensure the reliability and accuracy of IoT systems. For this reason, a novel approach is introduced that can report the AQI without using any PM sensor to avoid this hassle. They predict the AQI based on real-time traffic data and extra parameters such as temperature and humidity. These are machine-learning-based approaches and use real-time data to predict the AQI as discussed in section 5.0.3 of Chapter 5.

### 5.0.3 Machine Learning and Traffic-Based AQI Estimation

So far, we have discussed the methods, including some hardware that can sense and report the AQI. As mentioned at the beginning of this thesis, hardware-based solutions require frequent maintenance and cleaning. Researchers worldwide have introduced a few novel approaches that can report the AQI without using any PM sensor to avoid this hassle. They predict the AQI based on real-time traffic data and extra parameters such as temperature and humidity. These are machine-learning-based approaches and use real-time data to predict the AQI.

There has been some work in recent years in the case of estimating air pollution with the help of traffic and meteorological data using ML paradigms [66, 67, 68, 69, 70]. [66] collected a dataset from weather and air stations, including wind data, temperature, relative humidity, air pollution data, and ten agents present in the air. Fixed video cameras obtained vehicle information to collect traffic data. Various ML models were tested on the features extracted from the dataset. However, this method limits the AQI calculation to specific areas due to video camera installation to get traffic data. [67] used ML models to predict roadside $PM_{2.5}$ and $PM_{10}$ values on the dataset collected at 19 air quality monitoring sites in London, while [68] used RF models to analyze the $PM_{10}$ trends for 31 air quality monitoring sites in Switzerland. [69] used an ML-based approach to determine the air pollution level in a typical street canyon. A dataset has been collected in Zagreb city (capital of Croatia) containing $PM_{10}$, $NO_2$, and other pollutants on a daily basis for approximately three years. However, instead of finding the AQI in categories, a real number using a regression-based approach is calculated. [70] used an ML-based approach to predict the roadside particle mass concentration ($PM_{2.5}$ and $PM_{10}$) and particle number counts based on traffic and meteorological data in London, UK. The dataset was obtained from an air quality monitoring site in London and sampled hourly for a period of seven years. In this work also, instead of calculating the AQI as a category, the value of all the pollutants has been calculated as real numbers using the regression approach.

In all the above articles, the data has been obtained using meteorological sites over a period of years. However, PM values are spatially sensitive and can differ by a good margin in nearby locations. Hence, in this article, the ground truth data such as $PM_{2.5}$, $PM_{10}$, feature values such as temperature, and relative humidity are collected through a dedicated PM monitoring node [71]. The nodes are placed in close proximity so that the values obtained are as accurate as possible to the respective location. This data collection process ensures that the sensor values are co-located and accurate. Secondly, this article aims to predict the AQI category instead of a real-valued number. Calculating a level for the AQI makes it more user-friendly and intuitive.

*Chapter 6*

# Concluding Remarks

The research carried out in this thesis shows the dense IoT PM monitoring network, an end-to-end low-cost IoT system developed and densely deployed in Indian urban settings for monitoring PM with fine spatial and temporal resolution. For evaluating the dense deployment, 49 calibrated devices were deployed covering a $4\,\mathrm{km}^2$ area in Hyderabad, the capital city of Telangana state and the fourth most populated city in India. For data visualization, a web-based dashboard was developed for the real-time interface of PM data. The measurements over the year clearly show a significant difference between the mean and variance of PM values across different locations and seasons. The mean values and the variance were significantly higher in winter than in the summer and the monsoon. The IDW-based spatial interpolation results in monsoon, winter and summer at three different times show significant spatial variations in PM10 values. Furthermore, variation in PM values before and after the bursting of firecrackers on the day of Diwali is clearly visible in the results. The results also show noticeable temporal variations, with PM10 values rising by 4-5 (AV64) times at the same spot in a few hours, coinciding with Diwali celebrations and identifying the hotspots in dense deployment, which is not noticeable in sparse deployment. It has been shown that the correlation coefficient among a set of devices in the area has low values demonstrating that the PM values across a small region may be significantly different. A 350 m distance has been estimated for optimal device deployment for this data set based on insights deduced from the correlation versus distance plot. Thus, there is a need for dense deployment to understand the effect of local pollutants in the air and for improved spatial and temporal resolution of the pollutant data.

# Related Publications

**Journal (First Author):**

- Ayu Parmar, Spanddhana Sara, Ayush Kumar Dwivedi, C. Rajashekar Reddy, Ishan Patwardhan, Sai Dinesh Bijjam, Sachin Chaudhari, K.S. Rajan, Kavita Vemuri "Development of End-to-End Low-Cost IoT System for Densely Deployed PM Monitoring Network: An Indian Case Study," *IEEE Access*, 2023. *(under review)*.

**Patent:**

- Ayu Parmar, Sachin Chaudhari, Rajashekar Reddy Chinthalapani, Ayush Dwivedi, Niranjan Keesara, Mahesh Murty "Air Quality Monitoring Device for Monitoring an Air Quality with Less Energy Consumption" India Patent Application N0. 202141021963, February, 2022.

**Conference Paper (Second Author):**

- Nitin Nilesh, Ayu Parmar, Jayati Narang, Sachin Chaudhari "IoT and ML-based AQI Estimation using Real-time Traffic Data" *IEEE 8th World Forum on Internet of Things (WF-IoT)*, Yokohama, Japan 2022

# Bibliography


[1] C. C. Sobin, "A Survey on Architecture, Protocols and Challenges in IoT," *Wireless Personal Communications*, 2020.

[2] The Mortality Effects of Long-Term Exposure to Particulate Air Pollution in the United Kingdom, *COMEAP Health Protection Agency Report*, 2010.

[3] M. Greenstone, C. Hasenkopf, and K. Lee, "Air quality life index," 2022.

[4] R. Fuller *et al.*, "Pollution and health: a progress update," *The Lancet Planetary Health*, 2022.

[5] R. Khot and V. Chitre, "Survey on air pollution monitoring systems," *IEEE Int. Conf. on Innov. in Inf., Embedded and Commun. Sys. (ICIIECS)*, 2017.

[6] Y. Ying *et al.*, "A survey of wireless sensor network based air pollution monitoring systems," *Sensors*, 2015.

[7] Provisional population totals, census of India 2011; cities having population 1 lakh and above, *Office of the Registrar General & Census Commissioner, India*, (Accessed 11 Apr. 2022).

[8] National air quality index, "Central pollution control board india," (Accessed 11 Apr. 2022). [Online]. Available: https://app.cpcbccr.com/AQI_India/

[9] J. Apte *et al.*, "High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data," *Environmental Science & Technology*, 2017.

[10] C. e. a. Zhao, "Estimating the contribution of local primary emissions to particulate pollution using high-density station observations," *J. of Geophys. Res.: Atmos.*, 2019.

[11] B. Montrucchio *et al.*, "A densely-deployed, high sampling rate, open-source air pollution monitoring WSN," *IEEE Trans. on Veh. Technol.*, 2020.

[12] T. Becnel et al., "A distributed low-cost pollution monitoring platform," *IEEE Internet of Things Journal*, 2019.

[13] P. H.S et al., “The potential of commercial sensors for short-term spatially dense air quality monitoring: Based on multiple short-term evaluations of 30 sensor nodes in urban areas in korea,” *Aerosol Air Qual. Res*, 2019.

[14] R. Kamal, *Internet of Things: Architecture and Design Principles*. Mc Graw Hill Education, 2017.

[15] P. Lea, *Internet of Things for Architects*. Birmingham, UK: Packt Publishing Ltd, 2018.

[16] D. Norris, *The Internet of Things*. Mc Graw Hill Education, 2015.

[17] A. Bahgya and V. Madisetti, *Internet of Things: A Hands-on Approach*. Universities Press, 2015.

[18] “Internet Of Things (IoT),” accessed 23 Jan 2023. [Online]. Available: https://www.gartner.com/en/information-technology/glossary/internet-of-things

[19] “Internet of Things Global Standards Initiative,” accessed 23 Jan 2023. [Online]. Available: https://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx

[20] N. M. Kumar and P. K. Mallick, “The internet of things: Insights into the building blocks, component interactions, and architecture layers,” *Procedia Computer Science*, 2018.

[21] *”IoT 101 An Introduction to Internet of Things”*. Leverage, 2018.

[22] D. B. O. Hersent, *”The Internet of Things: Key Applications and Protocols”*. John Wiley and Sons Ltd, 2012.

[23] A. Al-Fuqaha et al., “Internet of things: A survey on enabling technologies, protocols, and applications,” *IEEE Communications Surveys Tutorials*, 2015.

[24] P. P. Ray, “A survey of iot cloud platforms,” *Future Computing and Informatics Journal*, 2017.

[25] L. M. Dang, M. J. Piran, D. Han, K. Min, and H. Moon, “A survey on internet of things and cloud computing for healthcare,” *Electronics*.

[26] J. Dizdarević, F. Carpio, A. Jukan, and X. Masip-Bruin, “A survey of communication protocols for internet of things and related challenges of fog and cloud computing integration,” *ACM Comput. Surv.*, 2019.

[27] J. Markendahl, S. Lundberg, O. Kordas, and S. Movin, “On the role and potential of IoT in different industries: Analysis of actor cooperation and challenges for introduction of new technology,” in *Internet of Things Business Models, Users, and Networks*, 2017.

[28] L. D. Xu, W. He, and S. Li, “Internet of Things in Industries: A Survey,” *IEEE Transactions on Industrial Informatics*, 2014.

[29] "20 Examples of Wearables and IoT Disrupting Healthcare," Dec 2020.

[30] R. K. Kodali, G. Swamy, and B. Lakshmi, "An implementation of IoT for healthcare," in *IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, 2015.

[31] "A systematic review of IoT in healthcare: Applications, techniques, and trends," *Journal of Network and Computer Applications*, 2021.

[32] D. Metcalf, S. T. Milliard, M. Gomez, and M. Schwartz, "Wearables and the Internet of Things for Health: Wearable, Interconnected Devices Promise More Efficient and Comprehensive Health Care," *IEEE Pulse*, 2016.

[33] F. John Dian, R. Vahidnia, and A. Rahmati, "Wearables and the Internet of Things (IoT), Applications, Opportunities, and Challenges: A Survey," *IEEE Access*, 2020.

[34] X.Krasniqi and E.Hajriz, "Use of IoT Technology to Drive the Automotive Industry from Connected to Full Autonomous Vehicles," *17th IFAC Conference on International Stability, Technology and Culture TECIS*, 2016.

[35] B. V. Philip, T. Alpcan, J. Jin, and M. Palaniswami, "Distributed Real-Time IoT for Autonomous Vehicles," *IEEE Transactions on Industrial Informatics*, 2019.

[36] D. Minovski, C. Åhlund, and K. Mitra, "Modeling Quality of IoT Experience in Autonomous Vehicles," *IEEE Internet of Things Journal*, 2020.

[37] M. Tagle et al., "Field performance of a low-cost sensor in the monitoring of particulate matter in Santiago, Chile," *Environmental Monitoring and Assessment*, vol. 192, no. 171, p. 18, Feb. 2020.

[38] P. Pant et al., "Monitoring particulate matter in India: Recent trends and future outlook," *The Verge*, vol. 12, no. 1, pp. 45–58, Jan. 2019.

[39] *"Long-Term, Time-Bound, National Level Strategy to Tackle Air Pollution-National Clean Air Programme (NCAP)*, accessed 2023, https://pib.gov.in/PressReleasePage.aspx?PRID=1655203.

[40] *WHO Global Air Quality Guidelines*, accessed 20 Feb. 2023, https://www.who.int/news-room/questions-and-answers/item/who-global-air-quality-guidelines.

[41] *WHO Global Air Quality Guidelines*, accessed 20 Feb. 2023, https://www.unep.org/.

[42] *European Environment Agency*, accessed 20 Feb. 2023, https://www.eea.europa.eu/.

[43] *US Environmental Protection Agency*, accessed 20 Feb. 2023, https://www.epa.gov/.

[44] *Asian Development Bank*, accessed 20 Feb. 2023, https://www.adb.org/what-we-do/themes/environment/main.

[45] R. C et al., “Improving spatio-temporal understanding of particulate matter using low-cost iot sensors,” in *Int. Symp. Personal, Indoor and Mobile Radio Commun.*, 2020.

[46] “London air quality network—summary report 2018,” *Techinal Report Environmental Research Group, King's College London: London, UK*, 2018.

[47] R. N. Murty et al., “CitySense: An Urban-Scale Wireless Sensor Network and Testbed,” in *IEEE Conference on Technologies for Homeland Security*, 2008.

[48] C.-H. Lee, Y.-B. Wang, and H.-L. Yu, “An efficient spatiotemporal data calibration approach for the low-cost pm2.5 sensing network: A case study in taiwan,” *Environment International*, 2019.

[49] Y. Lu, G. Giuliano, and R. Habre, “Estimating hourly pm2.5 concentrations at the neighborhood scale using a low-cost air sensor network: A los angeles case study,” *Environmental Research*, 2021.

[50] A. Anjomshoaa et al., “City Scanner: Building and Scheduling a Mobile Sensing Platform for Smart City Services,” *IEEE Internet of Things Journal*, 2018.

[51] J. Apte *et al.*, “High-resolution air pollution mapping with google street view cars: Exploiting big data,” *Environmental Science & Technology*, 2017.

[52] A. Rebeiro-Hargrave et al., “MegaSense: Cyber-Physical System for Real-time Urban Air Quality Monitoring,” in *15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2020.

[53] *SDS011 Nova Sensor Specifications*, accessed 23 July 2021, http://www.inovafitness.com/en/a/chanpinzhongxin/95.html.

[54] SHT21 Sensirion Temprature and Humidity Sensor, (Accessed 18 Dec. 2021). [Online]. Available: http://www.farnell.com/datasheets/1780639.pdf

[55] TTGO T-Call ESP32 module Specifications, (Accessed 18 Dec. 2021). [Online]. Available: https://docs.ai-thinker.com/_media/esp32/docs/esp32-sl_specification.pdf

[56] QoSIM UICC Assure eSIM Specification, (Accessed 18 Dec. 2021). [Online]. Available: https://sensorise.net/products/qosim-m2m-connectivity/qosim/

[57] J. C. E. Robert E. Burgan, Roberta A. Hartford, *"Using NDVI to Assess Departure from Average Greenness and Its Relation to Fire Business"*. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 1996.

[58] S. Chakraborti, S. Bhatt, and S. Jha, *"Changing dynamics of urban biophysical composition and its impact on urban heat island intensity and thermal characteristics: the case of Hyderabad City, India"*, 2017, (Accessed 13 Oct 2022). [Online]. Available: https://doi.org/10.1007/s40808-017-0324-x

[59] *Airveda Outdoor Air Quality Monitor*, accessed 23 July 2021, https://www.airveda.com/outdoor-air-quality-monitor.

[60] Y. Dodge, *"The Concise Encyclopedia of Statistics 1st ed. New York"*, 2008, (Accessed 13 Oct 2022). [Online]. Available: http://site.ebrary.com/id/10229059

[61] Aeroqual Series 500 – Portable Air Quality Monitor, (Accessed 8 Feb. 2022). [Online]. Available: https://www.aeroqual.com/products/s-series-portable-air-monitors/series-500-portable-air-pollution-monitor

[62] Aeroqual- Series 500, (Accessed 8 Feb. 2022). [Online]. Available: https://www.aeroqual.com/products/s-series-portable-air-monitors/portable-particulate-monitor#product-overview

[63] I. Patwardhan, S. Sara, and S. Chaudhari, "Comparative evaluation of new low-cost particulate matter sensors," in *Int. Conf. on Future Internet of Things and Cloud (FiCloud)*, 2021.

[64] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, *"Geographic Information Systems and Science"*. Wiley, 2005.

[65] S. Boslaugh, *"Statistics in a Nutshell"*. O'Reilly Media, Inc., 2012.

[66] C. Rossi, A. Farasin, G. Falcone, and C. Castelluccio, "A machine learning approach to monitor air quality from traffic and weather data." in *AmI (Workshops/Posters)*, 2019.

[67] A. Suleiman, M. Tight, and A. Quinn, "Applying machine learning methods in managing urban concentrations of traffic-related particulate matter (pm10 and pm2. 5)," *Atmospheric Pollution Research*, 2019.

[68] S. K. Grange, D. C. Carslaw, A. C. Lewis, E. Boleti, and C. Hueglin, "Random forest meteorological normalisation models for swiss pm 10 trend analysis," *Atmospheric Chemistry and Physics*, 2018.

[69] I. Šimić, M. Lovrić, R. Godec, M. Kröll, and I. Bešlić, "Applying machine learning methods to better understand, model and estimate mass concentrations of traffic-related pollutants at a typical street canyon," *Environmental pollution*, 2020.

[70] A. Suleiman and M. Tight, "Hybrid neural networks and boosted regression tree models for predicting roadside particulate matter," *Environmental Modeling & Assessment*, 2016.

[71] A. Parmar et al., "Development of low-cost IoT device for densely deployed PM monitoring network: An indian case study," Under Review in IEEE Access, https://bit.ly/3xGsmG8.