University of Bergen
Department of informatics

Western Norway University of Applied Sciences
Department of Computing, Mathematics and Physics

An Indoor/Outdoor Air Quality Relationship Analysis Using Internet of Things

Author: Lars Henrik Haug

Academic Supervisor: Rogardt Heldal

Western Norway University of Applied Sciences

June, 2019

Abstract

Exposure to high levels of air pollution is a significant cause of premature mortality. In Norway, people spend about 90% of their time indoors. Therefore, the relationship between indoor and outdoor concentrations of air pollution is important for the understanding of potential health effects from pollution. In addition, the indoor climate plays a critical role in the comfort of building occupants. Niche companies, with heavy-duty analyzers can be used to monitor the air quality in buildings. However, low powered and cheap IoT-devices with a long communication range offer the potential for more continuous monitoring and analysis of buildings air quality, while covering a larger geographical area.

In this research, we present an air quality monitoring system, based on a LoRaWAN network with low cost sensors. The dashboard web application includes all features needed to help increase air quality awareness in buildings. We analyze the relationship between indoor and outdoor air quality in different building types (office, campus, and residential), and investigate the difference between old and new buildings. We measured the concentrations of indoor and outdoor particulate matter ($PM_{2.5}$ and PM_{10}), carbon dioxide (CO_2), temperature, and humidity at six different buildings in Bergen, Norway. The results show that the overall air quality in the monitored buildings are within the set climate recommendations. In addition, our results indicate that there is a stronger relationship between indoor air quality and outdoor pollution in old mechanically ventilated buildings.

Acknowledgements

First, I would like to thank my supervisor Rogardt Heldal, and Øystein Thomassen at Rainfall for all the help and support throughout this thesis. You have always been available for questions and guidance whenever I needed.

I would also like to thank all the students at the Department of Informatics, especially the group at JAFU for all the support during late nights at the reading hall, laughs, and early morning coffee breaks.

Finally, I would like to thank my family, friends, and girlfriend for all the motivating words, encouragement, and advice over the years, and for stocking up the fridge once in a while when visiting. I would especially like to thank my brother, Ole, who helped improve the structure and language of this thesis.

Lars Henrik Haug
03 June, 2019

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

By the time you read the first sentence of our thesis, you will have taken at least two breaths of air. While the air may have felt ”fresh”, it contains pollutants. These pollutants are hazardous to you and others around you. An adult human being breathes about 10,000 liters of air each day. Therefore, air quality is of great importance to people’s health and why governments and the general population has become more alert to this issue in the last decades.

Recent reports from the World Health Organization (WHO) show that air pollution is the cause of premature mortality in several ways, contributing to death from heart diseases, stroke, and lung cancer among other things. In 2016, about 12.5% of all deaths worldwide were due to indoor and outdoor air pollution, making it the second leading cause of overall death after heart diseases [32]. Norwegians spend around 90% of their time inside, and during weekdays 1/3 of that time is spent at work [69]. This makes the indoor air quality of importance to the occupants. In Bergen, Norway, local newspapers publish articles yearly about the risk of high air pollution in the city center, during the coldest days of winter. The highest exposure of pollutants is, therefore, to be considered throughout the colder months of the year, with the highest variation throughout the spring and fall.

In this thesis, we will investigate how we can use Internet of Things (IoT) to monitor if the indoor air quality correlates to outdoor pollution. This will be enabled using different sensors in a long-range wireless network, where we use a web application to visualize the result. Our research will be conducted during the spring, where we will be looking at different locations in Bergen, with different attributes, to see whether or not there is a difference. We will look into the Indoor and Outdoor (I/O) relationship, analyze how well newer buildings withstand the outdoor pollution compared to older buildings, and if they are following the Norwegian laws on indoor climate. This research will be done in collaboration with the local consulting firm, Rainfall.

1.2 Goals and Research Question

In this thesis, the main goal is to analyze whether or not there is a correlation between indoor and outdoor air quality in different buildings. We want to conduct this research gathering observational data from sensors (IoT-devices), created in collaboration with Rainfall. In addition, we will build a system architecture for analyzing the data collected. In order to achieve these goals, we attempt to answer the following research questions:

- Can low powered, cheap IoT-devices with a long range be used to monitor the relationship between indoor and outdoor air quality?
- Is there a correlation between indoor air quality and outdoor pollutants in buildings around Bergen?
- Is there a stronger relationship between indoor air quality and outdoor pollutants in older buildings?
- How well do buildings in Bergen hold up to the indoor climate norms set by the Norwegian Institute of Public Health?

In addition, we have defined three sub-goals. (1) Use cheap sensors, and still get accurate data, (2) create an easy-to-use web application to visualize air quality information to building occupants, and (3) find out whether people reflect on indoor air quality. Achieving these sub-goals will help us get a better result of our prototype.

1.3 Related Work

Research on the topic of indoor air quality has emerged and researchers around the world have studied the relationship between indoor and outdoor air quality. The major focus has been on the I/O relationship of air quality in commercial buildings and schools. In 2014, A. Challoner and L. Gill wrote a research article on the indoor/outdoor air pollution relationship in commercial buildings. They monitored the concentrations of nitrogen dioxide and particulate matter inside and outside ten commercial buildings in Dublin, Ireland. Their research showed that naturally ventilated shops had the highest concentrations of particulates, but that mechanically ventilated buildings show a stronger correlation between indoor and outdoor nitrogen dioxide concentrations [19].

Blondeau et al. (2004) studied the relationship between outdoor and indoor air quality in eight French schools. Their results showed that the I/O ratio of nitrogen oxide and nitrogen dioxide were found to vary in a range from 0.5 to 1, and from 0.88 to 1, respectively. On the contrary, I/O ratios of ozone vary in a range from 0 to 0.45 and seem to be strongly influenced by the building air-tightness [15].

J. Wichmann et al. (2011) monitored the particulates, soot, and nitrogen dioxide indoor and outdoor relationships at homes, pre-schools, and schools in Stockholm Sweden. Their results concluded that the three major indoor environments occupied by children offer little protection against combustion-related particles and gases in the outdoor air [71]. These results show that there is a large number of buildings, both mechanical and natural ventilated, that has issues keeping pollution out.

There have been a number of different studies the last decades investigating the indoor/outdoor relationship all over the world among the sources mentioned above, in places such as Europe, Asia, and North America [37, 50, 36, 67]. A different point of view is to look further into the relationship between indoor air quality and outdoor pollutants in newer buildings, compared to old buildings, and look for an increase in standard when the ventilation is up to date. Researchers from Dalian University of Technology, China, and Northeast Petroleum University in Daqing, China, investigated the correlation between indoor and outdoor particulate matter of different building types in Daqing, China. Their result showed that during the summer, 80% of the indoor particles came from outside in naturally ventilated buildings, showing a significant positive correlation among indoor and outdoor particles concentration, and that there is a difference between the category of buildings [34].

Research on the topic of indoor air quality has been conducted using heavy-duty analyzers, such as Aerocet 531 aerosol profiler [37], or filters, such as a teflon filter [71], for the measurements of pollutants. These methods use expensive hardware or filters that have to be analyzed manually, instead of getting readings every hour. There has been less research completed with simple hardware devices, such as small sensors and IoT. In 2016, researchers from Molde University College in Norway and the University of Zilina in Slovakia did a systems analytics approach using wireless sensor network technologies and big data visualization for continuous assessment of air quality in a workplace environment [39]. We want to use this idea of wireless sensors to monitor the correlation between indoor and outdoor pollution.

After we started our research, two papers have been published on technology solutions for a monitoring system of air quality. Their studies are based on different proposals and prototypes, supporting our choice of Long Range technology and sensors for our prototype. Thu et al. (2018) propose a smart air quality end-to-end system as a case study in Yangon, Myanmar, collecting humidity, temperature, dust, and carbon dioxide. This is used to monitor the real-time status in the city [68]. Candia et al. (2018) present an experience in La Plata, Argentina where an air quality monitoring system for urban areas was assembled. They tested three different models of sensors, and used a scalable platform for IoT in the cloud, where they process the received data and monitor the network [17]. In our research, we receive our sensors observations through an IoT platform, to process the data before visualizing it.

1.4 Thesis Outline

Chapter 1: Introduction

In Chapter 1 we introduce the background of the problem statements, related work, and our motivation behind why we are looking further into our stated problems. We also present our goals and research questions behind the thesis.

Chapter 2: Background

In Chapter 2 we describe the theoretical background of air pollution, including indoor air quality, an air quality index, and air pollution sources. Furthermore, we clarify the Norwegian laws and social norms of indoor air quality, and explain Internet of (IoT) and Low-Power Wide Area Networks (LPWANs).

Chapter 3: Research Methods

In Chapter 3 we introduce the research methodology, design science research. We explain how the sampling and collection of data occurred. Furthermore, we describe how we use statistical analysis to analyze the indoor/outdoor relationship of air quality in buildings, and how we compare indoor climate. In addition, we present how the different sub-goals are achieved.

Chapter 4: System Architecture

In Chapter 4 we provide relevant information about the system architecture and used technologies in our prototype. We look further into the Perception Layer, Network Layer, and Application layer of our three-layered IoT architecture, including technologies such as LoRaWAN, The Things Network, and Amazon Web Services.

Chapter 5: Design and Implementation of Prototype

In Chapter 5 we describe how we design and implement the last part of our system architecture; the prototype. We further detail and explain the hardware used in the perception layer, and demonstrate the dashboard web application.

Chapter 6: Results and Evaluation

In Chapter 6 we present the evaluation of our collected data, web application, and people's reflections on indoor air quality. We visualize and discuss the results from our data collections, comparing the different sub-sets in our population.

Chapter 7: Conclusion

In Chapter 7 we discuss our findings, summarize the result, and conclude. In addition, we present future work that can be conducted to enhance the research.

Chapter 2

Background

In this chapter, we define air pollution and the different pollutants, particulate matter and carbon dioxide. We present the background of the laws of indoor air quality in Norway. In addition, in order to understand the design and implementation of our prototype, we give a background of the Internet of Things and Low-Power Wide-Area Networks.

2.1 Air Pollution

Air pollution occurs when harmful substances are introduced to the atmosphere. It can be harmful to both human health and the environment when a mix of particles and gases reach a high concentration. The terminology of air pollution splits into the matter of particulates and gases [70]. Air pollution is a local, pan-European and hemispheric issue, where pollutants released in one country may result or contribute to poor air quality elsewhere. This happens when the pollutants are transported in the atmosphere. On a general basis particulate matter, ground-level ozone and nitrogen dioxide are considered having the most effect on human health. However, carbon dioxide is the most common of the greenhouse gases, followed by methane [4]. Air pollution concentrations are expressed as either μg / m^3 (milligrams per cubic meter) or ppm(vol.) (parts per million by volume) [70].

2.1.1 Indoor Air Quality

Indoor Air Quality (IAQ) is the quality of the air within and around buildings and structures. IAQ can be affected by both gases or masses, like carbon dioxide and particulate matter. Poor air quality inside buildings affects occupants health and comfort, and understanding how to control these pollutants will reduce the exposure and lower the risk of health effects [8].

We spend a lot of our time indoors at work, school, and home. Many of us spend up to 90% of our day indoors, according to Norwegian Institute for Air Research (NILU) [33]. The indoor climate is important for health, well-being, productivity, and learning. Poor indoor climate can contribute to diseases and increased health problems. Extra vulnerable are children and people with respiratory diseases, allergies, and hypersensitivities. Healthy people may also experience repeated respiratory infections, headaches, fatigue, sore mucous membranes, and reduced concentration and workability when they stay in buildings with poor indoor climate for a long time [33]. Most common ailments and diseases are skin and mucosal irritation, headaches and odor complaints, respiratory diseases and allergic reactions, and worsening of respiratory infections [26].

In order to have a good indoor climate, there are not just the different pollutants that have to be at certain levels, temperature and humidity are important factors as well [13]. Temperature affects humans in two ways, it creates a comfortable and healthy living environment and it may accelerate mold and bacteria growth. Humidity also affects humans in two ways. High air humidity will increase chances of mold and feeling damp, while low air humidity is associated with dry throat, dry skin, and chapped lips. According to the Norwegian Labour Inspection Authority, room temperature should be around 22°C and the ideal relative air humidity below 60% in occupied buildings [13].

2.1.2 Air Pollution Sources

Outdoor air is often referred to as ambient air. The common sources of outdoor air pollution are emissions caused by combustion processes from motor vehicles, solid fuel burning, and industry. Other pollution sources include smoke from bushfires, windblown dust, and biogenic emissions from vegetation [44]. There are four categories within air pollution sources:

Mobile sources, stationary sources, area sources, and natural sources. Mobile sources include cars, buses, planes, and other moving sources. Stationary sources are man-made sources such as industrial facilities, factories, power plants, and oil refineries. Area sources can include agricultural areas, cities, and wood burning fireplaces. Natural sources include wind-blown dust, wildfires, and volcanoes [5].

Indoor air pollution, just as outdoor pollution, can come from sources outside the home, such as emissions from transport or smoke from neighboring wood heaters, and from sources within homes. The indoor sources are affected by building materials, people, work activities, cleaning, maintenance, and ventilation. The importance of the factors will vary both in time and from building to building [5].

2.2 Pollutants

Pollutants are substances which spreads into the air, water or soil. A pollutant may cause long- or short-term damage by causing health effects, or interfering with human comfort or amenities. Substances spread by nature itself, e.g., volcanic eruptions, is not considered as pollution. In Norway, the Pollution Control Act is established to protect the environment and nature against pollution [11].

2.2.1 Common Indoor and Outdoor Pollutants

Researching the relationship between indoor air quality and outdoor pollutants, we have to look further into what type of pollutants are the most interesting. There are a number of different sources of air pollution, which creates different types of pollutants. The most common outdoor pollutants are Particulate Matter (PM), ground-level Ozone (O_3), Nitrogen dioxide (NO_2), and Sulfur dioxide (SO_2) [44]. Some of the major indoor pollutants that are reducing the indoor air quality are Volatile organic compounds (VOCs), mold, Carbon monoxide (CO), and secondhand smoke [8]. Carbon dioxide (CO_2) is also a factor to bad indoor climate. Indoor levels that are unusually high can cause drowsiness, headaches, or make occupants function at lower activity levels. In this thesis, we will monitor PM and CO_2 as pollutants for our air quality analysis, as these substances are common within cities, and affect people's health and comfort.

Particulate Matter

PM is a mixture of solid or liquid matter at a microscopic level that gets into the air. These particles, once inhaled, can affect our hearts, lungs and cause serious health effects. Sources of PM can be natural or caused by human action. PM is usually divided into different subtypes like suspended particulate matter, thoracic and respirable particles and inhalable coarse particles [9].

Inhalable coarse particles is a type of PM that cause most problems to our health. This subtype is particles that are less than 10 micrometers in diameter and cause great problems as they can get deep into our lungs and might end up in our bloodstream. We divide them into fine particles, $PM_{2.5}$ (diameter less than 2.5 μm), and coarse particles, PM_{10} (diameter between 2.5 μm and 10 μm). There are many different types of PM, and it can be made up of hundreds of different chemicals. Where some might be directly from a source, others are a result of different chemical reactions. PM can be divided into locally generated and long-transported PM [51]. In Norway, the most common sources to PM are traffic, wood heaters, and industry [43].

According to the World Health Organization (WHO), exposure to PM in the air has been linked to several different health outcomes in Europe. Effects related to short-term exposure are lunge inflammatory reactions, respiratory symptoms, adverse effects on the cardiovascular system, and an increase in medication usage, hospital admissions, and mortality. Effects related to long-term exposure are an increase in lower respiratory symptoms, reduction in lung function in children, an increase in chronic obstructive pulmonary disease, reduction in lung function in adults, and reduction in life expectancy. Studies have reported significant associations between $PM_{2.5}$ concentrations and adverse health effects [46].

Carbon Dioxide

Carbon dioxide is the long-lived greenhouse gas that is the dominant contributor to climate emissions worldwide. CO_2 is an essential ingredient in photosynthesis, the process where plants make energy and occur naturally in the atmosphere. Rapid, deep, and persistent cuts in CO_2 and other long-lived greenhouse gases are necessary to stabilize global temperature rise in the long term, where up to 60% of CO_2 can stay in the atmosphere for over 100 years, and 25% for over 1,000 years. According to WHO, CO_2 is the largest contributor to climate change [45]. In this thesis, we are not researching the effect CO_2 has on climate change, rather than its direct negative effect on human health.

There are different sources to CO_2, as it occurs naturally. According to the International Energy Agency (IEA), emissions have more than doubled since the early seventies and increased by around 40% since 2000. Most of these increases are linked to increased economic output. Electricity and heat generation were the largest sources of emissions in 2016, accounting for 42% of the global total, while transport accounted for one-quarter of total emissions in 2016 [7]. Norway's emissions of CO_2 is mainly caused by the combustion of oil products, gas, and coal [38].

CO_2 is a major greenhouse gas and cause climate change, but it can also impact human health. The levels of CO_2 in the air will differ by the air exchange rate and sources, and peoples reaction to the exposure will be different. When the CO_2 levels rise, it becomes more difficult to breathe. It can cause irritation to humans by causing headaches and drowsiness by lower levels of exposure. Higher levels of exposure are associated with headaches, sleepiness, and stagnant, stale, stuffy air. Poor concentration, loss of attention, increased heart rate, and slight nausea may also be present. Toxicity or oxygen deprivation occurs at extreme levels of exposure [29].

2.2.2 Air Quality Index

Pollutant	**Index level** (based on pollutant concentrations in mg/m3)				
	Good	**Fair**	**Moderate**	**Poor**	**Very poor**
Particles less than 2.5 μm ($PM_{2.5}$)	0-10	10-20	20-25	25-50	50-800
Particles less than 10 μm (PM_{10})	0-20	20-35	35-50	50-100	100-1200
Nitrogen dioxide (NO_2)	0-40	40-100	100-200	200-400	400-1000
Ozone (O_3)	0-80	80-120	120-180	180-240	240-600
Sulphur dioxide (SO_2)	0-100	100-200	200-350	350-500	500-1250

Table 2.1: European Air Quality Index - Bands of concentrations and index level. [6]

The European Environment Agency (EEA) created a European Air Quality Index (AQI). This AQI was created to allow users to understand more about current air quality where they live, work, or travel [6]. In 2005, WHO published an updated version of their Air Quality Guidelines. It offers global guidance on thresholds and limits for key air pollutants that pose health risks. The European AQI is based on the guideline values of five key pollutants presented in Table 2.1.

EEA member countries measure, collect, and reports official data every hour. This data is used by the index to create an 'up-to-date' air quality index, for European citizens. The index corresponds to the poorest level for any of the five pollutants according to the following scheme [6].

2.3 The Norwegian Laws on Indoor Climate

To better understand what laws the employers and building owners have to follow in terms of indoor climate, we need to better understand the relevant Norwegian laws that cope with indoor climate; the Workplace Regulations and the Working Environment Act. The purpose of the Working Environment Act is to ensure safe working conditions among workers and to ensure that the working environment forms a basis for a health-promoting and meaningful work situation. This applies to all employees and contains provisions about employers and employees' obligations with respect to ensuring an acceptable working environment. Enterprises are required to have safety delegates and working environment committees, and some enterprises are required to have a corporate health service where necessary [24]. The purpose of the workplace regulation is "to ensure that employees health, safety, and welfare are safeguarded by adapting and designing workplaces and work premises for the work that is performed, to the individual employee and to special risk factors" [14].

The Working Environment Act is not definite when it comes to indoor climate issues. In § 4-4, we can read that factors such as indoor climate shall be "fully satisfactory with regard to the employees' health, environment, safety and welfare" [12]. According to the Workplace Regulations § 2-14, "work premises must be designed and furnished so that each individual workplace, personnel rooms, etc. have a satisfactory climate with regard to temperature, humidity, air quality and unpleasant odors, and protection against toxic or hazardous substances. When assessing the climate and air quality, consideration shall be given as necessary to the physical strains that the employees are exposed to". There are not any set limits or recommended values for pollutants according to Norwegian laws. The Norwegian Labor Inspection Authority has, therefore, published a guide to climate and air quality at the workplace, which is based on WHO guidelines to indoor air quality [13].

2.3.1 Social Norms

Norwegian Institute of Public Health (NIPH) has produced several recommended norms for some of the most common and important air pollutants. The Norwegian Ministry of Health and Care Services gave NIPH the task of revising the indoor air quality standards to harmonize these with the health-based air quality criteria's that apply to outdoor air. The norms are not formally established by the ministry, but NIPH has a high professional standing and is, therefore, a provider of the set norms [25].

Pollutants		
$PM_{2.5}$	25 μg / m^3 (24-hour mean)	10 μg / m^3 (annual mean)
PM_{10}	50 μg / m^3 (24-hour mean)	20 μg / m^3 (annual mean)
NO_2	200 μg / m^3 (1-hour mean)	40 μg / m^3 (annual mean)
SO_2	500 μg / m^3 (10-minute mean)	20 μg / m^3 (24-hour mean)
O_3	100 μg / m^3 (8-hour mean)	
CO_2	1800 μg / m^3	

Table 2.2: Indoor climate recommendations by NIPH [25].

2.4 Internet of Things

Internet of Things (IoT) is a concept of connecting any device with an on/off switch to the internet. This includes everything from lamps, and washing machines, to cellphones and wearable devices. It simply means taking all the "things" that you want and connect them to the internet. It can be classified as a collection of devices equipped with sensors and processors that can communicate with each other to serve a purpose [52].

IoT could help us in overcoming top global challenges; The aging population by helping our increasingly digital population live at home longer, safer and healthier. Climate change, by real-time and more accurate environment monitoring. Rapid urbanization, by providing better insight into human mobility patterns and air pollution levels. Depletion of energy, resources by increasing efficiency in energy usage and conserving energy. Food and water shortages, by revolutionize farming and increase the availability of fresh drinking water [35]. To realize this vision, "things" need to sense their environment and share this information to a human-made system that enables intelligent decision-making [52].

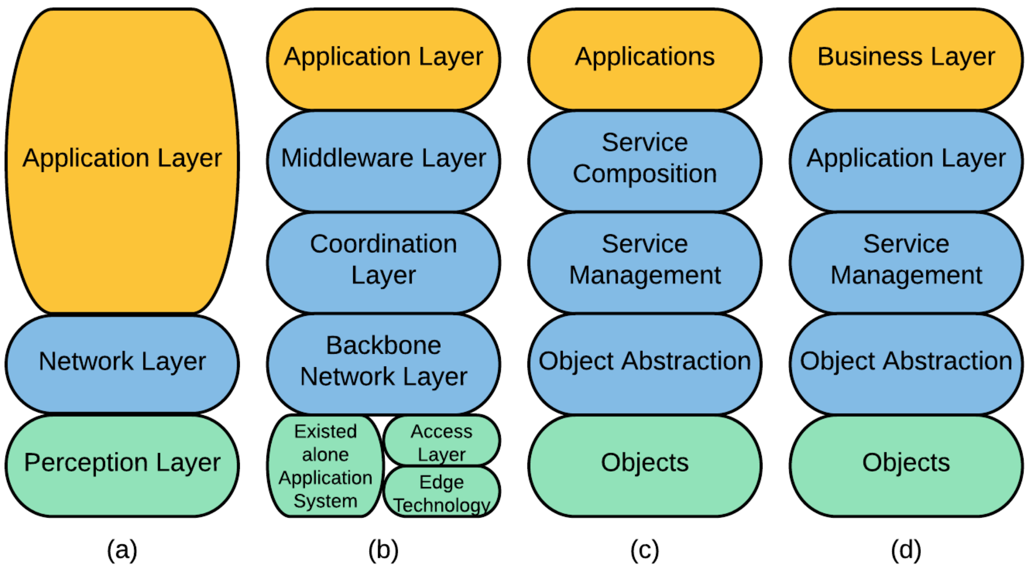

Figure 2.1: Types of IoT architecture [27]: (a) Three-layer. (b) Middle-ware based. (c) SOA based. (d) Five-layer.

The typical IoT architecture is a layered-oriented architecture, which intends to structure components and subsystems in regard to their task within the overall system. There are a number of different models, where the most common architectures range from three to five layers. Four of the most common architectures are illustrated in Figure 2.1, including Three-layer, Middle-ware based, Service-oriented architecture (SOA) based, and Five-layer architecture. The general concept of the architecture is to provide a proven guideline for the development of an IT application [27]. For our prototype, we adapt the three-layer architecture. The perception layer contains the physical properties of things around us that are part of the IoT, the network layer is responsible for processing the received data from the perception layer, and the application layer uses the processed data from the previous layer [1].

There are a number of options when connecting to IoT. Wi-Fi, Bluetooth, Radio Frequency Identification (RFID), and Low-Power Wide-Area Network (LPWAN) are some of the possible ways to connect your IoT devices to the internet. In our research, the IoT devices we use to monitor air pollution is based on a LPWAN technology, LoRaWAn, which is explained further in Chapter 4. According to Al-Sarawi et al. (2017), the perfect connectivity option has the optimal trade-off between power consumption, data rate, range, and security [55].

2.4.1 Low-Power Wide Area Network

Low-Power Wide Area Network (LPWAN) is a wireless area network technology that interconnects low bandwidth, battery-powered devices with low bit rates over long ranges. These sets of features are what makes LPWAN one of the most common technologies used for Machine-to-Machine (M2M) and IoT networks. Approximately one-fourth of an overall of 30 billion IoT/M2M devices are to be connected to the Internet using LPWANs. There is a variety of applications where LPWAN can be used. These sectors include but are not limited to smart city, personal IoT applications, smart grid, and logistics [52].

LPWANs are unique compared to traditional technologies like Bluetooth and Wi-Fi, which are adequate for consumer-level IoT applications. IoT applications in industrial and commercial deployments benefit from LPWAN tradeoffs of low power consumption, geographical coverage, and scalability. LPWANs range spreads from a few to tens of kilometers, with a battery life of ten years and beyond. LPWAN technologies are promising for low-power, low-cost, and low-throughput IoT. Long range LPWAN technologies enable devices to spread and move over a large geographical area, which creates the opportunity to turn on and off devices anywhere at any time. Achieving long range, scalability, and low power operations is at the expense of low data rate [52].

Design and Techniques

There are several competing LPWAN technologies. These technologies use a variety of techniques in order to achieve a set of features, such as low power consumption, geographical coverage, and scalability. In order to achieve a great signal strength while covering a wide area up to tens of kilometers, LPWAN technologies exploit Sub-GHz bands and special modulation schemes. Lower frequency signals experience less attenuation and multipath fading caused by obstacles, and are less congested than 2.4 GHz, that is used by most wireless technologies in the house. The result is higher reliability that enables long range and low power communication. Two modulation techniques, narrowband and spread spectrum, have been adopted by different LPWAN technologies in order to enable long range. By assigning each carrier a very narrow band, narrowband modulation makes the overall spectrum to be shared very efficiently between multiple links, resulting in simple and inexpensive transceiver design. Spread spectrum techniques spread a narrowband signal over a wider frequency band but with the same power density. More processing gained is, however, required on the receiver side to decode the signal that is received [52].

Low power consumption is key when using LPWAN technologies. In order to achieve a battery lifetime of 10 years or more, the technologies remove complexity from end-devices, use lightweight access control and duty cycling, while excluding mesh topology. The arrangement of nodes in LPWAN communication is by connecting devices directly to base stations, removing the need to process data from other nodes as done in mesh networks. Duty cycling allows LPWAN end-devices to turn off their transceivers when not required, and turned on only when data is transmitted or received. In this thesis, our LPWAN IoT devices use ALOHA as access protocol. The simplicity of this protocol makes each device simple and low cost, by offloading complexity from the end-devices. Creating more complex backend systems and base stations helps to create simple and low-cost end-devices. Using diversity techniques to accommodate as many devices as possible, is also a common technique in LPWANs to cope with the low power consumption of end-devices and achieve great scalability [52].

Chapter 3

Research Methodology

In this chapter, we take a closer look at our research method, design science research. We get a better understanding of how we sample and collect data. Including time series and observations, site locations, and the optimal placement of sensors.

We present how the statistical analysis is executed, with analyzing the I/O relationship using correlation and I/O ratio, and how we compare the different buildings indoor climate measurements compared to the set guidelines. We also present how the sub-goals are researched and evaluated.

3.1 Design Science Research

In this thesis, we adopt a design science approach as our research methodology. Design science research is a method that establishes and operationalizes research when the desired goal is an artifact or a recommendation. In addition, research based on design science can be performed in an academic environment and in an organizational context. Design science research aims to study, research, and investigate an artifact and its behavior, from an academic and organizational standpoint. It is a rigorous process of designing artifacts to solve problems, evaluate what was designed or what is working, and to communicate the results [23].

Hevner et al. (2004) define seven criteria that should be considered by researchers, in order to assist them while conducting design science research. The criteria for conducting design science research are design an artifact (1), problem relevance (2), design evaluation (3), research contribution (4), research rigor (5), design as a research process (6), and communication of the research (7). These criteria are essential because design science research demands the creation of a new artifact (criterion 1) for a specific problem (criterion 2). Once this artifact is proposed, its utility should be explained and the artifact must be adequately evaluated (criterion 3). The research contributions should be clarified for professionals interested in solving organizational problems and for the academic community to increase knowledge of the area (criterion 4). To ensure the validity of the research and expose its reliability, it is essential that investigations are conducted with an appropriate amount of rigor to demonstrate that the constructed artifact is suitable for its proposed use and that it has satisfied the criteria for its development (criterion 5). To construct or evaluate the artifact, it is essential that the researcher conducts research to understand the problem and to obtain potential problem-solving methods (criterion 6). The research results should be properly communicated to all interested parties (criterion 7) [30].

Performing design science research in information systems need a set framework in order for it to produce a successful artifact. Peffers et al. (2007) describe an effective framework for conducting design science research for use in information systems research. This framework is based on six steps [49]

- **Problem identification and motivation:** Define the specific research problem and justify the value of a solution.
- **Definition of objectives for a solution:** Infer the objectives of a solution from the problem definition and knowledge of what is possible and feasible.
- **Design and development:** Create the artifact. Such artifacts are potentially constructs, models, methods, or instantiations.
- **Demonstration:** Demonstrate the use of the artifact to solve one or more instances of the problem.
- **Evaluation:** Observe and measure how well the artifact supports a solution to the problem.
- **Communication:** Communicate the problem and solution to others.

In this thesis, we follow the structure presented roughly. In Chapter 1 and 2, we identified and explained our motivation for the stated problem. In this thesis, the artifact is our prototype of building a cheaper air quality monitoring system using IoT. The next chapters will cover objectives for a solution before we cover the design and implementation of our prototype, and a demonstration. At last, we present our results and evaluation of our research and prototype.

3.1.1 Sampling and Data Collection

Quantitative research methods are used when something needs to be measured, while, qualitative methods are used when a question needs to be described and investigated in some depth [66]. In this thesis, we, therefore, combine quantitative and qualitative research into a mixed-method research.

While performing a mixed-method research, the sampling and data collection needs to be carefully planned, in order to measure accurate data. The process of gathering and measuring information of interest is called a data collection. In order to analyze and visualize information to the user, we have to collect data. A method of data collection is simply a technique that is used to collect empirical research data, where the most important factor is to ensure that the measurements are accurate and honest. There are many methods used to collect or obtain data for statistical analysis, most commonly by observations, experiments, or surveys. Observations are the collection of information of a subject in their usual environment without altering with that environment, an experiment is a controlled study in which the researcher attempts to understand cause-and-effect relationships, and a survey solicits information from people [42]. In our thesis, we use surveys as self-administered questionnaires to evaluate our web application, and interviews to understand building occupants opinions around indoor air quality. The questionnaire and interview template is further explained in Appendix A. We use two different sensors to gather raw data as observations, and experiments comparing the relationship between different buildings.

The process of selecting a sample is known as sampling and refers to the selection of a subset in a population. The population is a complete set of elements (persons or objects) that possess some common characteristics defined by the sampling criteria established by the researchers, where the number of elements in the sample is the sample size [42]. There is a lot of different limitations when choosing a sampling method, such as time and money. Broadly, sampling methods are classified as probability samples and non-probability samples. The ideal form of sampling is a probability method called simple random sampling, where you have access to the whole population, but it is an unlikely scenario [2]. For this thesis, we use a selective non-probability sampling method, diversity sampling, to specifically seek differences between subgroups. Rather than constrain the targeting to limited groups and areas, diversity sampling spread the net as wide as possible to gain a wide range of subjects and views. It is a type of sampling where you deliberately choose members so that all views are represented, therefore, our sample consists of one old and one new building, within three different building categories.

Site Locations

In our study, we try to identify differences in a set of buildings within the city of Bergen. Rather than constrain the target to a limited set, we spread the sample wide to include new and old buildings composed of residential, campus, and office buildings. A broad study needs to identify a wide set of subjects, where diversity studies seek to discover and understand variation, while narrow studies seek to eliminate variation. The buildings and experimental rooms were selected to ensure diversity regarding the outdoor environment, building construction, and the ventilation system.

Urban air quality varies by location and depends on multiple factors like traffic and weather. In order to get the most exposure to pollutants around the measured buildings and to see the difference in different types of buildings, we decided to choose locations within the city center and spreading the sample wide. Within each category, we selected two buildings that had the same features, but with a prominent difference in year built or renovated. The locations are mapped out in Figure 3.1, and Table 3.1 explain different features and details of each building.

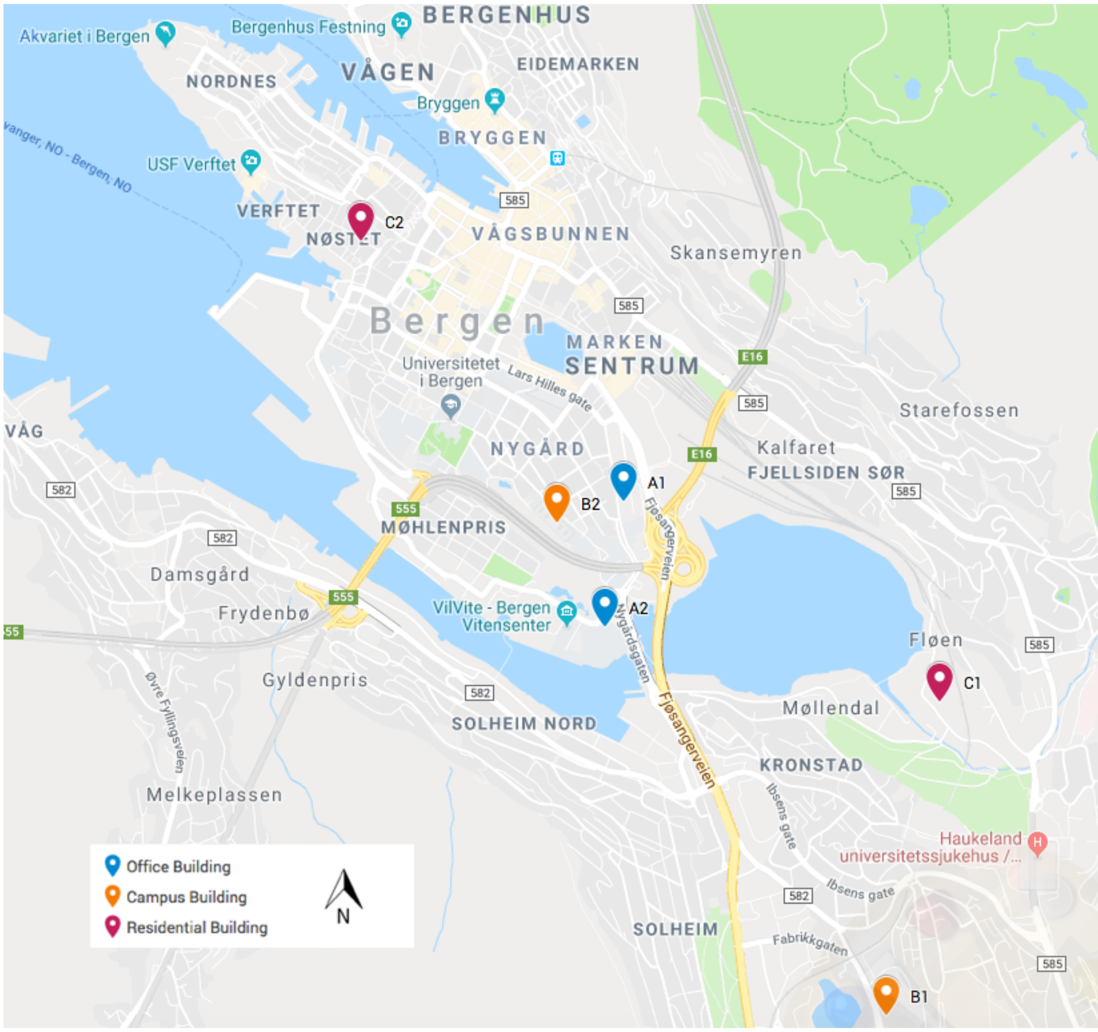

Figure 3.1: Selected locations.

Credit: Google Maps. (Retrieved: 05/1/2019).

Site no.	Building type	Ventilation	Year	Outdoor	Indoor
A1	Office building	Mechanical	2017	Roof	Hallway
A2	Office building	Mechanical	1989	Roof	Hallway
B1	Campus building	Mechanical	2014	Roof	Hallway
B2	Campus building	Mechanical	1977	Roof	Hallway
C1	Residential home	Natural	2014	Ground	Livingroom
C2	Residential home	Natural	1990	Ground	Livingroom

Table 3.1: Details and features of the monitored sites selected.

Placement of Sensors

During the collection of data, the most important factor is to ensure that the measurements are accurate and honest. In order to achieve this, we have to make sure the data is not manipulated by factors affecting the observation. Therefore, we have to carefully find the optimal solution for placement in each building.

The optimal sensor placement measuring air quality is away from factors manipulating the air while still in the range of the connected base station. When collecting data from the air inside buildings, keeping the sensors away from overcrowded rooms, windows, and ventilation holes, will prevent the data to accidentally be manipulated. Inappropriate conditions for air quality sensors are close to windows, near the entrance door, close to ventilation outlets, in areas with limited air circulation, in overcrowded rooms, and in areas where various chemicals may be presented, such as cleaning supplies.

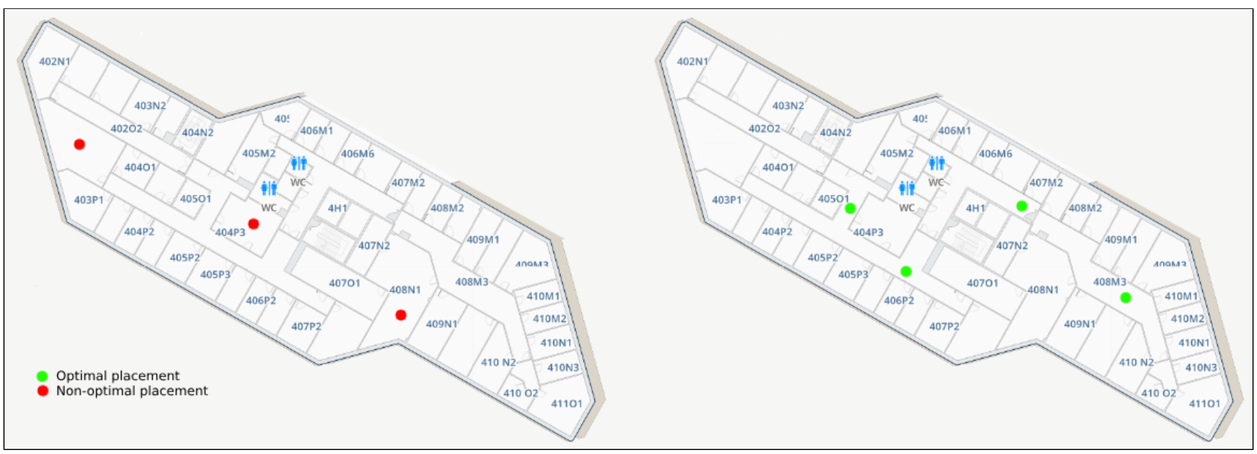

Figure 3.2: Example of optimal vs. non-optimal sensor placement in site location A2.

Credit: MazeMap. (Retrieved: 05/1/2019).

In Figure 3.2, we visualize an example of optimal vs. non-optimal sensor placement in site location A2 (office building). As all the different office spaces in the building have windows, ventilation vents, and is occupied frequently by people for long periods. Placing a sensor in these rooms can manipulate the actual air quality within the building. Therefore, we place the sensors in the building hallways, where the representation of air quality is more honest for each building. The same placements of sensors are set at the other office building, A1, and the two campus buildings, B1 and B2. The outdoor sensors were placed on top of the roof or on a roof terrace, multiple stories above the ground as close as possible to the ventilation intake.

The residential buildings, C1 and C2, has natural ventilation and therefore the sensors are placed differently. In both residential buildings, the indoor sensors were placed within a two-bedroom apartment. It was placed away from windows, and between bedrooms to be as far away from the most crowded parts of the apartments. The outdoor sensors were placed outside the windows of the apartments, in order to obtain closest air quality to the intake of where the apartments where ventilated.

Measurements

In Chapter 2, we explained what type of pollutants that cause health issues, affects humans and affects the environment. To look at how well the buildings in Bergen are withstanding air pollution from outside. We have to look at gasses and particulates that are common outdoors, impacts human health, and that will cause issues for the building occupants.

When time and money comes to short, we decided to focus our study on one of the most important greenhouse gases that also affects humans, CO_2, and one of the worst pollutants impacting human health, PM. These pollutants are common within cities, exposing buildings to high levels, making them more interesting to analyze. As the primary source of CO_2 levels in buildings is the respiration of the building occupants, the sensors need to be placed away from crowded rooms in order to see the correlation between indoor and outdoor air quality. We also measure temperature and humidity in order to analyze and compare the different building's climate.

The EEA use data from around European cities to create a European Air Quality index. Data is officially reported every hour by the countries, where the different gases are measured for hourly concentrations and particulates are based in the last 24 hours [6]. For this thesis, each building has been monitored for seven days, 24-hours, 10-12 times every hour, depending on delay from each sensor. This is in order to get a better overview of how the climate within the building is, and how it is affected by the outdoor air quality. Measuring twelve times as often as the reference station will increase the chance of accurate and honest data, and removing false positives that can manipulate our data collections.

3.2 Statistical Analysis

Statistics is a tool for converting data into information, where statistical analysis is a component of data analytics. It is the science of collecting, exploring and presenting large amounts of data to discover underlying patterns and trends. In this thesis, we want to find out if there is a stronger underlying pattern between the relationship of indoor/outdoor air quality in older buildings, and to see if the indoor climate of buildings in Bergen is within the set guidelines.

3.2.1 Analysis of Indoor/Outdoor Relationship

There are multiple different methods for analyzing the relationship between two variables. In order to find out if there is a relationship between indoor and outdoor pollution, we want to use two different methods. We want to measure the I/O ratio between the different pollutants, and also see how well each building's measurements correlate. This will give us a good indication of whether or not new buildings have a better possibility to withstand air pollution from outdoors.

Indoor/Outdoor Ratio

The indoor/outdoor ratio of the different pollutants can be used to measure the relationship between the I/O air quality. A ratio is a comparison between a pair of observations, and while you can usually obtain it by direct measurement, you might have to do some calculations to make it useful. In this case, we will look at the ratio of indoor pollutants to outdoor pollutants.

The I/O ratio can vary depending on multiple factors, and to estimate the contribution of outdoor particles to indoor levels, a physical-statistical model can be used. According to, Cyrys et al. (2004) [20], the model assumes that the amount of pollutants that enter the home without apparent indoor sources equals that leaving the home. Expressed as:

$$C_{IN} = \frac{PaC_{Out}}{a+k},$$

where C_{IN} and C_{OUT} are the indoor and outdoor concentrations, respectively ($\mu g/m^3$), P is the penetration efficiency, a is the ventilation rate (h^{-1}) and k is the deposition rate (h^{-1}). Thus, the I/O ratio (C_{IN}/C_{OUT}) is a function of the penetration efficiency and of particle losses from exfiltration and deposition (a and k) [20].

With our goal of measuring this without any additional factors, using only cheap sensors, we decided to use an easier way of presenting the ratio, without looking further into the concentration indoors. In order to compare the dynamics between indoor and outdoor air pollution in different buildings, the ratio of indoor/outdoor concentrations has been calculated according to Challoner et al. (2014) [19]:

$$I/O = \left(\frac{\sum_1^n C_{in}}{n}\right) \div \left(\frac{\sum_1^n C_{out}}{n}\right),$$

where C_{in} and C_{out} are the indoor and outdoor concentrations, respectively ($\mu g/m^3$), and n is the number of time steps.

Correlation

The simplest statistical technique for analyzing causal effects is a correlation analysis. Correlation is the process of establishing a relationship or connection between two or more variables. There are several types of correlation coefficient formulas, and we are measuring Pearson's correlation coefficient analysis of the relationship. Pearson's is one of the most commonly used formulas for calculating the correlation [54].

Correlation analysis measures the extent to which two variables vary together, including the strength and direction of their relationship. We calculated Pearson correlation coefficients (r) and 95% confidence interval (CI) between the indoor and outdoor relationship of gases and particulate matter. Correlation coefficient formulas are used to find how strong a relationship is between data. The formulas return a value between -1 and 1, where 1 indicates a strong positive relationship, -1 indicates a strong negative relationship, and a result of zero indicates no relationship at all [54]. Pearson's formula:

$$r_{x,y} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}},$$

where n is the number of pairs of scores, $\sum xy$ is the sum of the products of paired scores, $\sum x$ is the sum of x scores, $\sum y$ is the sum of y scores, $\sum x^2$ is the sum of squared x scores, and $\sum y^2$ is the sum of squared y scores.

3.2.2 Analysis of Indoor Climate

In Chapter 2 we talk about temperature and relative air humidity as important factors to create a good indoor climate. As it can create uncomfortable and inhabitable buildings, if these factors are not ideal. Therefore, we collect data compare the measurements from indoors, regarding pollutants, temperature, and humidity, and compare the different buildings.

Comparison of Distributions

In order to evaluate the different buildings indoor climate, we compare the frequencies of temperature and humidity in a violin plot. A violin plot is a combination of a box plot and a kernel density plot. Violin plots are useful for comparing distributions. It creates greater flexibility for plotting variation than box plots and its easier to directly compare data types [56]. According to Hintze and Nelson, the violin plot pools the best statistical features of different alternatives for graphical representation of data [31]. In our case, we compare each subgroup within. We compare frequencies below the set limits of temperature and humidity, and the max level of exposures to pollutants in a 24-hour mean.

3.3 Sub-goals

At the beginning of our thesis, we defined three sub-goals we wanted to achieve in order to create a well working prototype. These sub-goals will help us build a better prototype for the user, by being easier to use and understand, and it will make sure the data visualized is reliable.

In order to legitimize that our results from collecting the data are reliable, we have to compare our prototype to a reference point that is considered to deliver accurate and honest data. In order to test this, we decided to measure the air quality in the same room, comparing our devices observations. This will make sure that they are equally calibrated, measuring honest data. In addition, we compare the measurements to an air quality reference station used by countries to update the EEA air quality index [6], to measure our the accuracy of our sensors. In our case, we will use Bergen kommune's reference station at Klosterhaugen. It is centrally located in the city, as well as being secluded from the main traffic.

To evaluate the web application, we created a questionnaire based on the Likert scale to interview people in the buildings analyzed. The most important parts of a dashboard application are the user experience and how easy it is to understand/navigate. In order to test this, we decided to add five different statements: "The web application serves its purpose", "The web application is easy to navigate through", "The color combination and color scheme is good", "The web application is understandable, and "There are no unnecessary features on the web application".

Lastly, we want to find out whether people reflect on indoor air quality. In order to achieve this, we conducted an interview with people working in the buildings we did our analysis. We asked every participant the two questions: "Do you believe the air quality is better in newer buildings?", and "Do you think about the air quality when you enter a building?".

Chapter 4

System Architecture

In this chapter, we will provide relevant information about the system architecture including technologies in the perception layer, network layer, and application layer used in the architecture of our application. We will also present the concepts of how these technologies work, and how they are used to measure and analyze data.

4.1 The Architecture of The Application

The architecture of our application is divided into six different areas, including connected sensors, gateway, network server, IoT Hub, database, and web application, as seen in Figure 4.1, based on a three-layer hierarchical IoT architecture.

The perception layer consists of the connected sensors and our gateways. For communication between the connected sensors and our gateways, we use an application based on an LPWAN standard called Long Range Wide Area Network (LoRaWAN). From these gateways, our data is transmitted to the Network Layer using TCP/IP, to our network server, The Things Network. From the Network Layer, our data is transmitted using TCP/IP to the Application layer. The first component that processes the data is our IoT HuB, Amazon IoT Core. The data is added to a database service, Amazon DynamoDB, where our web application reads and visualize the data from our database.

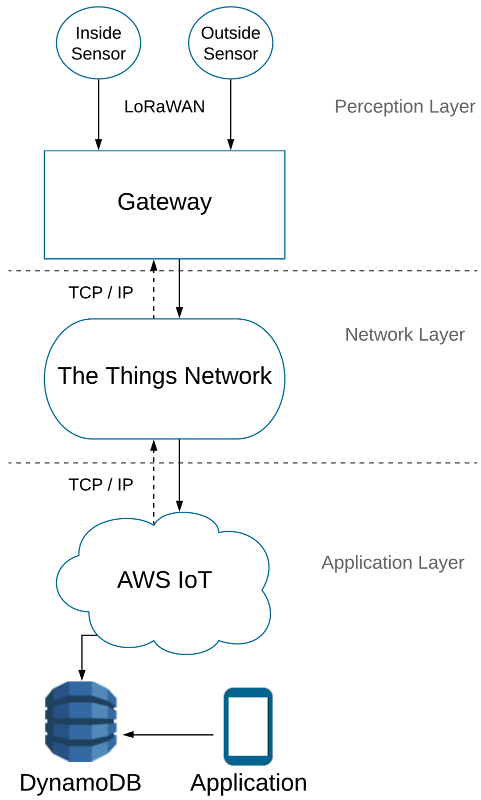

Figure 4.1: System architecture of the prototype.

In this thesis, we are working with the local consulting firm, Rainfall. We teamed up in order to help them create a monitoring use-case. In return, they supplied us with the necessary hardware needed to create our prototype. Rainfall wants to use IoT for air quality analysis, and pilot an end-to-end system used as a case study to excite future customers.

4.2 The Perception Layer

The first layer of a three-layer IoT architecture is the perception layer, which has sensors for sensing and gathering information. It senses some physical parameters or identifies other smart objects in the environment [65]. In our prototype, this layer is based on different devices, that has multiple integrated sensors and gateways that these devices are connected to.

Our devices have sensors that are used to measure parameters of its physical environment, in our case pollutants, humidity, and temperature, through SPS30 and SCD30. These measurements are transmitted to the next layer of the architecture through the Long Range (LoRa) modulation and LoRaWAN. It was through cooperation with Rainfall, that we ended up using LoRaWAN. LoRaWAN defines the communication protocol and system architecture for the network, while the LoRa physical layer enables the long-range communication link. The gateways are communicating with the devices and retransmits its data to connected systems. Data transmitted by a node is typically received by multiple gateways. Each gateway will forward the received packet from the end-node to the cloud-based network server [10].

4.2.1 LoRa

LoRa is a proprietary radio modulation technology used in LPWANs. The technology defines the radio part of the communication (physical layer) and is responsible for providing low-power and long-range communication. LoRa is based on chirp spread spectrum, which has been used in military and space communication for decades due to the long communication distances that can be achieved and its robustness to interference. These features are also the main advantages of LoRa, compared to other LPWAN technologies [10].

LoRaWAN

On top of the physical layer implemented on LoRa chipsets is the LoRaWAN protocol. It defines the communication protocol used by the application and for each specific class. Unlike LoRa, LoRaWAN is an open protocol, and the infrastructure is based on four elements. As shown in Figure 4.2, the infrastructure of LoRaWAN is based on end nodes, gateways, a network server, and an application [10].

The most important features with a LPWAN technology is coverage, power consumption, and scalability. LoRaWAN offers long-range communication, but the technology has limitations that needs to be clearly understood. Coverage is handled by receiving and transmitting data over a low frequency. Energy consumption is one of the most constraint requirements for the design and implementation of sensors communicating. Usually, IoT applications need the sensor nodes to operate reliably for an extended period of time. Each task consumes an amount of power for a period of time, making energy consumption crucial in choosing what protocols to use. LoRaWAN uses the ALOHA method for access when there is communication from and to a node. The node communicates when they have data ready to send whether event-driven or scheduled, to consume less power [10]. Scalability is an attribute that describes the ability of a process, network, or software to grow and manage increased demand. LoRaWAN scales by removing processing and complexity on the end devices, and use gateways to cooperate with devices.

LoRaWAN Infrastructure

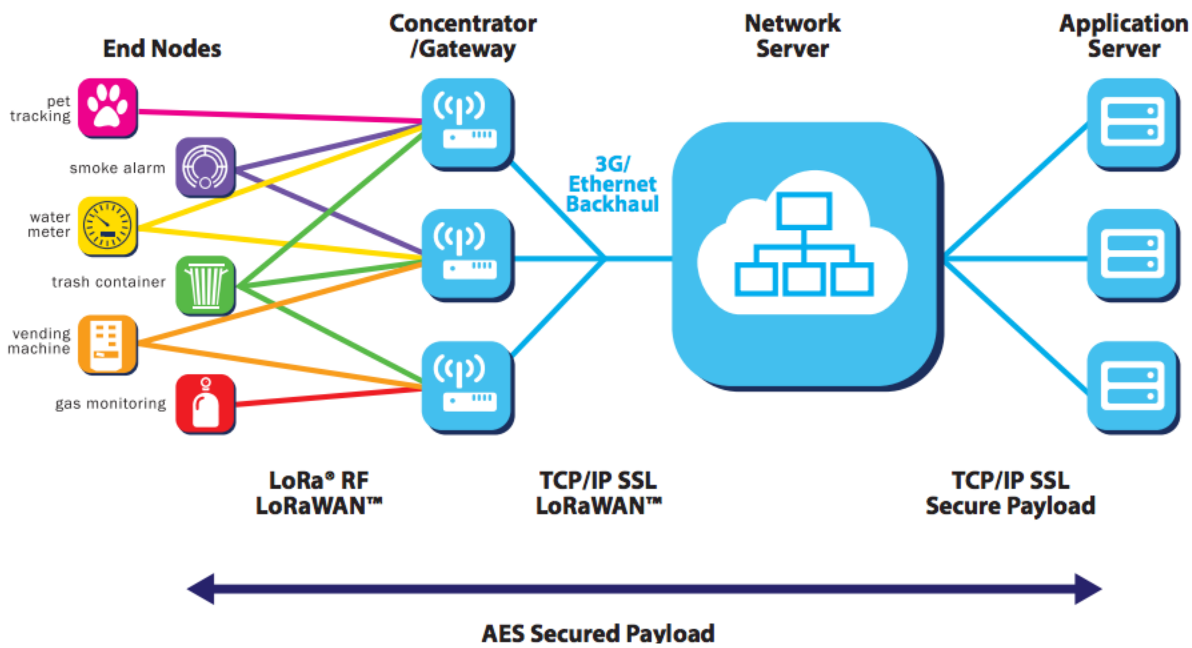

Figure 4.2: The LoRaWAN infrastructure [10].

LoRaWAN infrastructure is based on long-range star architecture as shown in Figure 4.2. In a LoRaWAN network end nodes are associated with a specific gateway, instead of another node to preserve battery lifetime, network capacity, and reduce complexity. A LoRa end node consists of a radio module with antenna and a microprocessor that process data. These devices are often battery powered and are equipped with a wireless transceiver. The end nodes are commonly connected to different sensors, such as GPS, temperature, and humidity, and transmits this data to the connected gateway [10]. LoRa gateways consist of a radio module with antenna and a microprocessor to process data received. A single gateway can serve thousands of devices [10]. The network and application server of LoRaWAN is open for different options compatible with the technology. The network server decodes the packets sent by the devices, performing security checks and adaptive data rate, thus generating the packets that should be sent back to the devices [10].

LoRaWAN Frequency Band and Security

LoRaWAN operates in the unlicensed radio spectrum. The frequency varies slightly from region to region based on the different regional spectrum allocations and regulatory requirements. In Europe and North America, LoRaWANs spectrum is defined, while other regions are still being processed. The LoRaWAN frequency in Europe is 867-869MHz.

LoRaWAN security is designed to fit LoRaWAN design criteria: low power consumption, low implementation complexity, low cost, and high scalability. It specifies three security keys, NwKSKey, AppSKey, and AppKey, where all keys have a length of 128 bits. The security design adheres to use of standard, well-vetted algorithms, and end-to-end security. The fundamental properties that are supported in the security are mutual authentication, integrity protection, and confidentiality [3].

The security mechanisms mentioned previously rely on the well-tested and standardized AES cryptographic algorithms. AES is used primitive combined with several modes of operation, such as CMAC for integrity protection and CTR for encryption. The applications payload is always encrypted end-to-end between the end-device and the application server. Integrity protection is provided in a hop-by-hop nature [3].

4.3 The Network Layer

The second layer of a three-layer IoT architecture is the network layer, which is responsible for connecting to other smart devices, network devices, and servers. Its features are also used for transmitting and processing sensor data [65]. In our prototype, the network layer is based on The Things Network, a long-range, low-power IoT data network that transmits and process data to an IoT Hub. As we decided on using LoRaWAN as our communication protocol, we had to choose a network server that is built for LoRaWAN communication, to transform LoRaWAN to IP, and that can integrate third-party integrations built for IoT applications. The sensors in our perception layer send data every 5-6 minutes depending on the delay. This indicates a dataset of 1680-2016 data points in each collection, that our network layer transmits to the application layer.

There are multiple network servers, both open-source and commercial, such as LoRa Server, Loriot, The Things Network (TTN), and ResIOT. We investigated different types of these platforms supporting LoRaWAN, and decided to test The Things Network and Loriot. These platforms supported a starting kit we used to test the network servers while supporting the different features we needed. When testing the two servers, TTN had a less complicated setup for integrating third-party integrations, making communication complete in just a couple of minutes, while Loriot never completed the communication setup while testing. Thus, we chose to use TTN as our network server for the prototype.

4.3.1 The Things Network

The Things Network (TTN) is a global, crowdsourced, open, free, and decentralized IoT network. It is a contributor member of the LoRa Alliance, a non-profit association collaborating to drive the LoRaWAN protocol as the leading standard. TTN is created to enable low power devices to use long-range gateways, connected to a network to exchange data with an application. They also provide The Things Industries, a commercial version of TTN, for companies to build global, long-range, and low power data networks [40]. In our case, there was no need to build such a big network of end-devices, as they bundle from 1,000 nodes and we use less than 10.

For our thesis, the network server works as a decoder and transceiver, transmitting and receiving communications. The devices used in our prototype encodes the data observations measured into a byte array, in order to send smaller packets of data. In the network server, the payload sent from a device, through a gateway, is decoded by a TTN Decoder. A decoder works as a function accepting the payload as a byte array and returns an object containing the decoded values. In our case, how the source code is encoding the values is proprietary, as Rainfall owns the source code. The encoded data is received by TTNs decoder and decoded by reversing the operations of the payloads source code. The payload is then forwarded with the TTN application payload, metadata, and gateway payload to the application layer.

4.4 The Application Layer

The last layer of a three-layer IoT architecture is the application layer, which is responsible for delivering application-specific services to the user. It defines various applications in which IoT can be deployed, for example, smart homes, smart cities, and smart health [65]. One of the arguments to choose TTN to be our network server was the possibility to integrate third-party integrations. Integrations are the easiest way to connect our devices to an application, bringing LoRaWAN to different IoT platforms, and TTN has connections to external IoT platforms, such as Azure IoT Hub, AWS IoT, and IBM Watson IoT.

In our prototype, we decided to use Amazon Web Services (AWS) as our integration, as a result of the detailed quick start guides TTN contribute, in order to set up communication to AWS IoT, and in regard to AWS free tier for all the features our application needs. The application layer is, therefore, based on different AWS services, such as AWS IoT Core and Amazon DynamoDB, and an application as shown in Figure 4.1. The application is the last part of our system architecture and is further explained in Chapter 5, where we cover the design and implementation of our prototype.

4.4.1 Amazon Web Services

Amazon Web Services (AWS) is an on-demand cloud computing platform, created for individuals, companies, and governments, where pricing is on a pay-as-you-go basis. For our prototype, we use two AWS services, AWS IoT Hub and Amazon DynamoDB. These services are used in order to process, manage, and store our measurements for the application to analyze and visualize it to the users [60].

AWS IoT Core

One of the features AWS offer is called AWS IoT Core. It is a cloud service that lets connected devices easily and securely interact with cloud applications and other devices. AWS IoT consists of multiple components, such as message broker and rules engines. It works by connecting and managing your devices while securing and processing data received. In our prototype, we use features for managing and acting [62].

AWS IoT Core allows us to connect and manage our devices with a secure device connection. Our network server, TTN, has an application where all our devices are registered. When we link this application to AWS IoT Core we can manage these devices in our platform. The last feature that is used in our prototype is to process and act upon the data sent from our devices. The feature let you filter, transform, and act upon data, based on a set of rules we define. This is used by sending the data received to Amazon DynamoDB. This data is not transformed, but the defined payload has been filtered to our choice.

Amazon DynamoDB

Amazon DynamoDB is a fully managed proprietary NoSQL database service that supports key-value and document data structures. DynamoDB supports some of the world's largest scale applications, with no servers to provision, patch, or manage, and enables you to build business-critical applications at scale. More than 100,000 AWS customers have chosen DynamoDB as their database for mobile, web, gaming, IoT, and other applications that need low-latency data access at any scale [61].

For our prototype, we need a database for storing measurements over time, as a part of our application's backend. All data from our measurements are stored in a DynamoDB table, where we add certain values from the device payload, and add columns telling what type of building we are measuring, and if the sensor is indoors or not.

Chapter 5

Design and Implementation of Prototype

To understand how our prototype works, we need to get a better understanding of the perception layer, and the last part of the application layer. In this chapter, we describe how we designed and implemented our working prototype, and the type of hardware used. We also present the approach of how this is designed, and we give a demonstration of the web application.

5.1 Design

Designing our prototype was conducted in two steps, designing the devices and designing the application. In order to design our prototype, we adopted a process called design thinking. In 2012, Joseph Pistrui points out that however big the data or thorough the quantitative analysis is, making real sense of this requires human insight. Reaching that level of insight requires a different approach, one that makes room for intuition and creativity alongside analysis. For many, that alternative approach is design thinking [28].

Design thinking is generally defined as an analytic and creative process that engages a person in opportunities to experiment, create and prototype models, gather feedback, and redesign [53]. According to Brown, the design thinking process is best described as a system of spaces rather than a predefined series of orderly steps. These spaces are labeled as ”inspiration”, for motivating the search for solutions, ”ideation”, for the process of generating, developing, and testing ideas, and ”implementation”, for the charting of a path to market [16].

We designed our hardware devices together with Rainfall, as they were the ones creating the final product. Our first vision was for the devices to use LoRaWAN communication, sending a payload of different pollutants, temperature, and humidity. The result ended up with a device monitoring CO_2, $PM_{2.5}$, PM_{10}, temperature, and humidity. The issue with this was that it was not able to repel water or have the operating temperature below 0 °C when stored outside. We also wanted it to measure the O_3 and NO_2 gases, but this was not possible for the production line. The end result was a device able to repel water and operate below 0 °C.

The application's design started with a piece of paper, drawing the layout of our web application. It was inspired by smart homes visualizing humidity and temperature using graphs. First, the application had the ability to add buildings and devices connected to an already added building. We focused on producing working software, and our first version was showing a graph over all the measurements in a time series. As we ended up with four monitoring devices, we measured data in different buildings using the same devices. The result ended up as Figure 5.1, a dashboard web application with the possibility to look at each pollutant by itself as a time series and scatter plot, displaying the I/O ratio and correlation coefficient value.

Dashboard Web Application

A dashboard application can enable a user to quickly view data, and in some cases, data from multiple different areas. We decided to make a simple and understandable dashboard, to present our analyzed data. After working agile with the design of our application, producing working code, we ended up with the web application displayed in Figure 5.1.

There are many reasons why our application is based on a dashboard method. The format of the dashboard application is adaptable to a variety of devices and platforms, making it easy to access and adapt to the needs of end users. There are four main driving needs for the use of dashboards, according to Pauwels et al. (2009), with regard to marketing that can be applied to information sharing. [48]. (1) poor organization of the many pieces of potentially decision-relevant data, (2) managerial biases in information processing and decision making, (3) the increasing demands for marketing accountability, that can be applied to information sharing, and (4) the need for cross-departmental integration in performance reporting practices and for resource allocation [48].

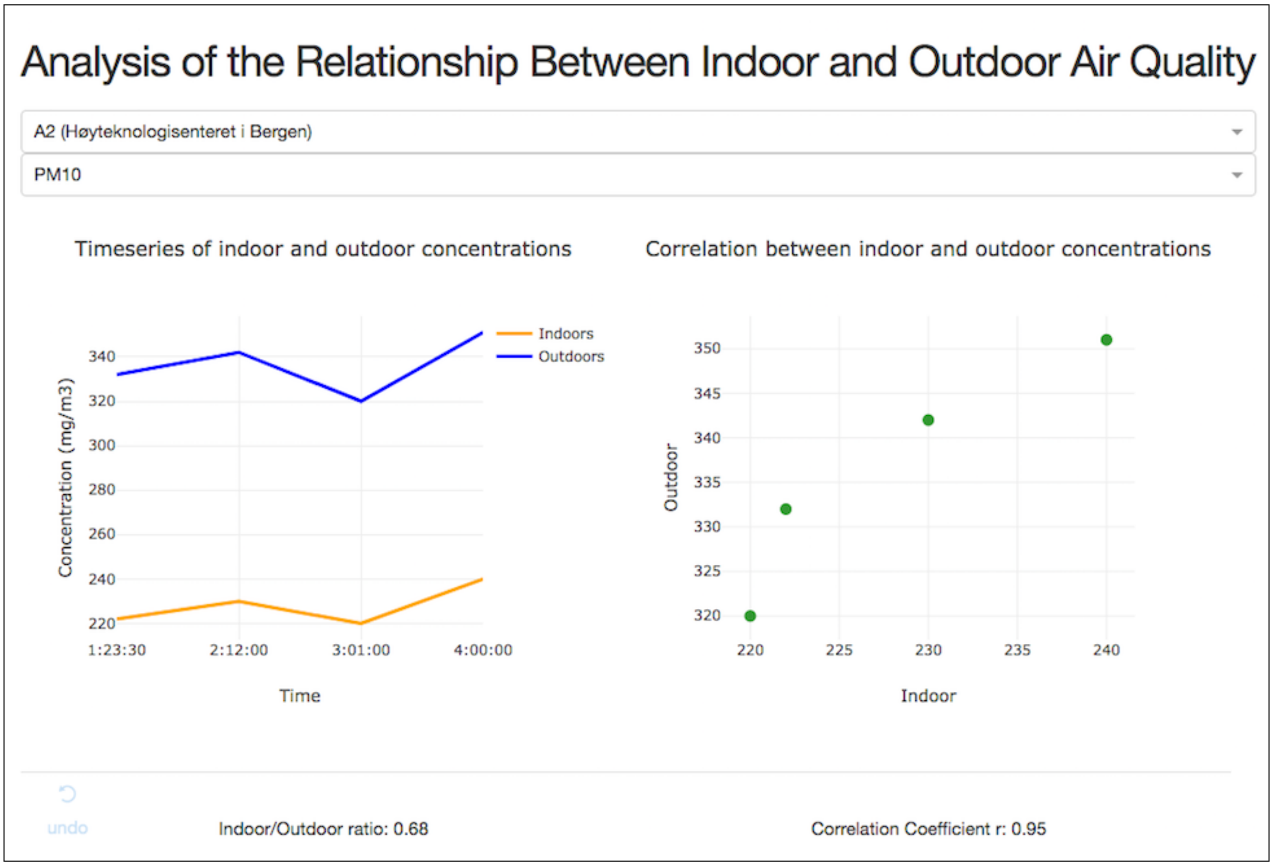

Figure 5.1: Our last version of the dashboard web application (not real data).

5.2 Hardware

In this thesis, we are working with Rainfall in order to get the hardware needed to build our prototype. Hardware includes the physical, tangible parts and components used in our prototype. The hardware components are built up by multiple devices, each including two different sensors, and gateway base stations.

Measuring reliable data is important, therefore the calibration of the different sensors is of great importance. It is critical to remember that no sensor is perfect, and calibrating these sensors as equal and accurate as possible is the main reason for teaming up with a consulting firm. Sensors that are subjected to factors such as, heat, humidity, and cold while being stored or shipped may show a change in response. Differences in sensor design may mean that two different sensors may respond differently to similar conditions. Creating equal sensors, calibrated together, and shipped/stored combined will make sure the data measured is honest, and comparing these to reference stations will show how accurate the data is.

End Nodes

In each device, as seen in Figure 5.2, we have included two different sensors. One sensor is monitoring temperature, humidity, and CO_2, while the other sensor is measuring dust levels of $PM_{2.5}$ and PM_{10}. These sensors are made by Sensirion, a company from Switzerland, and is called SPS30 and SCD30. Sensirion is the leading manufacturer of high-quality sensors and sensor solutions for the measurement and control of humidity, gas, and liquid flows [57].

Figure 5.2: Example of one of the IoT-devices used for our research.

SPS30 is a Particulate Matter sensor, measuring dust levels in the air. The sensor represents a new technological breakthrough in optical PM sensors. Its measurement principle is based on laser scattering and makes use of Sensirion's innovative contamination-resistance technology. This technology, together with high-quality and long-lasting components, enables accurate measurements from the device's first operation and throughout its lifetime of more than eight years [58]. SCD30 is a CO_2, humidity and temperature sensor. Due to the dual-channel principle for the measurement of carbon dioxide concentration, the sensor compensates for long-term drifts automatically by design. The very small module height allows easy integration into different applications [59]. These sensors are observing the change in air quality and help us analyze the relationship between indoor and outdoor air quality and between different buildings indoor climate.

Different studies on related research use heavy-duty analyzers, in order to measure different pollutants. An example is the Aerocet 531 particle mass profiler, to measure particulate matter [37]. This devices cost around $2,000.00 - $3,000.00. Our devices cost $200.00 - $300.00, fully calibrated, professional standard with an IP67 rating. This indicates that the unit is resistant to dust, and can be dropped into a body of water up to a meter deep for half an hour.

Gateways

Kerlink base station is a robust, performant and highly reliable outdoor LoRaWAN gateway for smart IoT networks. Based on LoRa technology, this gateway is a leading reference for the deployment of resilient public and private operated networks. According to The Things Network, this type of gateway station is mainly for industrial use, because it is suitable to be mounted outside, and you need technical skills in order to update and manage them [41].

In our prototype, the standard configurations work properly with our network layer, getting its power by PoE (Power over Ethernet), and being connected to the internet through Ethernet. It is possible to connect the Kerlink to a GPRS/3G connection if there is a need for this when mounted outside. The software is open source and proprietary. The hardware is proprietary, however, this was not a problem as it only worked as a data transmitter in our prototype.

5.3 Software

The software side of our prototype is built from multiple Python libraries. They are collections of pre-built compiled code which can be used to extend the application's features. The application is based on the following libraries:

- Pandas: For data manipulation and analysis.
- Boto3: Amazon Web Services SDK for Python.
- Flask: Python micro web framework.
- Plotly: Interacting graphing library.
- Dash: For building analytical web applications.

Data retrieval

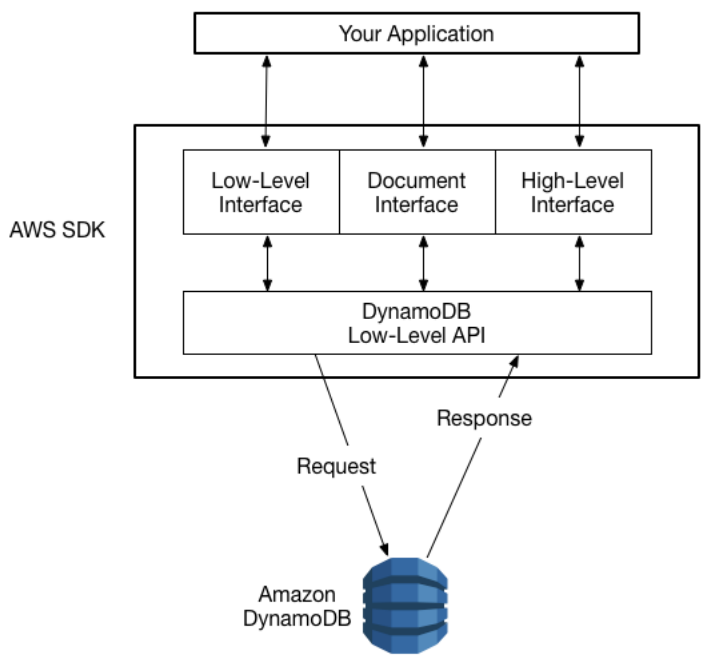

Figure 5.3: High level overview of AWS SDK support for DynamoDB.

Credit: Amazon Web Services [64].

Measurements of pollutants and observations of the weather are stored in Amazon DynamoDB, as explained in Chapter 4. Retrieving this data to our prototype is done by using

AWS own software development kit (SDK) called Boto3. Boto is an SDK designed to improve the use of the Python programming language in AWS [63]. Boto3 makes it easy to integrate your Python application, library, or script with AWS services including Amazon S3, Amazon EC2, Amazon DynamoDB, and more. The interface has two distinct levels of APIs. Client (or "low-level") APIs provide one-to-one mappings to the underlying HTTP API operations. Resource APIs hide explicit network calls but instead provide resource objects and collections to access attributes and perform actions [63].

In our prototype, we use Boto3 in our web application to query data from our DynamoDB database and return it as a JSON object using the helping class created by AWS. This class is using libraries such as *json*, to encode and decode python objects, and *decimal*, to support correctly-rounded decimal floating points arithmetic, to convert the data. An overview of how Boto3 supports DynamoDB is illustrated in Figure 5.3.

Data Analysis

After the data retrieval has been completed successfully, data visualized to the users has to be manipulated in order to become informative. In order to structure the data, we use Pandas. Pandas is an open-source, BSD licensed library written for Python programming language. It stands for "Python Data Analysis Library", and is used for manipulation and analyzing data. It takes data files, such as CSV, JSON, and XLSX, and creates a Python object with rows and columns, called data frames. The library features multiple commands, such as data filtration, column insertion, and deletion [47].

Pandas have made it easier to analyze data with Python, and we are using it for all analysis and data manipulation in our prototype. The data we get from calling DynamoDB using boto3, is structured by manipulating the data using Pandas. When the user changes pollutant or building, Pandas will use the variables chosen by the user and calculate the indoor/outdoor ratio, and correlation coefficient value.

Data Visualization

The last part of our web application is visualizing all the information created. This is done using three libraries, Flask as a web server, Plotly for graphing, and Dash for the interface.

Dash is a Python framework created to make dashboard applications and is built on top of different frameworks, such as Flask, Plotly.js, and React.js.

Dash is created for dashboard applications, and applications are cross-platform and mobile-ready. [22]. Dash apps are composed of two parts. The first part is the for the layout of the app and it describes what the application looks like. The second part describes the interactivity of the application and provides classes for all of the HTML tags, and the keyword arguments describing the HTML attributes [21]. In this thesis, our web application is divided vertically, where information is presented as two different plots. One showing the time series and I/O ratio of the chosen pollutant, and the other showing a scatter plot to look at the correlation.

5.4 Demonstration of Web Application

As described in step 4 in our research methodology, we will now illustrate the artifact in a way that was presented for our participants that evaluated the prototype.

Starting the Application

The first page a user gets presented is the applications start-up page, illustrated in Figure 5.4. Here the user has the possibility to choose what building that they want information about, and the pollutant they want to have presented.

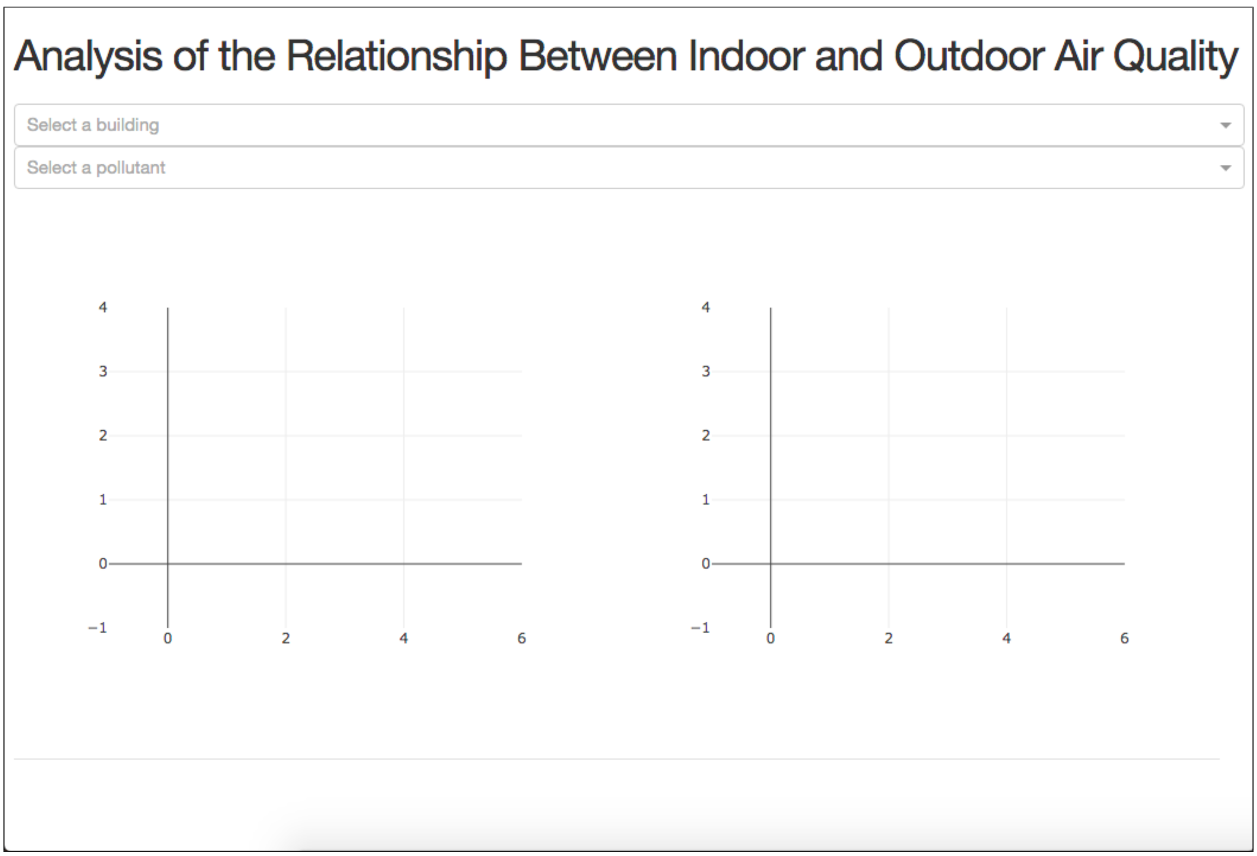

Figure 5.4: Step 1 - The start page of our dashboard web application.

Choosing Building and Pollutant

The dashboard application has two different drop-down menus, where the user choose a pollutant/weather element and the building of their choice, as illustrated in Figure 5.5. In this demonstration, the user chooses to look further into Humidity in Building A1.

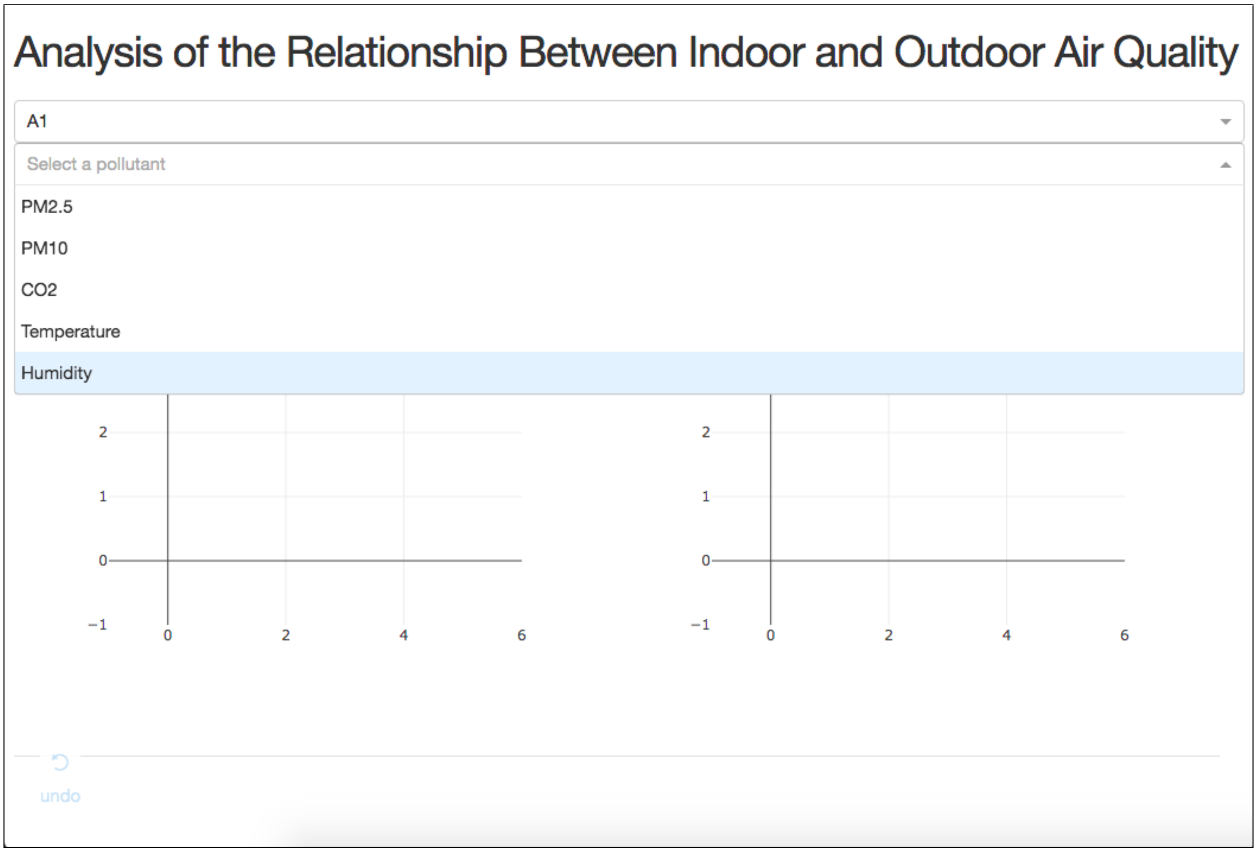

Figure 5.5: Step 2 - Choosing what type of pollutant and building to be visualized.

Visualizing the data collections as information

After the user chooses building A1 and Humidity, the web application will visualize the relationship of Humidity indoors and outdoors, as presented in Figure 5.6. The relationship is presented by a time series to visualize the ratio between indoor and outdoor concentration, and a scatterplot to visualize the correlation. Underneath the graphs, the application calculates the I/O ratio and correlation coefficient of the element chosen in drop-down menu 2.

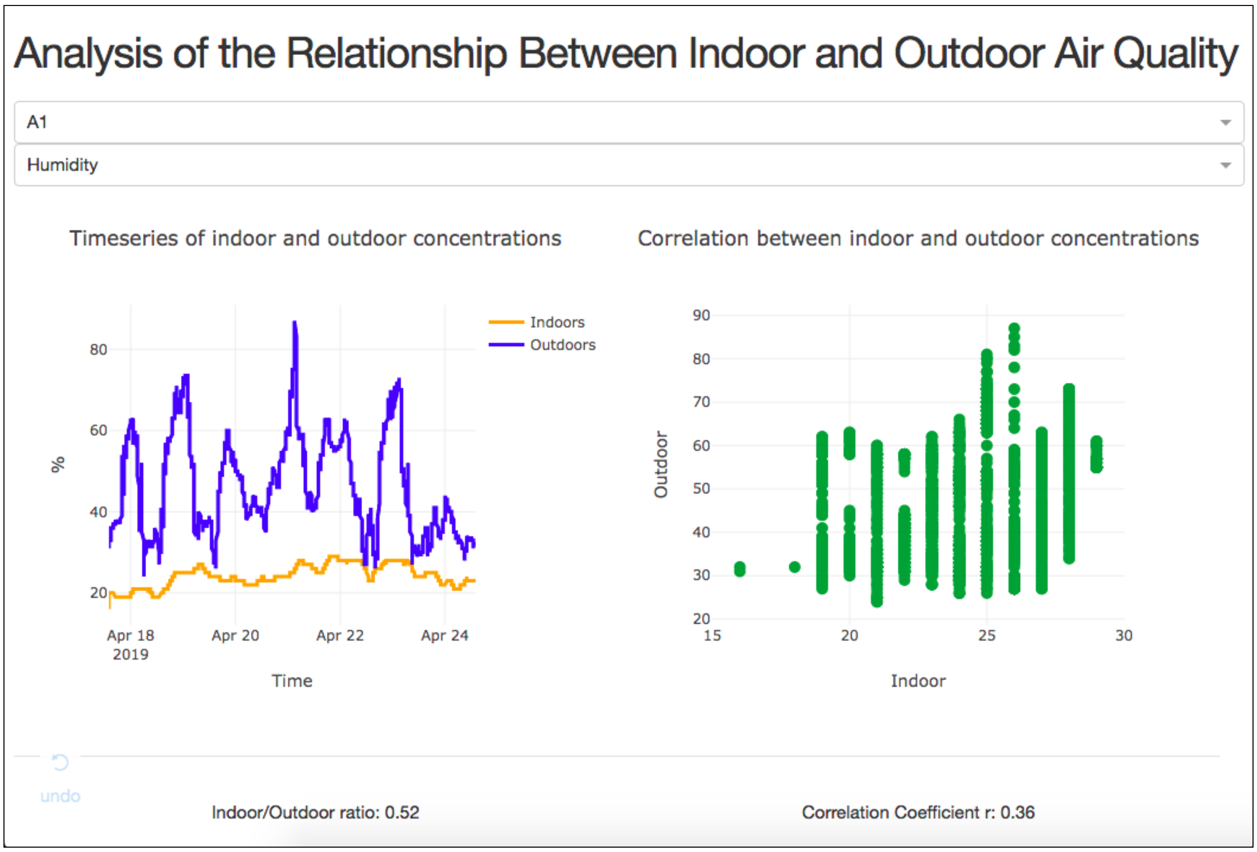

Figure 5.6: Step 3 - The application is visualizing information about building A1 and Humidity.

Changing Between Different Pollutants/Buildings

Every user has the opportunity to change the choice in each drop-down menu, visualizing another element and/or building. Demonstrated in Figure 5.7, the application is presenting the relationship between the indoor/outdoor PM_{10} concentrations in building A2 after the user change drop-down menu.

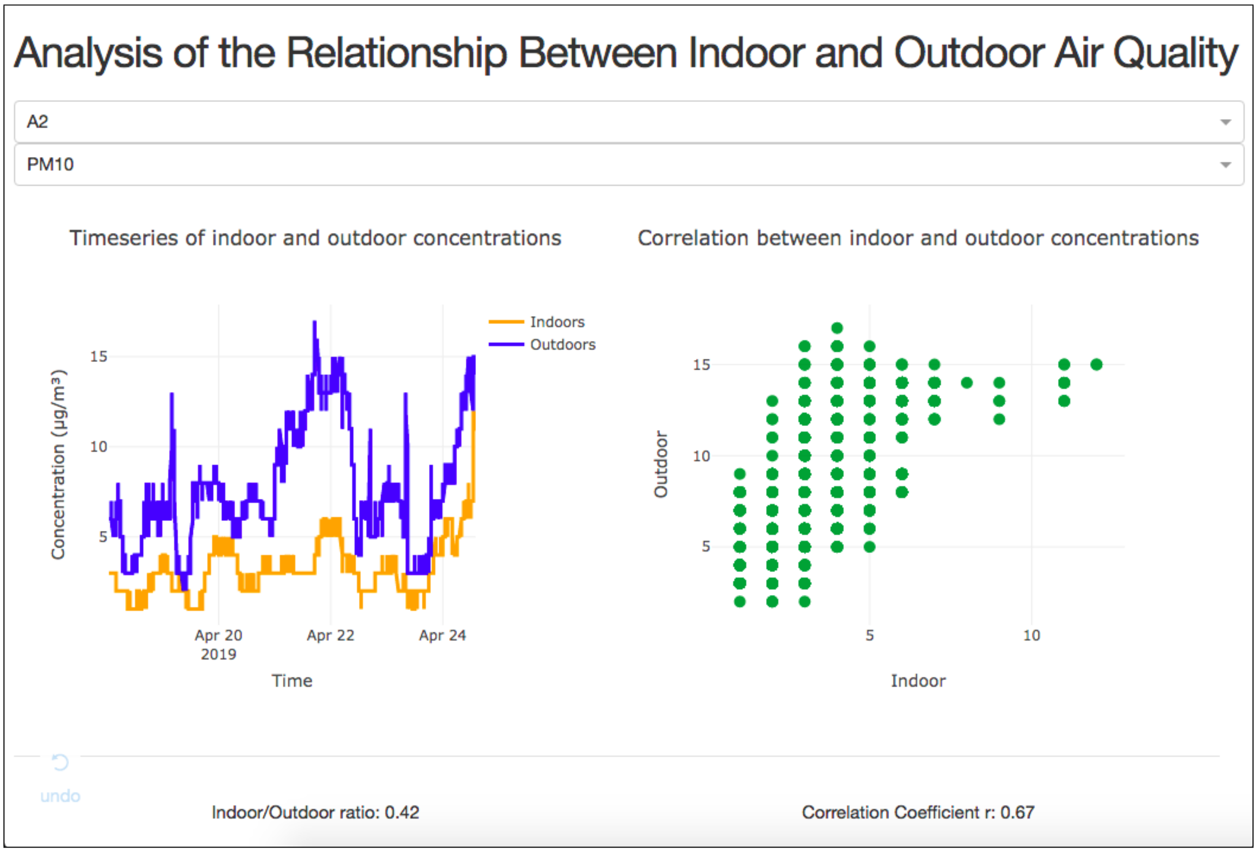

Figure 5.7: Step 4 - The user change building and pollutant.

Different Features Provided by Dash

In the last step of this demonstration, we want to explain some of the features Dash includes in our application, marked with red in Figure 5.8. There are a number of different features, such as zoom, pan, hover, and undo. This gives the user the opportunity to look further into details, and to use the graphs more interactive.

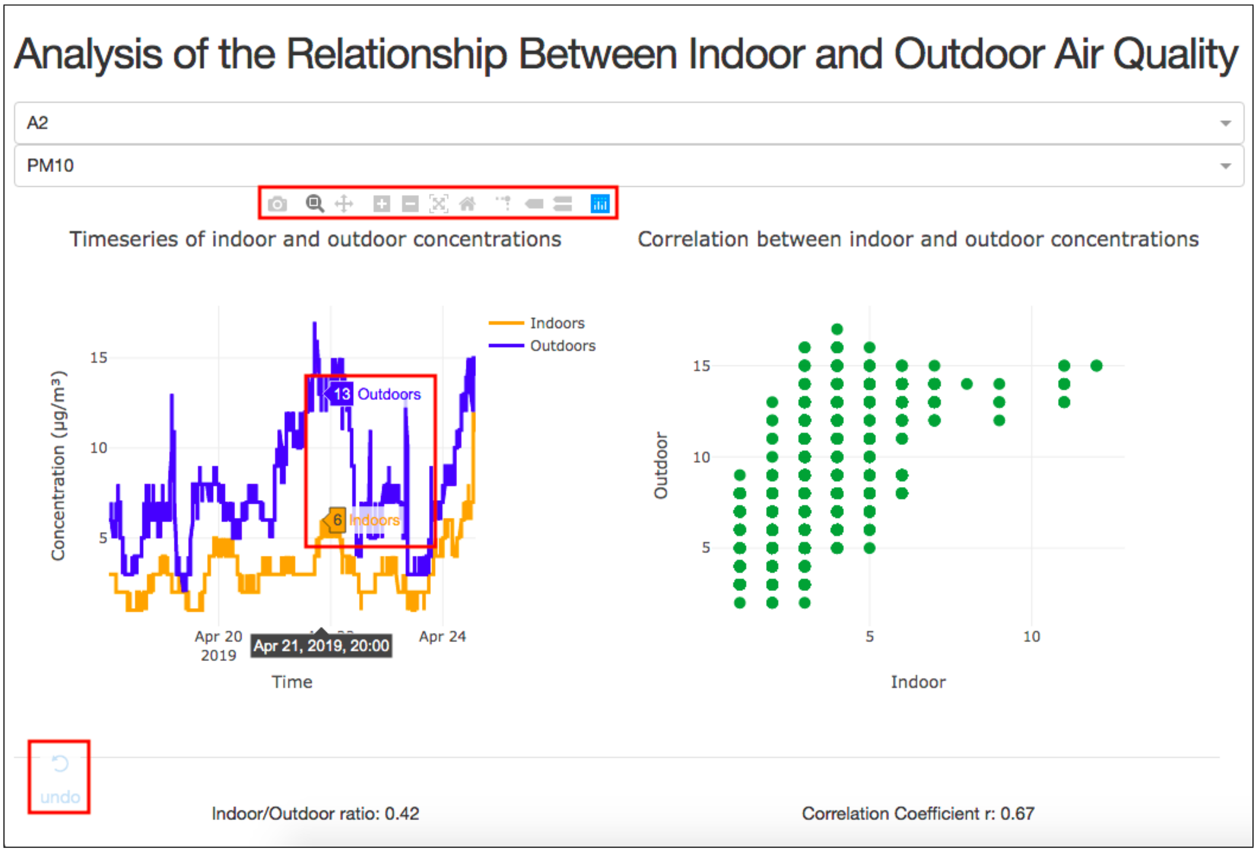

Figure 5.8: Step 5 - Different features provided by Dash.

Chapter 6

Results and Evaluation

In Chapter 3, we explain our research methodology, and one of the steps in that methodology is to evaluate our artifact. In this chapter, we will present our results and evaluate our prototype in three steps. (1) We present the results from the evaluation of the dashboard web application. (2) We visualize the results from our sensor reliability test, before visualizing the end results from the indoor/outdoor relationship analysis. (3) At last, we present the results of the indoor climate analysis and compare the results to peoples reflections of indoor air quality.

6.1 Evaluation of Web Application

In this research, we want to use cheap sensors to gather collections of air quality data, and increase building occupants indoor air quality awareness. To get a better result of our prototype, we defined a goal to create an easy-to-use web application to visualize air quality information. In order to visualize this in a proper manner, we have created a dashboard web application. Creating a well-functioning, understandable web application is important when you are presenting it as a dashboard application. In order to evaluate our application, we asked different occupants at the monitored sites to answer a questionnaire after a prototype demonstration, as shown in Chapter 5.

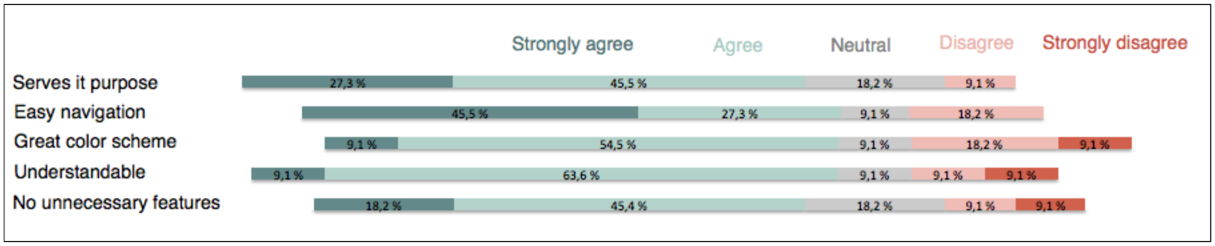

Figure 6.1: The results from the web application questionnaire.

In Chapter 5, we explained what makes a good dashboard web application. Our questionnaire represents this and is further explained in Appendix A. The questionnaire gave us a mixed result of answers, presented in Figure 6.1. The questionnaire is based on a 5-level Likert scale, where the user specifies their level of agreement or disagreement on a symmetric agree-disagree scale for a series of statements. Thus, the range captures the intensity of their feelings for a given item [18]. 11 occupants answered the questionnaire after they completed a demonstration of the web application. The result indicates that the majority of users agree on all statements. On average, seven out of 10 participants (69%) agree or strongly agree with the statement presented to them.

6.2 Indoor/Outdoor Relationship

In Chapter 3, we explained the two different statistical analysis methods we are using to calculate the I/O relationship of different pollutants at our monitored sites. Our analysis is done by calculating the I/O ratio in our monitored sites, and the correlation coefficient (r) value between indoor and outdoor pollutants. In addition, we want to use low-cost hardware, while measuring reliable collections of data.

Reliable Collections of Data

For our results to be considered as research, we have to make sure that the collected data is accurate and honest. Our devices were calibrated in the same environment, before they were stored, and shipped. During transportation, the devices could be exposed to factors that can manipulate their accuracy when measuring. In order to achieve honest data collections, we compare two of our devices observations (S1 and S2) during a 7-hour time series. In our case, we measured during the night in a closed off room, making sure there was no manipulation of data from outside factors. The results are represented in Figure 6.2. The two plotted lines regarding PM_{10} are covering the observations of $PM_{2.5}$. This is a result of similar concentrations

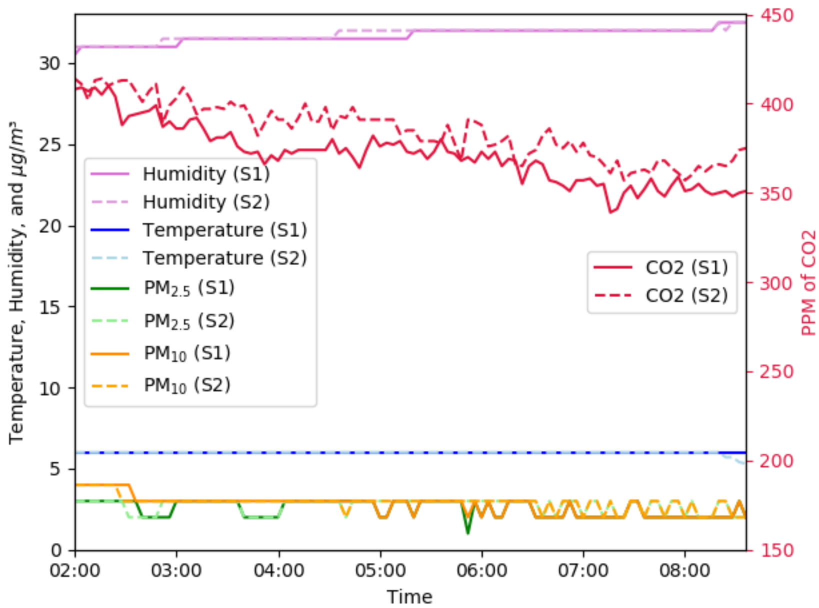

Figure 6.2: A comparison of two devices that is placed in a closed off environment.

The results show that our devices are measuring almost identical values for $PM_{2.5}$, PM_{10}, temperature, and humidity throughout the time series. The correlation coefficient (r) value range from 0.5-0.9, indicating that the sensors are measuring honest data. For CO_2, the sensors are showing a slight but consistent difference throughout the time series, with an r-value = 0.89, indicating a strong positive correlation.

The placement of sensors is also important in order to achieve honest results, as mentioned in Chapter 2. We have, therefore, created a 7-hours example of sensor placement, mounting two sensors on opposite sides of a building and comparing their observations to a reference station at Florida, Bergen [1]. *Sensor One* is placed in the sun, while *Sensor Two* is placed in the shade. The results are presented in Figure 6.3.

[1]Reference station data is gathered from NRK/Meteorologisk institutt at www.yr.no.

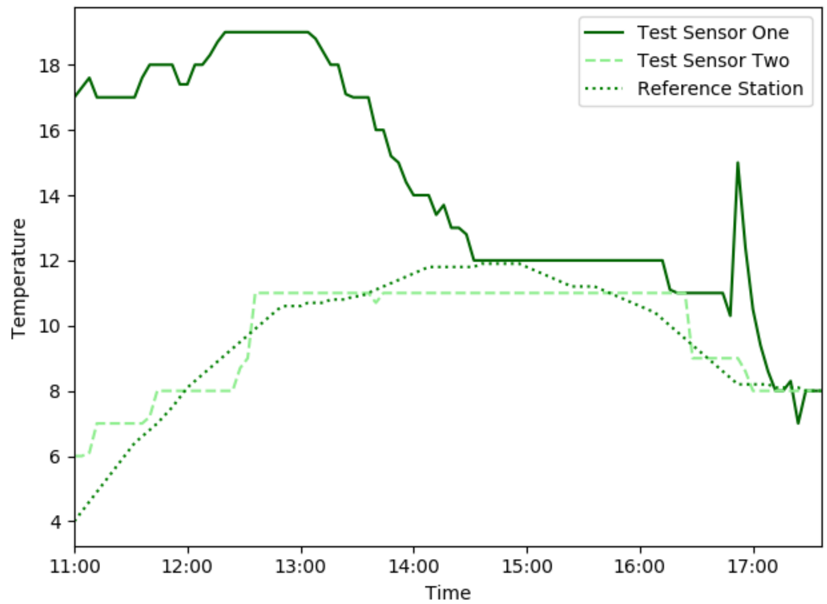

Figure 6.3: Comparison of sensor placements.

Credit: Reference station data is gathered from NRK/Meteorologisk institutt at www.yr.no.

The results show that during the midday, and until *Sensor One* is in the shade (around 17:00), the temperature measured is higher than the observations from *Sensor Two* and the reference station. *Sensor Two* is fairly accurate to the reference station, showing that measuring the temperature outside should be done away from the sun. The rest of our measurements are, therefore, done in the shade, if possible. In naturally ventilated buildings, sensor placement is outside the windows used for ventilation, causing the placement in some cases to be exposed to the sun.

After we established that the data collected was honest, we had to make sure that the different measurements were accurate. To further legitimize that our results from collecting the data can be used for research, we have compared the collected data of particulate matter to a reference station at Klosterhaugen, Bergen [2]. There is no reference station in Bergen that is measuring CO_2. Therefore, in regards to CO_2, we have not compared the sensors.

[2]Reference station data is gathered from Bergen kommune/Statens vegvesen at www.luftkvalitet.info.

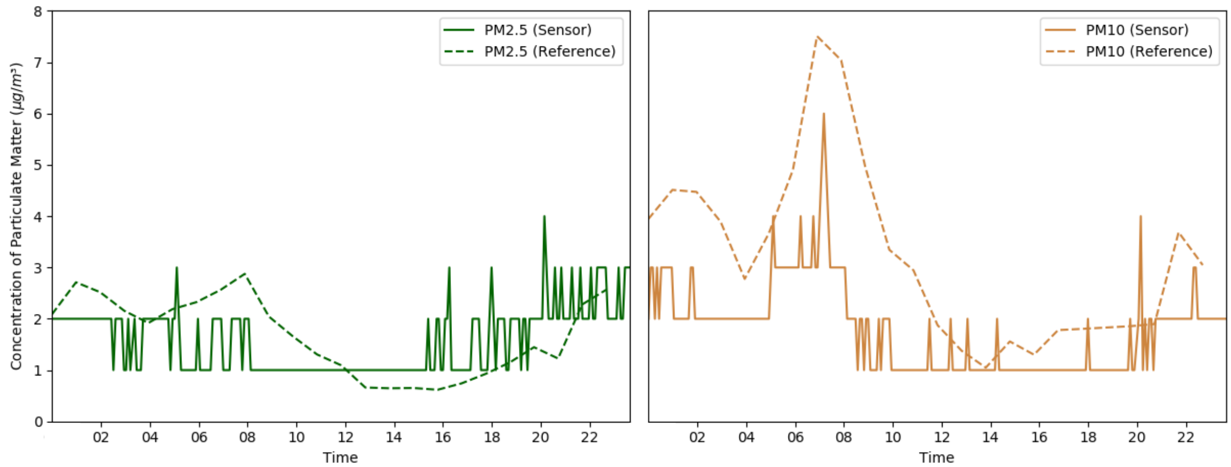

Figure 6.4: Comparison between reference station and sensors.

Credit: Reference station data is gathered from Bergen kommune/Statens vegvesen at `www.luftkvalitet.info`.

The results, as presented in Figure 6.4, show that for both $PM_{2.5}$ and PM_{10}, the data observed is fairly accurate compared to Bergen kommune's reference station. The data is collected during a 24-hour time series, where the reference station sends the average data values every hour, and our sensors send the average data value every fifth minute. Our results indicate that our sensors are measuring accurate data, with the correlation coefficient (r) value to be 0.39 in regard to $PM_{2.5}$ and 0.54 in regard to PM_{10}.

Indoor/Outdoor Ratio

In Figure 6.5 and Figure 6.6, we visualize an overview of the different pollutants concentration, temperature, and humidity, indoors and outdoors, in office building A1. The same time series are presented in Figure 6.7 and Figure 6.8 for building A2. For all other buildings monitored, the overviews are presented in Appendix B. These graphs are created in order for the reader to get a better understanding and overview of the ratio between the different pollutants.

In building A1, the time series show the concentration of PM_{10} indoors (purple line), while the concentration of $PM_{2.5}$ indoors (yellow line) is not showing. This is a result of the $PM_{2.5}$ line is plotted behind PM_{10} line. For all time series regarding the concentration of pollutants, CO_2 concentrations are presented at the second y-axis, while the other pollutants refer to the first y-axis.

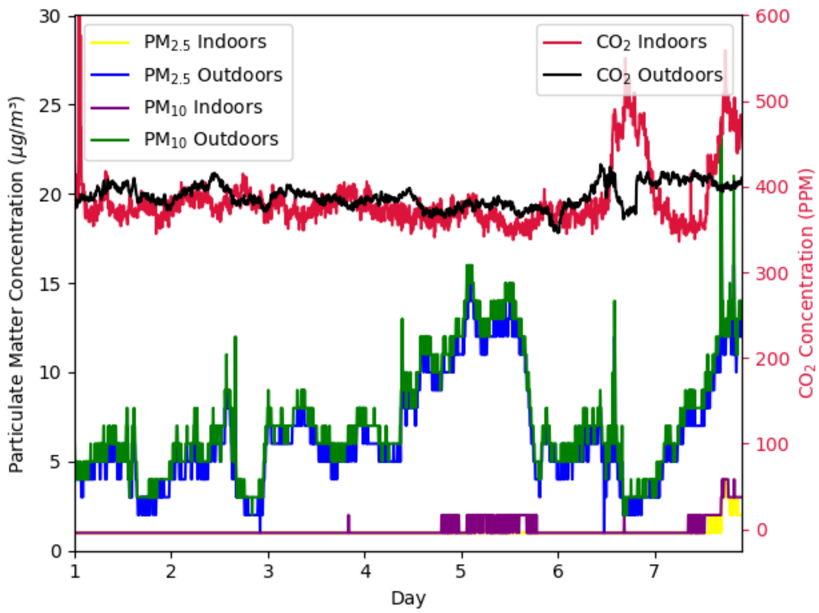

Figure 6.5: Time series comparison between pollutants in office building A1.

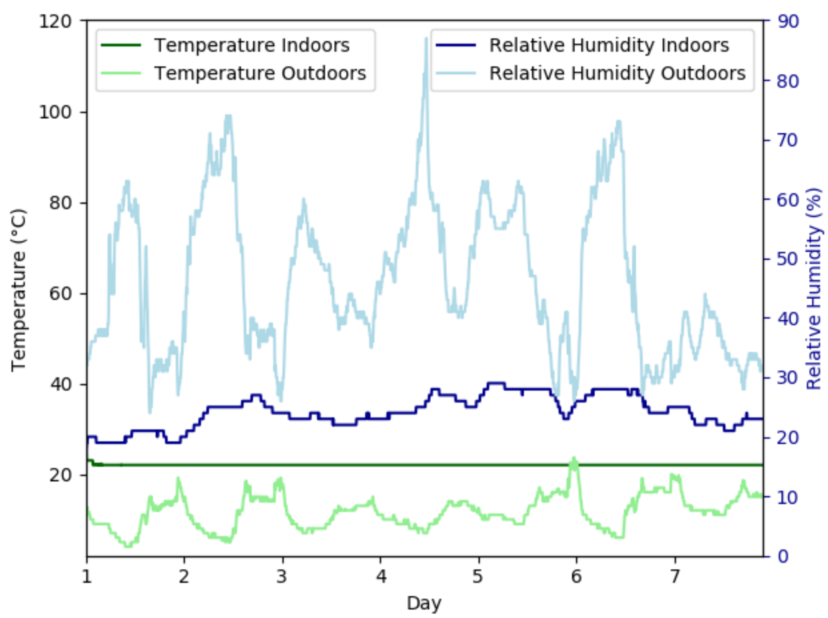

Figure 6.6: Time series comparison between temperature and humidity in office building A1.

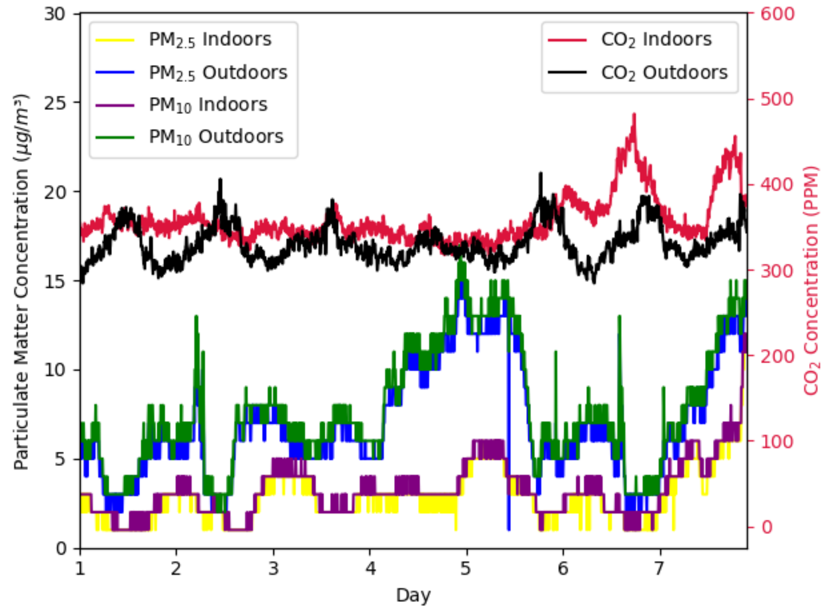

Figure 6.7: Time series comparison between pollutants in office building A2.

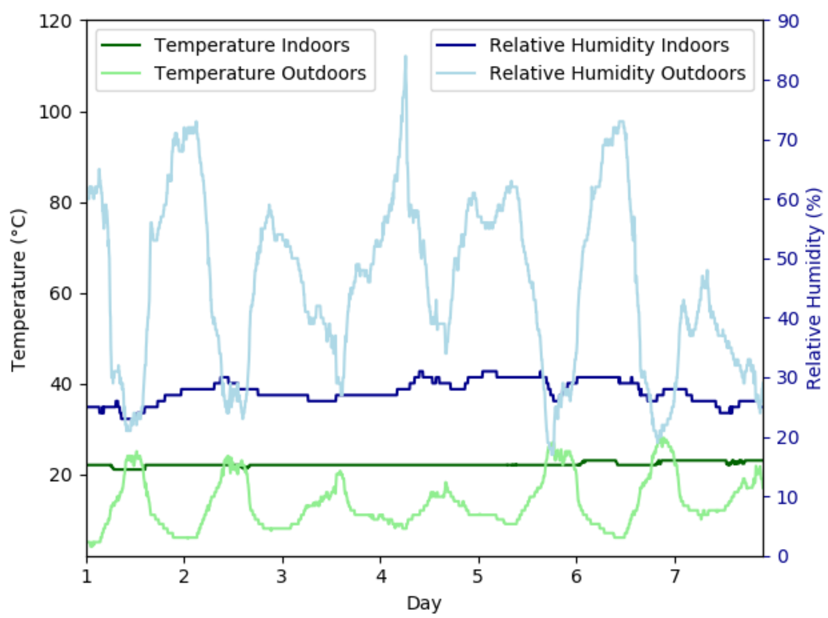

Figure 6.8: Time series comparison between temperature and humidity in office building A2.

In order to analyze the I/O ratio, we have to look back at the equation explained in Chapter 3, using building A1 as an example. In A1 there were 1864 observations collected in our dataset (n = 1864). When looking at pollutant $PM_{2.5}$, the sum of C_{in} = 1 990 and C_{out} = 12 136, therefore, in building A1 the I/O Ratio = 0,164 for $PM_{2.5}$.

	Office		**Campus**		**Residential**	
	A1	A2	B1	B2	C1	C2
$PM_{2.5}$	0.164	0.401	0.526	0.561	0.675	1.206
PM_{10}	0.161	0.417	0.395	0.441	0.462	1.043
CO_2	0.987	1.083	0.993	1.347	1.268	1.904
Temperature	1.888	1.641	5.487	4.071	2.506	2.356
Humidity	0.524	0.602	0.395	0.401	0.534	0.683

Table 6.1: Summary of I/O ratios in the monitored buildings.

The results of the I/O ratios are shown in Table 6.1 for $PM_{2.5}$, PM_{10}, CO_2, temperature, and humidity, where buildings with the number 1 are the new buildings, e.g., A1, and buildings with 2 are the old buildings, e.g., A2. From a health perspective, an I/O ratio value below 1 is obviously preferable, indicating that the building provides some form of protection from outside pollutant exposure.

In regard to the I/O Ratio of $PM_{2.5}$ and PM_{10}, the mechanically ventilated buildings (Office and Campus), and the naturally ventilated buildings (Residential), are found to be higher in older buildings. This indicates that the newer buildings provide a better form of protection from outside pollutant exposure. However, the CO_2 measurements found are all closer to 1, or higher, for all monitored buildings, showing no form of protection in any of the monitored buildings. This might be a result of external factors, such as building occupants.

Correlation

In Figure 6.9 and Figure 6.10, we visualize a scatter plot of the concentration of the different pollutants in office buildings, in regard to the concentration indoors and outdoors. The same plots are presented for all building types in Appendix B. These plots are presented for the reader to get a better understanding of the correlation between the different pollutants. Each plot shows the pollutant concentration in either a new building, an old building, or both, with the regression line for either the new building (Regression Line - New) or the old building (Regression Line - Old).

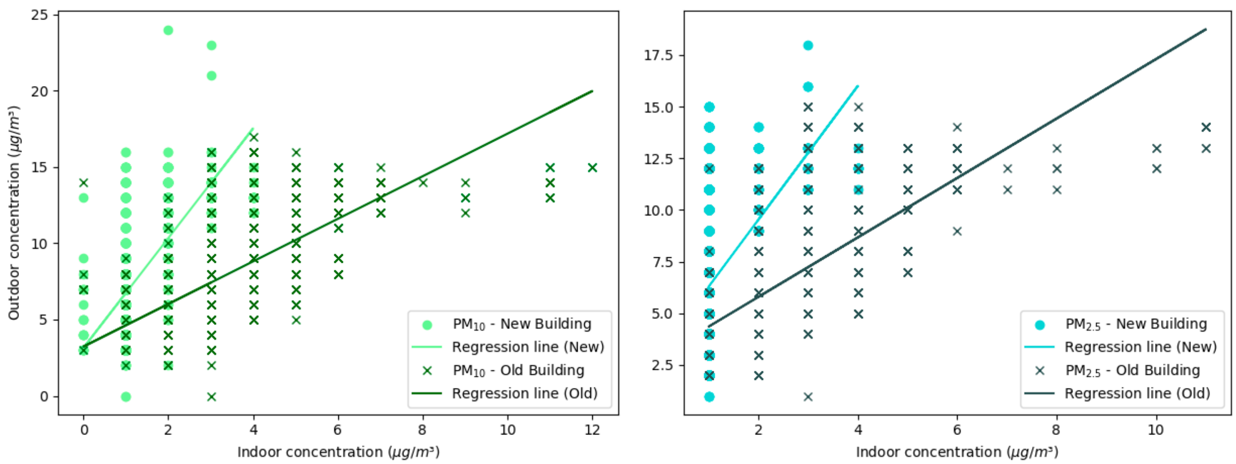

Figure 6.9: The correlation of particulate matter in office buildings.

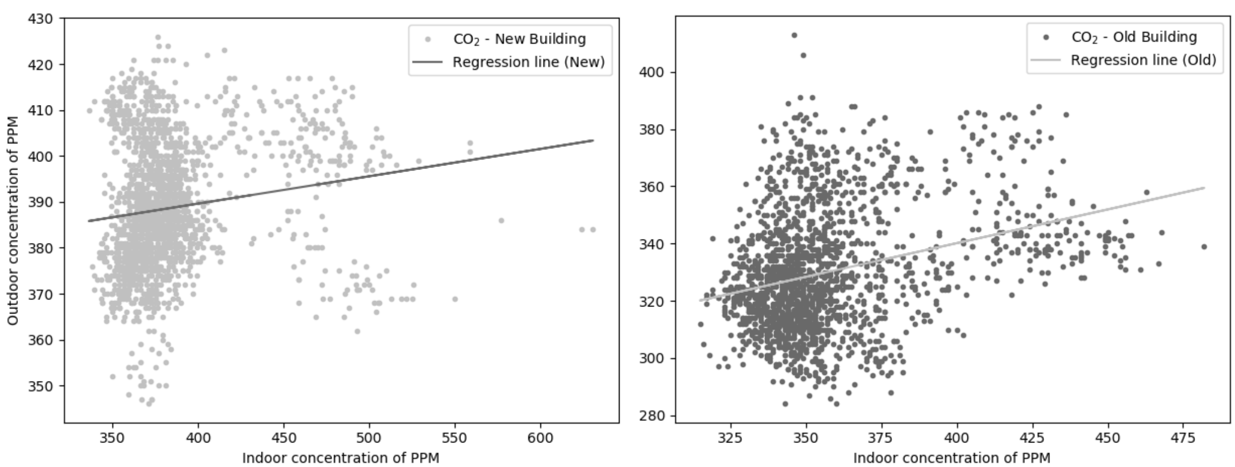

Figure 6.10: The correlation of carbon dioxide in office buildings.

The results from measuring the correlation coefficient (r) are presented in Table 6.2 for $PM_{2.5}$, PM_{10}, CO_2, temperature, and humidity. The correlation coefficient indicates the nature and strength of the relationship between x and y. Values of r range from -1 to +1.

	Office		Campus		Residential	
	A1	A2	B1	B2	C1	C2
$PM_{2.5}$	0.356	0.666	-0.017	0.528	0.558	0.114
PM_{10}	0.569	0.661	-0.012	0.682	0.449	0.096
CO_2	0.153	0.306	-0.327	0.327	0.040	-0.251
Temperature	-0.011	-0.074	0.378	0.186	0.107	0.484
Humidity	0.355	0.345	0.751	0.627	0.212	0.578

Table 6.2: Summary of correlation coefficient (r) in the monitored buildings.

In the mechanically ventilated buildings, the r-value of $PM_{2.5}$ and PM_{10} are found to be closer to 1 in older office and campus buildings than the new buildings. This indicates that there is a stronger positive correlation between indoor and outdoor concentrations of particulates in old buildings. CO_2 measurements are found to be close to 0 or with a weak negative correlation in new buildings, while old buildings has a weak positive correlation. This indicates that the relationship between the two variables are weak or none existing.

For naturally ventilated buildings, there is a stronger correlation between pollutants in newer buildings, compared to old buildings. This was calculated during high levels of particulates outside C2, as presented in Figure B.11. The levels measured of CO_2 results in a correlation close to 0 and a weak negative correlation, indicating that there is no relationship regarding CO_2 in naturally ventilated buildings.

6.3 Indoor Climate

In our last research question, we ask how well the buildings in Bergen hold up to the indoor climate norms set by the Norwegian Institute of Public Health in Norway. According to NIPH and the Norwegian Labour Inspection Authority, in an occupied building the temperature should be around 22°C and relative humidity below 60%. Building occupants should not be exposed to higher levels than 25 $\mu g/m^3$ of $PM_{2.5}$ and 50 $\mu g/m^3$ of PM_{10} over a 24-hour mean, and 1800 $\mu g/m^3$ (1000 PPM) of CO_2, see Table 2.2. We compare the results of the research to peoples reflections in regard to indoor air quality.

Comparison of Distributions

In Figure 6.11 and Figure 6.12, we represent the frequency of temperature and humidity measured in each sub-set of our population, in order to compare the indoor climate in each building to one another. The frequencies are color-coded to differentiate between the new (A1, B1, and C1) and old (A2, B2, and C2) buildings within each building type. For all measured buildings, the concentrations of pollutants indoors are lower than the indoor climate recommendations by NIPH. As different pollutants are not the only factor when creating a good indoor climate, we want to compare the indoor climate based on humidity and temperature in the different buildings.

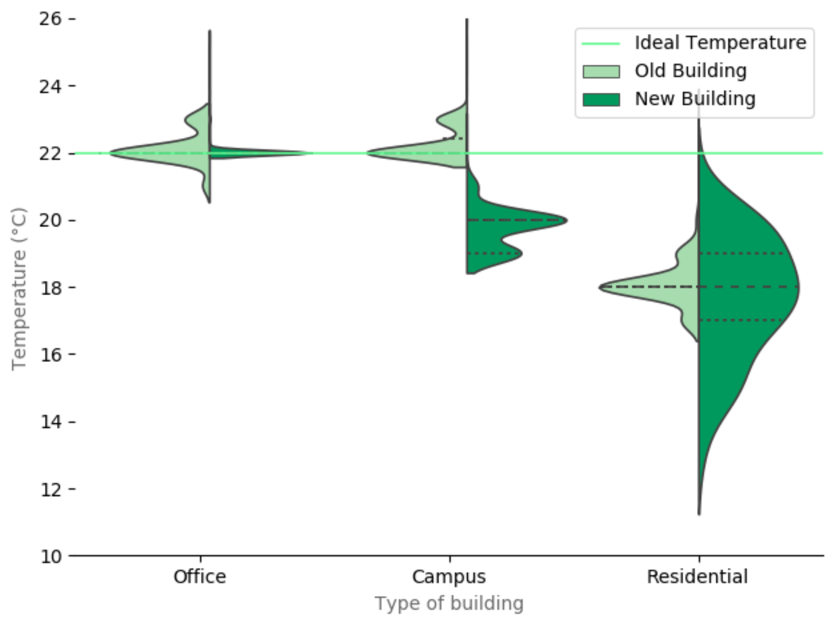

Figure 6.11: Comparison between the temperature in the different building types.

Comparing the different buildings within each category shows us that there is not a big difference between new and old buildings, towards temperature. In mechanically ventilated buildings, most rooms have thermostats that can be changed by users of the room. Thus, the user has the opportunity to change the factors within rooms in mechanically ventilated buildings.

In campus building B1, the thermostat in the area of our IoT-device was set to 20 °C, while the other mechanically ventilated buildings thermostat was set to 22 °C. The results from office and campus buildings indicate that there is no difference between new and old mechanically ventilated buildings, and the residential buildings indicate that none of the naturally ventilated buildings had ideal temperature throughout our monitoring.

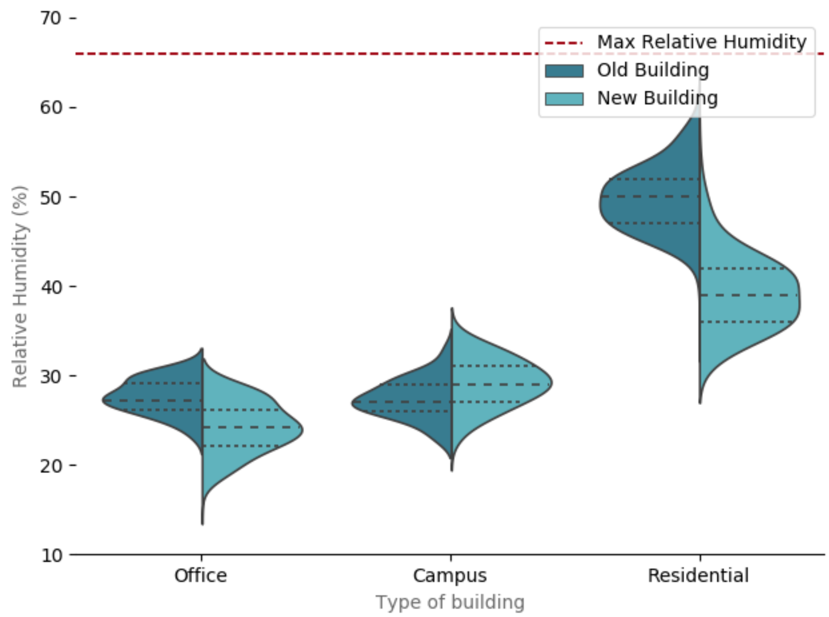

Figure 6.12: Comparison between the relative air humidity in the different buildings.

Relative humidity levels below 20% are recognized as a low level, while above 60% is recognized as a high level. High air humidity will increase chances of mold and feeling damp, while low air humidity is associated with a dry throat, dry skin, and chapped lips. As shown in Figure 6.12, office buildings, old and new, has usual levels of relative humidity between 30 and 20%. This is far below the max level of relative humidity according to NIPH, the same results are indicated for our campus buildings. The residential buildings had a much higher relative humidity during our observations, as these buildings were naturally ventilated.

Building Occupants Opinions

The research conducted in this thesis is focused on indoor air quality awareness, making it a higher priority for building occupants. In order to create a better prototype, and to understand more about peoples reflections of indoor air quality, we interviewed occupants in the different building sites monitored. We asked them questions regarding their instant thoughts on air quality when entering a building, and if they believe that the air quality is better in newer buildings. The results are presented in Figure 6.13.

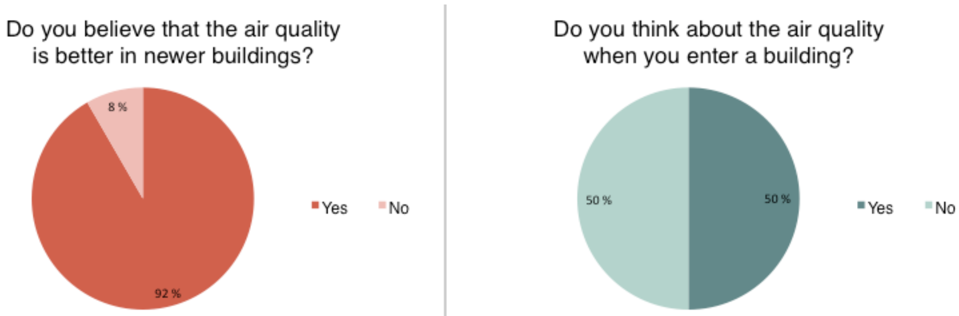

Figure 6.13: Representation of building occupants mindset around air quality.

To create a better prototype and understand what to visualize, we need to get a better understanding of the different building occupants opinion around indoor air quality. After completing our interview process with 12 different building occupants, the results show that 92% of the occupants believe that the air quality in a new building is better. This tells us that the air quality should be better in newer buildings, in line with the results of our analysis of the indoor climate. The fact that just half of the people interviewed "thinks about air quality when entering a building", indicates that informing building occupants can help to express awareness around indoor air quality. Our dashboard web application can, therefore, help to address this issue.

6.4 Sources of Error

The results of a research project will always be accompanied by certain sources of error. The sources of error can occur both in the planning, collection, and processing of the data. Such errors are well documented, and a lot of time and effort is invested to reduce the possibility of their occurrence when conducting research. In this thesis, sources of error have occurred during different times throughout the collection of data.

During the planning process of our research, the design of our IoT-devices did not go as planned. First, there were problems with the operating temperature, that the devices could not handle temperatures below 0, and the first batch of sensors did not handle rainy weather. These errors delayed our process in monitoring the different sites by months. Our plan was to find areas within each building sub-set that was similar, e.g., hallway or apartment. This was a source of error when analyzing the indoor climate in the new campus building, as the thermostat in the area of monitoring was set to 20 $^{\text{o}}$C compared to 22 $^{\text{o}}$C in the old building.

Making sure that the data collection was reliable, did not go as first planned. At one of the sites, it was hard to keep the sensors away from the sun. In one of the residential buildings, the apartment monitored was on ground-level, while the other was at the 4^{th} floor. Scenarios like these can affect the results. We also had issues connecting the gateways and setting up sensors. These issues occurred due to problems with internet permission granted from the buildings IT-department, and for allowing to set up sensors by the Facility Management. This resulted in half a day missing for comparing the campus buildings, compared to the other buildings monitored. Our first collection of data was comparing the different sensors to each other and a reference station. As there is no reference station monitoring CO_2, we were not able to legitimize that the measurements of CO_2 was completely accurate.

The last step before analyzing our results was to process the raw payloads read by our database. Depending on the sensor, data was transmitted every 5-6 minute. This indicates a data collection of 1680-2016 data points within 7 days. In some cases, the sensors sent false positives as data. In such cases, we had to remove the data points, making some of the collections of data smaller than first anticipated. Sources of error occurred throughout the whole process of this thesis, and even though the path to our final goal was not straight forward, our collections of data, analysis, and prototype ended up as a success.

Chapter 7

Conclusion

In this chapter, we conclude by summarizing how our results answers in regard to the research questions and goals listed in Chapter 1. We also present a proposal for future work, with the potential of creating a better prototype for air quality analysis in buildings.

Sub-Goals

In order to create a better prototype when conducting our research, we defined three sub-goals and how to achieve them. We begin by concluding the results from our tests, presented in Chapter 6.

- *"Create an easy-to-use web application to visualize air quality information."*
  The answers submitted show that the majority of users agree on all of the statements presented to them. The application we created has all the requirements we want in a dashboard application. This indicates that as a dashboard web application, it was a success. And thus, our application includes all features needed to help increase air quality awareness in buildings.

- *"Use cheap sensors and still get accurate data."*
  The sensors used for research are cheap compared to methods used in related research. Therefore, we tested our sensors reliability to each other and a reference station. The results show that the sensors' observations are close to equal, with a high correlation, to both each other and the reference station. This indicates that our inexpensive sensors are observing accurate and honest data.

- *"Find out whether people reflect about indoor air quality."*
  After interviewing multiple building occupants, only half of the participants think about air quality when entering a building, while 92% believe the air quality is better in newer buildings. The results show that our participants indicate that the indoor air quality is better in newer buildings. This is accurate compared to our results from analyzing the indoor climate. In addition, it indicates that our application can increase air quality awareness in buildings.

Reasearch Questions

In this thesis, we have created a functional air quality analysis prototype that can harvest collections of data from various buildings around Bergen, including a designed dashboard web application and IoT devices for data observations. The prototype helps researchers by simplifying the process of collecting air quality data, visualizing, and analyzing the data. After working through all parts of this thesis we are able to answer the research questions presented in Chapter 1:

- *"Can low powered, cheap IoT-devices with a long range be used to monitor the relationship between indoor and outdoor air quality?"*
  We created a solution for an architecture analyzing the data collected using IoT-devices. We uncovered that using these devices based on two different sensors, with LoRaWAN as communication, can be used to monitor the relationship between indoor and outdoor air quality. The sensors are low powered and cheap, covering up to 15 km. This indicates that the solution is a reliable option in comparison to devices used to monitor in related research.

- *"Is there a correlation between indoor air quality and outdoor pollutants in buildings around Bergen?"*
  Our results show that there is a weak to moderate correlation between indoor air quality and outdoor pollutants in buildings around Bergen. There is a stronger correlation in new naturally ventilated buildings, while old mechanically ventilated buildings has a stronger positive correlation.

- *"Is there a stronger relationship between indoor air quality and outdoor pollutants in older buildings?"*
  In mechanically ventilated buildings, the I/O ratio are found to be higher and have a stronger positive correlation in older buildings, in regard to $PM_{2.5}$ and PM_{10}. In naturally ventilated buildings the I/O ratio are found to be higher in old buildings, while having a stronger correlation in new buildings, in regard to $PM_{2.5}$ and PM_{10}. This indicates that there is a stronger relationship between indoor air quality and outdoor pollutants in older mechanically ventilated buildings.

- *"How well do buildings in Bergen hold up to the indoor climate norms set by the Norwegian Institute of Public Health?"*
  According to NIPH, the indoor climate norms include safe concentrations of pollutants, a certain range of temperature, and a maximum level of relative humidity. The whole population monitored follows the indoor climate norms by NIPH. However, there is a difference between the population.

People spend most of their time indoors, at work, in bed, or inside restaurants. Air pollution is a major cause of premature deaths worldwide, and most buildings have no monitoring or visualization of the indoor climate. When the indoor air quality exceeds guideline limits, IoT could help us overcome these challenges. The prototype built in this thesis, using long-range communication offers the potential for more continuous monitoring of buildings. This will help increase air quality awareness and prevent occupants from long-term exposure to high concentrations.

7.1 Future Work

The presented research in this thesis is a good starting point for further air quality analysis research using IoT. Future conducted work would be to apply a machine learning model. The model should be trained to make predictions of parameters such that proactive actions can be taken to alleviate the impacts of air pollution. This can make buildings save money, and keep the indoor climate good for the building occupants by turning on/off the ventilation throughout the day.

In future work, the population in our thesis should be increased within each sub-group, making the population larger and more diversified. This will enrich the datasets, the statistical analysis, and help train the model. Creating multiple sets of calibrated sensors will increase the number of buildings within the population one can monitor at a time. This will make sure you get a more honest dataset when comparing the different buildings within each sub-set. If the researchers in future work have sensors for each building in the population, data can be gathered throughout multiple seasons that will differ in temperature and exposure to pollution. This will produce large datasets, that can be used for a larger analysis as well as a better trained model.

Glossary

ALOHA

A random access protocol, widely used for coordinating the access of large numbers of intermittent transmitters in a single shared communication channel.

Attenuation

The reduction of the amplitude of a signal, electric current, or other oscillation.

Chirp Spread Spectrum

A spread spectrum technique that use wide-band linear frequency modulated chirp pulses to encode information.

Greenhouse gas

Any gas that has the property of absorbing infrared radiation (net heat energy) emitted from Earth's surface and reradiating it back to Earth's surface.

Multipath fading

It occurs when signals reach a receiver via many paths, and their relative strengths and phases change.

PM_{10}

$PM_{2.5}$ are fine particles less than 10 μg.

$PM_{2.5}$

$PM_{2.5}$ are fine particles less than 2.5 μg.

List of Acronyms and Abbreviations

AQI	Air Quality Index.
AWS	Amazon Web Services.
CO	Carbon monoxide.
CO_2	Carbon dioxide.
EEA	European Environment Agency.
I/O	Indoor and Outdoor.
IAQ	Indoor Air Quality.
IEA	International Energy Agency.
IoT	Internet of Things.
LoRa	Long Range.
LoRaWAN	Long Range Wide Area Network.
LPWAN	Low-Power Wide-Area Network.
M2M	Machine-to-Machine.
NILU	Norwegian Institute for Air Research.
NIPH	Norwegian Institute of Public Health.
NO_2	Nitrogen dioxide.
O_3	Ozone.
PM	Particulate Matter.
RFID	Radio Frequency Identification.
SO_2	Sulfur dioxide.
SOA	Service-oriented architecture.
TTN	The Things Network.
VOCs	Volatile organic compounds.
WHO	World Health Organization.

Bibliography

[1] Mohammed Abdmeziem, Djamel Tandjaoui, and Imed Romdhani. "Architecting the Internet of Things: State of the Art". In: *Robots and Sensor Clouds*. Springer, 2016, pp. 55–75. ISBN: 978-3-319-22168-7.

[2] Anita S. Acharya, Anupam Prakash, Pikee Saxena, and Aruna Nigam. "Sampling: Why and how of it". In: *Indian Journal of Medical Specialties* 4.2 (2013), pp. 330–333.

[3] Actility, Gemalto, and Semtech. *White paper: LoRaWAN™ SECURITY. Full End-To-End Encryption For IoT Application Providers*. Tech. rep. Feb. 2017.

[4] European Environment Agency. *Air pollution*. 2017. URL: `https://www.eea.europa.eu/themes/air/air-pollution-sources` (visited on 01/13/2018).

[5] European Environment Agency. *Air pollution sources*. 2017. URL: `https://www.eea.europa.eu/themes/air/air-pollution-sources` (visited on 01/13/2018).

[6] European Environment Agency. *Air quality index*. 2018. URL: `https://www.eea.europa.eu/ds_resolveuid/6d9fb51e062f467d950048cc7d5919b5` (visited on 01/10/2019).

[7] International Energy Agency. *CO2 Emissions Statistics*. 2018. URL: `https://www.iea.org/statistics/co2emissions/` (visited on 03/20/2019).

[8] United States Environment Protection Agency. *Introduction to Indoor Air Quality*. 2018. URL: `https://www.epa.gov/indoor-air-quality-iaq/introduction-indoor-air-quality` (visited on 01/10/2019).

[9] United States Environment Protection Agency. *Particulate Matter (PM) Basics*. 2018. URL: `https://www.epa.gov/pm-pollution/particulate-matter-pm-basics` (visited on 11/15/2018).

[10] LoRa Alliance. *White paper: LoRaWAN™. What is it? A technical overview of LoRa and LoRaWAN™*. Tech. rep. Nov. 2015.

[11] Gisle Andersen. *Forurensning*. 2017. URL: https://snl.no/forurensning (visited on 01/10/2019).

[12] Arbeidsmiljøloven. *Lov om arbeidsmiljø, arbeidstid og stillingsvern mv. (arbeidsmiljøloven) (LOV-2005-06-17-62)*. 2005. URL: https://lovdata.no/dokument/NL/lov/2005-06-17-62 (visited on 10/20/2018).

[13] Arbeidstilsynet. *Veiledning om Klima og luftkvalitet på arbeidsplassen*. 2016. URL: https://www.arbeidstilsynet.no/contentassets/3f86f6d2038348d18540404144f76a22/luftkvalitet-pa-arbeidsplassen.pdf (visited on 10/20/2018).

[14] The Norwegian Labour Inspection Authority. *The Workplace Regulations*. 2018. URL: https://www.arbeidstilsynet.no/en/laws-and-regulations/regulations/the-workplace-regulations/ (visited on 10/21/2018).

[15] Patrice Blondeau, Vlad Iordache, Olivier Poupard, and Francis Allard. "Relationship between outdoor and indoor air quality in eight French schools". In: *Atmospheric Environment* 39.11 (2005), pp. 2071–2080. ISSN: 1352-2310. DOI: /10.1016/j.atmosenv.2004.12.016.

[16] Tim Brown. "Design Thinking". In: *Harvard Business Review* 86.6 (2008), pp. 84–92.

[17] Agustin Candia, Soledad Represa, Daniela Giuliani, Miguel Luengo, Andres Porta, and Luis Marrone. "Solutions for SmartCities: proposal of a monitoring system of air quality based on a LoRaWAN network with low-cost sensors". In: IEEE, 2018, pp. 1–6. DOI: 10.1109/CACIDI.2018.8584183.

[18] Charles T Carroll. "Tool Training: Voice of the Customer". In: *Six Sigma for Powerful Improvement: A Green Belt DMAIC Training System with Software Tools and a 25-Lesson Course*. CRC Press, 2013, p. 20. ISBN: 9781466564695.

[19] Avril Challoner and Laurence Gill. "Indoor/outdoor air pollution relationships in ten commercial buildings: PM2.5 and NO2". In: *Building and Environment* 80 (2014), pp. 159–173. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2014.05.032.

[20] Josef Cyrys, Mike Pitz, Wolfgang Bischof, Heinz-Erich Wichmann, and Joachim Heinrich. “Relationship between indoor and outdoor levels of fine particle mass, particle number concentrations and black smoke under different ventilation conditions”. In: *Journal Of Exposure Analysis And Environmental Epidemiology* 14.4 (2004), pp. 275–283. ISSN: 1559-064X. DOI: 10.1038/sj.jea.7500317.

[21] Dash. *Dash Layout*. URL: https://dash.plot.ly/getting-started (visited on 04/20/2019).

[22] Dash. *Introduction to Dash*. URL: https://dash.plot.ly/introduction (visited on 04/20/2019).

[23] Aline Dresch, Daniel P. Lacerda, and Jose A. V. Antunes Jr. “Design Science Research”. In: *Design Science Research: A Method for Science and Technology Advancement*. Springer, 2015, pp. 67–102. ISBN: 9783319073736.

[24] The Working Environment and Safety Department. *The Working Environment Act*. 2008. URL: https://www.regjeringen.no/en/topics/labour/the-working-environment-and-safety/innsikt/the-working-environment-act/id447107/ (visited on 10/21/2018).

[25] Folkehelseinstituttet. *Anbefalte faglige normer for inneklima*. 2015. URL: https://www.fhi.no/globalassets/dokumenterfiler/rapporter/2015/anbefalte-faglige-normer-for-inneklima-pdf.pdf (visited on 11/01/2018).

[26] Folkehelseinstituttet. *Inneklima og helseplager*. 2018. URL: https://www.fhi.no/ml/miljo/inneklima/fremhevede-artikler-inneklima-og-helse/inneklima-og-helseplager-ny/ (visited on 11/01/2018).

[27] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash. “Internet of Things: A survey on enabling technologies protocols and applications”. In: *IEEE Communications Surveys & Tutorials* 17.4 (2015), pp. 2347–2376. ISSN: 1553-877X. DOI: 10.1109/COMST.2015.2444095.

[28] MaryAnne M Gobble. “Design Thinking”. In: *Research-Technology Management* 57.3 (2007), pp. 59–61. ISSN: 0895-6308. DOI: 10.5437/08956308X5703005.

[29] Wisconsin Department of Health Services. *Carbon Dioxide*. 2018. URL: https://www.dhs.wisconsin.gov/chemical/carbondioxide.htm (visited on 03/20/2019).

[30] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. "Design science in Information Systems research". In: *MIS Quarterly* 28.1 (2004), pp. 75–105. ISSN: 0276-7783. DOI: 10.2307/25148625.

[31] Jerry L. Hintze and Ray D. Nelson. "Violin Plots: A Box Plot-Density Trace Synergism". In: *American Statistician* 52 (1998), pp. 181–184. ISSN: 1537-2731. DOI: 10.2307/2685478.

[32] Mark Z. Jacobson. "100% Clean, Renewable Energy and Storage for Everything". Unpublished. 2019. URL: https://web.stanford.edu/group/efmh/jacobson/Articles/I/AirPollutionDeath.pdf.

[33] Norsk institutt for luftforskning. *Innendørs luftkvalitet*. 2018. URL: https://www.nilu.no/forskning/urban-luftkvalitet/innendors-luftkvalitet/ (visited on 12/02/2018).

[34] Yang Lv, Haifeng Wang, Shanshan Wei, Lei Zhang, and Qi Zhao. "The Correlation between Indoor and Outdoor Particulate Matter of Different Building Types in Daqing, China". In: *Procedia Engineering* 205 (2017), pp. 360–367. ISSN: 1877-7058. DOI: 10.1016/j.proeng.2017.10.002.

[35] Ieva Martinkenaite. *5 global challenges that could be solved by the Internet of Things (IoT)*. Telenor. 2018. URL: https://www.telenor.com/5-global-challenges-that-could-be-solved-by-the-internet-of-things-iot/ (visited on 03/01/2019).

[36] David Massey, Jamson Masih, Aditi Kulshrestha, Mahima Habil, and Ajay Taneja. "Indoor/outdoor relationship of fine particles less than 2.5μm (PM2.5) in residential homes locations in central Indian region". In: *Building and Environment* 44.10 (2009), pp. 2037–2045. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2009.02.010.

[37] Aonghus McNabola, Andrew McCreddin, Laurence Gill, and Brian Broderick. "Analysis of the relationship between urban background air pollution concentrations and the personal exposure of office workers in Dublin, Ireland, using baseline separation techniques". In: *Atmospheric Pollution Research* 2.1 (2011), pp. 80–88. ISSN: 1309-1042. DOI: 10.5094/APR.2011.010.

[38] Miljødirektoratet. *Karbondioksid (CO2)*. 2018. URL: https://www.miljostatus.no/tema/klima/norske-klimagassutslipp/co2-utslipp/ (visited on 03/20/2019).

[39] Judith Molka-Daniels, Veronika Olešnaníková, Peter Šarafín, Róbert Žalman, and Per Engelseth. "Systems Analytics Approach Using Wireless Sensor Network Technologies and Big Data Visualization for Continuous Assessment of Air Quality in a Workplace Environment". In: *NOKOBIT* 24.1 (2016). Paper presented at NOKOBIT 2016, Bergen, 28-30 Nov. ISSN: 1894-7719.

[40] The Things Network. *Building a global open LoRaWAN™ network*. 2019. URL: `https://www.thethingsnetwork.org` (visited on 03/20/2019).

[41] The Things Network. *Kelrink LoRa IoT Station*. 2019. URL: `https://www.thethingsnetwork.org/docs/gateways/kerlink/` (visited on 04/20/2019).

[42] Shun-Chen Niu. "Data Collection and Sampling". University Lecture. unpublished. 2000. URL: `https://www.utdallas.edu/~scniu/OPRE-6301/documents/Data_Collection_and_Sampling.pdf`.

[43] Luftkvalitet i Norge. *Lokal luftforurensning*. URL: `https://luftkvalitet.miljostatus.no/artikkel/170` (visited on 11/15/2018).

[44] World Health Organization. *Ambient air pollution: Pollutants*. 2018. URL: `https://www.who.int/airpollution/ambient/pollutants/en/` (visited on 01/10/2019).

[45] World Health Organization. *Climate risks from CO2 and short-lived climate pollutants*. 2019. URL: `https://www.who.int/sustainable-development/housing/health-risks/climate-pollutants/en/` (visited on 03/20/2019).

[46] World Health Organization. *Health effects of particulate matter. Policy implications for countries in eastern Europe, Caucasus and central Asia*. 2013. URL: `http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2013/health-effects-of-particulate-matter.-policy-implications-for-countries-in-eastern-europe,-caucasus-and-central-asia-2013` (visited on 11/11/2018).

[47] Pandas. *Python Data Analysis Library*. 2019. URL: `https://pandas.pydata.org/` (visited on 04/20/2019).

[48] Koen Pauwels, Tim Ambler, Bruce Clark, Pat LaPointe, David Reibstein, Bernd Skiera, Berend Wierenga, and Thorsten Wiesel. "Dashboards as a Service. Why, What, How, and What Research Is Needed?" In: *Journal of Service Research* 12.2 (2009), pp. 175–189. DOI: `10.1177/1094670509344213`.

[49] Ken Peffers, Tuure Tuunanen, Marcus A Rothenberger, and Samir Chatterjee. "A Design Science Research Methodology for Information Systems Research". In: *Journal of Management Information Systems* 24.3 (2007), pp. 45–77. ISSN: 1557-928X. DOI: 10.2753/MIS0742-1222240302.

[50] Andrea Polidori, Mohammad Arhami, Constantinos Sioutas, Ralph Delfino, and Ryan Allen. "Indoor/Outdoor Relationships, Trends, and Carbonaceous Content of Fine Particulate Matter in Retirement Homes of the Los Angeles Basin". In: *Journal of the Air & Waste Management Association* 57.3 (2007), pp. 366–379. DOI: 10.1080/10473289.2007.10465339.

[51] Prabhat Rai. "Particulate Matter and Its Size Fractionation". In: *Biomagnetic Monitoring of Particulate Matter*. Elsevier, 2016, pp. 1–13. ISBN: 978-0-12-805135-1.

[52] Usman Raza, Parag Kulkarni, and Mahesh Sooriyabandara. "Low Power Wide Area Networks: An Overview". In: *IEEE Communications Surveys & Tutorials* 19.2 (2017), pp. 855–873. ISSN: 1553-877X. DOI: 10.1109/COMST.2017.2652320.

[53] Rim Razzouk and Valerie Shute. "What Is Design Thinking and Why Is It Important?" In: *Review of Educational Research* 82.3 (2012), pp. 330–348. DOI: 10.3102/0034654312457429.

[54] Ronald Rousseau, Leo Egghe, and Raf Guns. "Statistics". In: *Becoming Metric-Wise: A Bibliometric Guide for Researchers*. Chandos Publishing, Jan. 2018, pp. 67–90. ISBN: 9780081024744.

[55] Shadi Al-Sarawi, Mohammed Anbar, Kamal Alieyan, and Mahmood Alzubaidi. "Internet of things (IoT) communication protocols: Review". In: IEEE, 2017. DOI: 10.1109/ICITECH.2017.8079928.

[56] National Institute of Scinece and Technology. *VIOLIN PLOT*. 2015. URL: https://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/violplot.htm (visited on 04/20/2019).

[57] Sensirion. *About Us*. 2019. URL: https://www.sensirion.com/en/about-us/ (visited on 04/04/2019).

[58] Sensirion. *Datasheet SPS30*. 2018. URL: https://cdn.sparkfun.com/assets/2/d/2/a/6/Sensirion_SPS30_Particulate_Matter_Sensor_v0.9_D1__1_.pdf (visited on 04/04/2019).

[59] Sensirion. *Interface Description Sensirion SCD30 Sensor Module*. 2018. URL: https://cdn.sparkfun.com/assets/d/c/0/7/2/SCD30_Interface_Description.pdf (visited on 04/04/2019).

[60] Amazon Web Services. *About AWS*. 2019. URL: https://aws.amazon.com/about-aws/ (visited on 03/20/2019).

[61] Amazon Web Services. *Amazon DynamoDB*. 2019. URL: https://aws.amazon.com/dynamodb/ (visited on 03/20/2019).

[62] Amazon Web Services. *AWS IoT Core*. 2019. URL: https://aws.amazon.com/iot-core/ (visited on 03/20/2019).

[63] Amazon Web Services. *AWS SDK for Python (Boto3)*. 2019. URL: https://aws.amazon.com/sdk-for-python/ (visited on 04/20/2019).

[64] Amazon Web Services. *Overview of AWS SDK Support for DynamoDB*. 2019. URL: https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/Programming.SDKOverview.html (visited on 04/20/2019).

[65] Pallavi Sethi and Smruti Sarangi. "Internet of Things: Architectures, Protocols, and Applications". In: *Journal of Electrical and Computer Engineering* (2017), p. 25. ISSN: 2090-0147. DOI: 10.1155/2017/9324035.

[66] Linda Shields and Alison Twycross. "The difference between quantitative and qualitative research". In: *Paediatric nursing* 15.9 (2003), p. 24.

[67] Marianne Stranger, Sanja S. Potgieter-Vermaak, and Rene Van Grieken. "Particulate matter and gaseous pollutants in residences in Antwerp, Belgium". In: *Science of The Total Environment* 407.3 (2009), pp. 1182–1192. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2008.10.019.

[68] Min Thu, Wunna Htun, Pyone Schwe, and Nay Tun. "Smart Air Quality Monitoring System with LoRaWAN". In: IEEE, 2018, pp. 10–15. DOI: 10.1109/CACIDI.2018.8584183.

[69] Odd F. Vaage. "Utendørs $2\frac{1}{2}$ time - menn mer enn kvinner". In: *Samfunnsspeilet* 26.4 (2012), pp. 37–42.

[70] Daniel Vallero. "The State of the Atmosphere". In: *Fundamentals of Air Pollution*. Vol. 4. Academic Press, Aug. 2014, pp. 5–7. ISBN: 978-0-12-401733-7.

[71] Janine Wichmann, Tomas Lind, Måns A. Nilsson, and Tom B. Bellander. "PM 2.5, soot and NO 2 indoor–outdoor relationships at homes, pre-schools and schools in Stockholm, Sweden". In: *Atmospheric Environment* 44.36 (2010), pp. 4536–4544. ISSN: 1352-2310. DOI: 10.1016/j.atmosenv.2010.08.023.

Appendices

A Interview and Questionnaire

Indoor Air Quality Interview of Building Occupants

In order to evaluate and answer our sub-problem regarding peoples opinion on air quality in new buildings compared to old. We asked each participant two questions:

1. Do you think about the air quality when you enter a building?

2. Do you believe that the air quality is better in newer buildings?

Evaluation of Prototype

In order to evaluate our prototype, we had participants try our web application before answering their level of agreement to five different statements. They got told that: ”In this evaluation, they will be presented with a Dashboard Web Application designed to visualize the relationship between indoor and outdoor air quality. The purpose of this web application is to get a better understanding of how the air quality in your buildings are withstanding the pollution and weather elements from outside.” Each participant was told how to navigate, and what type of features that are available to the user.

Statement: The web application serves it purpose.
Level of agreement: Strongly Disagree — Disagree — Neutral — Agree — Strongly Agree

Statement: The web application is easy to navigate through.
Level of agreement: Strongly Disagree — Disagree — Neutral — Agree — Strongly Agree

Statement: The color combination and color scheme is good.
Level of agreement: Strongly Disagree — Disagree — Neutral — Agree — Strongly Agree

Statement: The web application is understandable.
Level of agreement: Strongly Disagree — Disagree — Neutral — Agree — Strongly Agree

Statement: The web application has no unnecessary features.
Level of agreement: Strongly Disagree — Disagree — Neutral — Agree — Strongly Agree

B Figures

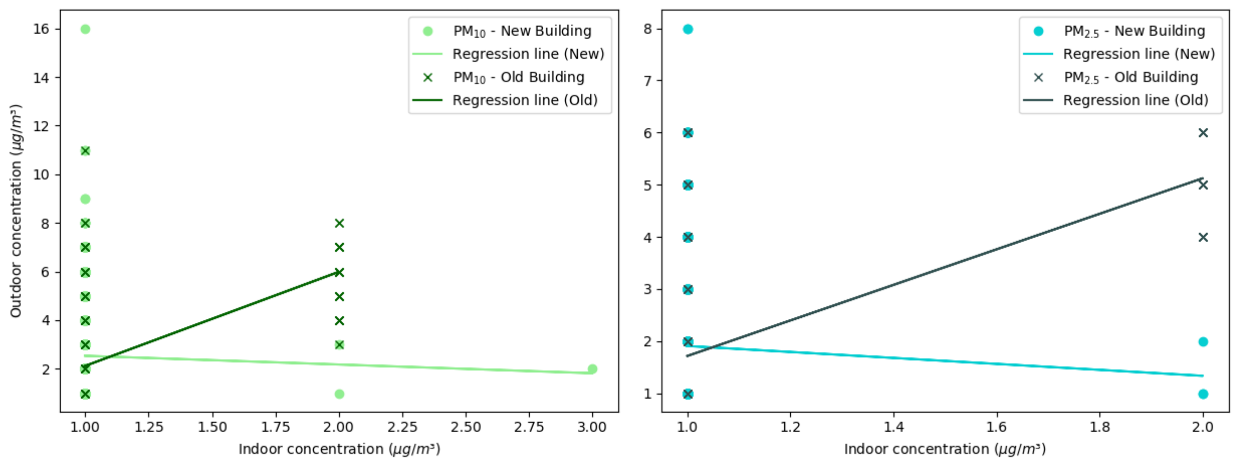

Figure B.1: The correlation of particulate matter in campus buildings

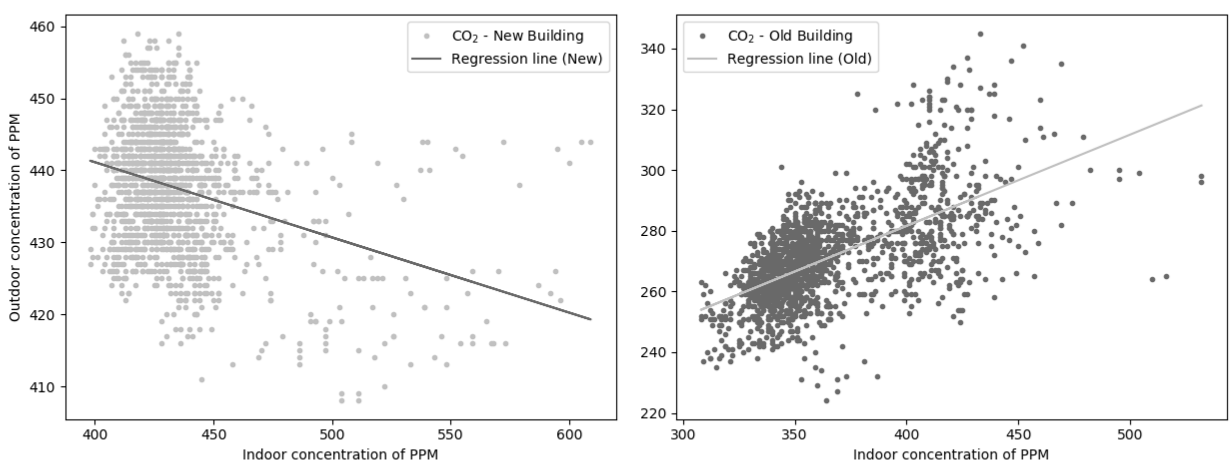

Figure B.2: The correlation of carbon dioxide in campus buildings

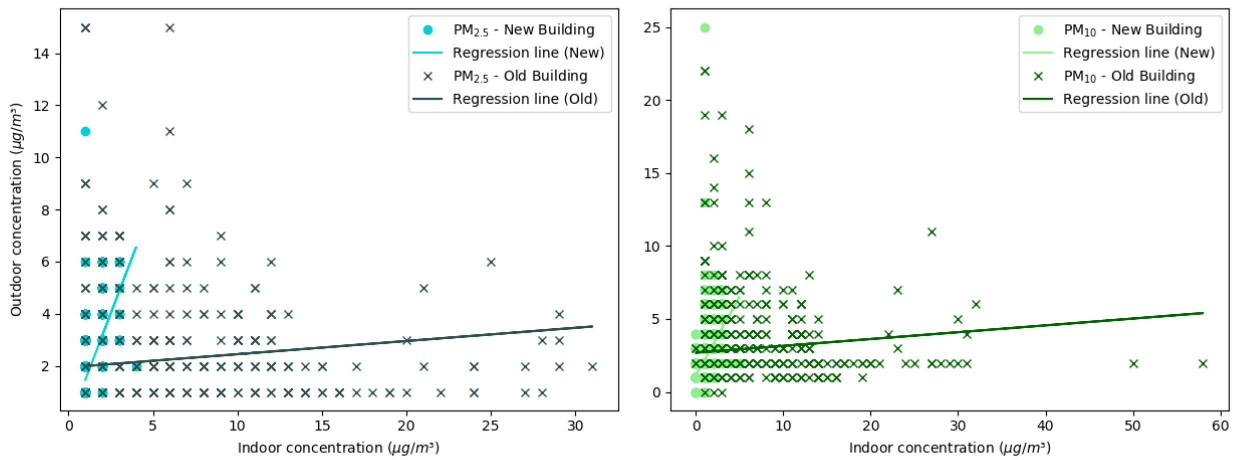

Figure B.3: The correlation of particulate matter in residential buildings

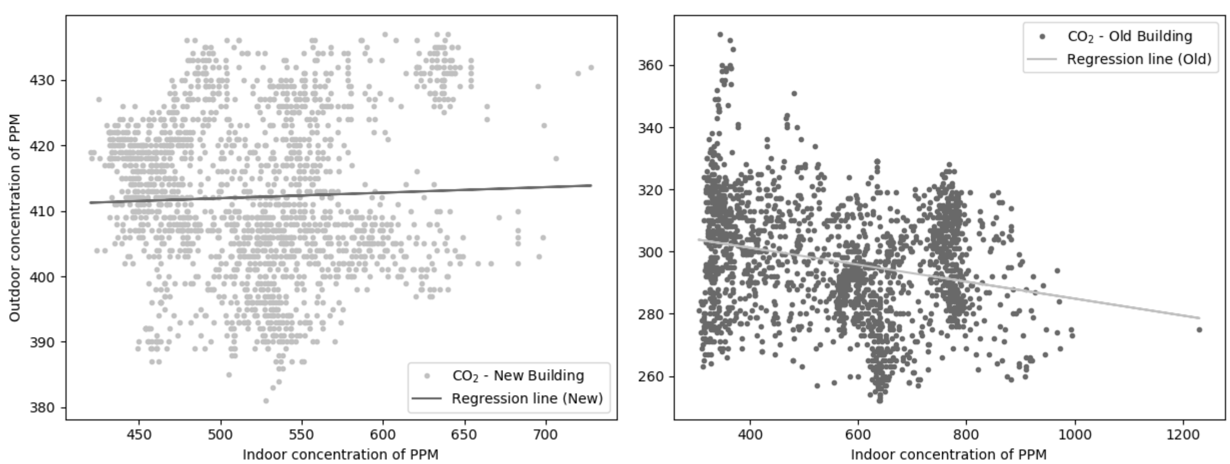

Figure B.4: The correlation of carbon dioxide in residential buildings

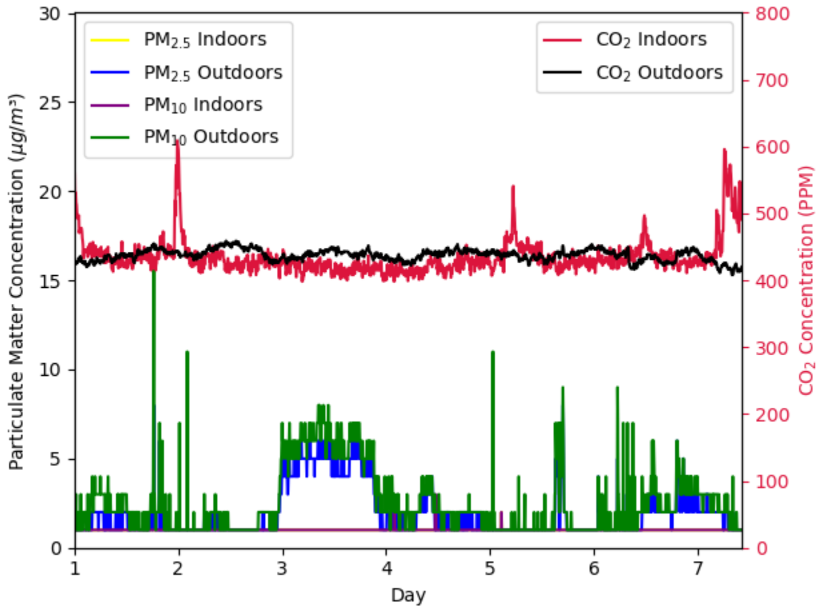

Figure B.5: Time series comparison between pollutants in campus building B1

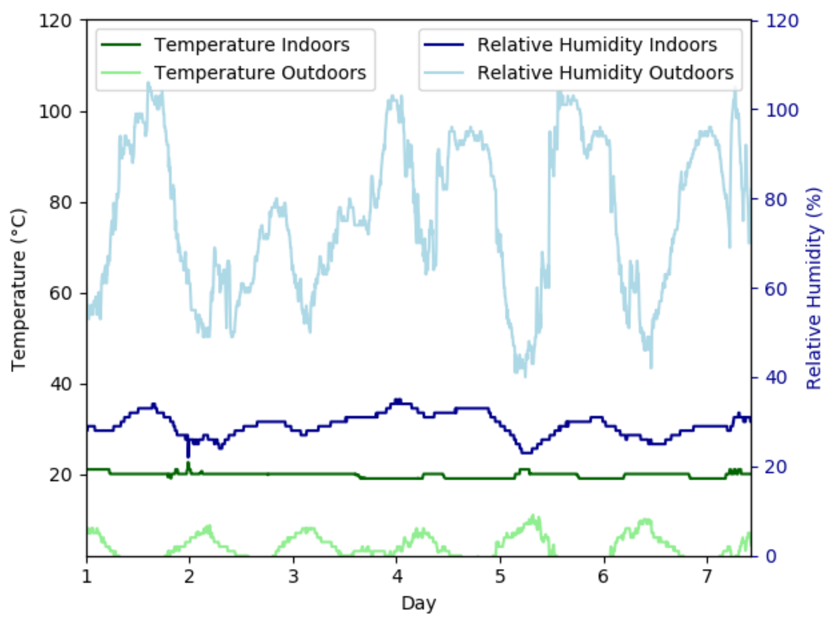

Figure B.6: Time series comparison between temperature and humidity in campus building B1

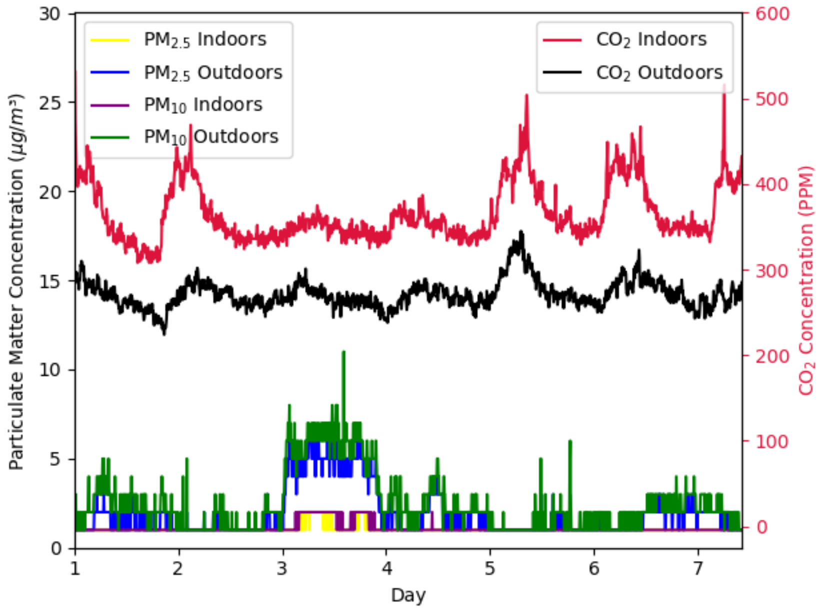

Figure B.7: Time series comparison between pollutants in campus building B2

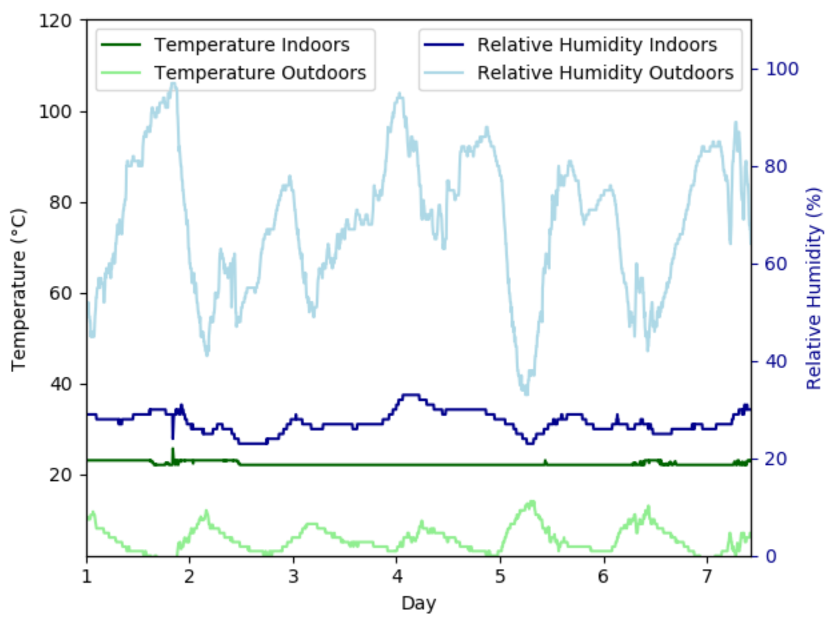

Figure B.8: Time series comparison between temperature and humidity in campus building B2

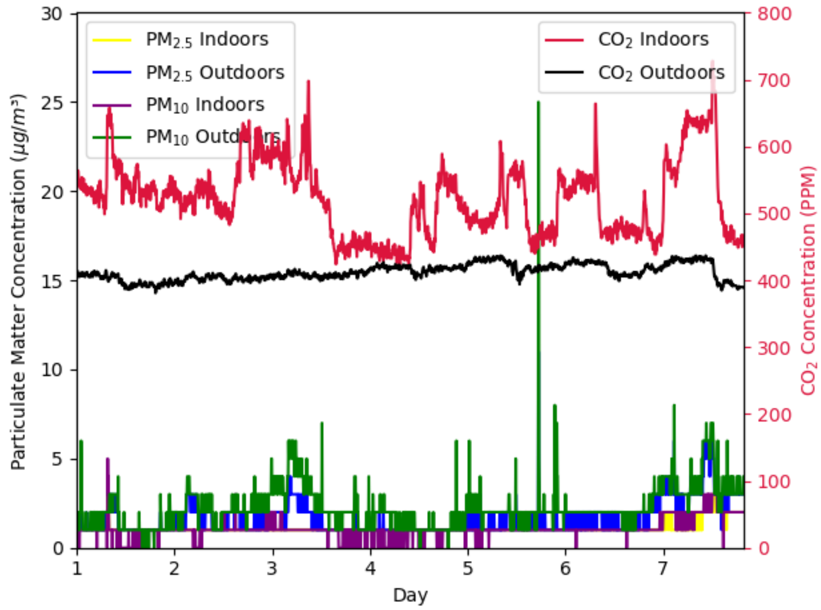

Figure B.9: Time series comparison between pollutants in residential building C1

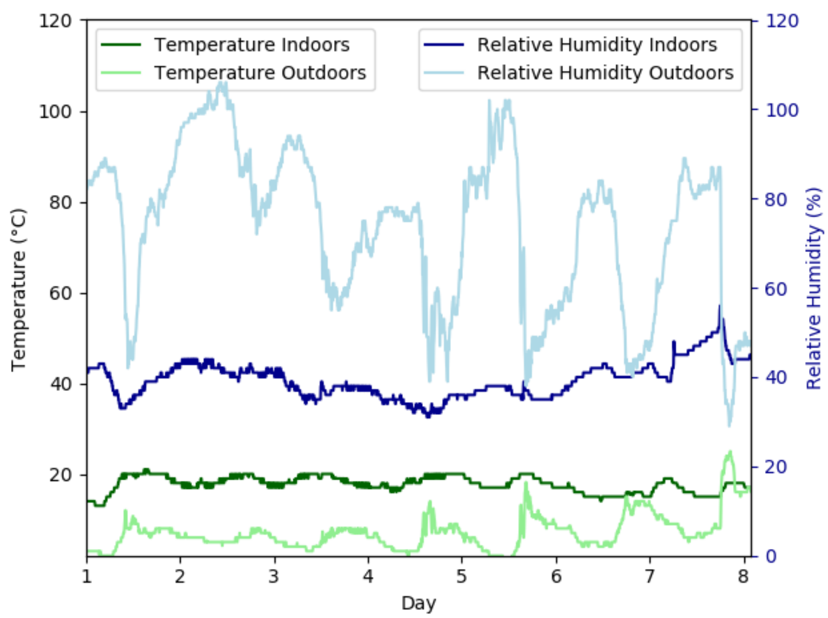

Figure B.10: Time series comparison between temperature and humidity in residential building C1

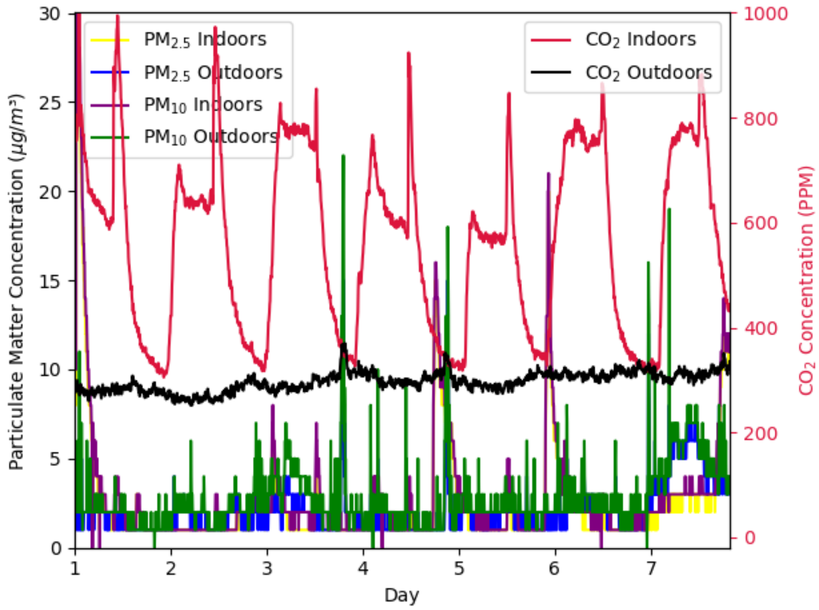

Figure B.11: Time series comparison between pollutants in residential building C2

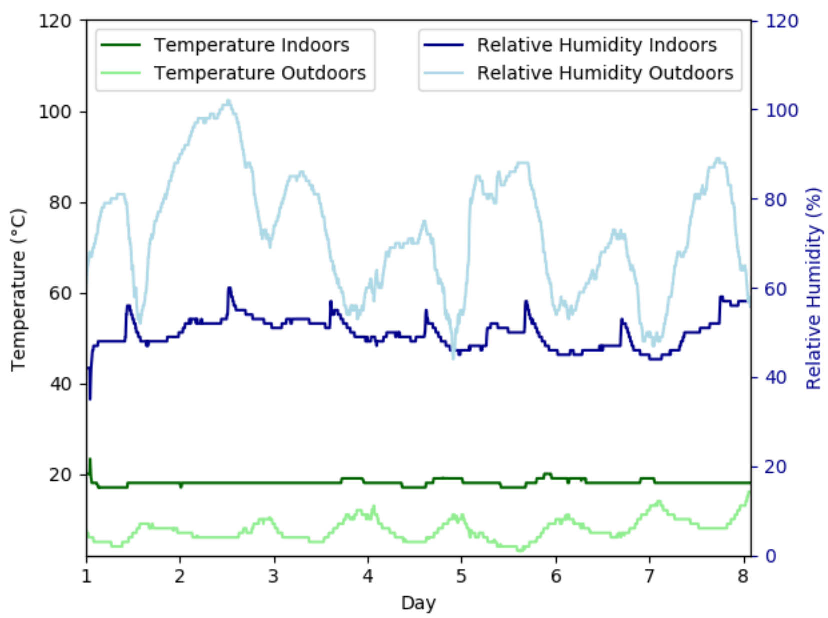

Figure B.12: Time series comparison between temperature and humidity in residential building C2