CHALMERS
UNIVERSITY OF TECHNOLOGY

Making sense of low-cost sensors

Air quality monitoring in Gothenburg, Sweden

Master's thesis in Industrial Ecology

MEHDI GHALEBANI

Department of Space, Earth and Environment
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2020

Master's thesis 2020

Making sense of low-cost sensors

Air quality monitoring in Gothenburg, Sweden

MEHDI GHALEBANI

Department of Space, Earth and Environment
Division of Microwave and Optical Remote Sensing
Chalmers University of Technology
Gothenburg, Sweden 2020

Making sense of low-cost sensors
Air quality monitoring in Gothenburg, Sweden
MEHDI GHALEBANI

Supervisor: Johan Mellqvist, Department of Space, Earth and Environment
Supervisor: Rikard Mattsson, Ericsson AB
Examiner: Johan Mellqvist, Department of Space, Earth and Environment

Master's Thesis 2020
Department of Space, Earth and Environment
Division of Microwave and Optical Remote Sensing
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

Making sense of low-cost sensors
Air quality monitoring in Gothenburg, Sweden
MEHDI GHALEBANI
Department of Space, Earth and Environment
Division of Microwave and Optical Remote Sensing
Chalmers University of Technology

Abstract

Over the recent decade, Emergence and commercialization of the new, low-cost, sensor technologies have created the possibility of major paradigm shifts in air quality monitoring. Their price of three orders of magnitude lower than standard/reference instruments provides the opportunity for new applications such as higher geographical and temporal resolutions of the measurements. There have been studies on the performance of a network of these sensors, however, their individual reliability is still questionable. This study aimed to evaluate the performance of one of the most common low-end sensors available on the market, SDS011, as well as a middle-end sensor, SDS019 under different circumstances such as temperature and humidity. The main research questions were: how reliable are these sensors and what are the causes of errors for these sensors and is it possible to find correction factors based on meteorological data? To address the research questions, a range of experiments in different environments, including field and laboratory, have been conducted under several humidity and temperatures. The results of the experiments illustrated a high linear correlation between the SDS011 and SDS019 sensors with the reference sensor(Optical Particle Sizer) at laboratory experiments. The data were fitted to the reference sensor using a linear regression model. additionally, a multiple linear regression was applied to include the temperature and relative humidity as additional input parameters to the regression model. The results of the multiple and normal regression were compared and discussed under different circumstances for both SDS011 and SDS019 sensors. The field experiments showed significant differences between the SDS011 and reference instruments and these could not be explained by humidity alone. They were not significantly reduced when applying laboratory correction factors either. A three week comparison of the SDS011 against the golden standard for PM in air quality monitoring (TEOM) showed periods with both decent and poor agreement, illustrating that the SDS011 sensors respond to PM but that they are rather unreliable when used as single devices. Further research work is needed to understand this. Nevertheless the sensors are suitable for operation in a network to obtain spatial air quality information , both as stationary and mobile.

Keywords: air quality monitoring, low-cost sensors , multiple linear regression, signal processing, aerosols , particulate matter , PM2.5, particles , dust, human health.

Acknowledgements

Along the way I have met and collaborated with so many kind soles who have contributed to my thesis in one way or another. I would like to thank Ericsson who gave me this opportunity to learn and grow in their company. special thanks to Lars-Erik Lindberg, Jonas Wilhelmsson, Magnus Persson and Rikard Mattsson for great support especially with the programming and IoT service and infrastructure. Additionally I would like to thank Lindholmen Science park for providing the majority of the finance as well as inclusion in the workshops and meetings for GOTCHA project, thanks to Gunnar Ohlin and Rodrigue Alfahel. It is worth mentioning the support from Miljöförvaltningen and providing the oppportunity to perform some experiments at Nordstan Air quality station. In my academic journey, I would of course like to direct gratitude towards my supervisor Johan Mellqvist who has always supported me in the obstacles with visionary advice. Also would like to thank Vladimir Conde and Department of Space, Earth and Environment, Division of Microwave and Optical Remote Sensing for providing the support and equipment for the laboratory experiments and part of the equipment for the city installations. And last but not least, friends and family without whom I would never be able to accomplish my goals.

Mehdi Ghalebani, Gothenburg, April 2020

Contents

List of Figures **xi**

List of Tables **xv**

1 Introduction **1**
1.1 Particulate matter, definition, sources, risks 1
1.2 The Changing Paradigm of Air Pollution Monitoring 4
1.3 Goal and Scope of the study 4
1.4 Low-cost Sensors and Internet of Things 6

2 Theory **9**
2.1 Physical Principles of Particle Sensors 9
2.2 Optical Particle Sensors 11
2.3 The Effect of Humidity on Optical Particle Sensors 13

3 Methods **15**
3.1 Laboratory Experiments 15
3.1.1 Materials and Requirements 17
3.1.2 Operational Procedures 17
3.2 Stationery measurement in the city with a Van 19
3.2.1 Materials and Requirements 19
3.2.2 Operational Procedures 19
3.3 Femman air quality station 20
3.3.1 Materials and Requirements 20
3.3.2 Operational Procedures 20
3.4 Data processing and Analyses 21
3.4.1 Interpolations 22
3.4.2 Moving average 23
3.4.3 Correlation Analysis and Regression 26

4 Results and Discussions **27**
4.1 Size Channels 27
4.2 General Explanation of results 28
4.3 Data Set A (Laboratory experiment) 29
4.4 Data Set B (Laboratory experiment) 35
4.5 Data Set C (Laboratory experiment) 41
4.6 Data Set D (Field experiment with van) 47

4.7 Data set E (Femman air quality station) 55
4.8 Analysis of the results and Discussions 58

5 Conclusion **61**
5.1 Conclusions 61
5.2 Recommendations for future Studies 61

References **63**

A Maximum values for size channel of OPS, Field vs. laboratory data **I**

List of Figures

1.1 Burden of disease attributable to 20 leading risk factors in 2010, expressed as a percentage of global disability-adjusted life-years [3] . . . 1
1.2 Size comparison of PM2.5 and PM10 against the average diameter of a human hair (∼70 μm) and fine beach sand (∼90 μm) [12] 2
1.3 Prototypical size distribution of particles, their sources and pathways of formations; dashed line corresponds approximately to 2.5 μm of diameter [16] 3
1.4 Scope of this study, the figure is adapted from Morawska, Lidia, et al who illustrated a general framework for air pollution monitoring [21] . 5
1.5 Low-cost Sensors and Internet of Things 8

2.1 Methods and instruments for PM measurement[32]. 10
2.2 components of an optical particle sensor [36] 11
2.3 SDS011 and SDS019 tow examples of low-cost optical sensors, OPS an example of a reference optical sensor 13

3.1 A schematic diagram of the instruments and setup for the laboratory experiments 16
3.2 The instruments and setup for the laboratory experiments 16
3.3 Stationery Measurement in the City with a Van. 19
3.4 The left picture: Long term data gathering by SDS011 sensors at Femman wather station. The right picture: comparison of OPS and TEOM. 20
3.5 Data processing steps when running different sensors simultaneously . 21
3.6 An example of MATLAB Interpolation function (Interp1). 22
3.7 An example of MATLAB Interpolation function (Interp1). 23
3.8 Comparison of the response time for different sensors before filtering, Red: OPS, Blue: SDS019, Black: SDS011 24
3.9 Comparison of the response time for different sensors after filtering, Red: OPS, Blue: SDS019, Black: SDS011 24
3.10 rmax for differnet time domains 25
3.11 rmax for differnet time domains, a closer look to the peak of figure 3.13: the red curve is where the trend changes (in this case 94 seconds) 25

4.1 Mean values for size channel of OPS, Field vs. laboratory data 27
4.2 Temperature, Data set A 29
4.3 Relative Humidity, Data set A 30

4.4 PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set A 30
4.5 PM2.5 mass concentrations for SDS011 and OPS excluding the first 3 channels, Data set A 31
4.6 SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set A 31
4.7 Relative errors for SDS019 and SDS011 vs. OPS, Data set A 32
4.8 Relative errors for SDS011 vs. OPS before and after applying the correction factors based on Linear regression, Data set A 33
4.9 Linear Regression plots, Data set A 34
4.10 Temperature, Data set B 35
4.11 Relative Humidity, Data set B 36
4.12 PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set B 36
4.13 SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set B 37
4.14 Relative errors for SDS019 and SDS011 vs. OPS, Data set B 38
4.15 Relative errors for SDS011 vs. OPS before and after applying the correction factors based on Linear regression, Data set B 39
4.16 Linear Regression plots, Data set B 40
4.17 Temperature, Data set C 41
4.18 Relative Humidity, Data set C 42
4.19 PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set C 42
4.20 SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set C 43
4.21 Relative errors for SDS019 and SDS011 vs. OPS, Data set C 44
4.22 Relative errors for SDS011 vs. OPS before and after applying the correction factors based on Linear regression, Data set C 45
4.23 Linear Regression plots, Data set C 46
4.24 Temperature, Data set D 47
4.25 Relative Humidity, Data set D 48
4.26 PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set D 48
4.27 PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set D, between the first and second spike 49
4.28 SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set D, between the first and second spike 50
4.29 Linear Regression plots, Data set D, between the first and second spike 51
4.30 PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set D, between the second and third spike 52
4.31 SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set D, between the second and third spike 53
4.32 Linear Regression plots, Data set D, between the second and third spike 54
4.33 SDS011 measurements for 3 weeks in parallel with TEOM at Femman air quality station 55
4.34 The absolute error and RH for SDS011 measurements for 3 weeks in parallel with TEOM at Femman air quality station 56

4.35 OPS measurements for 2 days (48 hours) in parallel with TEOM and SDS011 at Femman air quality station 57

A.1 Maximum values for size channel of OPS, Field vs. laboratory data . II

List of Tables

2.1 OPS size channels based on OPS product specifications (tsi.com) . . 12
2.2 product specifications of SDS011 (inovafitness.com) 12
2.3 product specifications of SDS019 (inovafitness.com) 13

4.1 Data Set A, 22 degrees (room Temperature) 29
4.2 Data Set A, 22 degrees (room Temperature) 34
4.3 Data Set B, 25 degrees 35
4.4 Data Set B, 25 degrees 40
4.5 Data Set C, 22 degrees (room Temperature) 41
4.6 Data Set C, 22 degrees (room Temperature) 46
4.7 Data Set D, ambient air with an average Temperature of 7 degrees . . 47
4.8 Data Set D, ambient air with an average Temperature of 7 degrees, between the second and third spike 51
4.9 Data Set D, ambient air with an average Temperature of 7 degrees, between the first and second spike 54

1

Introduction

1.1 Particulate matter, definition, sources, risks

Air pollution is a process that exposes humans, other living organisms, and the natural environment to possibly harmful substances in the atmosphere[1]. Some air pollutants can cause severe health effects even at relatively low concentrations due to their high exposure risk both in indoor and outdoor environments[2]. Among the 20 leading risk factors contributing to the burden of disease in, expressed as a percentage of global disability-adjusted life-years (DALYs), household air pollution from solid fuels and ambient particulate matter pollution were ranked 3rd and 9th with $3 \cdot 5$ million ($2 \cdot 6$ million to 4.4 million) premature deaths and 4.3% ($3 \cdot 4$–$5 \cdot 3$) of global DALYs in 2010 and $3 \cdot 1$ million ($2 \cdot 7$ million to $3 \cdot 5$ million) deaths and $3 \cdot 1$% ($2 \cdot 7$–$3 \cdot 4$) of global DALYs respectively[3], see Figure 1.1. The overall contribution of these two air pollution risk factors in 2010 is 6.6 million deaths. Airborne Particulate matter (PM) is a heterogeneous mixture of solid and liquid particles with small mass which allows them to suspend in the air. The suspension gives them mobility and the ability to stay in the air long enough to increase the exposure possibility. The particles possess varying chemical composition and size in space and time[4].

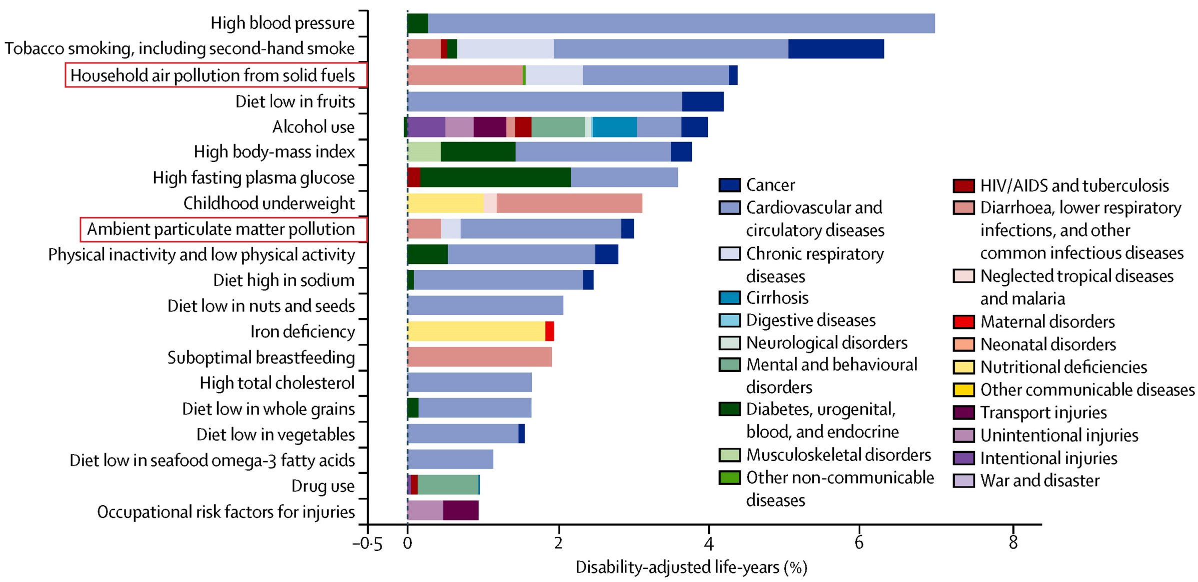


Figure 1.1: Burden of disease attributable to 20 leading risk factors in 2010, expressed as a percentage of global disability-adjusted life-years [3]

Several studies are providing scientific evidence for the health disrupting impacts that PM can cause, e.g. nonfatal heart attacks, irregular heartbeat [5], cardiovascular diseases [6], respiratory diseases, asthma[7], lung cancer[8], and premature death[3]. In 2013, the International Agency for Research on Cancer (IARC) classified Particulate Matter (PM) from outdoor air pollution as carcinogenic to humans[9]. The finest particles are dangerous since they penetrate deep into humans' lungs and even blood streams. The mobility and inhalability of the PM are the determining characteristics of their risk, and these two parameters are very much dependent on the aerodynamic diameter of a particle which therefore is usually used to categorize PM pollution[10]. USEPA categorizes the PM into coarse (PM10) and fine particles (PM2.5) [11]. Figure 1.2[12] illustrates the size of coarse and fine particles compared to human hair and fine beach sand. PM2.5 could be composed of chemicals such as sulfate, nitrate, ammonium, hydrogen ion, elemental carbon; organic compounds, PAH, metals, Pb, Cd, V, Ni, Cu, Zn, particle-bound water and biogenic organics. Whereas PM10 is usually composed of resuspended dust, soil dust, street dust, coal and oil fly ash, metal oxides of Si, Al, Mg, Ti, Fe, Calcium carbonate, Sodium chloride, sea salt, pollen, mold spores, and plant part[13]. since the fine particles are smaller and generally lighter than coarse particles, they can suspend more in the air and have a lifetime of days to weeks, while that of coarse particles is in the order of minutes to hours[13]. Due to the same reason, fine particles have more mobility and they can travel as far as 100 to 1000 kilometers while the longest travel distance for coarse particles is 10 kilometers[14]. The main sources of fine particles are combustion of coal, oil, gasoline, NOx, SO2, and organics including biogenic organics, e.g., (terpenes); high temperature processes; smelters, and steel mills. Whereas coarse particle usually are caused by resuspension of soil tracked onto roads and streets; Suspension from disturbed soils, e.g., farming, mining, resuspension of industrial dust, construction, coal and oil combustion and ocean spray[14].

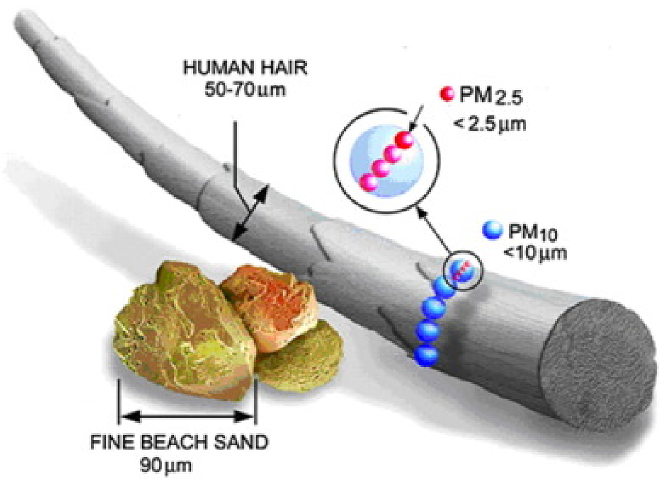


Figure 1.2: Size comparison of PM2.5 and PM10 against the average diameter of a human hair (~70 μm) and fine beach sand (~90 μm) [12]

For each size range of the particles, there are specific processes that lead to formation of particles. mechanical processes such as the break-up of larger solid particles are mostly responsible for the formation of coarse particles whereas the processes leading to the formation of fine particles are more diverse and include chemical processes as well as mechanical processes such as coagulation and aggregation of smaller particles with aerodynamic diameters of less than 0.1 μm (known as ultrafine particles) and subsequent formation of the fine particles [15]. The ultrafine particles are either emitted due to combustion or formed by nucleation [16]. Nucleation is a process in which a nucleus provides a surface on which low-vapor-pressure substances, formed by chemical reactions in the atmosphere or high temperature vaporization, can condense. The nucleation and condensation growth of primary particles (ultrafine) leads to particles in the accumulation mode with aerodynamic diameters between 0.1 and 1.0 μm, with a typical size of 0.3 μm. Condensation growth happens on one ultrafine particle and leads to the formation of one fine particle whereas coagulation is a process that combines a number of ultrafine particles and leads to the formation of fine particles. Condensation growth is therefore most efficient when the surface area is high while coagulation is most efficient at high particle concentrations. Figure 1.3 illustrates the size distribution and formation modes of particles[17].

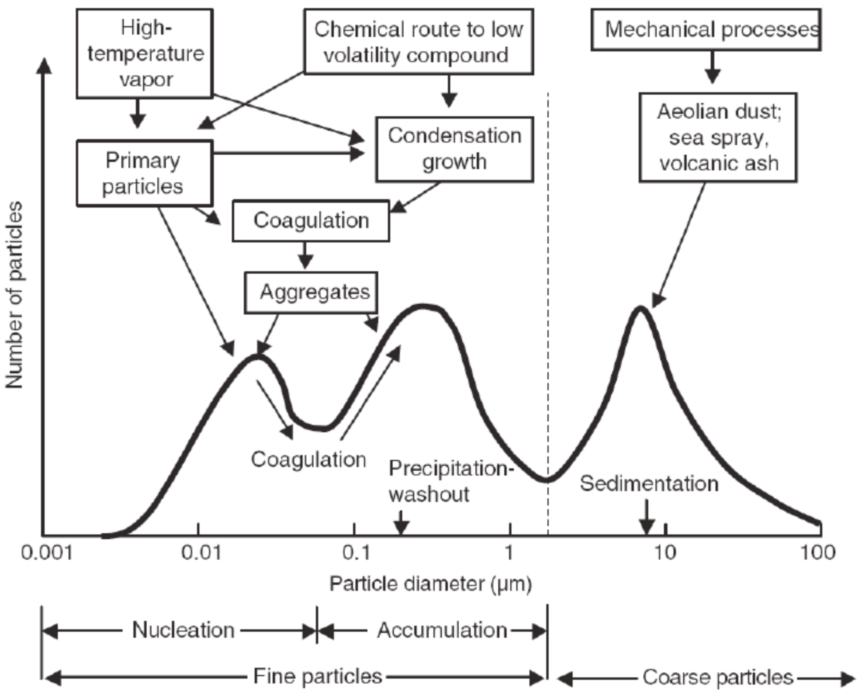


Figure 1.3: Prototypical size distribution of particles, their sources and pathways of formations; dashed line corresponds approximately to 2.5 μm of diameter [16]

The described features of particles and their posed health risks raises the concerns about their pollution levels. To ensure that the risks attributed to these factors is reduced, the World Health Organization (WHO) has provided air quality guideline values for different substances including PM. Guideline values for PM2.5 in terms of annual mean and 24-hour mean are 10 and 25 μg/m3[18].

1.2 The Changing Paradigm of Air Pollution Monitoring

The traditional air pollution monitoring approach mostly use expensive, complex, stationary equipment[19] which restricts the measurements to only few instruments and locations . In addition, in the developing world they are generally too expensive. These paradigms are changing fast due to recent technological advances in portable low-cost air pollution sensors which can report data in near-real time at a high-temporal and spatial resolution, using wireless communication/infrastructure and enhanced visualization and computational capabilities[20].

Implementation of networks of mobile and stationary low-cost sensors increases spatial and temporal resolution of monitoring coverage which enables immediate access of information to the general public as well as possibility of supplementary data air quality modelling to assess personal exposure of the pollutants and health effects. However, all the mentioned benefits and applications are valid given the quality of the data measured by the low-cost sensors meets the requirements of monitoring purpose in terms of precision, accuracy, sensitivity, etc. [20].

1.3 Goal and Scope of the study

In the context of low-cost sensors data quality has a vital role since data of poor or unknown quality is worse decision support than no data. There are many low-cost sensors already available in the market and implemented whose performance has not been evaluated under ambient conditions, taking into account environmental factors such as different levels of concentration and humidity and temperatures. Therefore, this study aims to evaluate the performance of two of the common and wide-spread commercial low cost PM sensors under different conditions including simulated pollution in the laboratory as well as ambient air measurment in the city of Gothenburg.

The main research questions posed are: how reliable are these sensors and what are the causes of errors for these sensors and is it possible to find correction factors based on meteorological data? How is the relationship between the performance of the SDS011 and SDS019 low-cost sensors and parameters such as humidity and temperature? How does the performance of these two sensors compare to each other and to a reference sensor under different environmental circumstances?

To address the research questions, a range of different activities and experiments was performed and the scope of these activities during the course of this project are illustrated in Figure 1.4. The activities include selection and assessment of the sensors in the laboratory and field in different ambient circumstances. In addition, they include preparation of hardware and software for deployment of the sensors for long term measurements in the field.

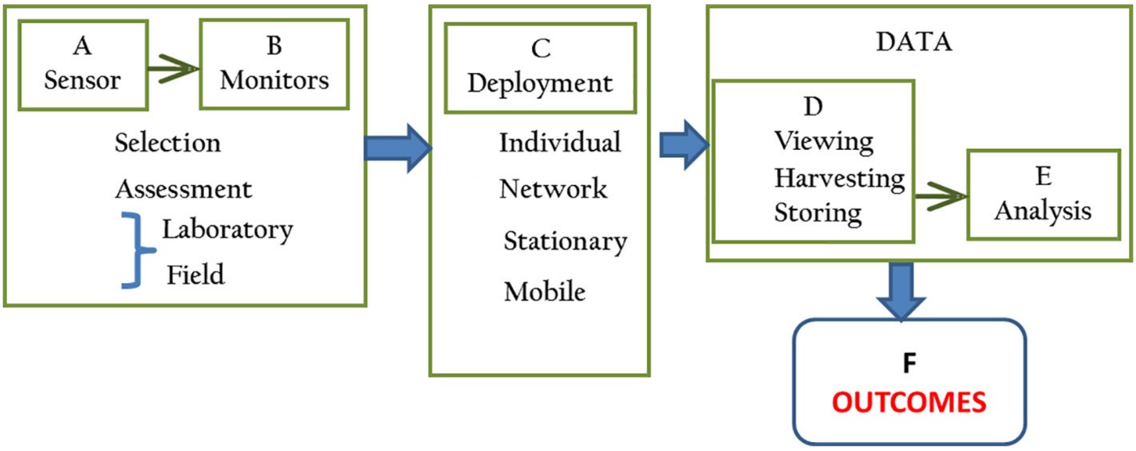


Figure 1.4: Scope of this study, the figure is adapted from Morawska, Lidia, et al who illustrated a general framework for air pollution monitoring [21]

.

1.4 Low-cost Sensors and Internet of Things

The low-cost sensors are usually part of Internet-of-Things (IoT). IoT corresponds to a network of devices able to exchange data. The components of the network can consist of different devices including sensors and embedded electronics and network connectivity, which enables the components to connect and collaborate with each other[22]. Although IoT has been around since the 1980s[23], it has first gained attention in recent times thanks to the 5G network. The 5G network technology provides the infrastructure for the implementation of IoT in a very large geographical scale in urban areas with an unprecedented bandwidth and rate of data communication between the objects, enabling the cities to possess a highly collaborative environment whose components have interactions based on real-time data[24].

The IoT has already been used in different applications. Automobiles, city infrastructures such as smart lighting, power, cooling, water and alarm systems are some examples of current IoT applications in the cities[25–28]. However, 5G provides an infrastructure which brings the application of IoT to the next level[29]. The combination of IoT and 5G is a tool for smart and sustainable city planning and maintenance. With such a combination, the interconnection between the objects in the cities provides the cities more efficiency in terms of energy use and resource efficiency in general[28]. This combination of technologies can contribute to sustainable mobility systems, enabling the city planners to do active interventions on both private and public transportation vehicles based on real-time data from all components of the city, including the vehicles[30]. Interventions such as variable speed limits in certain streets according to the traffic flow, hindering the vehicles to exceed certain speeds by a central command. Another example of such interventions could be for hybrid vehicles and making them switch to electricity in areas with high pollution levels. Yet another example, a dynamic and real-time guideline for self-driving vehicles in the future, to optimize the traffic flow and prevent accidents. One key player for such purposes is to have a larger coverage and a higher resolution of air quality measurements in urban areas.

Air pollution is a significant problem in many major cities in the world. It has impact on both human health and the global environment. There are a lot of regulations for the limits on the concentration of the different air pollutants in the air and more restrict ones may be passed by the EU in the years to come. For the societies and authorities to be able to improve the air quality, they need to have a realistic sense of the air quality and the spatial distribution, and reliable measurements should take place and be available and reported regularly. Such data can provide decision support for authorities and information whether they comply with newly passes regulations and EU directives with respect to air quality. However, today in the major cities usually measure the air quality in only a few locations in the city center and consequently, their reported data does not necessarily represent the air quality of the whole city. This provides a scientific incentive to investigate the expansion of the coverage of the measurements in the big cities.

In addition, in residential areas few air quality measurements are carried out, and, such data is important also to raise the awareness of people and get consensus about new environmental legislation or directives. This provides the social incentive for exploring the expansion of the coverage of the measurement stations. In the case of Gothenburg, currently there exist five reliable air quality stations, run by municipality. The air quality data reported from these stations represent the real air quality only in the vicinity of 1 kilometer or less around the location. It goes without saying that every large-scale project is coupled with high investment costs. There has been a plethora of studies and projects in this field, trying to provide low-cost solutions for a better coverage of air quality measurements in different cities around the world. Some examples of these projects are the following: Opensense project, which is conducted in Switzerland doing real-time air pollution monitoring with sensors on city buses[31]. GreenIoT project in Sweden, which utilizes IoT to measure air pollution level in the city center of Uppsala. LoV-IoT (*Luft- och vattenövervakning med internet of things*) -air and water monitoring with internet of things, is also another project in Sweden which has tested a range of different air quality sensors *(loviot.se)*. The common idea behind these and other projects is to minimize the costs by implementing a large number of very cheap sensors and accounting for the overall data from the array of the sensors.

This master thesis project was conducted at Gothenburg in Sweden and the raw sensors and data acquisition system were funded as part of a project called ElectriCity while the other hardware (such as protective boxes, power) and installation costs at bus-stops were funded by internal funding at Chalmers. Fifteen partners from industry, academy, and society are working together to develop, test and demonstrate new solutions for sustainable mobility *(electricitygoteborg.se)*. The partners are listed as The *Volvo Group, Västra Götalandsregionen, Västtrafik, The city of Gothenburg, Chalmers University of Technology, Swedish Energy Agency, Johanneberg Science Park, Lindholmen Science Park, Göteborg Energi, Keolis, Älvstranden Utveckling, Akademiska Hus, Chalmersfastigheter, and Ericsson.* This wide range of different partners from the producers of electrified vehicles to the municipality and public transport authorities provides the possibility of designing and testing prototypes.

This work was the prototype of application of IoT for air quality monitoring in the city of Gothenburg. The work includes a literature review on the applications of IoT for sustainability and low-cost aerosol sensors, laboratory experiments on a small number of low-cost aerosol sensors, field experiments on the performance of a combination of different low- and medium-cost aerosol sensors as both stationary and mobile measurement devices using IoT. In this project I have built together 20 small sensors and one mobile one on a bus, with help from Ericsson and Chalmers. The sensors were installed in the end of this masters project and they have started to produce data. Their data processing is not part of this work. The idea was to calibrate the smaller sensors with the medium quality one which was mobile. This work is an attempt to build stepping stones of a high coverage system of stationary and mobile sensors in the city of Gothenburg by deploying low-cost sensors in the bus stations and calibrate them against a medium costly sensors on top of an electric bus from public transport, Figure 1.5. The work is also a case study of application of IoT and performance of sensors in the city of Gothenburg and provides a platform for a higher connectivity as soon as 5G is commercialized.

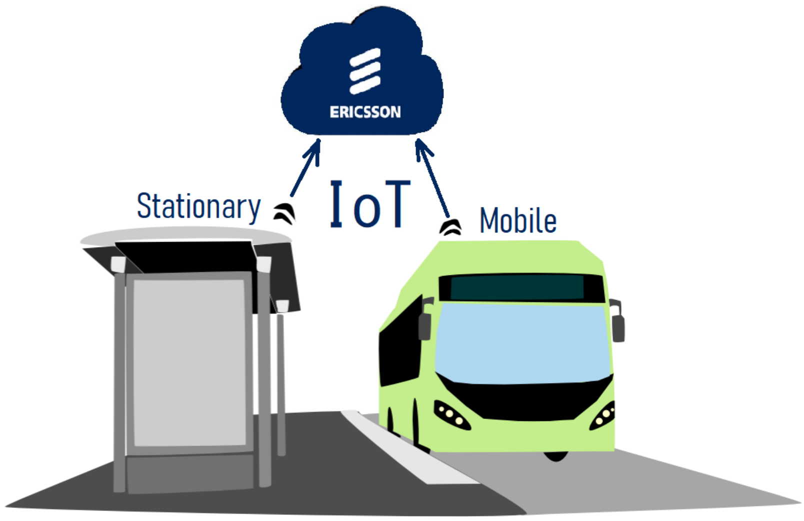


Figure 1.5: Low-cost Sensors and Internet of Things

2

Theory

2.1 Physical Principles of Particle Sensors

For most of the air pollutants, knowing the concentration of the substance is usually enough for having an understanding of the air quality. However, when it comes to particulate matters, an additional parameter plays a key role: Particle size distribution. Smaller particles are able to remain in the atmosphere for a longer period and penetrates deeper into to humans. Therefore, the combination of the particle size distribution and particle concentration provides more complete information on the air quality[32]. There are several physical principles for measuring the particle concnetration (by number or mass) and size distribution[32]: This includes for instance utilizing the principles of gravimetry, optical light scattering, light absorption, electrical and aerodynamic mobility . Many of the instruments count particles (number concentration) in different size bins and then the particle mass is obtained by assuming the density and shape of the particle[33]. The gravimetric instruments, on the other hand, measures the mass of PM10 or PM2.5. particles collected on a filter by microbalance technique and they are the most accurate. However they are slow , bulky and do not provide the size distribution. The most common particle measuring instruments for measuring concentration and size distribution are summarized in Figure 2.1.

Amaral et al. performed experiments on different instruments working with different principles and they were evaluated based on the following characteristics[32]: the ability to sample particles in real-time; need to dilute gas flow before collection; detection limit of the equipment; size range; the accuracy of the equipment. And it was found out that instruments measuring based on the light scattering principle (optical devices) have the highest accuracy of all except for the microbalance instruments, which has better accuracy and are usually used as reference measurement instruments. However, there are two drawbacks to the microbalance instruments; First is their high price, which makes them not an affordable choice for using in a large number of sensors for reasonable coverage of the city area. And the second drawback is their size and mass which makes the mobility of the sensors more difficult, compared to the optical sensors.

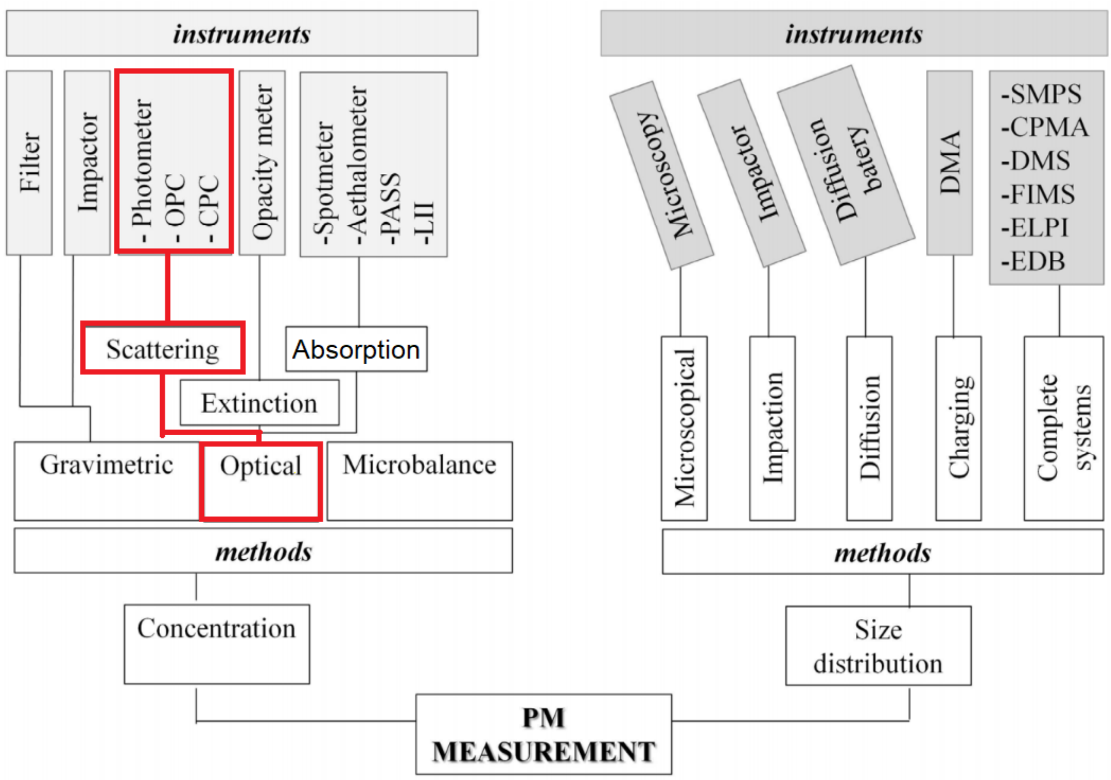


Figure 2.1: Methods and instruments for PM measurement[32].

This thesis mostly focuses on optical particle sensors and compares two of the commercially available low-cost optical sensors (SDS011 and SDS019) to a reference optical particle sizer (TSI 3300 OPS). At some point, performance of the mentioned optical sensors was also compared with another reference sensor Tapered Element Oscillation Microbalance (TEOM) which works on another physical principle (microbalance). The TEOM instrument is used to detect aerosol particles in real-time by measuring PM mass concentration due to the accumulation of particles in a sampling filter and based on the alteration of the resonance frequency of a tapered quartz wand [32].

2.2 Optical Particle Sensors

Optical sensors consist of a light source which lights up the particles passing the detection chamber. part of the light is absorbed by the particles and turned to other types of energy such as heat[34]. The other part of beam is irradiated to different directions (scattered). and by summation of scattering and absorption, extinction of the light can be calculated[35]. These three features can be used as a working principle for optical instruments to count the particles and determine their mass concentration. All of the mentioned instruments for this study (OPS, SDS011, SDS019) determine the concentrations based on scattering of the light beams in the orthogonal direction. The similarities and differences between them are elaborated in the next section. Figure 2.2 illustrates different components inside the detection chamber of TSI 3300 OPS.

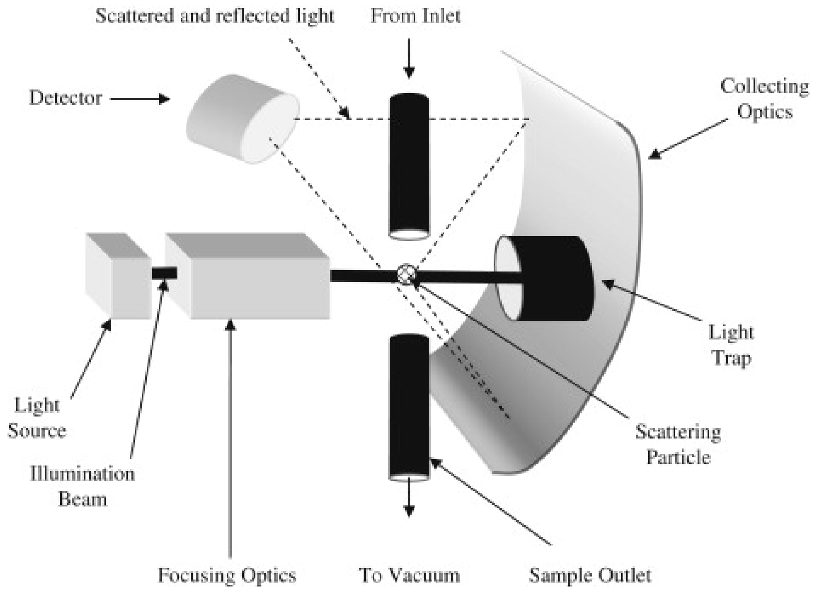


Figure 2.2: components of an optical particle sensor [36]

The OPS instrument has a built in pump which determines the flow rate of the particles into the detection unit and controls it in a way that single particles is counted and depending on the intensity of the scattered light for each one, the counted particles are binned into 16 different size channels ranging from 0.3 to 10 μm. Table 2.1 shows the cut points of the OPS size channels[34]. To the contrary, the low-end optical sensors are not able to measure the scattering from single particles and inside obtains the concentration based on the combined scattering of all particles passing through the sensor. Furthermore, the performance of these sensors is affected when the compositions of PM2.5 vary greatly. Thus, the majority of the low-cost instruments are used indoors or in situations where particle composition is constant[37]. The SDS011 has a small computer-like fan with a varying flow rate into its detector, which affects the accuracy of the measurement. The SDS019 is based on four SDS011 sensors connected to an air pump with controlled flow rate and it also corrects for pressure and relative humidity. It is however 10 times more expensive. TSI OPS is 100 times more costly than SDS019. The low-end sensors work in certain size and concentration ranges, see Table 2.2 and 2.3. SDS011 measures PM2.5 and PM10. SDS019 additionally measures PM100 *(inovafitness.com)*. Tables 2.2 and 2.3 provide a piece of information on product specifications of SDS011 and SDS019. And Figure 2.3 shows SDS011, SDS019 and OPS.

Table 2.1: OPS size channels based on OPS product specifications (tsi.com)

Size channel cut point number	size channel cut point (μ m)
Bin 1	0.3
Bin 2	0.374
Bin 3	0.465
Bin 4	0.579
Bin 5	0.721
Bin 6	0.897
Bin 7	1.117
Bin 8	1.391
Bin 9	1.732
Bin 10	2.156
Bin 11	2.685
Bin 12	3.343
Bin 13	4.162
Bin 14	5.182
Bin 15	6.451
Bin 16	8.031
Bin 17	10

Table 2.2: product specifications of SDS011 (inovafitness.com)

Property	Value
Measurement parameters	PM 2.5, PM10
Concentration Range	0.0-999.9 μg /m^3
Humidity Range	Max 70%
Corresponding Time	1s
Counting Efficiency	70% @ 0.3 μm and 98% @ 0.5 μm
Minimum resolution of particle	0.3 μm

Table 2.3: product specifications of SDS019 (inovafitness.com)

Property	Value
Measurement parameters	PM2.5, PM10, PM100
Concentration Range	0.0-1999.9 μg /m^3
Humidity Range	Max 98%
Corresponding Time	1s
Counting Efficiency	70% @ 0.3 μm and 98% @ 0.5 μm
Minimum resolution of particle	0.3 μm

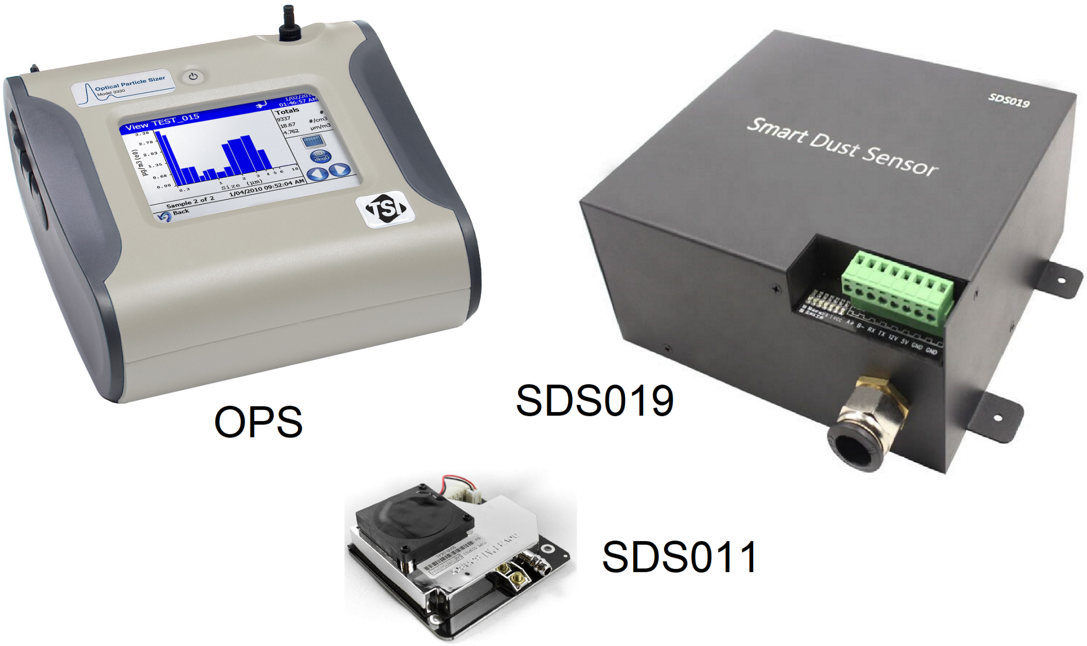


Figure 2.3: SDS011 and SDS019 tow examples of low-cost optical sensors, OPS an example of a reference optical sensor

2.3 The Effect of Humidity on Optical Particle Sensors

Research has revealed that light scattering is strongly affected by high Relative Humidity (RH)[38–43]. There are two main phenomenon causing the error caused by humidity in optical sensors: condensed fog droplets of similar size as PM-relevant particles[44] and hygroscopic growth of saline particles[45].

Deliquescence, the process by which a substance absorbs moisture from the atmosphere until it dissolves in the absorbed water and forms a solution. The process takes place as long as the vapour pressure of the formed solution does not exceed the partial pressure of water vapour in the air. Therefore, the more humidity in the air, the more partial pressure of water vapor and the more chance for deliquescence.

Ammonium sulfate and ammonium nitrate, which are major semi-volatile components of PM2.5, respectively have deliquescence points of 62% and 80%. Thus, at RH is above 62%, particle-bound water increases. When RH is above 65% mass of particles grows and this growth is intensified when RH exceeds 80%, which eventually leads to the overestimation of PM2.5[43].

In order to assess the influence of RH on measurement results, the hygroscopic growth factor has been included, which is the ratio of a particle's diameter under high RH conditions to that under dry conditions. At high RH, especially above 90%, a small alteration in RH can result in considerable changes in the hygroscopic growth factor. For instance, the hygroscopic growth factor of particles in the range of 200 nm to 1 μm is about 1.5 at 90% of RH and reaches nearly 2.0 when RH is 95%[46]. To cope with this effect, Petters, M. D., & Kreidenweis, S. M. [47] suggested this approach to describe the relationship between particle dry diameter and cloud condensation nuclei (CCN) activity using a single hygroscopicity parameter k, which is based on *Köhler* theory[48]. This can be done by computing a multicomponent hygroscopicity parameter by weighting component hygroscopicity parameters by their volume fractions in the mixture, given that the composition data and the hygroscopicity parameter of each component are available. Additionally, The hygroscopicity parameter can be used as an input to model the CCN activity of atmospheric particles, including those containing insoluble components[49].

However, calculation of the hygroscopicity parameter requires data on the composition of the mixtures of the particles and their their individual hygroscopicity parameters which can also vary city by city. Additionally there are several studies that focus on the sources with another approach to do indirect measurements and estimate the amount of particle pollution by emission factors. In this study the attempt was to have authentic ambient situation and deal with the data as it is without any physical addition to the system such as dryers or chemical absorbents. Also, investigating the possibility of a relation between the reported RH and temperature and the performance of low-cost sensors by comparing them to reference instruments. However, also for the reference instruments the inclusion of water vapour in the particle mass is partly present. This has not been accounted for in the study.

3
Methods

A combination of different approaches was implemented for this investigation. The methods can be divided into two main categories: theoretical and experimental. The theoretical part consists of reviewing the most recent literature on the topic and the experimental part includes the following activities: laboratory and field experiments, designing and prototyping of some of the instruments for the experiments, design of experiments and setups, computer programming for using instruments, and computer programming for analyzing the results. The materials and requirements and the methods are elaborated in the following sections of this chapter. The project started with a literature review to understand the research field and methods and to find relevant research questions. The literature includes both peer-reviewed papers and reports

The experimental part of the project can be divided into laboratory and field experiments. The laboratory experiment consists of a setup including different sensors and a test chamber which was aimed to simulate the real-world circumstances of the ambient air. The field tests were conducted in different locations and occasions with different time spans ranging from a couple of hours to months. The laboratory and field tests were designed for different purposes whose combination serves the overall aim of the study. A field experiments was carried out at Femman air quality station operated by the Environmental department of the municipality of the city of Gothenburg (*Miljöförvaltningen*). Another field expedition was carried out with a van in the city of Gothenburg.

3.1 Laboratory Experiments

A number of devices were used to simulate ambient air with particle pollution and variable meteorological circumstances to evaluate and compare the performance of the low-cost sensors compared to the reference sensor. The general setup of the laboratory experiment is illustrated in Figure 3.1. Here a certain temperature, relative humidity and number of ammonium sulphate particles were released into the test chamber. Several particle instruments were then used to measure the particle concentration inside the test chamber at the different environmental conditions. The general idea of the setup is sketched in a schematic diagram in Figure 3.1.

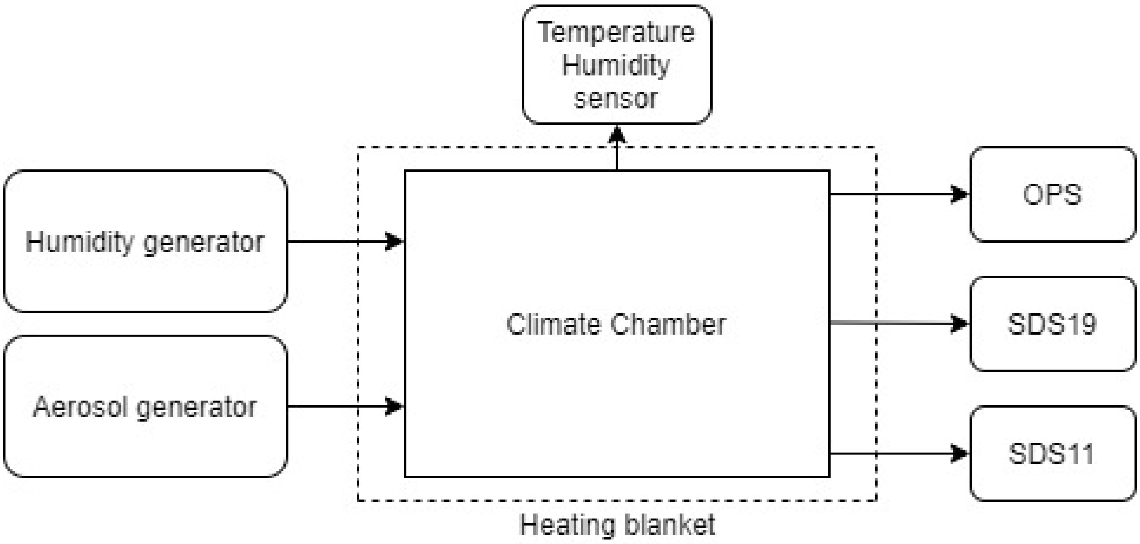


Figure 3.1: A schematic diagram of the instruments and setup for the laboratory experiments

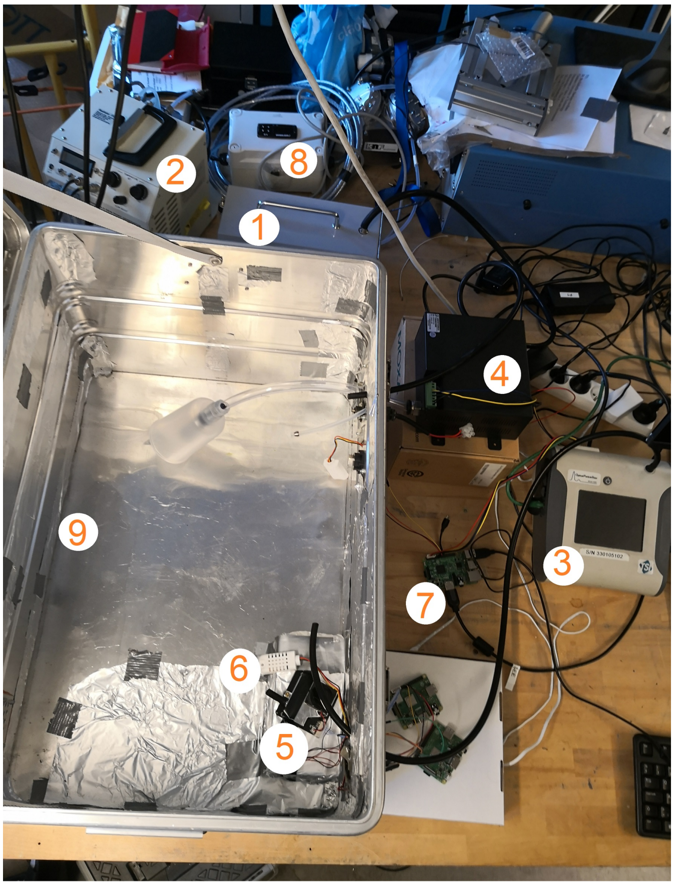


Figure 3.2: The instruments and setup for the laboratory experiments

3.1.1 Materials and Requirements

The instruments used for the experiments are illustrated in Figure 3.2:
1. PORTABLE ATOMIZER / AEROSOL GENERATOR / MODEL 3079A
2. LI-610 Portable Dew Point Generator
3. Optical Particle Sizer / Spectrometer / Model 3330
4. Nova fitness SDS019 sensor
5. Nova fitness SDS011 sensor
6. AM2302 digital temperature and humidity sensor, 5V, Adafruit (referred as T and RH)
7. Raspberry pi
8. Heating blanket
9. test Chamber (a metallic box of size of 70 x 48 x 40 cm) cube shaped with holes and hose

For each particle sensor a Raspberry Pi computer was used to log the particle concentration data, RH and temperature. A T and RH sensor can be read with the same Rasberry pi reading a particle sensor. The aerosol generator was used to generate ammonium sulphate aerosols inside the test chamber and the humidity generator was utilized to provide different levels of humidity inside the chamber. The heating blanket covered some of the sides of the chamber from outside to provide different operational temperatures inside the chamber.

3.1.2 Operational Procedures

The operational stages for the experiments are categorized into two main phases: Preparation phase and During the experiments.

Phase 1. Preparation:
Before each test day, The test chamber was cleansed and the sensors inside it were checked to see if the wire connections are tight and firm enough and the devices are placed in the correct positions. It took a couple of hours for the moisture to accumulate inside the chamber, therefore the first step was to run the humidity generator and place its outlet as an inlet to the chamber with a hose until the humidity inside the chamber reached 100 %. This step was, most of the time, done a day before the experiment to save time. The humidity was monitored by the T and RH sensor. There are some considerations regarding raising the humidity inside the chamber worth mentioning: It was made sure that the OPS and SDS019 were turned off while the humidity generator was operating. The reason is that both of the mentioned devices suck the air into their sampling unit and this would hinder the humidity inside the chamber to rise because of the continuous air suction for detection. While working with the humidity generator it also was made sure that the cooler and pump were switched on. On the test day, while waiting for the relative humidity to reach to 100 % the following activities were done:

First the time was calibrated and synchronized for each device. After that, the OPS was turned on. Then the logging programs of the different sensors were opened and ready to be run. To begin the experiment, the following actions were taken: Running the logger program for SDS011 and T and RH sensors on the raspberry pies, removing the tapes from the holes of the chamber and placing the hose of OPS inside the hole, turning the SDS019 on, starting to run the OPS and SDS019 logger programs. The last sensor to be turned on was SDS019 as its hose was inside the chamber and turning it on will immediately remove some humidity from the chamber.

Phase 2. During the experiments:
As mentioned before the SDS019 and OPS reduce the humidity when they measure the particles. Therefore, the test was designed to be conducted in a way that the relative humidity starts decreasing from around 100 % to the room humidity level, which usually was something between 30 to 50 %. Most of the experiments were conducted at room temperature and for some of them, the heating blanket was placed around the chamber to increase the temperature before and during the test. The humidity generator was kept running during the experiments so that the humidity reduces with a slower rate than as if it was turned off and there was not a source of moisture anymore during the experiment. The reason for the attempt to maintain the high humidity levels for a longer while was that the data measured in a relative humidity of 70 % and more are more valuable for the purpose of this research. Another action at this phase was particle injection to the chamber which would take place periodically to make some peaks in the particle concentration levels from the background level. The injections were conducted by placing a hose to the outlet of the aerosol generator and placing the other end of the hose to one of the inlet holes on the chamber for a second or a fraction of a second to produce some impulses. The time intervals between the impulses were determined based on the real-time data on the particle level in a way that each injection took place as soon as the particle levels reached back down to the background level.

The flow rate of the aerosol generator was 300 L/h(*www. tsi. com*), which was too high. Therefore, a T-shaped hose connector was utilized to branch out the majority of the flow rate of the aerosol generator and the other branch of the hose was used for the injections with a 90 % reduction in the flow rate. At the end of each test day, all of the sensors and instruments were turned off and the data was extracted and stored and documented. Sometimes due to the unrealistic values of particle levels, and the possibility of saturation of the sensors, a number of data sets were excluded from the study. The process, in general, was a semi-batch process with the test chamber as the volume control or the system under study and there was a constant input of moisture to the system as a periodic input of aerosols. There was not any intended outlet flow out of the chamber, however, the air was able to escape out through uncovered wires. Another mechanism that took place in such a setup is the decay of the concentration levels of the particles due to deposition[50].

3.2 Stationery measurement in the city with a Van

Figure 3.3: Stationery Measurement in the City with a Van.

3.2.1 Materials and Requirements

The applied sensors for conducting this test were the same as the laboratory experiments explained in the previous section.

3.2.2 Operational Procedures

The location was a parking house in the city of Gothenburg in Sweden which was called Olskroken P-hus at the address Lilla Olskroksgatan 1. All the required instruments were taken to the measurement location with a van. The van was parked inside the parking house in a way that its back door was opened and facing the ambient air by a close-by highway intersection, see Figure 3.3. The meaurements were done at the ground floor and the parking house with plastic shield making a semi-covered wall and making ít possible for the wind to flow through quite easily. On a couple of occasions cigarettes were lit up to inject a spike of particles into the sensor for test purposes. The weather during this experiment was rather windy (10 m/s) and part of the time it was raining. The wind blew from the nearby highway intersection of the E6 and E20 approximately 200 m away. The operational procedures were very similar to that of the laboratory experiments explained in the previous section with the difference that particle injection was not conducted and instead a couple of cigarettes were lit up to cause some peaks and deviations from the background levels.

3.3 Femman air quality station

Femman air quality station is operated by the Environmental department of the municipality of the city of Gothenburg (Miljöförvaltningen) and it is located at the rooftop of the Femman shopping mall, approximately 7 floors above the ground level. The station is equipped with high-end instruments to measure a wide range of substances as well as temperature and humidity and most importantly for this research, aerosols. The latter are measured by a TEOM reference instrument. The following data were gathered with the purpose given below: SDS011 sensors were operated during two months (only 3 weeks of data is available due to prototypings in the IoT code) in parallel with the TEOM, see Figure 3.4 (left side). Additionally an OPS was operated during two days in parallel with the TEOM, see Figure 3.4 (right side).

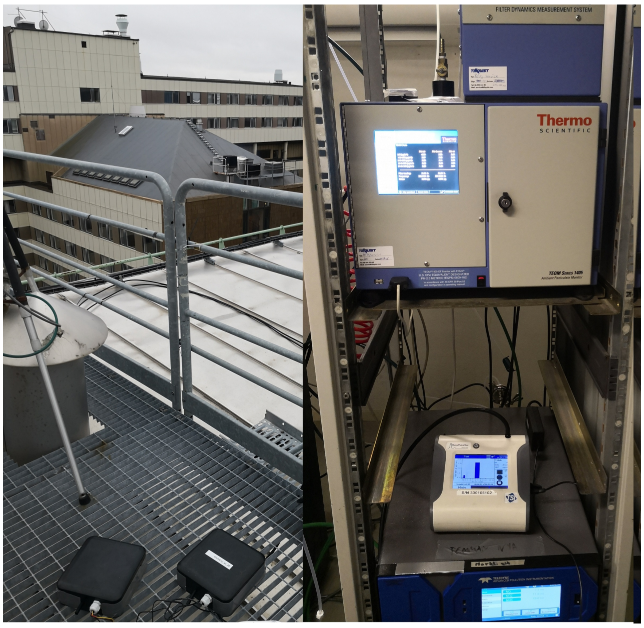


Figure 3.4: The left picture: Long term data gathering by SDS011 sensors at Femman wather station. The right picture: comparison of OPS and TEOM.

3.3.1 Materials and Requirements

1. OPS sensor with a computer for logging the data
2. SDS011 sensors, raspberry pies for logging the data, T and RH sensors, waterproof boxes, jumper wires for connections

3.3.2 Operational Procedures

A script was developed in python which starts the connection with WiFi, logs the data and saves locally to the SD-card of the raspberry pies and also sends the data using Message Queuing Telemetry Transport (MQTT) messages to the IoT broker in real-time.

3.4 Data processing and Analyses

When measuring with different sensors simultaneously, there are a number of issues that need to be addressed and for that, a number of concepts to be defined. The concepts to be defined are response time and sampling rate. The response time is the time it takes from the moment the substance reaches the detection unit to the sensor until it leaves the unit (residence time of substances inside the detection chamber). an indicator for response time of sensors is the time it takes from 10 % to 90 % of a logged signal. Another concpet to be defined is sampling rate which is the frequency in which the sensor logs a new measurement. and the combination of the two determines the time resolution of the data. Another concept is time lag which can be due to the time it takes for the substances to go through the hose and reach the sensor with a linear time shift which can easily be corrected or due to the effect of response time. The latter is not necessarily a linear shift in the recorded data. Its effect is rather a different resolution of data per the same time interval for 2 different sensors.

Figure 3.5 shows the data processing steps taken for the measurements. After logging and data extraction for each sensor, the timestamps of different sensors were transformed to the same type and the overlapping time span of all of them was identified. The next step was to fix the inconsistent sampling rates for each sensor as well as synchronizing the time resolution and sampling rates for different sensors by interpolations. To solve the problem of the different response times for different sensors, a moving average was applied to the signals. Finally, the time lags caused by different reasons were corrected and the correlation analyses were performed. More details about the mentioned data treatment steps are provided in the following of this chapter.

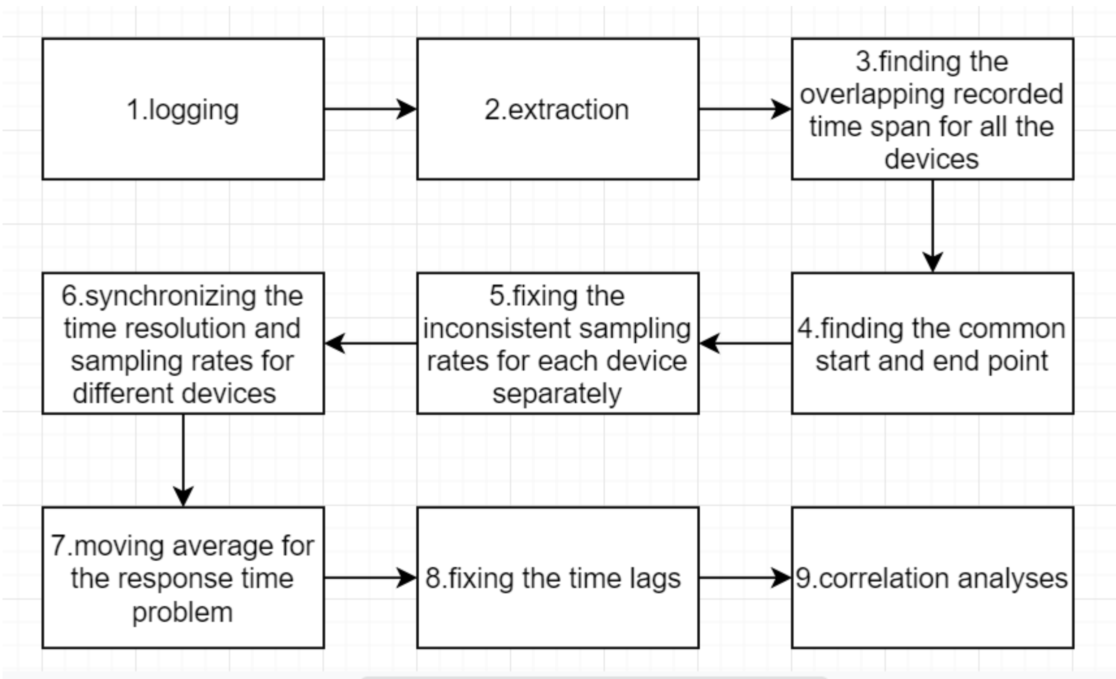


Figure 3.5: Data processing steps when running different sensors simultaneously

3.4.1 Interpolations

Different sensors had different start times and end times per test day. For each test day, the longest time interval where all the sensors had overlap was extracted and the rest was excluded. The OPS and SDS019 were programmed to log every second while the T and RH sesnors and SDS011 logged with 2 and 5 second time resolution, respectively. To solve problems of sensors time resolutions, a builtin function from MATLAB, called Interp1 (Copyright 1984-2018 The MathWorks, Inc.), was used. The functions provide different methods for interpolation. The implemented method for this study was called "nearest" which connects fills in between the existing data by connecting a curve to the nearest existing neighboring data point, see Figure 3.6.

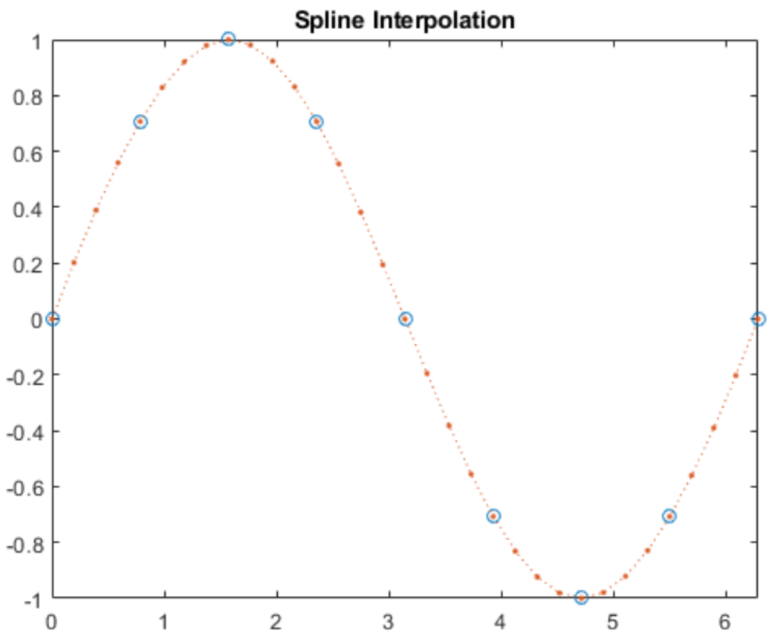


Figure 3.6: An example of MATLAB Interpolation function (Interp1).

The Interp1 function was applied to all of the data sets from all the sensors on all of the test days. As a sample of the implementation of the function the compression between the result of implementing this function and the actual data is illustrated in the Figure 3.7. As seen in the Figure 3.7, the interpolation (red) is matching well with the actual data (blue). The interpolation function requires the starting point and endpoint of the data set when all the sensors overlap in the measurement records. And the rest of the data was excluded. Finding the start and endpoint of the data and synchronizing the different sensors to the same time span automatically solved the time lag between the sensors.

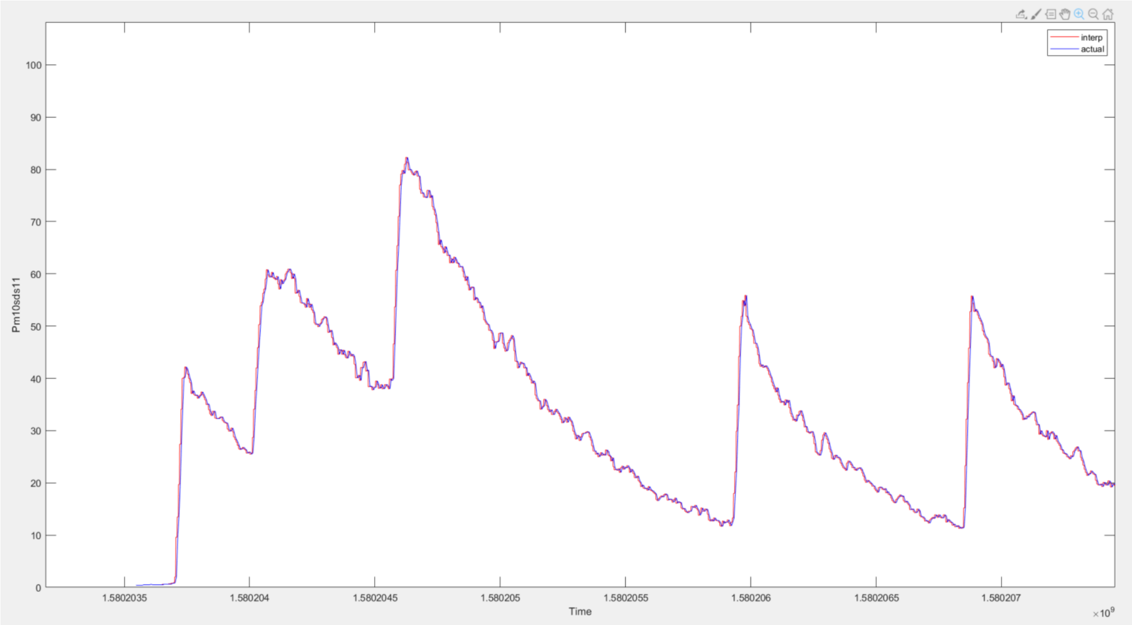


Figure 3.7: An example of MATLAB Interpolation function (Interp1).

3.4.2 Moving average

The tests in this study showed that SDS019 and SDS011 have similar response times in contrast to the OPS which was much faster, see Figure 3.8. The data for OPS(faster response time) and SDS011 and SDS019 are shown with red, black and blue colors respectively. To solve the response time problem, the data from the faster sensor, with short response time, had to be slowed down and time synchronized with the slower sensors, with long response time. This was done by applying a moving average on the fastest sensor in a way that the new data for each point is the weighted average of the last N seconds (time domain) before that point. The weighted average can be done linearly or with many different curvature assumptions. However, a Gaussian distribution seems to be more realistic according to the mechanism by which these sensors detect the aerosols. The average also needs to be normalized according to N seconds, so it does not change the magnitude of the data. There is a function in MATLAB which does this moving average. The function is called "Smoothdata" (Copyright 2016-2018 The MathWorks, Inc.). An example of its application on one of the laboratory data sets is shown in fugue 3.8, which can be compared with Figure 3.9. The Gaussian filter causes an additional time lag between the different sensors. A cross correlation was applied on OPS and SDS sensors. The MATLAB function xcorr(x,y) (Copyright 1988-2019 The MathWorks, Inc.) returns the cross-correlation of two discrete-time sequences. Cross-correlation measures the similarity between a vector x and shifted (lagged) copies of a vector y as a function of the lag [51]. The returned arguments are r and lags, which respectively represent the correlation and the different lags at which the correlations were calculated. By plotting the r against different lags, the absolute maximum of the r was selected as a criterion for determining the highest correlation for a given time domain of "Gaussian-filtered" OPS data, see Figure 3.10. The corresponding lag to that maximum r (rmax) was used to synchronize the the time lag between the OPS and the SDS sensor. By raising the time domain from 1 second to an extreme value of 1000, it was obswrved that the trend for the rmax is risisng up to some time domin and after that it starts falling, see Figure 3.11.

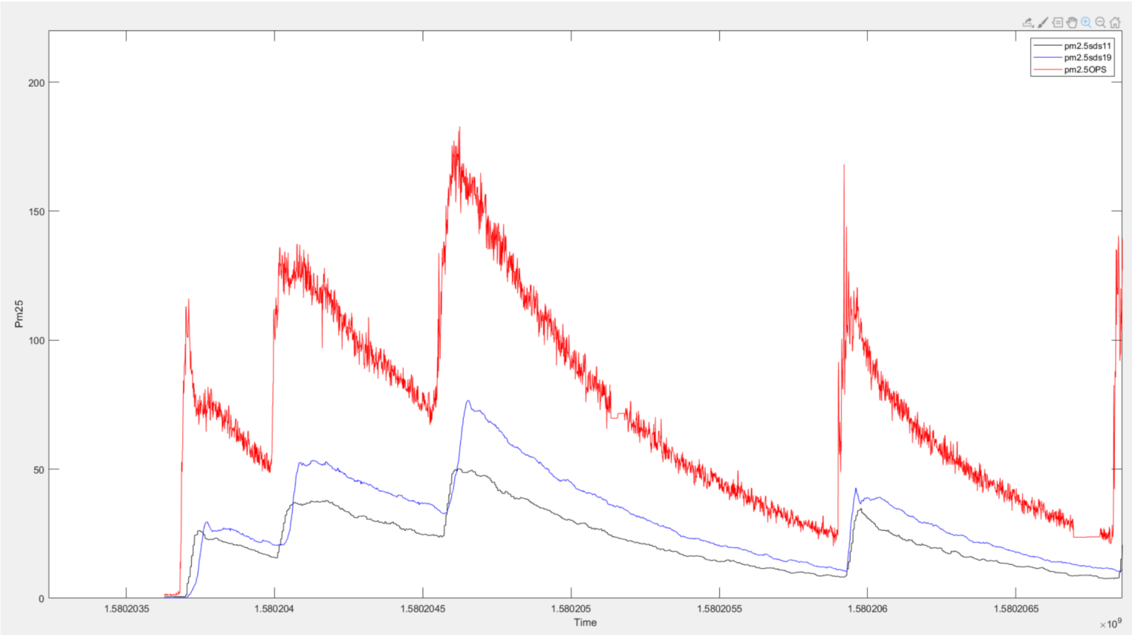


Figure 3.8: Comparison of the response time for different sensors before filtering, Red: OPS, Blue: SDS019, Black: SDS011

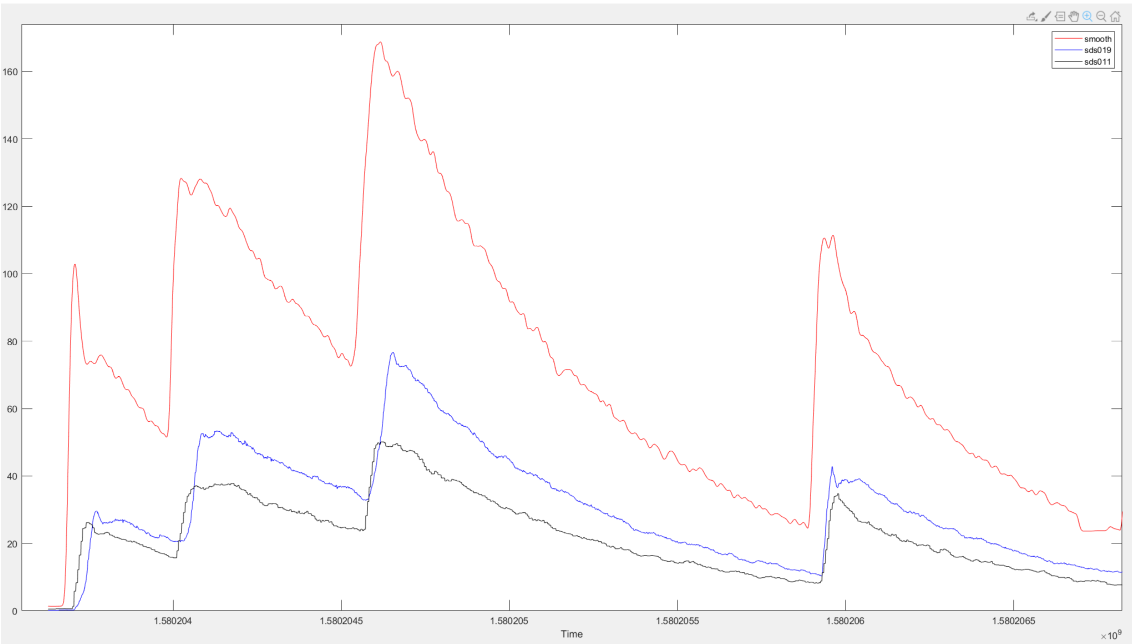


Figure 3.9: Comparison of the response time for different sensors after filtering, Red: OPS, Blue: SDS019, Black: SDS011

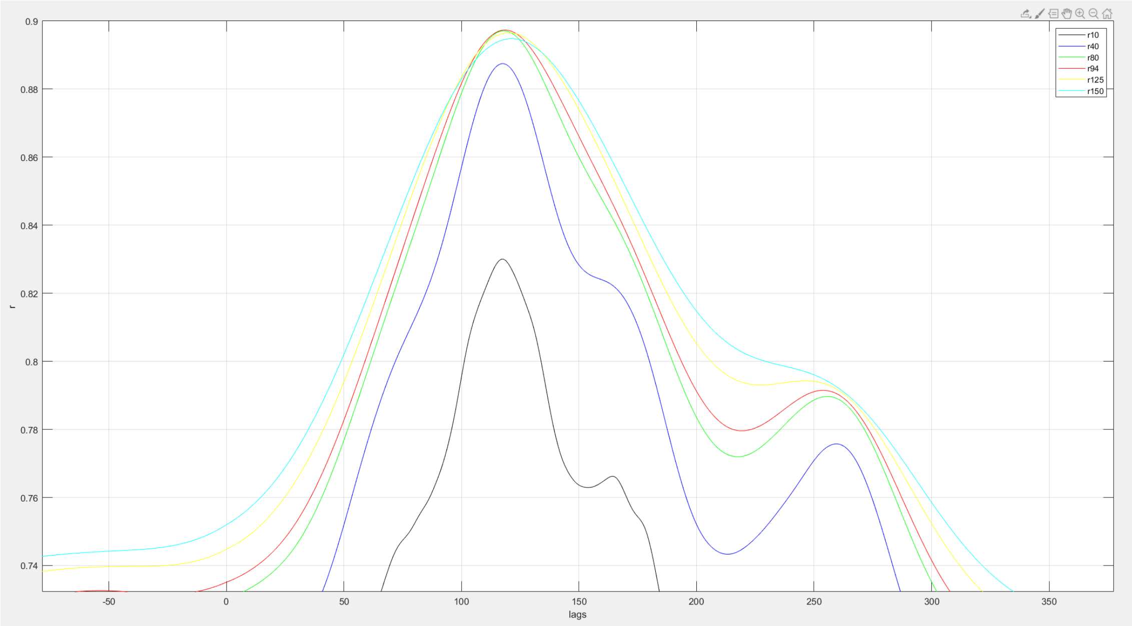


Figure 3.10: rmax for differnet time domains

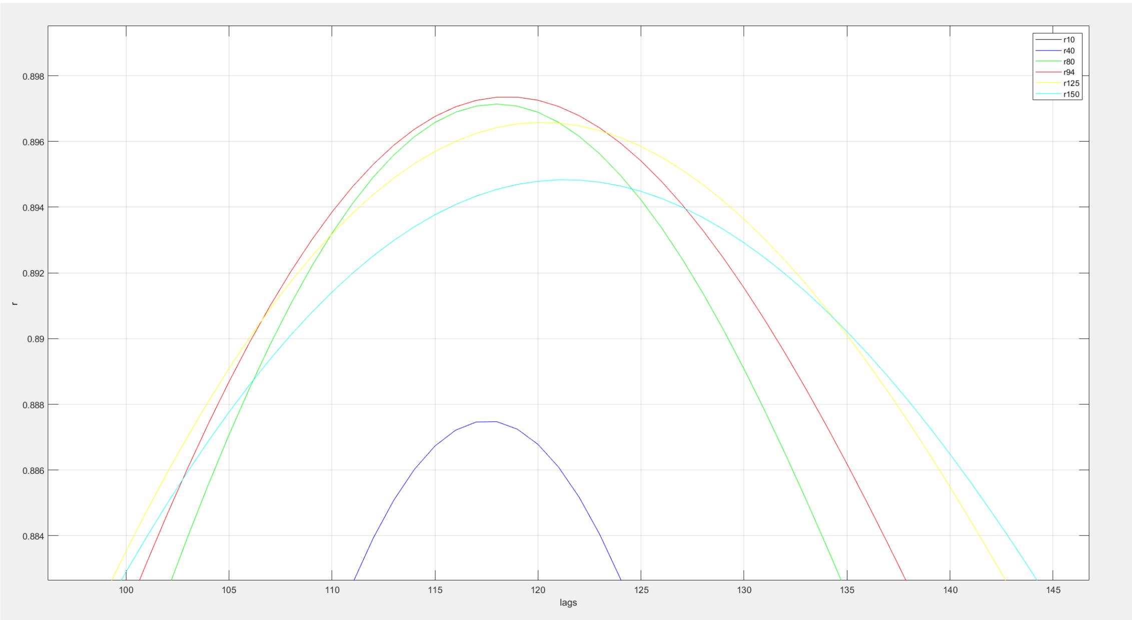


Figure 3.11: rmax for differnet time domains, a closer look to the peak of figure 3.13: the red curve is where the trend changes (in this case 94 seconds)

The maximum of rmax within different time domains was obtained by running a loop for time domains at the mentioned range. The time domain corresponding to the maximum rmax was chosen for each data set to apply with the Gaussian filter and the time lags corresponding to that optimal filter was used to fix the time lag of the OPS data caused by the optimal Gaussian filter. Finally, another interpolation was applied to all the data to synchronize the time series.

3.4.3 Correlation Analysis and Regression

To analyse the performance of the different sensors against the reference one a correlation analysis was performed. This was done after data processing, according the described in the previous subsections. The correlation analysis was carried out using a cross correlation MATLAB function called Corrcoef[51]. The function produces a matrix of sample correlation coefficients for a data matrix (where each column represents a separate quantity). The correlation coefficients range from -1 to 1, where bValues close to 1 indicate that there is a positive linear relationship between the data columns. Values close to -1 indicate that one column of data has a negative linear relationship to another column of data (anticorrelation). Values close or equal to zero suggest there is no linear relationship between the data columns.

After correlation analysis, the linear regression was performed and the slope and intercept of the lines were calculated. The most common type of linear regression is a least-squares fit, which can fit both lines and polynomials, among other linear models. Additionally, the possibility for a multiple regression by including T and RH as extra inputs to the regression problem was investigated. These different sets of regression methods were applied to the data sets and the R-squared values of different methods were compared to determine the method. The the mean relative error of SDS011 and SDS019 compared to OPS calculated based on the following equation.

$$Relative\ error = [\frac{SDS - OPS}{OPS}] \tag{3.1}$$

Equation for single linear regression model:

$$PM\,2.5\_corrected = PM\,2.5\_measured * Slope \tag{3.2}$$

Equation for single linear regression model with intercept:

$$PM\,2.5\_corrected = Intercept + PM\,2.5\_measured * Slope \tag{3.3}$$

Equation for Multiple linear regression model:

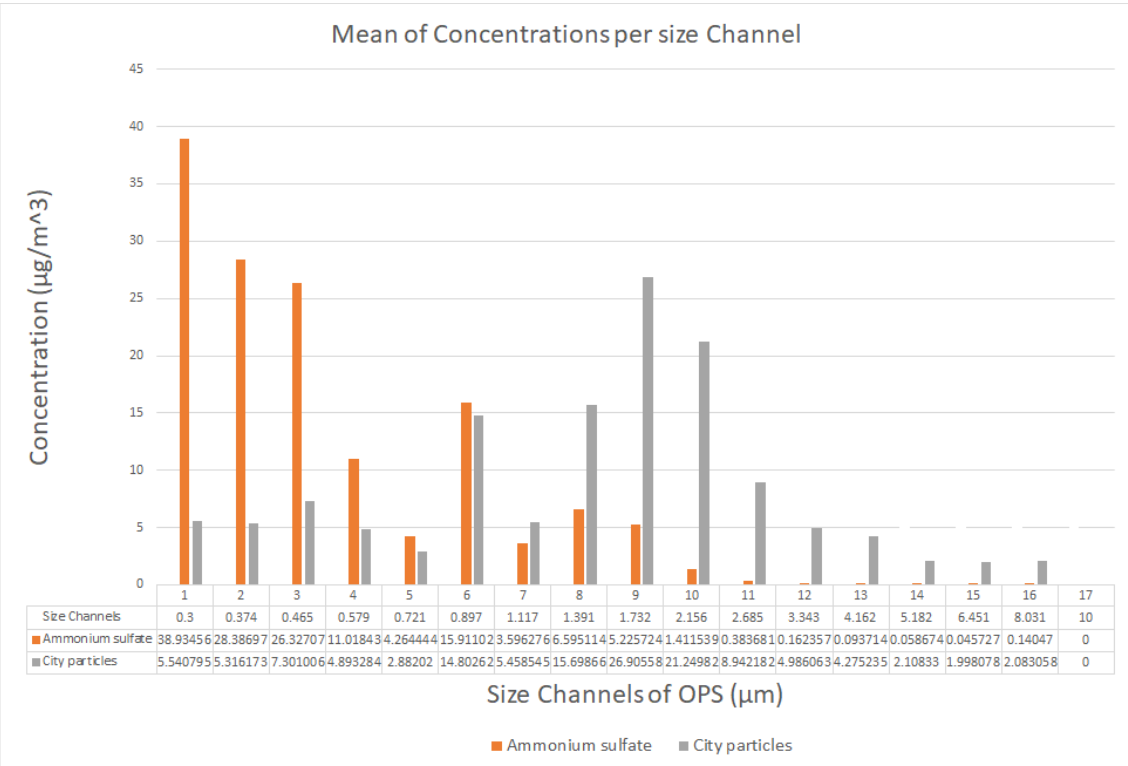
$$PM2.5\_corrected = Intercept + \Sigma parameter\_measured * parameter\_Slope \tag{3.4}$$

where parameter is: PM 2.5, T, RH.

4
Results and Discussions

4.1 Size Channels

A comparison between the measurements if OPS in 2 different environments were performed. Figure 4.1 shows the typical mass size distribution of particles for pure ammonium sulfate from laboratory experiments with orange color and the mixture of particles in the city ambient air with gray bars. It can be seen in Figure 4.1 that the size distribution of the generated ammonium sulfate particles is such that all particles are lower than 2.5 μm, hence corresponding to PM2.5 is under focus and PM10 is delimited. On the other hand, the gray bars are not negligible from channels 11 to 17 which means there are coarse particles in the measurement site in the city (Olskroken).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Size Channels	0.3	0.374	0.465	0.579	0.721	0.897	1.117	1.391	1.732	2.156	2.685	3.343	4.162	5.182	6.451	8.031	10
Ammonium sulfate	38.93456	28.38697	26.32707	11.01843	4.264444	15.91102	3.596276	6.595114	5.225724	1.411539	0.383681	0.162357	0.093714	0.058674	0.045727	0.14047	0
City particles	5.540795	5.316173	7.301006	4.893284	2.88202	14.80262	5.458545	15.69866	26.90558	21.24982	8.942182	4.986063	4.275235	2.10833	1.998078	2.083058	0

Figure 4.1: Mean values for size channel of OPS, Field vs. laboratory data .

4.2 General Explanation of results

Several days of experiments and iterations for the instruments were conducted, and here a few data sets are presented in this chapter with their results. Data sets A,B,C are the chosen data sets from the laboratory experiments whereas data set D is from a measurement in the city of Gothenburg. Data set E corresponds to the field tests at Femman air quality station. Data sets A to E are presented with their humidity and temperature plots over the time span of each data set as well as a table including the properties of the experiment and another table referred as results table including some statistics and errors. The first block of each table shows the mean relative error of SDS011 and SDS019 compared to OPS calculated based on the equation 3.1. The mean relative errors are also plotted. The second and third blocks of each result table show the coefficients for the single linear regression without and with intercept, which can be applied to the equations 3.2 and 3.3 respectively. And the last block shows the coefficients for multiple linear regression which can be used based on the equation 3.4. Additionally, the regression plots for all the data sets are presented. After comparing the R-squared values for the single regression models the one with the intercept was chosen as correction factor and applied to the same data set for each data set and the new relative error after applying the correction factors based on equation 3.3 and was plotted for each data set. The mean values of that calculation is presented in the last line of first block of the results tables for each data set and is compared to the second line in the same block of the table.

4.3 Data Set A (Laboratory experiment)

Comparison of the Figures 4.4 and 4.5 reveals that The offset different between the OPS and SDS011 can be corrected by excluding the first three size channels of OPS from the Pm2.5 calculations. This could be due to the low counting yield of SDS011 at those size ranges.

Table 4.1: Data Set A, 22 degrees (room Temperature)

properties	observations
Location	laboratory
Particle mixture	Pure Ammonium Sulfate
Duration (minutes)	207
Average Temperature	21.6
Maximum Temperature	22.7
Minimum Temperature	20.7
Average RH(%)	52.5
Maximum RH(%)	81.2
Minimum RH(%)	40.7
Average PM2.5	130
Maximum PM2.5	829
Minimum PM2.5	5

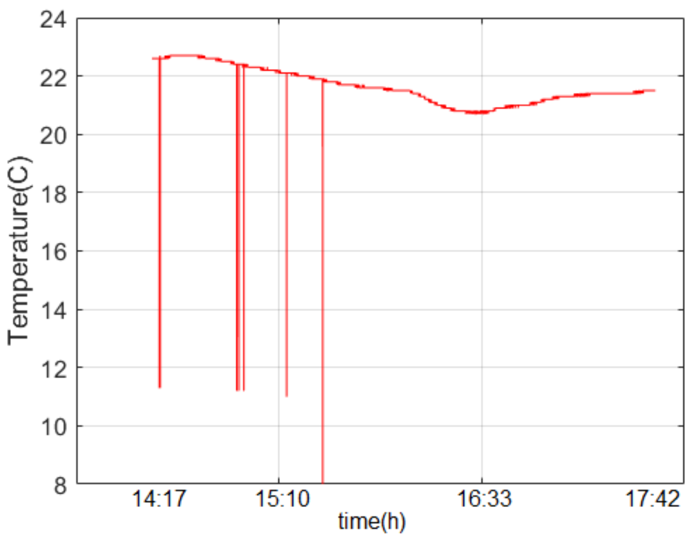


Figure 4.2: Temperature, Data set A

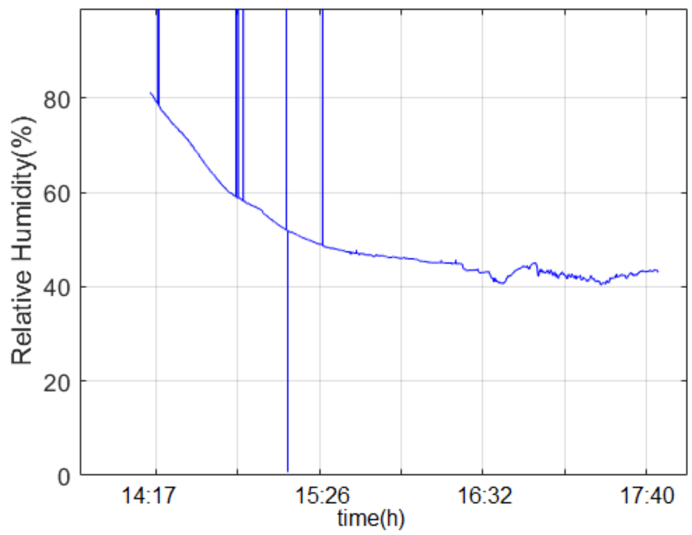


Figure 4.3: Relative Humidity, Data set A

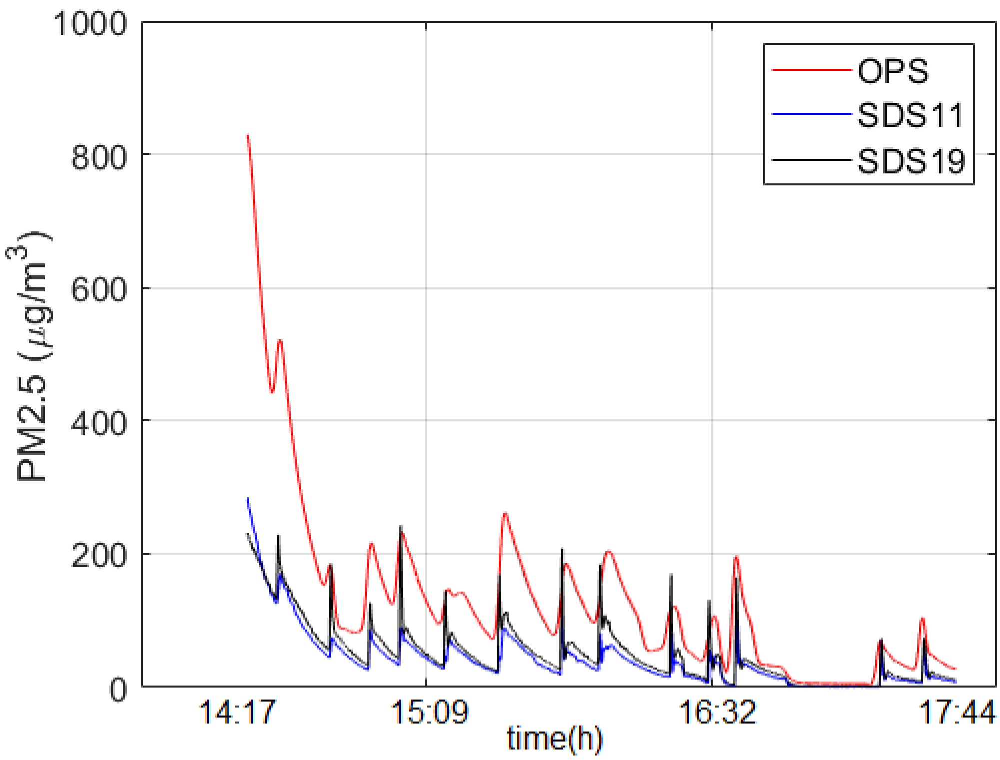


Figure 4.4: PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set A

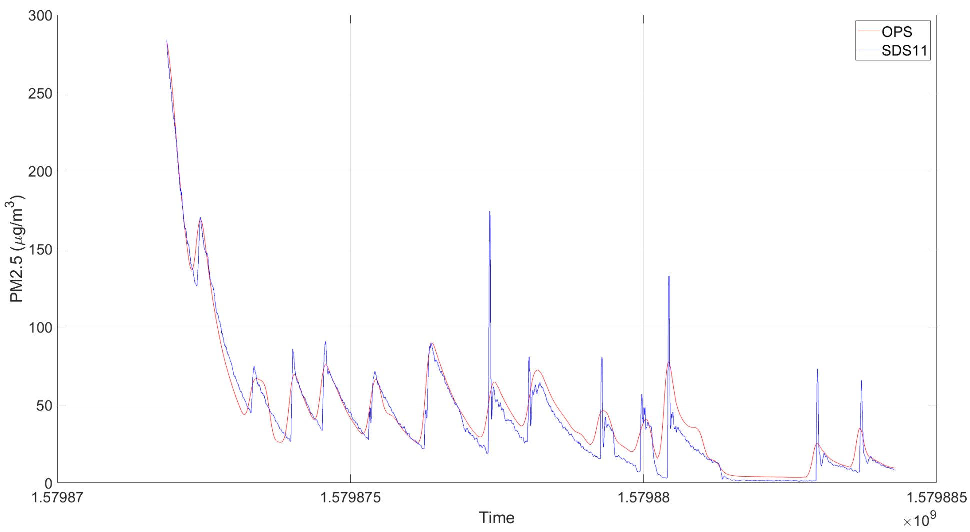


Figure 4.5: PM2.5 mass concentrations for SDS011 and OPS excluding the first 3 channels, Data set A

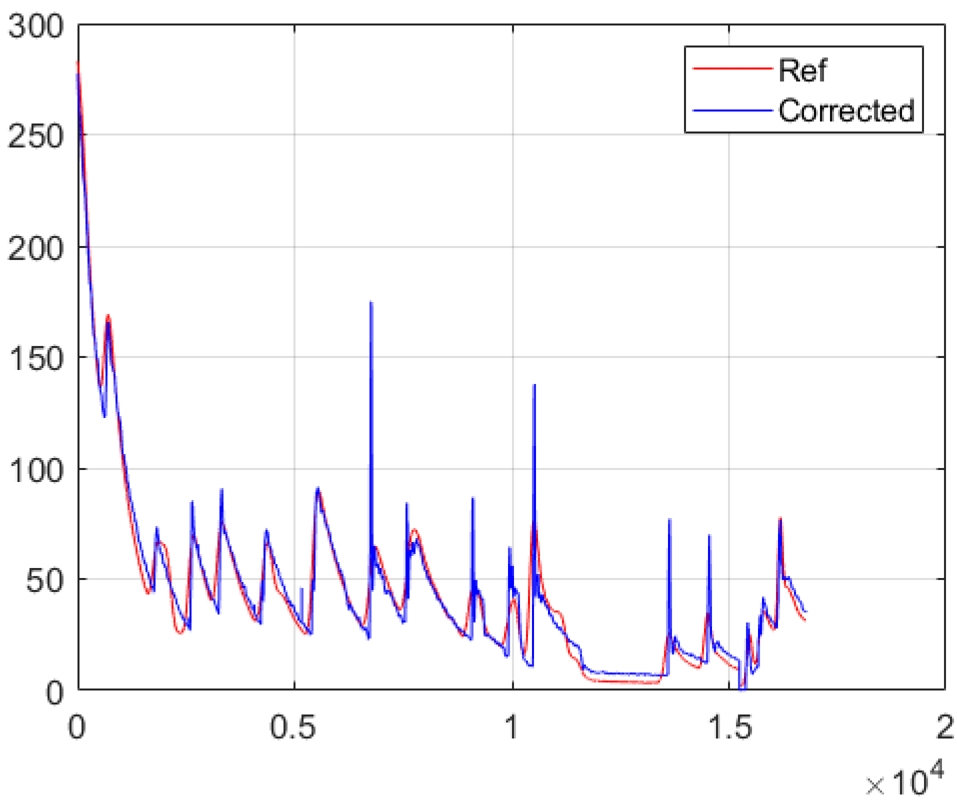


Figure 4.6: SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set A

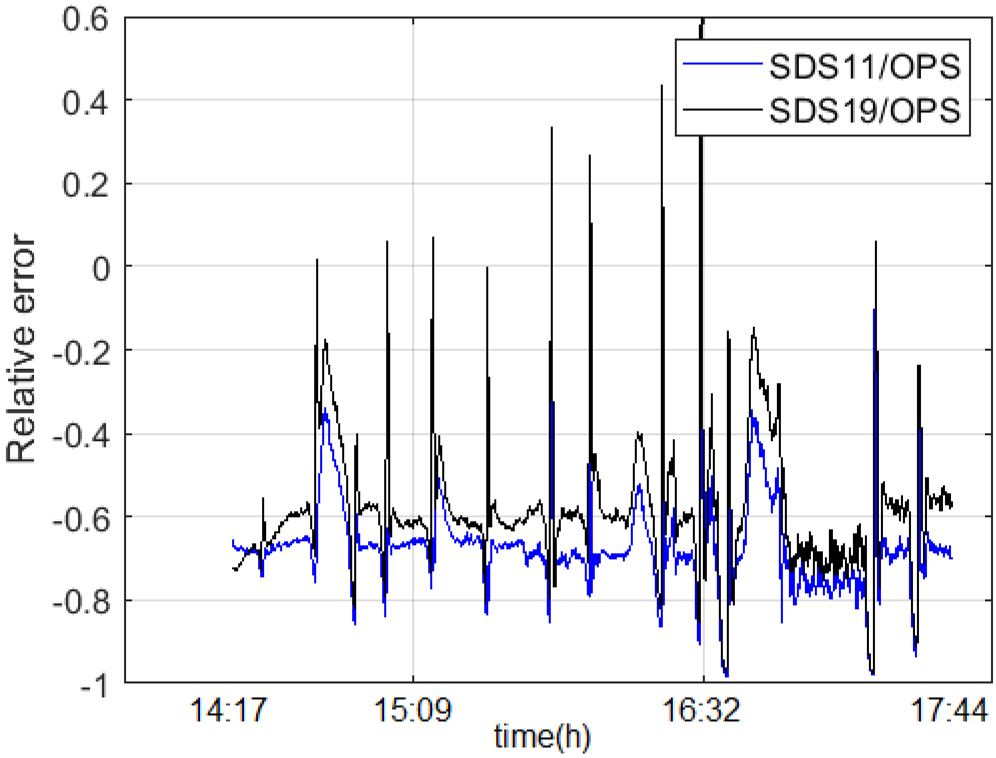


Figure 4.7: Relative errors for SDS019 and SDS011 vs. OPS, Data set A

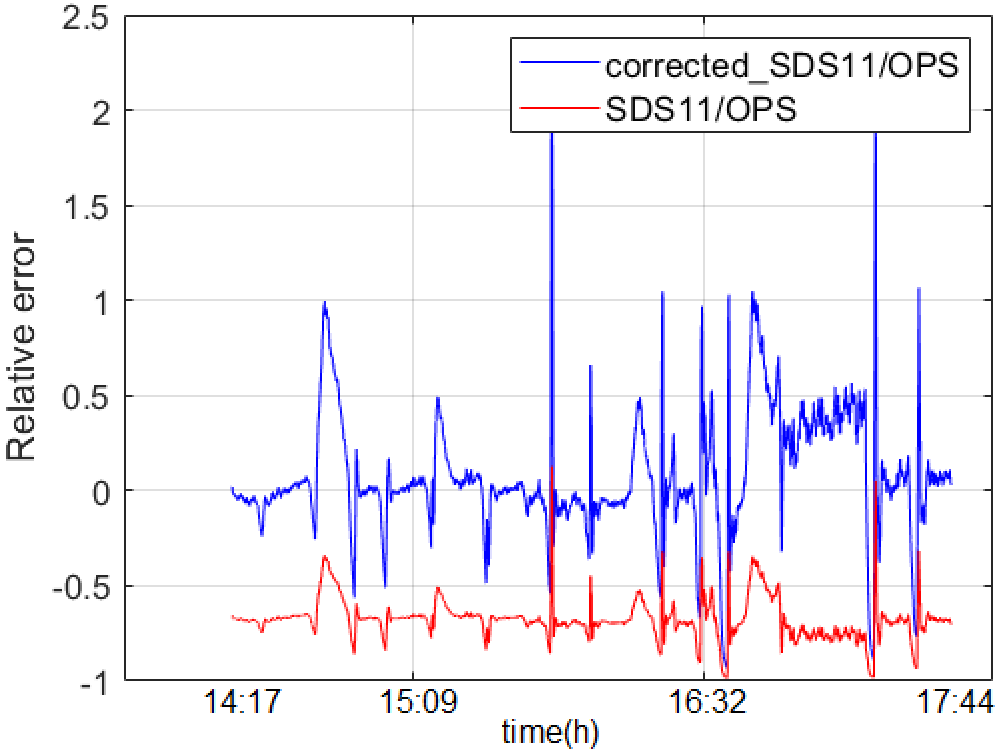


Figure 4.8: Relative errors for SDS011 vs. OPS before and after applying the correction factors based on Linear regression, Data set A

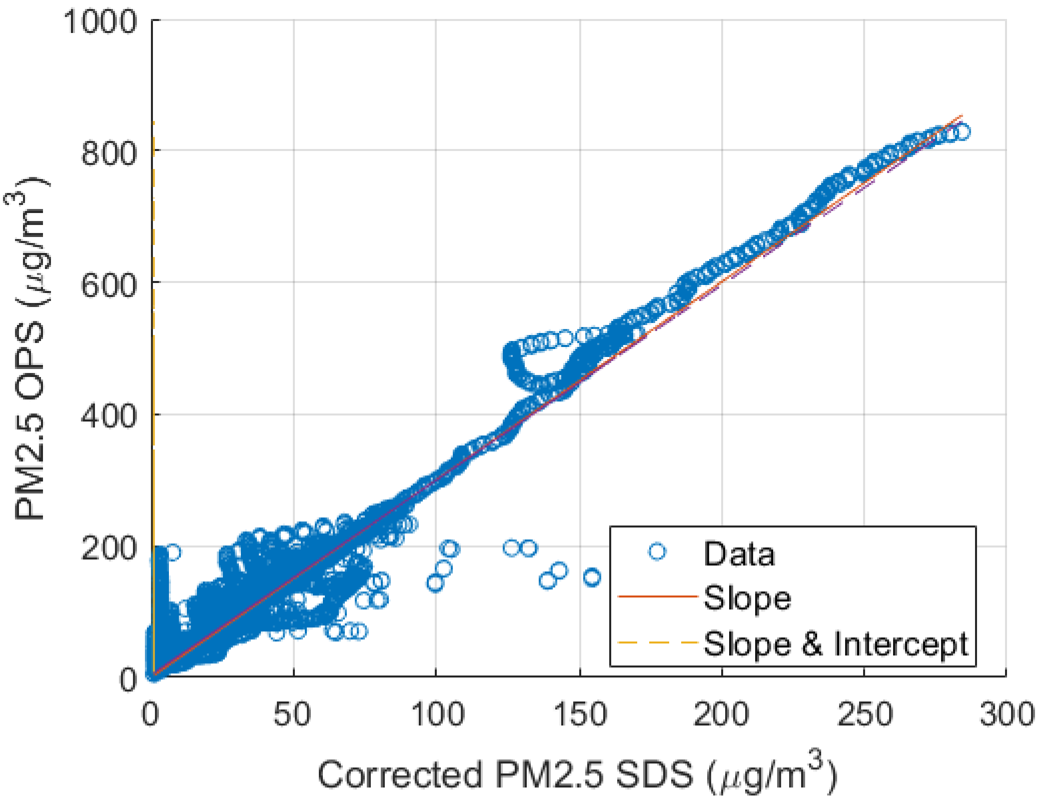


Figure 4.9: Linear Regression plots, Data set A

Table 4.2: Data Set A, 22 degrees (room Temperature)

properties	observations
Mean relative error of SDS019	-57.65 %
Mean relative error of SDS011	-67.41 %
Mean relative error of SDS011 after applying the correction factors	8.11 %
lope (Single Linear Regression)	3.004
r^2 (Single Linear Regression)	93.88 %
Slope (Single Linear Regression with intercept)	2.959
Intercept (Single Linear Regression with intercept)	3.936
r^2 (Single Linear Regression with intercept)	93.93 %
Intercept (Multiple Linear Regression)	59.34
PM2.5 factor (Multiple Linear Regression)	0.962
RH factor (Multiple Linear Regression)	-0.009
Temperature factor (Multiple Linear Regression)	-2.522
r^2 (Multiple Linear Regression)	93.90 %

4.4 Data Set B (Laboratory experiment)

Table 4.3: Data Set B, 25 degrees

properties	observations
Location	laboratory
Particle mixture	Pure Ammonium Sulfate
Duration (minutes)	182
Average Temperature	25.2
Maximum Temperature	26.3
Minimum Temperature	23.4
Average RH(%)	50.5
Maximum RH(%)	83.1
Minimum RH(%)	39.9
Average PM2.5	72.9
Maximum PM2.5	411.5
Minimum PM2.5	1.6

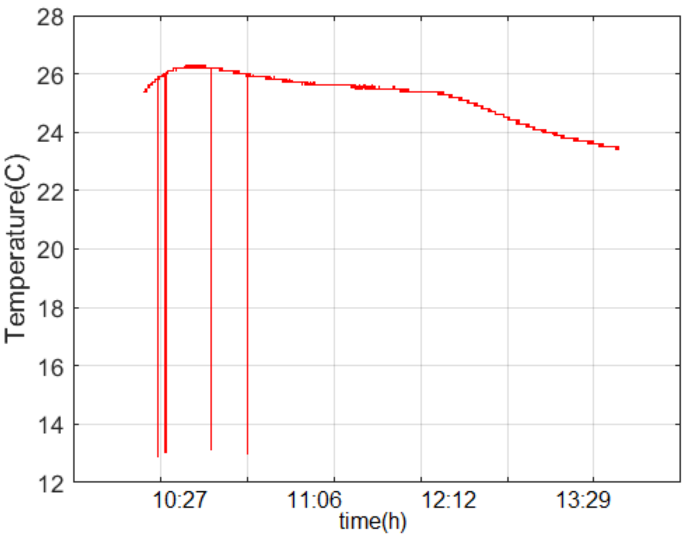


Figure 4.10: Temperature, Data set B

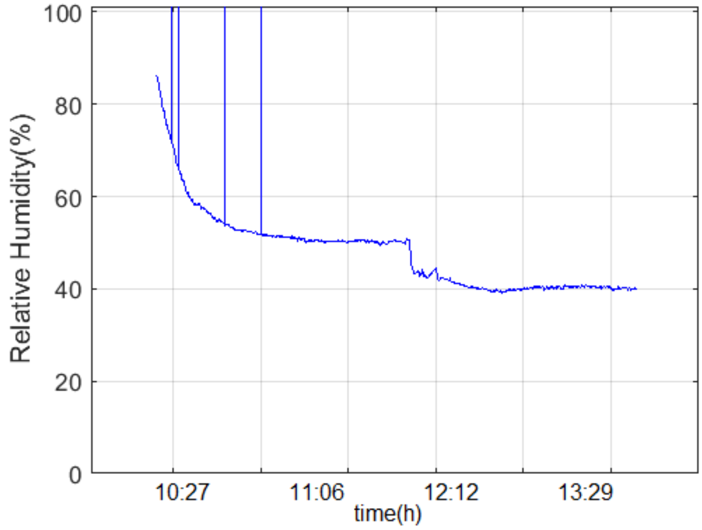


Figure 4.11: Relative Humidity, Data set B

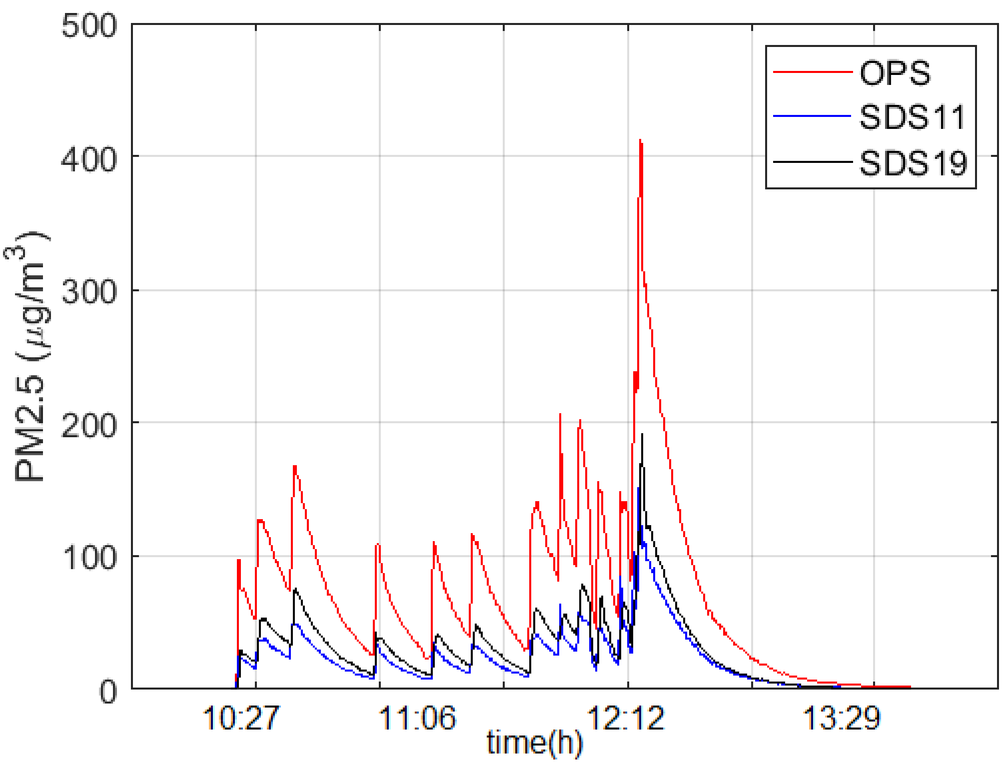


Figure 4.12: PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set B

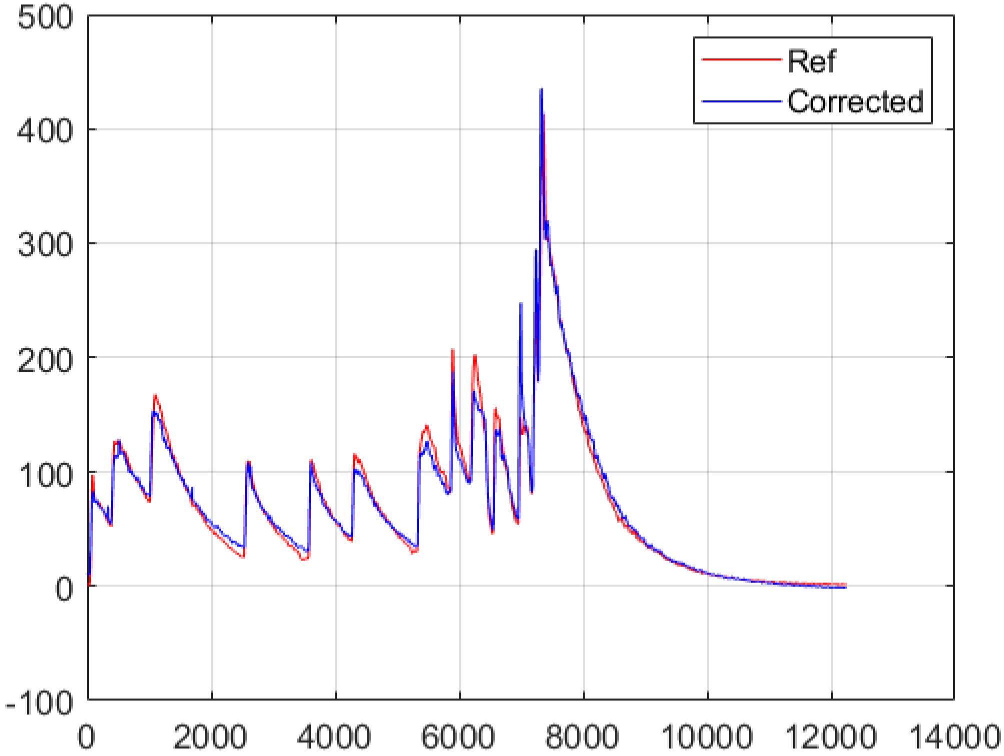


Figure 4.13: SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set B

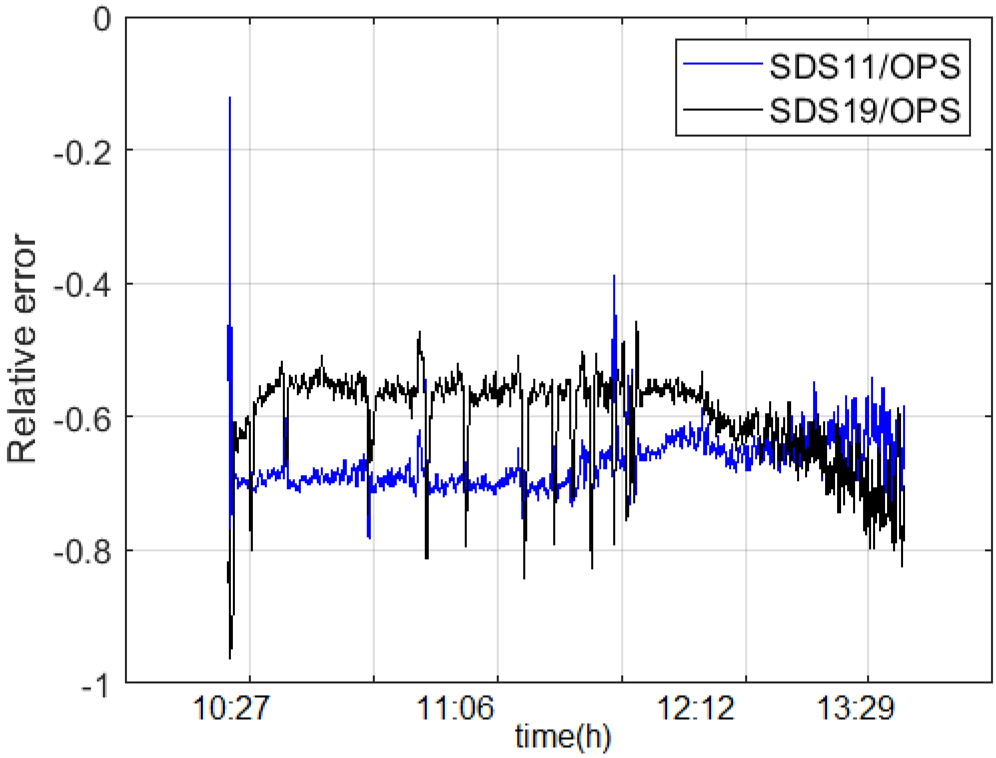


Figure 4.14: Relative errors for SDS019 and SDS011 vs. OPS, Data set B

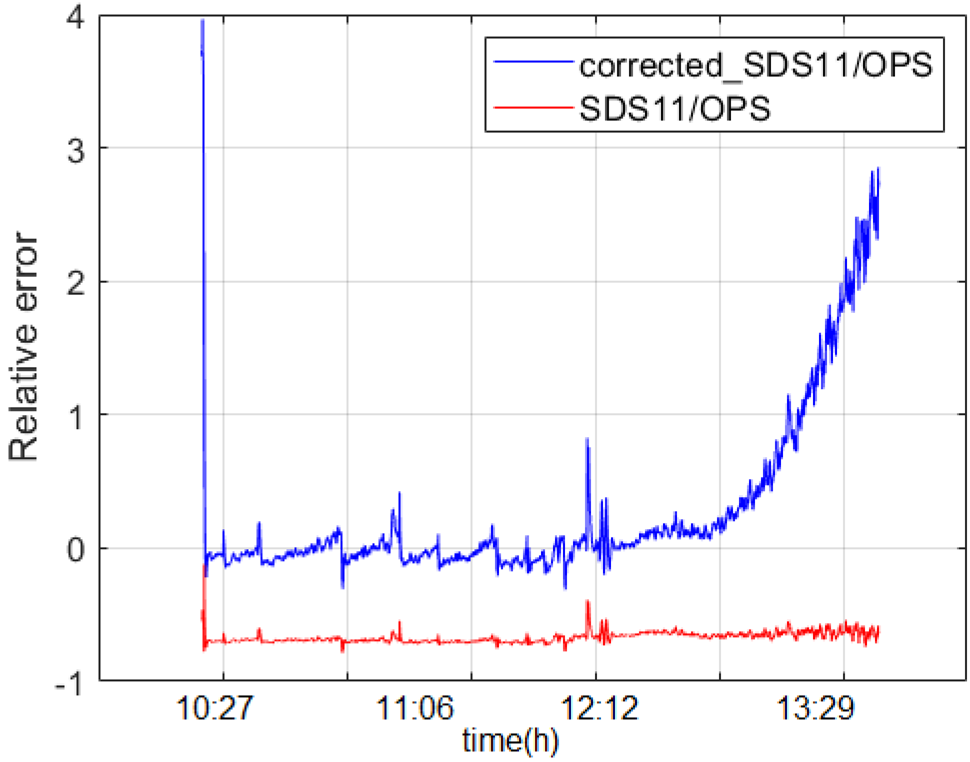


Figure 4.15: Relative errors for SDS011 vs. OPS before and after applying the correction factors based on Linear regression, Data set B

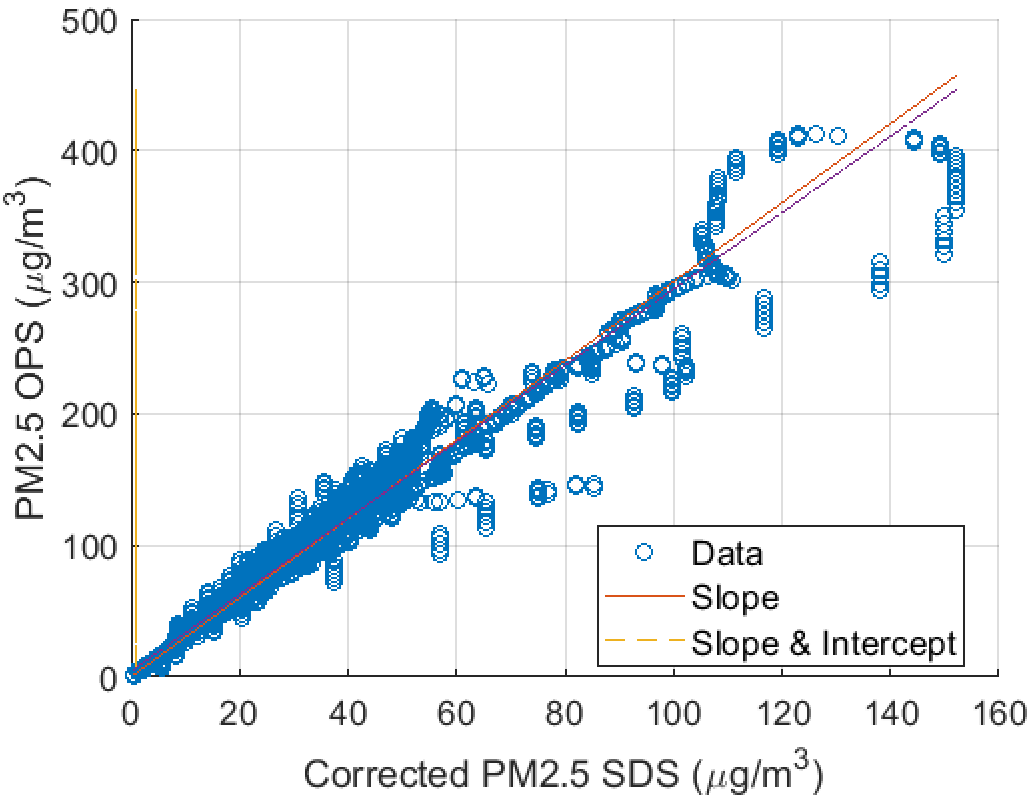


Figure 4.16: Linear Regression plots, Data set B

Table 4.4: Data Set B, 25 degrees

properties	observations
Mean relative error of SDS019	-60.27 %
Mean relative error of SDS011	-66.95 %
Mean relative error of SDS011 after applying the correction factors	27.87 %
Slope (Single Linear Regression)	3.003
r^2 (Single Linear Regression)	96.78 %
Slope (Single Linear Regression with intercept)	2.903
Intercept (Single Linear Regression with intercept)	4.429
r^2 (Single Linear Regression with intercept)	97.00 %
Intercept (Multiple Linear Regression)	-92.23
PM2.5 factor (Multiple Linear Regression)	2.84
RH factor (Multiple Linear Regression)	0.0186
Temperature factor (Multiple Linear Regression)	3/918
r^2 (Multiple Linear Regression)	97.32 %

4.5 Data Set C (Laboratory experiment)

Table 4.5: Data Set C, 22 degrees (room Temperature)

properties	observations
Location	laboratory
Particle mixture	Pure Ammonium Sulfate
Duration (minutes)	15
Average Temperature	22.7
Maximum Temperature	22.8
Minimum Temperature	22.6
Average RH(%)	71.4
Maximum RH(%)	85.7
Minimum RH(%)	62.1
Average PM2.5	107.4
Maximum PM2.5	259.6
Minimum PM2.5	3.3

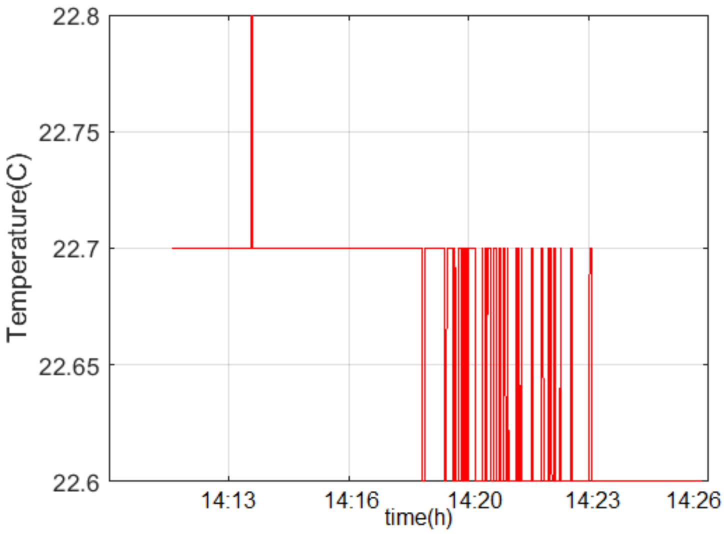


Figure 4.17: Temperature, Data set C

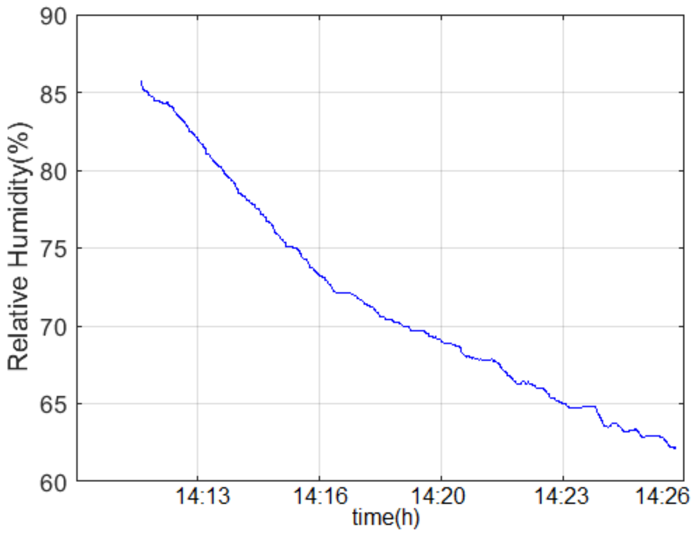


Figure 4.18: Relative Humidity, Data set C

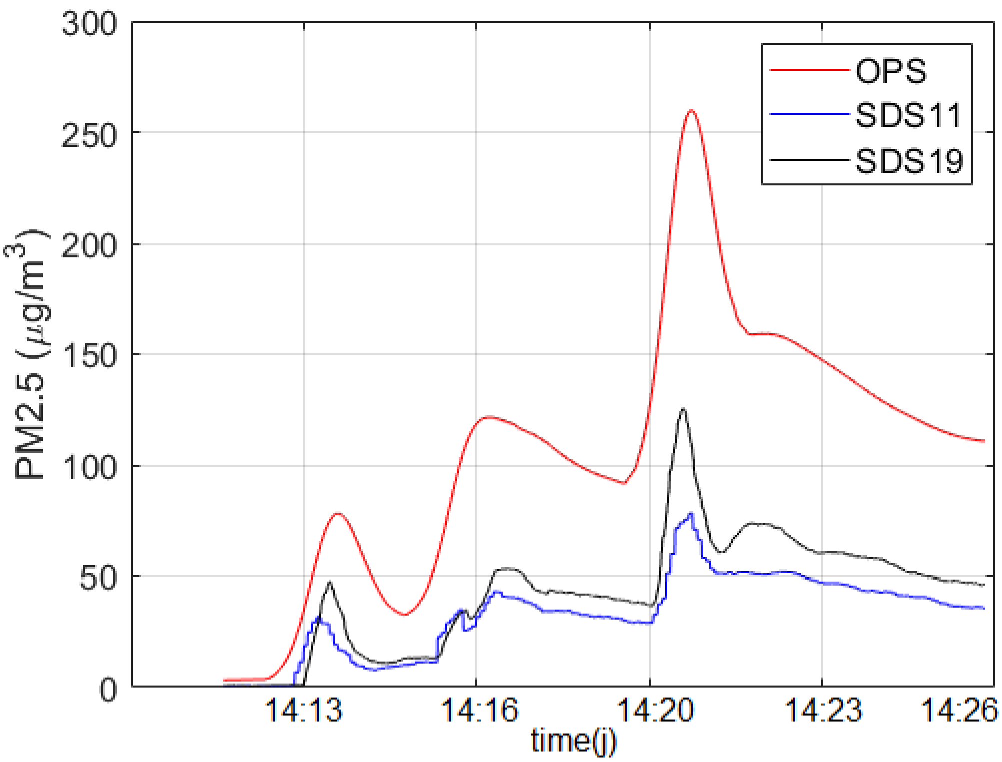


Figure 4.19: PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set C

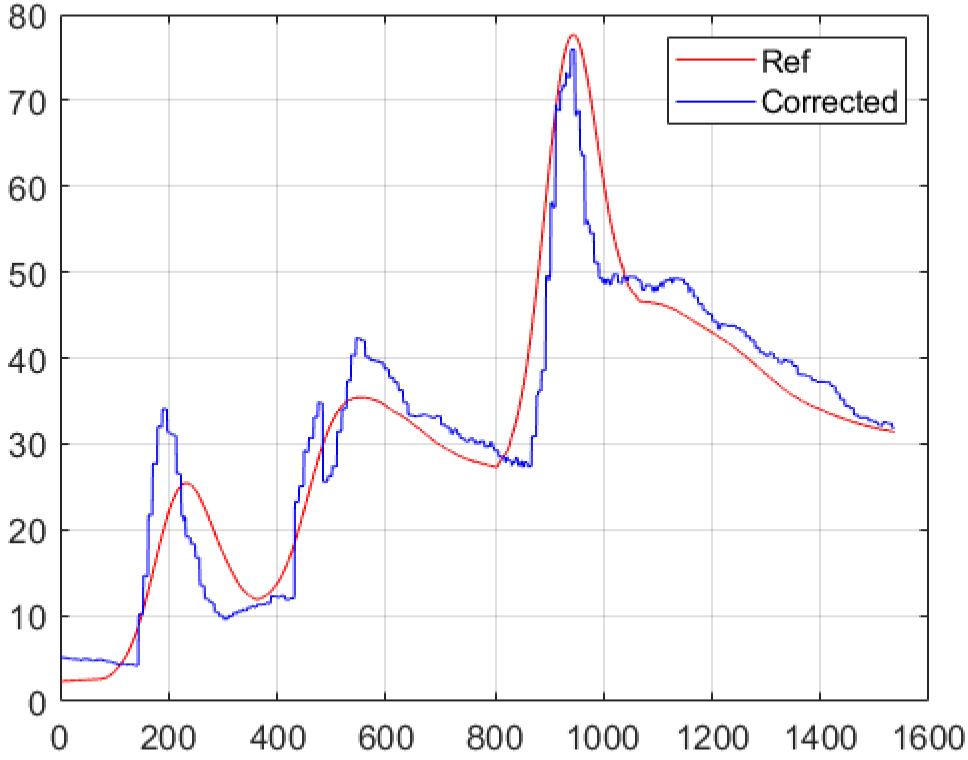


Figure 4.20: SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set C

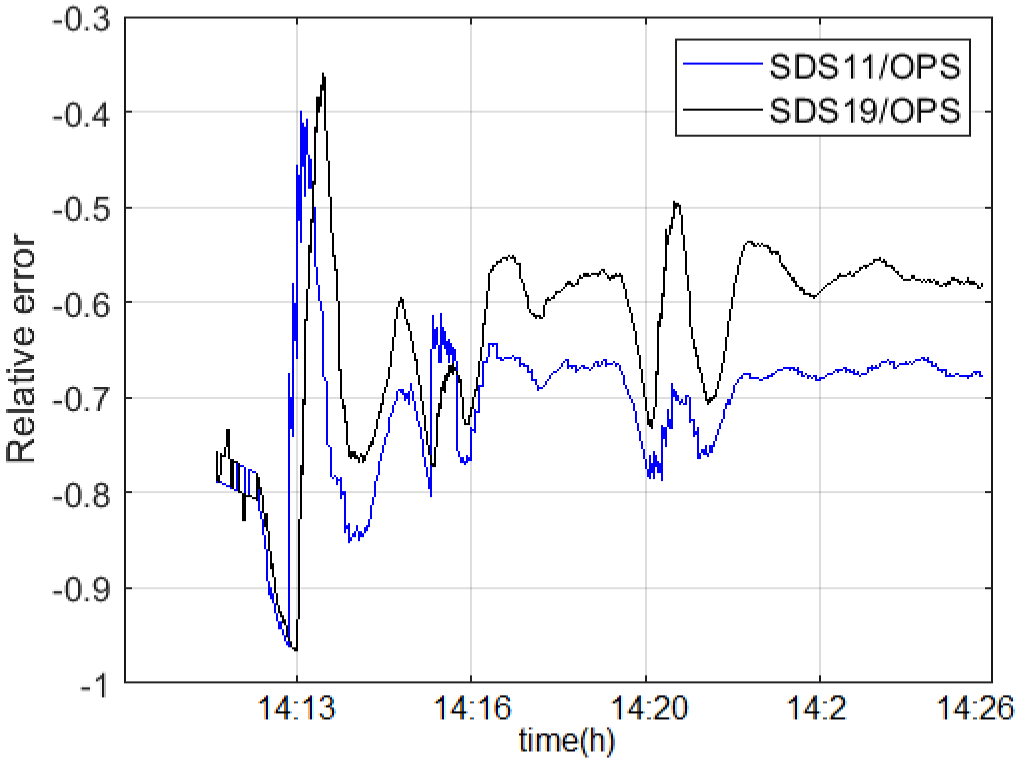


Figure 4.21: Relative errors for SDS019 and SDS011 vs. OPS, Data set C

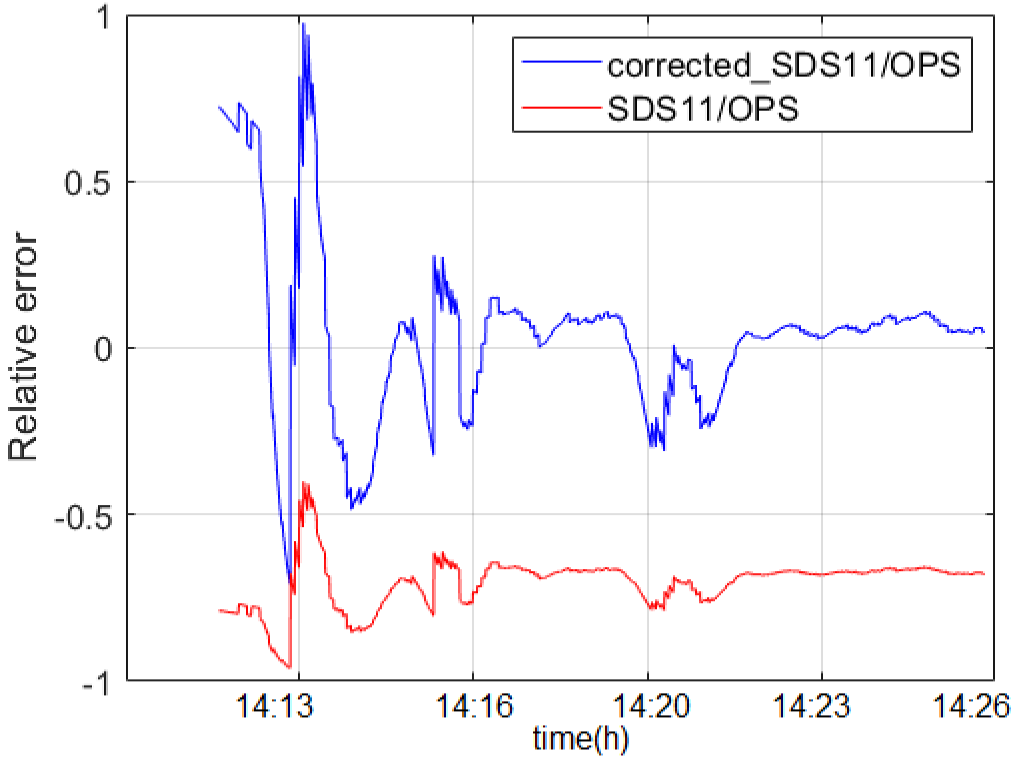


Figure 4.22: Relative errors for SDS011 vs. OPS before and after applying the correction factors based on Linear regression, Data set C

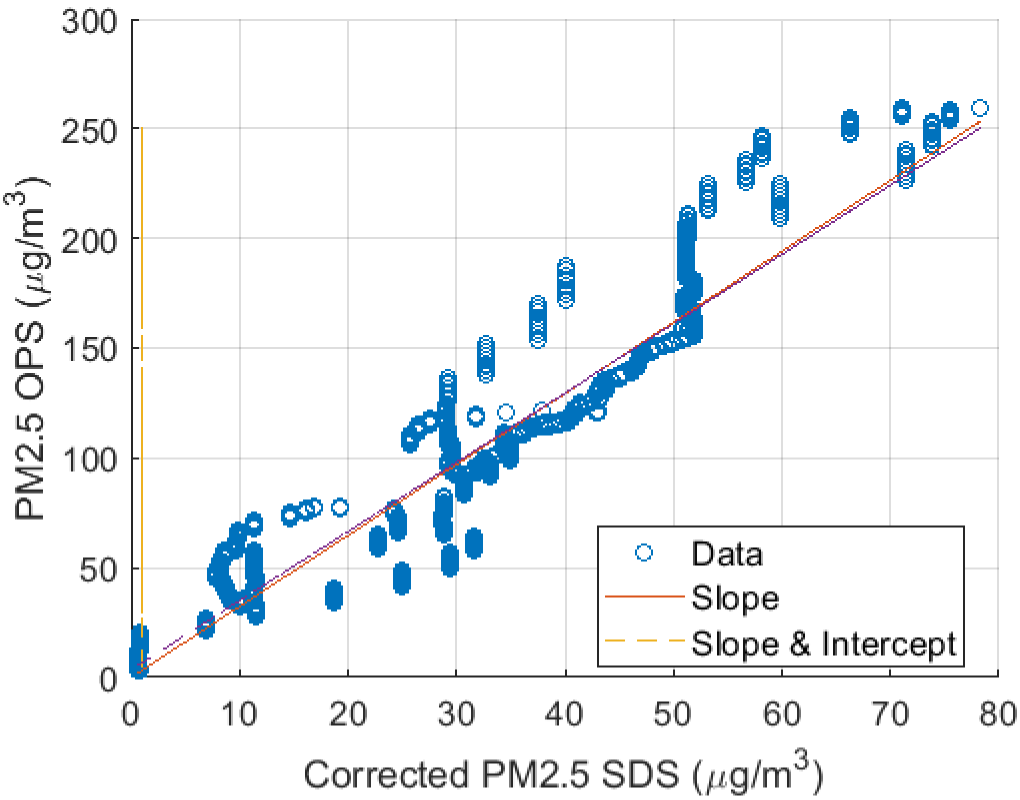


Figure 4.23: Linear Regression plots, Data set C

Table 4.6: Data Set C, 22 degrees (room Temperature)

properties	observations
Mean relative error of SDS019	-6314 %
Mean relative error of SDS011	-70.40 %
Mean relative error of SDS011 after applying the correction factors	4.12 %
Slope (Single Linear Regression)	3.23
r^2 (Single Linear Regression)	91.96 %
Slope (Single Linear Regression with intercept)	3.153
Intercept (Single Linear Regression with intercept)	3.476
r^2 (Single Linear Regression with intercept)	92.04 %
Intercept (Multiple Linear Regression)	102.08
PM2.5 factor (Multiple Linear Regression)	0.986
RH factor (Multiple Linear Regression)	0.358
Temperature factor (Multiple Linear Regression)	-5.650
r^2 (Multiple Linear Regression)	89.4 %

4.6 Data Set D (Field experiment with van)

Data set D is divided into to subsets and separate correction factors are calculated for it. Three big spikes are observed in Figure for data set D, which is split into two data sets in between the spikes. For this experiment, the difference (background concentrations) observed between the OPS and SDS sensors can be explained by OPS having a peak in its size distribution around 1.7 micrometer, see Figure 4.1. The peak is not typical for combustion, hence, it might be due to ocean spray or dust formation. SDS sensors were not sensitive to this but OPS showed a background concentration.

Table 4.7: Data Set D, ambient air with an average Temperature of 7 degrees

properties	observations
Location	A Parking lot in the city
Particle mixture	Ambient air in the city (Olskroken)
Duration (minutes)	404
Average Temperature	6.9
Maximum Temperature	9.8
Minimum Temperature	6.3
Average RH(%)	89.8
Maximum RH(%)	93.4
Minimum RH(%)	74.3
Average PM2.5	45.4
Maximum PM2.5	3030
Minimum PM2.5	31.8

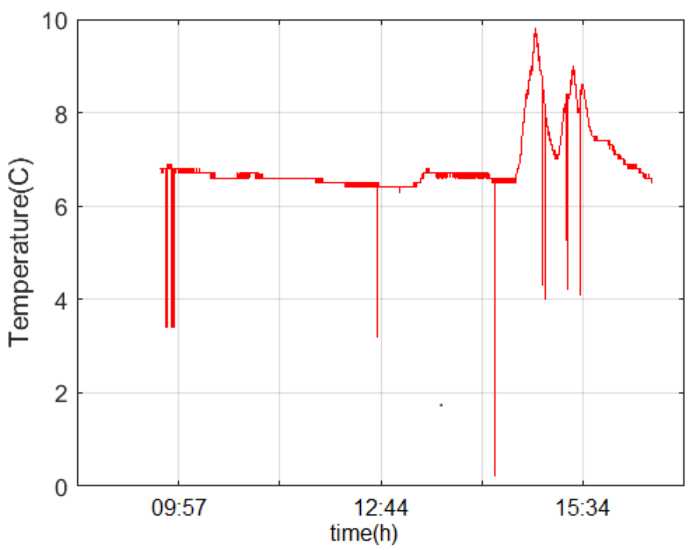


Figure 4.24: Temperature, Data set D

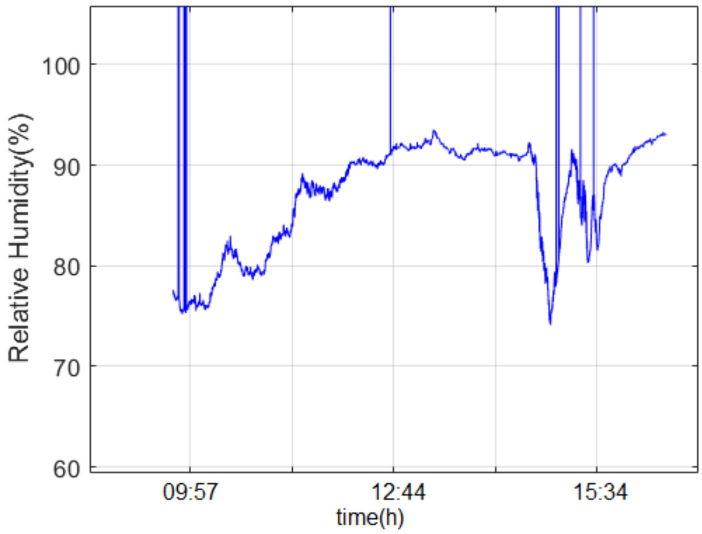


Figure 4.25: Relative Humidity, Data set D

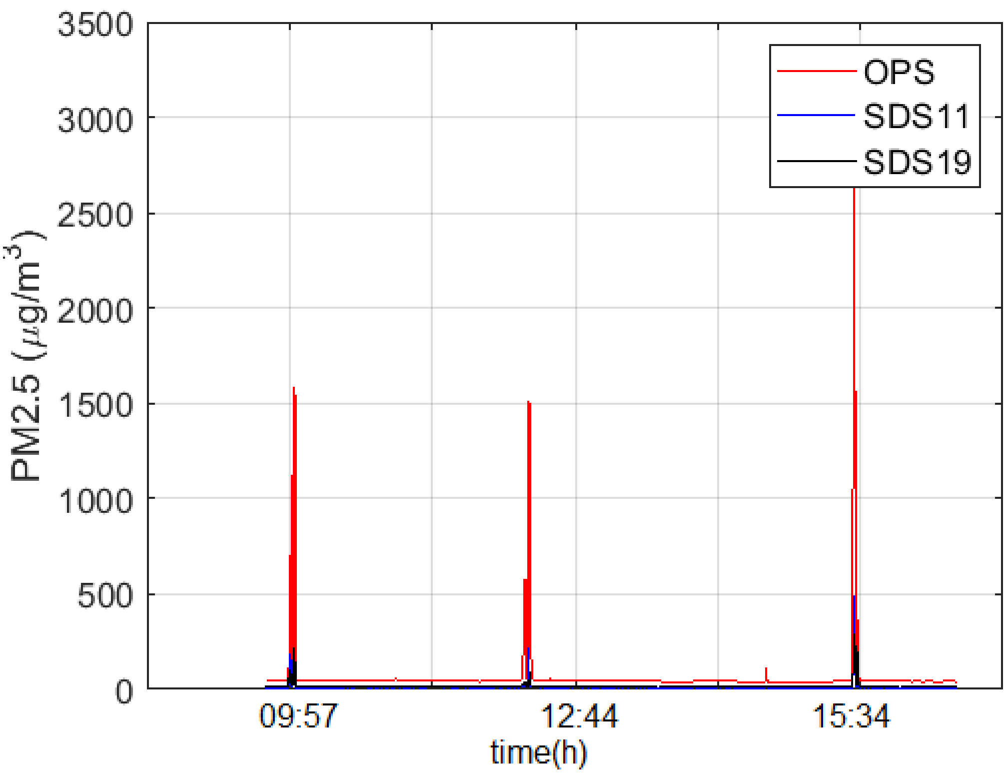


Figure 4.26: PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set D

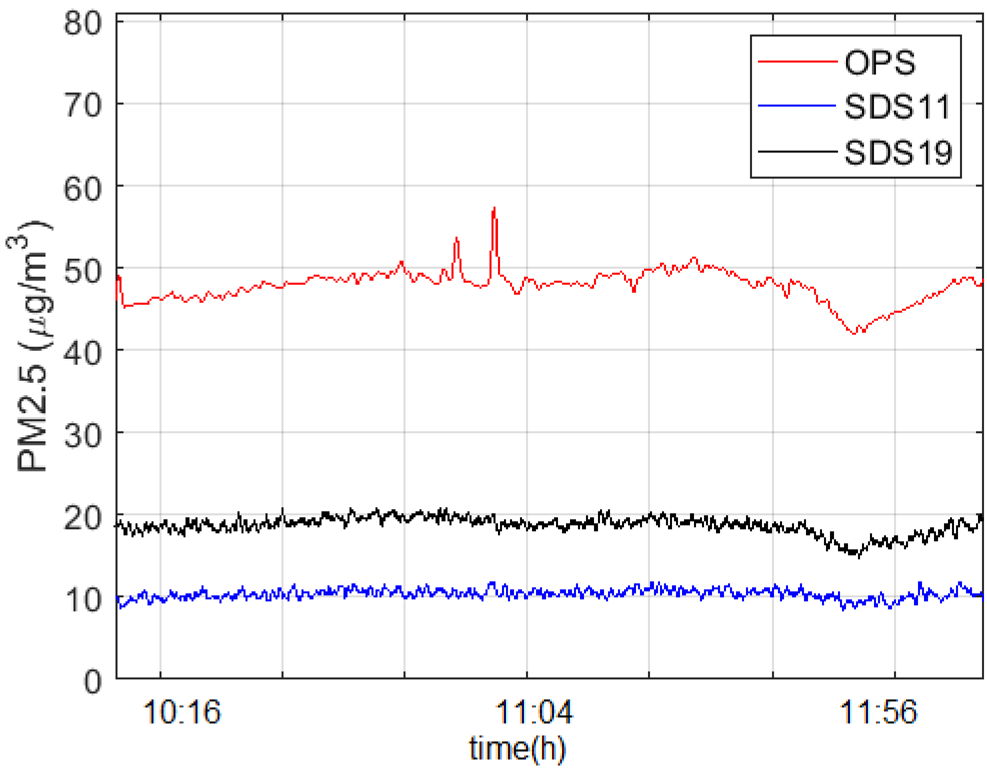


Figure 4.27: PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set D, between the first and second spike

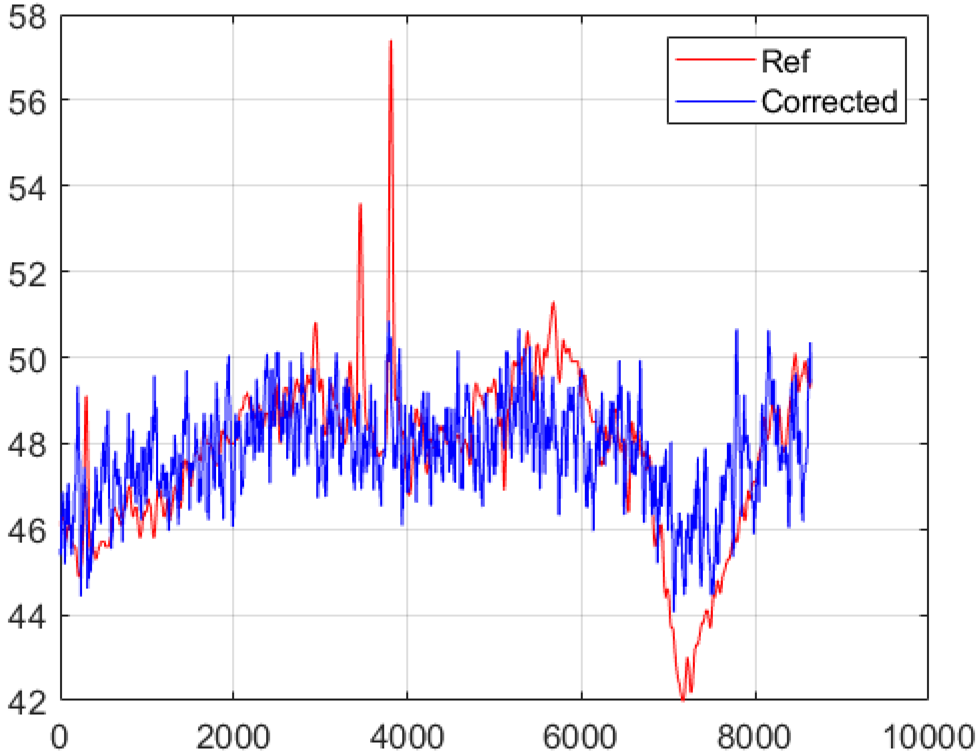


Figure 4.28: SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set D, between the first and second spike

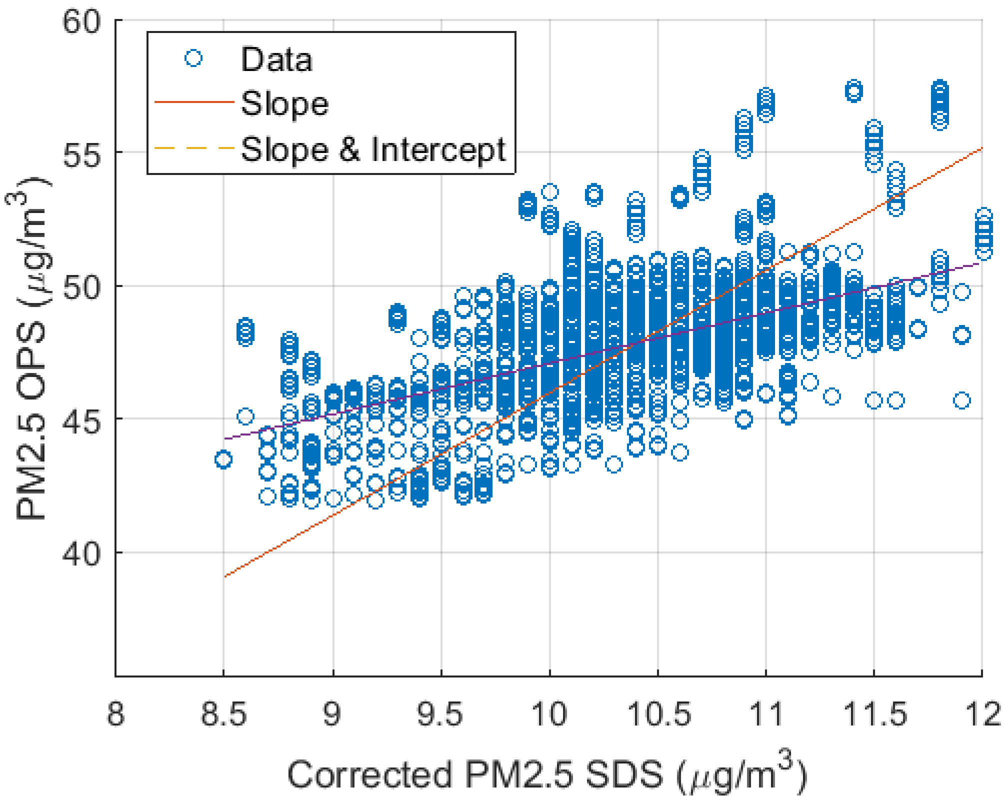


Figure 4.29: Linear Regression plots, Data set D, between the first and second spike

Table 4.8: Data Set D, ambient air with an average Temperature of 7 degrees, between the second and third spike

properties	observations
Mean relative error of SDS019	-64.57 %
Mean relative error of SDS011	-77.67 %
Mean relative error of SDS011 after applying the correction factors	0.32 %
Slope (Single Linear Regression)	4.511
r^2 (Single Linear Regression)	60.41 %
Slope (Single Linear Regression with intercept)	3.752
Intercept (Single Linear Regression with intercept)	7.3197
r^2 (Single Linear Regression with intercept)	63.02 %
Intercept (Multiple Linear Regression)	3.032
PM2.5 factor (Multiple Linear Regression)	3.72
RH factor (Multiple Linear Regression)	3.72
Temperature factor (Multiple Linear Regression)	0.69
r^2 (Multiple Linear Regression)	63.99 %

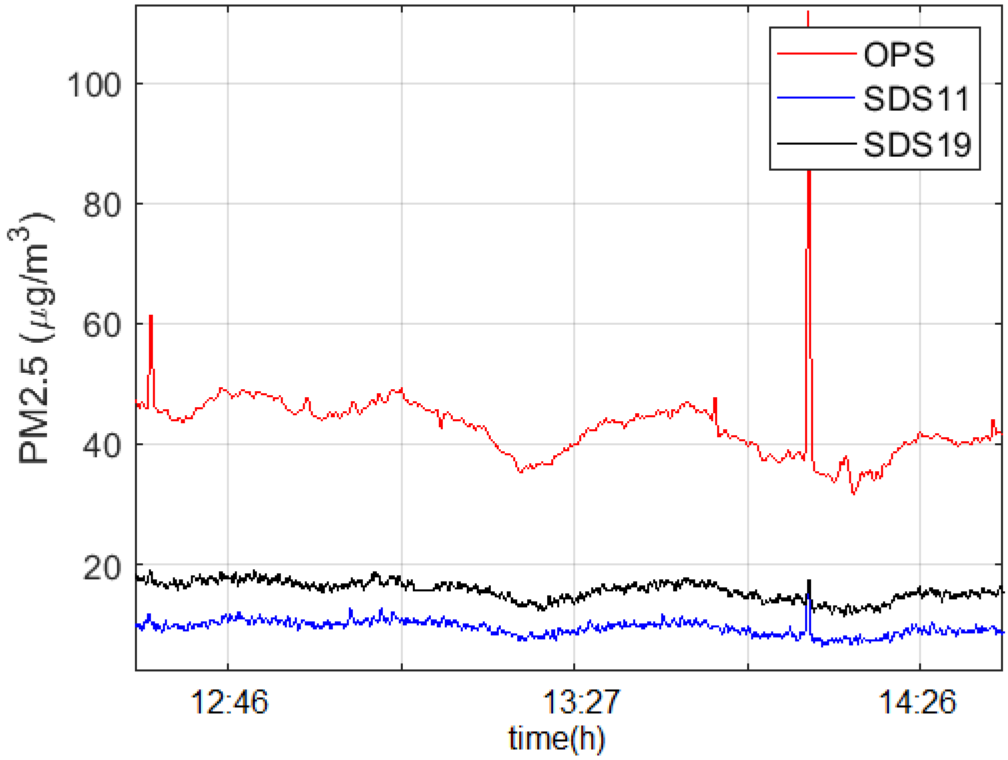


Figure 4.30: PM2.5 mass concentrations for OPS, SDS011, and SDS019, Data set D, between the second and third spike

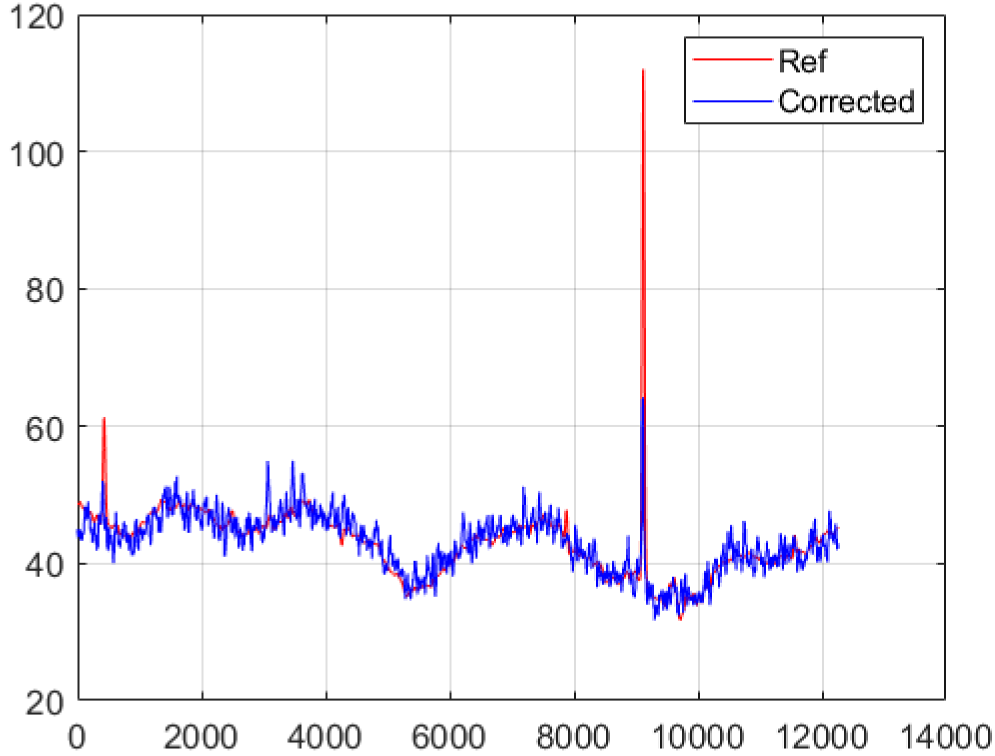


Figure 4.31: SDS011 Corrected with Multiple Linear Regression model compared to OPS, Data set D, between the second and third spike

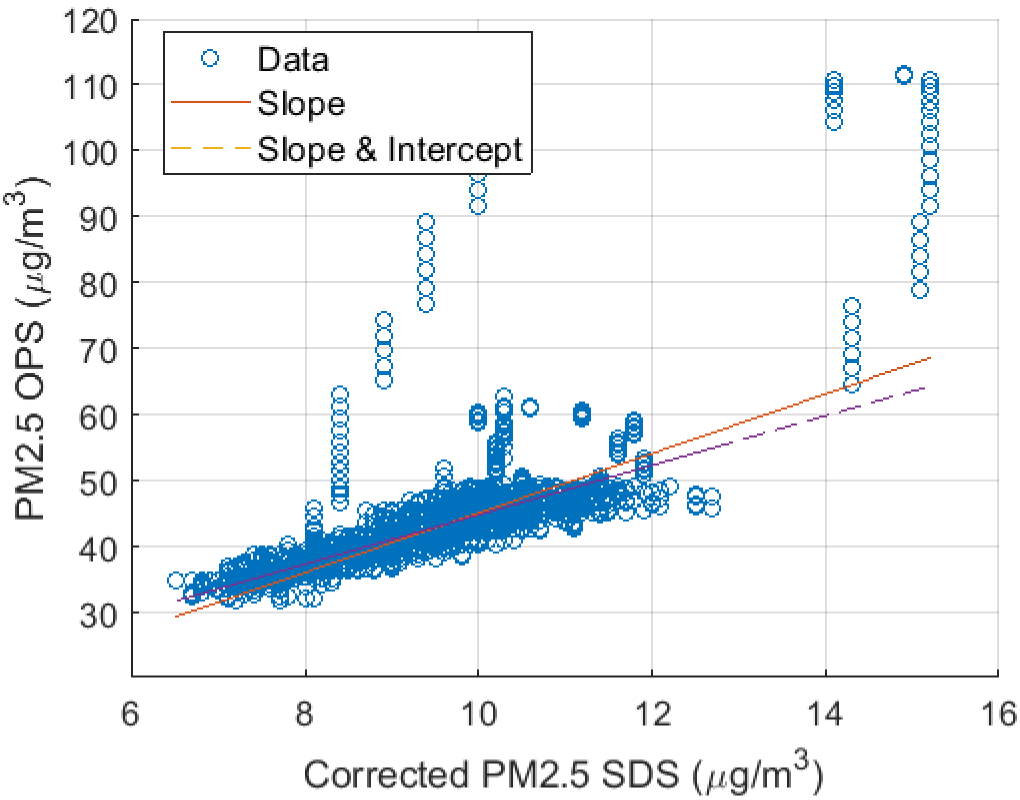


Figure 4.32: Linear Regression plots, Data set D, between the second and third spike

Table 4.9: Data Set D, ambient air with an average Temperature of 7 degrees, between the first and second spike

properties	observations
Mean relative error of SDS019	-63.49 %
Mean relative error of SDS011	-78.21 %
Mean relative error of SDS011 after applying the correction factors	0.10 %
Slope (Single Linear Regression)	4.598
r^2 (Single Linear Regression)	-36.97 %!
Slope (Single Linear Regression with intercept)	1.891
Intercept (Single Linear Regression with intercept)	28.165
r^2 (Single Linear Regression with intercept)	35.03 %
Intercept (Multiple Linear Regression)	41.96
PM2.5 factor (Multiple Linear Regression)	1.89
RH factor (Multiple Linear Regression)	-0.04
Temperature factor (Multiple Linear Regression)	-1.53
r^2 (Multiple Linear Regression)	35.43 %

4.7 Data set E (Femman air quality station)

SDS011 measurements for 3 weeks in parallel with TEOM at Femman air quality station is shown with the RH and T data in Figures 4.33. SDS011 shows an inconsistent performance about humidity. There are examples such as 3/2/2020 and 12/2/2020 where the humidity is below 80 % and SDS011 does not correlate with TEOM. On the other hand, there are times such as 8/2/2020 and 15/2/2020 where SDS011 shows a higher value due to high humidity (above 80 %) and correlates very well with TEOM. This is in contrary with the product specification of the SDS011 where it states that SDS011 is less likely to have humidity based errors at RH below 70 %. Howeveer, du to the inconsistency of the behaviour of SDS011 in this data set, the data set is not very conclusive and not so many statistical and regression analyses can be performed on this data set.

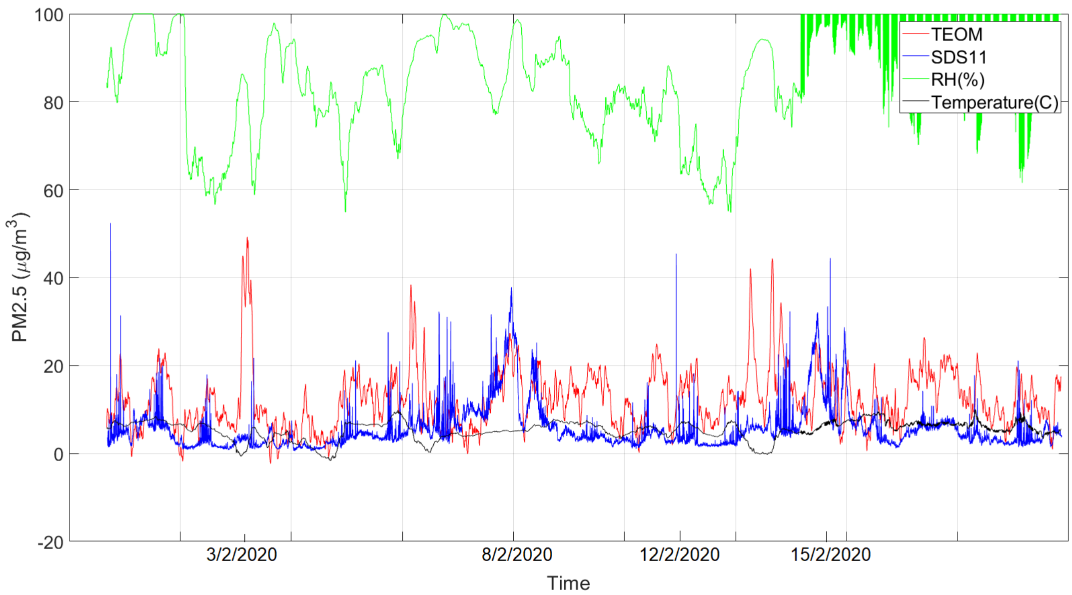


Figure 4.33: SDS011 measurements for 3 weeks in parallel with TEOM at Femman air quality station

An interesting observation for the same data set is seen in Figure 4.34, which shows the absolute error of SDS011 from TEOM compared to RH. The interesting finding here is that the absolute value of the relative humidity being above 80 % for a certain moment is not necessarily the cause of the error for such sensors, but the fluctuations in relative humidity correlates very well with the absolute error.

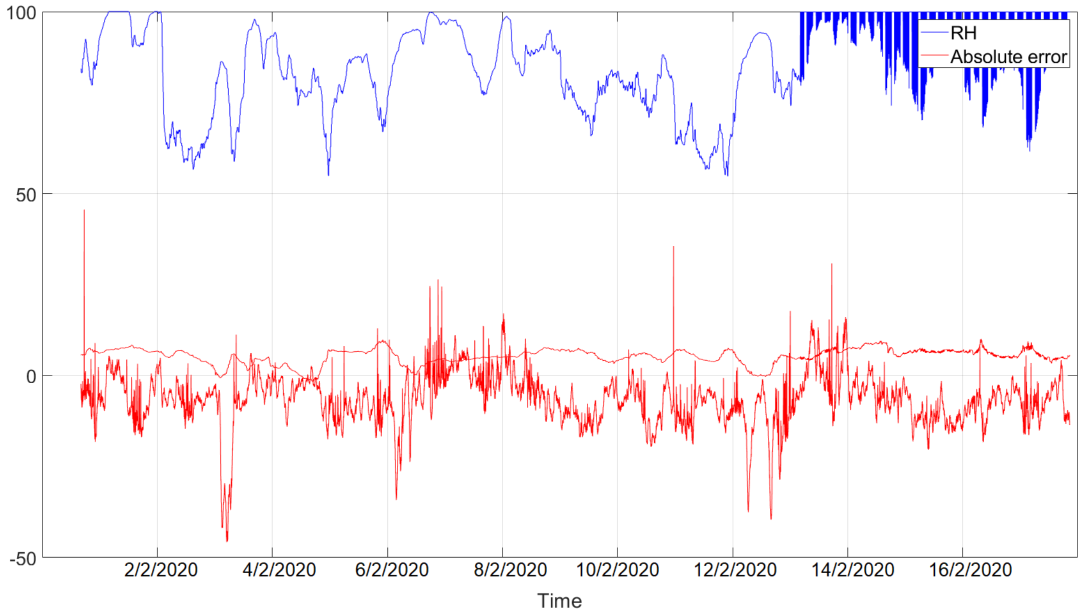


Figure 4.34: The absolute error and RH for SDS011 measurements for 3 weeks in parallel with TEOM at Femman air quality station

Figure 4.35 shows the OPS measurements for 2 days in parallel with TEOM and SDS011 at Femman air quality station. There are big spikes in the middle of the Figure where all of the three sensors react to something and show very high values. That interval cannot be physically explained. Excluding the spiky interval shows an absolute agreement between the sensors. However, the performance of OPS shows that it may not be as reliable as TEOM as a reference deice.

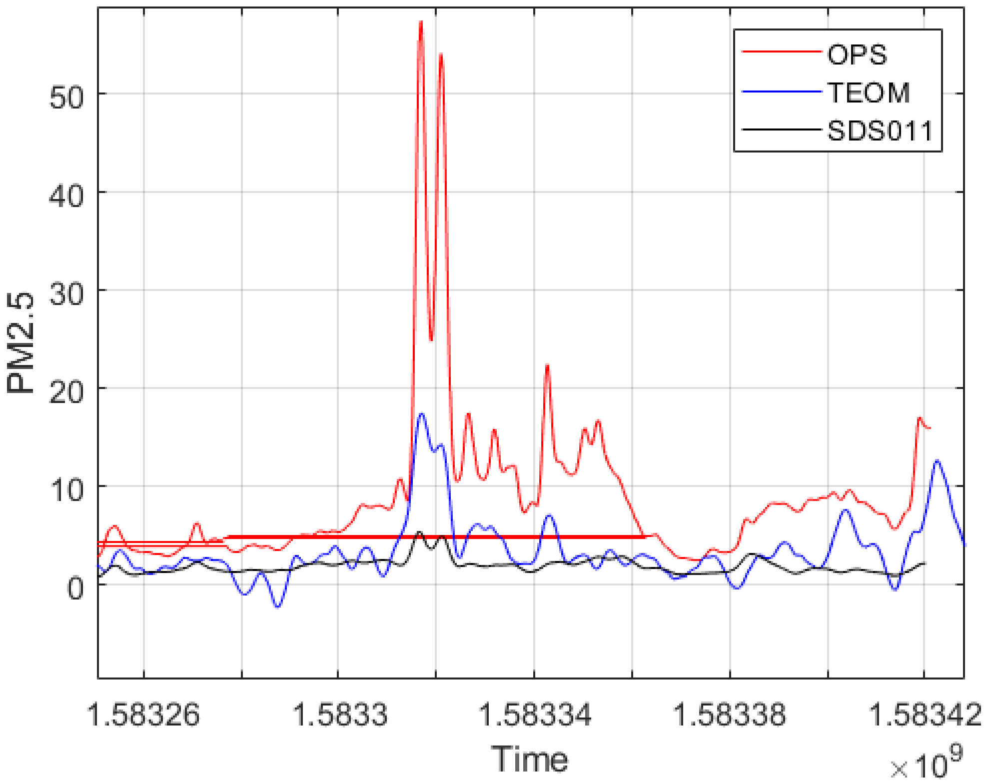


Figure 4.35: OPS measurements for 2 days (48 hours) in parallel with TEOM and SDS011 at Femman air quality station

4.8 Analysis of the results and Discussions

Looking at all data sets at concentration plots, implies the general impression that there is reasonable correlation between the SDS011 and the OPS in most cases. However, in terms of absolute values, OPS showed higher values in majority of the time. This is in contrary with the expectations based on section 2.3. This can be explained by the fact that OPS has higher counting yield in the particles smaller than 0.5 μm, and the effect of this phenomenon overcomes the exaggeration of the concentrations due to humidity by SDS sensors. It should also be noted that after completion of this work it was found that the OPS had a dirty inlet that possibly caused abstraction and lower flow, and after cleaning it increased the values by 58 %. Such an effects is consistent with a 37 % lower flow. Whether this was the case during my field and laboratory tests is uncertain but this will exaggerate the differences between OPS and SDS. Hence my conclusion are still valid. In addition the refractive index used by the OPS was the one of black carbon, but testing showed that this had very small impact on the retrieved results.

The same plots show a precise agreement between SDS011 and SDS019. The laboratory experiments reveal that SDS019 performs slightly better than SDS011 with an average relative error (from OPS) of 10 to 20 % lower than that of SDS011. The relative error plots show that the relative error for SDS011 ranges from 50 to 80 % in different data sets and this error can be improved and reduced down to below 10 % and in some cases below 1 % by applying the single linear regression correction factor to the laboratory data. Comparing the r-squared values of derived correction factors based on multiple linear regression with single linear regression did not show a significant improvement, and it turned out that taking into account the temperature and relative humidity as input parameters into a mathematical model without considering their physical aspects is not very useful. However, measuring and knowing the temperature and other meteorological parameters besides the PM concentrations may be useful in explaining the measurement artifacts in some cases. The plots for corrected SDS011 based on multiple linear regression compared to OPS are also presented. By looking at the last row of results tables it can be inferred that the multiple mode can provide a really good fit with r-squared values of up to 97 % for he laboratory experiments whereas in the case of field experiment for data set D the goodness of the fir is not as high. Thus, the performance of the multiple model for correction factor is not necessarily better and it differs case by case. However, the comparison of linear and multiple linear regression for SDS sensors showed that the error caused by high humidity cannot be fully corrected by taking into account the humidity as an input parameter to the model. Thew reason for this is that OPS, SDS have fundamental differences with TEOM, larger than the ones caused by the humidity only.

The most relevant comparison in this study is probably the one in data set E (Figure 4.33) comparing the SDS011 against TEOM (the golden standard for PM in air quality monitoring). This comparison shows periods with both good and poor agreement which partly seems related to humidity but it is hard to totally understand. The results hence indicate that the SDS011 is not performing well enough as a single monitor for air quality. When it comes to these kind of measurements there are several things that could possibly go wrong and cause errors. From electric wires and connections being disconnected to short circuits due to water in the field experiments as well as fluctuations in voltage and internet connection all of which can harm the instruments as well as inconsistent logging. There are also systematic and physical souses of error such as effect of humidity which was mentioned in section 2.3. Additionally sometimes there are some measurement artifact which cannot be explained such as field experiment at the air quality station illustrated in Figures 4.35. To eliminate the errors caused by humidity, some studies come up with new design trying to eliminate the moisture of the air sample before it reaches the detection unit of the sensors by heating or chemical absorbents.

In this project it has been demonstrated that the SDS011, which is designed for indoor application, can be used outdoors in a weatherproof box together with a Raspeberry Pi computer. And be operated in a network to get a sense of pollution levels in an area. As part of this project we were not fully able to corrects for humidity problems, which is reported to be a problem. Some papers suggest removing the moisture instead but this has not been explored here. The results here indicate that using the sensor alone for precise measurements must be done with caution.

Raspberry pi with wifi network as the electronics to complement these sensors is not a stable and reliable choice, because of the connection inconsistency observed during the field experiments. Instead, PyCom and 5G network is expected to be more reliable and will be tests later on as a continuation of this work.

5
Conclusion

5.1 Conclusions

Generally a good correlation between uncorrected SDS sensors and OPS was observed in the laboratory experiments. The correlation in the field experiments was poor on the other hand. Despite the correlations, a systematic differences in the absolute values was observed. This offset can be corrected with correction factors based on linear regression, However, the comparison of linear and multiple linear regression for SDS sensors showed that the error caused by high humidity cannot be fully corrected by taking into account the humidity as an input parameter to the model. Because the SDS sensors have fundamental differences with OPS and TEOM (reference) in sensing. SDS011 and SDS019 did not show a significant difference in accuracy, wherase their prices differ significantly. SDS011 is designed for indoor applications. However, with waterproof packaging similar to what was design throughout the course of this project and also with some correction factors, its error can be reduced and it can be used in ambient environment with a reasonable measurement quality to get a general sense of PM pollution levels, and subsequently a network of them can be used for a general sense of pollution level in a bigger area. However, it is not recommended to use them alone as an indicator of the air quality.

5.2 Recommendations for future Studies

At the end of this work, 5 sensors are permanently installed as stationary units (SDS011) in 5 bus stations in Gothenburg reporting data in real-time by IoT. additionally, one mobile unit (SDS019 and SDS011) is installed on a bus from public transport which visits these 5 stations frequently and the data is compared and calibrated. One SDS011 will be also installed at Nordstan air quality station besides the TEOM reference instrument for calibration purposes. Additionally, the frameworks and codes for data analysis as well as the prototyping and design experience for hardware and sensor packages are documented and available for future work.

To improve the correction factors, it is recommended to perform more experiments under more diverse circumstances such as temperature and humidity and background concentration levels. These experiments shall be a combination of laboratory and field experiments with different particles sources in laboratory other than ammonium sulfate to cover a greater range of particle mixtures. For field installations raspberry pi with wifi network as the electronics to complement these sensors is not a stable and reliable choice. Instead, PyCom and 5G network is expected to be more reliable.

References

[1] Kinney, P. L. Climate change, air quality, and human health. *American journal of preventive medicine*. 2008;35(5):459–467.

[2] Kim, K.-H., Kabir, E., Kabir, S. A review on the human health impact of airborne particulate matter. *Environment international*. 2015;74:136–143.

[3] Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet*. 2012;380(9859):2224–2260.

[4] Organization, W. H. et al. Health effects of particulate matter. *Policy implications for countries in eastern Europe, Caucasus and central Asia*. 2013.

[5] Davidson, C. I., Phalen, R. F., Solomon, P. A. Airborne particulate matter and human health: a review. *Aerosol Science and Technology*. 2005;39(8):737–749.

[6] Atkinson, R. W., Fuller, G. W., Anderson, H. R., Harrison, R. M., Armstrong, B. Urban ambient particle metrics and health: a time-series analysis. *Epidemiology*. 2010:501–511.

[7] Cadelis, G., Tourres, R., Molinie, J. Short-term effects of the particulate pollutants contained in Saharan dust on the visits of children to the emergency department due to asthmatic conditions in Guadeloupe (French Archipelago of the Caribbean). *PloS one*. 2014;9(3).

[8] Fang, Y., Naik, V., Horowitz, L., Mauzerall, D. L. Air pollution and associated human mortality: the role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present. *Atmospheric Chemistry and Physics*. 2013.

[9] Loomis, D., Huang, W., Chen, G. The International Agency for Research on Cancer (IARC) evaluation of the carcinogenicity of outdoor air pollution: focus on China. *Chinese journal of cancer*. 2014;33(4):189.

[10] Esworthy, R. Air quality: EPA's 2013 changes to the particulate matter (PM) standard. Library of Congress, Congressional Research Service. 2013.

[11] Quality, A. EPA's 2013 Changes to the Particulate Matter (PM) Standard. *URL: https://fas. org/sgp/crs/misc*;42934.

[12] Slezáková, K., Morais, S., Pereira, M. Indoor Air Pollutants: Relevant Aspects and Health Impacts. Feb. 2012. ISBN: 978-953-307-854-0. DOI: 10.5772/30364.

[13] Cheung, K., Daher, N., Kam, W., Shafer, M. M., Ning, Z., Schauer, J. J., et al. Spatial and temporal variation of chemical composition and mass closure of ambient coarse particulate matter (PM10–2.5) in the Los Angeles area. *Atmospheric Environment*. 2011;45(16):2651–2662.

[14] Srimuruganandam, B., Nagendra, S. S. Source characterization of PM10 and PM2. 5 mass using a chemical mass balance model at urban roadside. *Science of the total environment*. 2012;433:8–19.

[15] Oberdörster, G., Celein, R. M., Ferin, J., Weiss, B. Association of particulate air pollution and acute mortality: involvement of ultrafine particles? *Inhalation toxicology*. 1995;7(1):111–124.

[16] Wilson, W., Chow, J. C., Claiborn, C., Fusheng, W., Engelbrecht, J., Watson, J. G. Monitoring of particulate matter outdoors. *Chemosphere*. 2002;49(9):1009–1043.

[17] Slezakova, K., Morais, S., Carmo Pereira, M. do. Indoor air pollutants: relevant aspects and health impacts. *Environmental Health—Emerging Issues and Practice*. 2012.

[18] Organization, W. H. et al. *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment*. Tech. rep. Geneva: World Health Organization, 2006.

[19] Chow, J. C. Measurement methods to determine compliance with ambient air quality standards for suspended particles. *Journal of the Air & Waste Management Association*. 1995;45(5):320–382.

[20] Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S., et al. *The changing paradigm of air pollution monitoring*. 2013.

[21] Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., et al. Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? *Environment international*. 2018;116:286–299.

[22] Kopetz, H. Internet of things. *Real-time systems*. Springer; 2011, pp. 307–323.

[23] Suresh, P., Daniel, J. V., Parthasarathy, V., Aswathy, R. A state of the art review on the Internet of Things (IoT) history, technology and fields of deployment. *2014 International conference on science engineering and management research (ICSEMR)*. IEEE. 2014, pp. 1–8.

[24] Singh, S., Singh, P. Key concepts and network architecture for 5G mobile technology. Vol. 1. 5. 2012, pp. 165–170.

[25] Vlacheas, P., Giaffreda, R., Stavroulaki, V., Kelaidonis, D., Foteinos, V., Poulios, G., et al. Enabling smart cities through a cognitive management framework for the internet of things. *IEEE communications magazine*. 2013;51(6):102–111.

[26] Jin, J., Gubbi, J., Marusic, S., Palaniswami, M. An information framework for creating a smart city through internet of things. *IEEE Internet of Things journal*. 2014;1(2):112–121.

[27] Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M. Internet of things for smart cities, IEEE Internet Things J. 1 (1)(2014) 22–32. *Google Scholar Google Scholar Cross Ref Cross Ref*. 2014.

[28] Zhu, C., Leung, V. C., Shu, L., Ngai, E. C.-H. Green internet of things for smart world. *IEEE access*. 2015;3:2151–2162.

[29] Ijaz, A., Zhang, L., Grau, M., Mohamed, A., Vural, S., Quddus, A. U., et al. Enabling massive IoT in 5G and beyond systems: PHY radio frame design considerations. *IEEE Access*. 2016;4:3322–3339.

[30] Bibri, S. E. The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. *Sustainable cities and society*. 2018;38:230–253.

[31] Kaivonen, S., Ngai, E. C.-H. Real-time air pollution monitoring with sensors on city bus. *Digital Communications and Networks*. 2020;6(1):23–30.

[32] Amaral, S. S., De Carvalho, J. A., Costa, M. A. M., Pinheiro, C. An overview of particulate matter measurement instruments. *Atmosphere*. 2015;6(9):1327–1345.

[33] Kulkarni, P., Baron, P. A., Willeke, K. Introduction to aerosol characterization. *Aerosol measurement: principles, techniques, and applications*. 2011;3.

[34] Han, H.-S., Sreenath, A., Birkeland, N. T., Chancellor, G. J. Method for Combining Electrical Mobility and Optical Size Distributions for Wide Range Particle Size Distribution Measurement. *The 7th Asian Aerosol Conference*. 2011.

[35] Giechaskiel, B., Maricq, M., Ntziachristos, L., Dardiotis, C., Wang, X., Axmann, H., et al. Review of motor vehicle particulate emissions sampling and measurement: From smoke and filter mass to particle number. *Journal of Aerosol Science*. 2014;67:48–86.

[36] Welker, R. W. Size analysis and identification of particles. *Developments in Surface Contamination and Cleaning*. Elsevier; 2012, pp. 179–213.

[37] Shao, W., Zhang, H., Zhou, H. Fine particle sensor based on multi-angle light scattering and data fusion. *Sensors*. 2017;17(5):1033.

[38] Baynard, T., Garland, R., Ravishankara, A., Tolbert, M., Lovejoy, E. Key factors influencing the relative humidity dependence of aerosol light scattering. *Geophysical research letters*. 2006;33(6).

[39] Donateo, A., Contini, D., Belosi, F. Real time measurements of PM2. 5 concentrations and vertical turbulent fluxes using an optical detector. *Atmospheric Environment*. 2006;40(7):1346–1360.

[40] Sioutas, C., Kim, S., Chang, M., Terrell, L. L., Gong Jr, H. Field evaluation of a modified DataRAM MIE scattering monitor for real-time PM2. 5 mass concentration measurements. *Atmospheric Environment*. 2000;34(28):4829–4838.

[41] Chakrabarti, B., Fine, P. M., Delfino, R., Sioutas, C. Performance evaluation of the active-flow personal DataRAM PM2. 5 mass monitor (Thermo Anderson pDR-1200) designed for continuous personal exposure measurements. *Atmospheric Environment*. 2004;38(20):3329–3340.

[42] Karagulian, F., Belis, C. A., Lagler, F., Barbiere, M., Gerboles, M. Evaluation of a portable nephelometer against the Tapered Element Oscillating Microbalance method for monitoring PM 2.5. *Journal of Environmental Monitoring*. 2012;14(8):2145–2153.

[43] Cropper, P. M., Hansen, J. C., Eatough, D. J. Measurement of light scattering in an urban area with a nephelometer and PM2. 5 FDMS TEOM monitor: Accounting for the effect of water. *Journal of the Air & Waste Management Association*. 2013;63(9):1004–1011.

[44] Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., Morawska, L. The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog. *Atmospheric Measurement Techniques*. 2018;11(8):4883–4890.

[45] Budde, M., Schwarz, A. D., Müller, T., Laquai, B., Streibl, N., Schindler, G., et al. Potential and limitations of the low-cost SDS011 particle sensor for monitoring urban air quality. *ProScience*. 2018;5:6–12.

[46] Liu, P., Zhao, C., Göbel, T., Hallbauer, E., Nowak, A., Ran, L., et al. Hygroscopic properties of aerosol particles at high relative humidity and their diurnal variations in the North China Plain. 2011.

[47] Petters, M., Kreidenweis, S. A single parameter representation of hygroscopic growth and cloud condensation nucleus activity. 2007.

[48] Chen, J.-P. Theory of deliquescence and modified Köhler curves. *Journal of the atmospheric sciences*. 1994;51(23):3505–3516.

[49] Petters, M., Kreidenweis, S. A single parameter representation of hygroscopic growth and cloud condensation nucleus activity? Part 2: Including solubility. 2008.

[50] Manikonda, A., Zıková, N., Hopke, P. K., Ferro, A. R. Laboratory assessment of low-cost PM monitors. *Journal of Aerosol Science*. 2016;102:29–40.

[51] Stoica, P., Moses, R. L., et al. Spectral analysis of signals. 2005.

A

Maximum values for size channel of OPS, Field vs. laboratory data

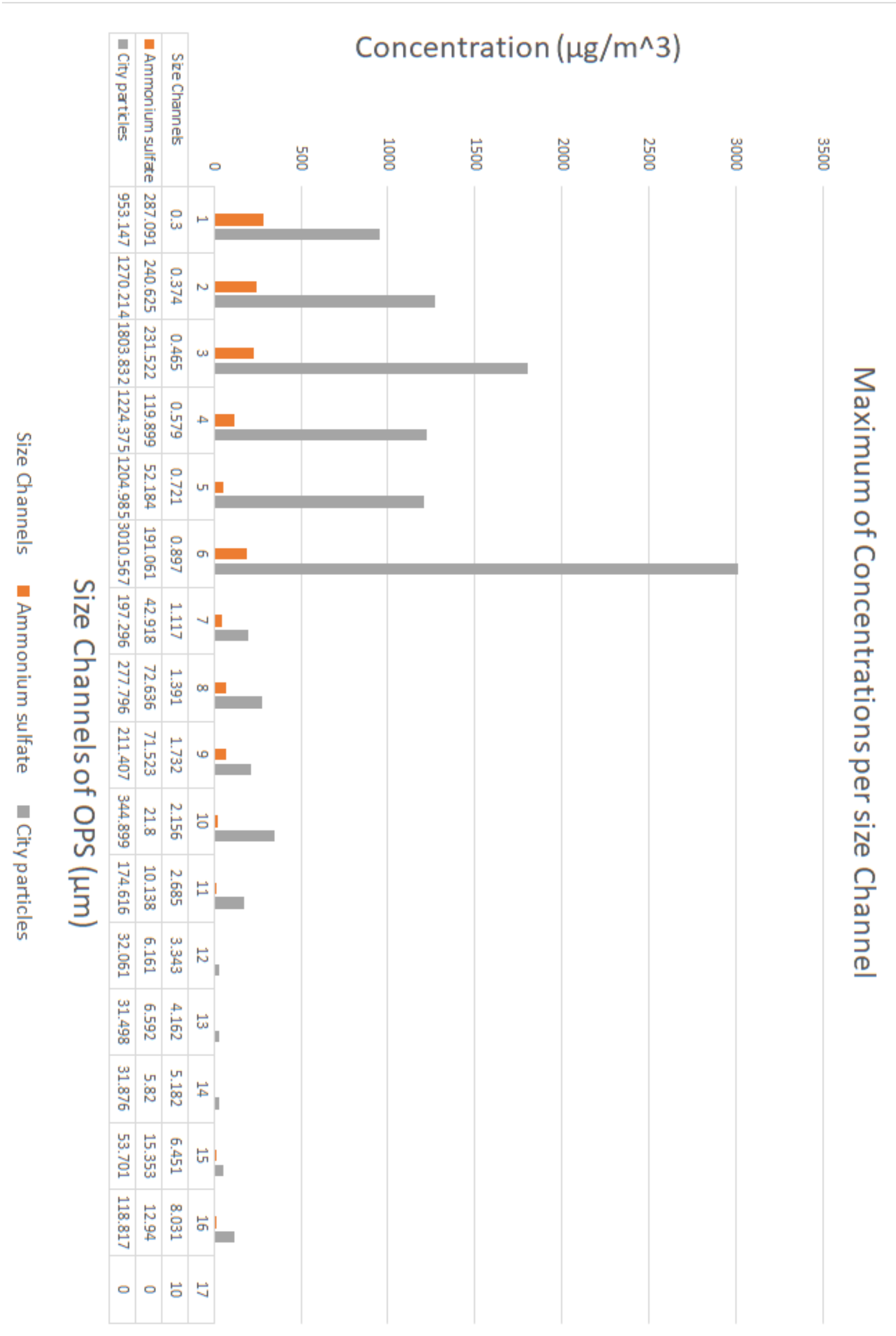


Figure A.1: Maximum values for size channel of OPS, Field vs. laboratory data .