UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACOLTÀ DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA MECCANICA

Sviluppo di un banco di calibrazione per sensori di misura della qualità dell'aria

Development of a calibration bench for air quality measurement sensors

Tesi di Laurea di:

Gianluca Sartini

Relatrice:

Prof. Ing. Sara Casaccia

Correlatori:

Dott. Ing. Nicole Morresi

Prof. Ing. Gian Marco Revel

Anno Accedemico 2023/2024

Acknowledgements

I’d like to thank professors Sara Casaccia, Nicole Morresi and Gian Marco Revel, who gave me the opportunity to carry out the internship on a topic that today plays a fundamental role in the community, for supporting me and for the patience that they had.

I’d also like to thank my family and my friends for always believing in me, I couldn't have done it without your support.

Gianluca Sartini

Abstract

The aim of this project is to develop a calibration bench for low cost particulate matter (PM) sensors and the development of a mathematical model that could be used to monitor the PM concentration. The main disadvantage of this technology is the low resolution offered, so we developed a chamber and an experimental procedure to test their performances. An atomizer was used to inject particulate matter into our chamber, the testing included the employment of different injection times (1, 5, 10, 20 and 30 seconds) using the same experimental parameters to verify if a diverse operating time resulted in a significant difference in the particle concentrations recorded. The analysis explained that a different injection time does not correspond to a significant variation in the concentrations recorded and the highest concentrations recorded were those arisen from an intermediate injection time (10 and 20 seconds), also with the lowest standard deviations recorded, between 2 and 3 $\mu g/m^3$ considering all experiments with 10 seconds. The mathematical model developed, exploiting Python and the SciPy curve-fit module, permitted the plotting of fitted curves against the experimental curves to understand how much the model comes close to the real-life testing results. The points from the model's fitted curve recorded standard deviations of the residuals between 3.6 and 5.8 $\mu g/m^3$, data analyzable as a good but not great prediction of the actual experimental curves, showing promises for a future development that could enhance the model to get closer to reality. Another possibility given by the model is, using the inverse equation, the calculation of the time needed to reach a certain fixed concentration inside the experimental chamber, giving the chance to estimate the concentration inside the box at any given point of the tests.

Index

List of figures

List of tables

1. Introduction

In the last decades the discussion around the importance of air quality has steadily assumed a bigger role, growing together with environmental awareness. To give some numbers that explain the magnitude of the topic, the European Union estimates that in Europe, every year, approximately 300000 deaths may be linked to air pollution [1]. Since the '80s the discoveries made on the dangers on human beings related to foul-up pushed the EU to implement strict measures to reduce air pollution, currently there are two directives, 2004/107/CE and 2008/50/CE. Directive 2008/50/CE introduced additional PM2.5 objectives targeting the exposure of the population to fine particles [2]. These objectives are set at national level and are based on the average exposure indicator (AEI). This is determined as a 3-year running annual mean PM2.5 concentration averaged over the selected monitoring stations in agglomerations and larger urban areas, set in urban locations to best assess the PM2.5 exposure of the population. The directive also sets the concentration level's limit for PM2.5, 20 $\mu g/m^3$ average in one year since 2020, with no permitted exceedances. For PM10, the limit is 50 $\mu g/m^3$ average in 24 hours, with 35 possible exceedances each year, and 40 $\mu g/m^3$ average in one year, with no permitted overcomings. There are some discussions to make the directives more restrictive from 2030 as an intermediate step towards achieving zero pollution goals by 2050 [3]. In the United States, the Environmental Protection Agency regulates the matter, but they differentiate the standards (as represented in the NAAQS table, National Ambient Air Quality Standards) in primary and secondary standards [4]. Primary standards provide public health protection, including protecting the health of "sensitive" populations such as asthmatics, children, and the elderly. Secondary standards provide public welfare protection, including protection against decreased visibility and damage to animals, crops, vegetation, and buildings. For PM2.5 the primary standard on a yearly average is 9 $\mu g/m^3$, the secondary is 15 $\mu g/m^3$, the standard average on a day is the same for both categories, 35 $\mu g/m^3$. For PM10 levels the standard is only for 24 hours, with no category differentiation, 150 $\mu g/m^3$. EPA has also developed an Air Quality Index (AQI) that is used to report air quality [5]. This AQI is divided into six categories indicating increasing levels of health concern. An AQI value over 300 represents hazardous air quality and below 50 the air quality is good. For the good category the limit of PM10 is 54 $\mu g/m^3$ over a day, thus indicating that the 150 $\mu g/m^3$ standard isn't very restrictive. The main air pollutants are: particulate matter (PM10 and PM2.5), nitrogen oxides (NO_x), sulfur dioxide (SO_2), carbon monoxide (CO) and carbon dioxide (CO_2). Having described the landscape and the increasingly restrictive requirements requested by EU and EPA, the

importance of monitoring air quality comes by itself. The main instrument to measure and evaluate air quality has always been the monitoring stations (*fig. 1.1*), spread throughout the world. These stations constantly measure the presence of air pollutants in different points of the cities. The data collected provides an overview of air quality and allows experts to evaluate the effectiveness of environmental policies and implement corrective measures when necessary. The negative aspects tied with these stations are high costs, their bulky dimensions which make them difficult to position, and the fact that they were designed to be installed in large areas and this does not allow to use them to monitor air quality in smaller areas, for example inside buildings or in even more restricted environments.

Fig. 1.1 – Air quality monitoring station

To overcome these limits, low-cost sensors (LCS) (*fig. 1.2*) are increasingly becoming popular and the technology behind them is getting better and more reliable. LCS have small dimensions, offering advantages for installation and can be used to create mobile monitoring stations. These sensors' shortcoming is the measurement uncertainty, due both to the quality of the sensors but also to the effect of parameters such as temperature and relative humidity that act as disturbing elements. The purpose of this project was to estimate the performances of LCS developing a calibration bench for particulate sensors, and the main objective was the prototyping and the development of a mathematical model that could assist us to measure the particle concentrations in order to get mathematical curves obtained from the tests conducted with the aim of estimating the performances of the sensors tested, since it's a relatively recent topic, there aren't many standards set so we tried to work out an experimental procedure that could help us along the way.

Fig. 1.2 – A low cost ethanol sensor for alcohol detection

This project was developed within the scope of the Ecosystem of Innovation VITALITY (*fig. 1.3*), which is supported by the National Recovery and Resilience Plan (NRRP). VITALITY aims to implement innovative solutions to increase the coverage of services, improve the quality, safety and sustainability of production systems, public administrations and living and working conditions [6].

Fig. 1.3 - The Vitality logo

1.1 State of the art

With the goal of knowing the state of the art and if similar topics were already discussed we started from a literature review, and there are many articles that deal with the topic. A study [7] tried to settle the best solution for the chamber in which to carry out the tests using low-cost particulate matter sensors, using CFD modelling and experimental measurements. Their recommended design for the ventilated chamber: parallel fan layout, fan pressure of 12 Pa, and inlet air flow rate between 10 and 15 L/min. The size of test chamber is 0.61 × 0.61 × 1.22 m^3. In the chamber (*fig. 1.4*), two fans are located at the top corners of the chamber to mix air and particle flows. The control system provides a stable air flow, constant room temperature (23 ± 0.5 °C) and

humidity (50 ± 3%) for the experiment. A humidifier was used inside the chamber to maintain the relative humidity (RH) to a desired value. The aerosol generation unit (to generate particulate matter) contains pressure regulator, atomizer, diffusion dryer, and MFC (mass flow controller).

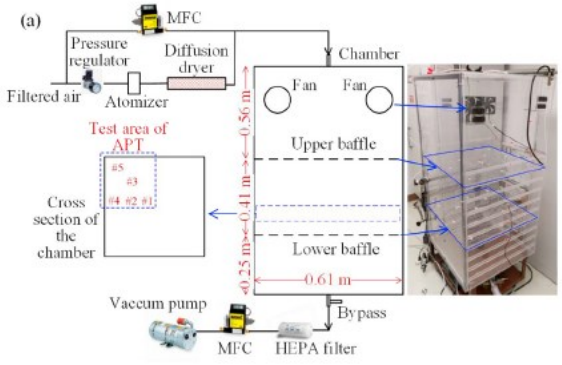

Fig. 1.4 – The experimental chamber

Another paper [8] deals with the same topic of developing a low-cost calibration chamber to run the experiments using low-cost sensors, they tried a different approach since their chamber (*fig. 1.5*) doesn't have the possibility to control temperature and relative humidity. Two different types of particles, alumina oxide and ammonium nitrate, were utilized for the calibration of the PM low-cost sensors. Alumina oxide represented a dry and non-cohesive dust, and this was generated by a Palas solid particle dispenser (RBG 1000-C). The ammonium nitrate particles were generated with a single-jet atomizer (TSI model 9302) and clean dry air. A compressor provided the clean, dry air, which was further cleaned using an Enmet air filtration panel that included a water collector bowl, a coalescer filter, a charcoal absorber, and an adsorber filter. Before introducing the aerosol flow to the calibration chamber, a diffusion dryer (TSI 3062) filled with dry silica gel reduced the humidity of the aerosol stream. In this study they developed and evaluated a cost-effective calibration chamber that provided spatially uniform concentrations to multiple PM sensors. The design and evaluation relied on a CFD model and a rigorous experimental evaluation. The computational model showed that the chamber can provide a uniform PM concentration to calibrate eight sensors at one time within 6% error, and the experimental results demonstrated the robustness of the calibration model to the sensor position within the chamber with excellent reliability (ICC > 0.771).

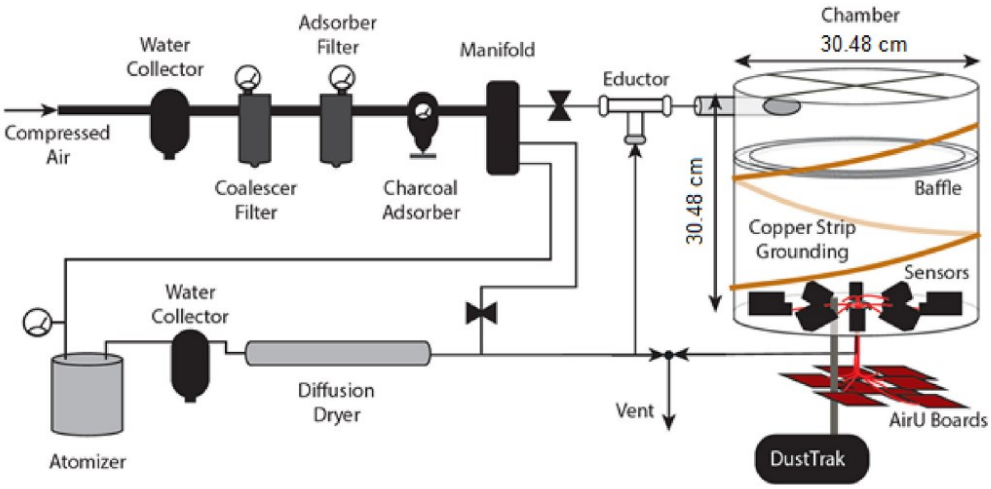

Fig. 1.5 – Schematic of the calibration system

Another article [9] tried to demonstrate a system for performance evaluation and calibration of low-cost gas sensors applied to air quality monitoring, they calibrated CO, SO_2 and O_3 sensors to install them in a drone, thus making it a flying air quality analyzer. For the calibration (*fig. 1.6*) they used a climatic chamber with a range of 0-60°C. A cylindrical inner chamber was inserted inside the climatic with a volume of 19.25 L, made entirely (body and lid) of AISI 304 steel and with all NPT (national pipe thread) threads. The primary gaseous standards of carbon monoxide (CO) and sulfur dioxide (SO_2) with concentrations of 2500 ± 1.85 ppm and 20 ± 1.80 ppm, respectively, were obtained from certified commercial gas mixtures, with balance in N_2 and concentrations obtained by the gravimetric method, secondary gas standard with different concentrations of ozone were obtained using an ozone generator present in the 146IQ multicalibrator, (Thermo Scientific), which produces it by exposing the air to light at 185 nm. The generation of purified air (zero gas), used in dilution of the primary gas standards, was performed by a zero-gas generation equipment (111iQ and 1150 CO Reactor, Thermo Scientific brand), to efficiently remove NO_2, SO_2, O_3 and hydrocarbons. This equipment can operate with flow rates of up to 20 L/min. The performance evaluation system implemented allowed the evaluation of three low-cost gas sensors (CO, O_3 and SO_2) using eight performance parameters: baseline drift, linearity, precision, bias, accuracy, response time, interference from other gases pollutants and long-term drift. The proposed system showed good performance in terms of generating and quantifying gaseous patterns, with measured concentrations that are stable, repeatable, and close to those pre-stipulated.

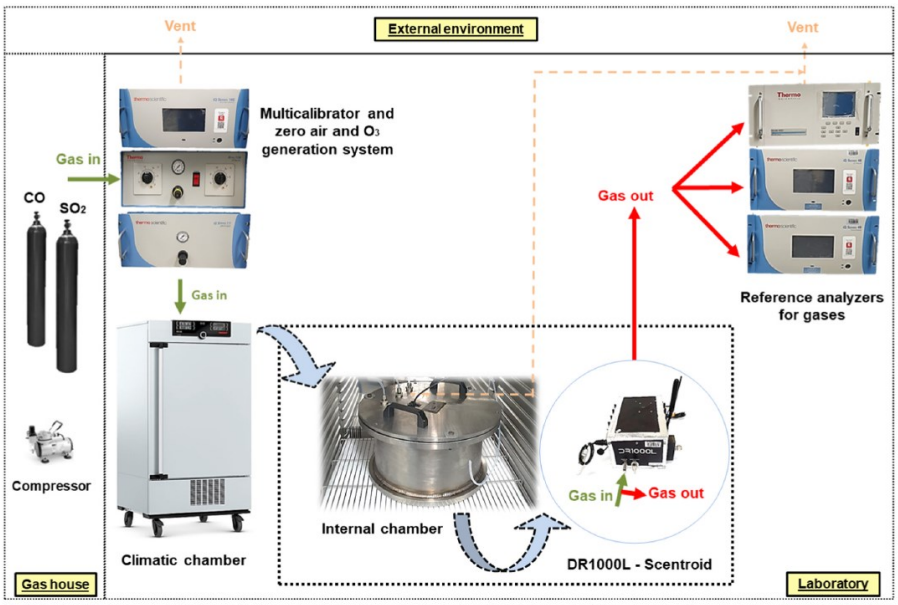

Fig. 1.6 - Schematic of the implemented portable gas sensors performance evaluation system

In another paper [10] the authors evaluated three different low cost sensors and compared their results to a standard device. For the experiment they used a cleanroom with the size of 5 × 2.4 × 2.25 m, was equipped with one mixing fan and cigarette smoke generator. The smoke generators created an environment with particulate concentrations of over 1000 µg/m^3 (PM2.5) to cover the entire measuring range of the sensors to be evaluated. This work was focused on the calibration of the sensors, all the tests performed were kept within the range of temperature of 15 ˚C - 30 ˚C and relative humidity of 50% - 85%. Extreme conditions such as very cold/high temperature and/or relative humidity may seriously reduce sensor accuracy. The one-point and two-point calibration methods are highly effective when the sensor errors are relatively linear with the standard measuring device. However, when the error between the sensor and the standard instrument is not linear, it is necessary to use the multi-point curve method to adjust the measurement results of the sensor. Another study [11] managed to find a low-cost calibration method for temperature, relative humidity and carbon dioxide sensors. The experiments were carried out inside a plastic box of 6.27 L closed with a plastic lid. These methods are based on simple techniques and procedures that allow temperature calibration in the range of 0 - 50 °C, relative humidity from 0 to 90%, and CO_2 between 0 and 1100 ppm. For these experiments, the average air T and RH were 27 °C and 64%. The calibration of the temperature parameter of the AM2315 sensors was performed in two stages: the cooling of the ambient air inside a Styrofoam box filled with smaller pieces of Styrofoam, leaving only a small volume for the sensors and the reference glass thermometer using a Peltier cooler so that the temperature decreases from room (26 °C) down to 0 °C. The heating of ambient air in a

closed plastic box by placing a mica heater inside the box that is coupled to a thermostat so that the temperature rises from room T to 50 °C, six solutions with glycerol concentrations of 99, 80, 70, 50, 40, and 0 vol% were prepared to perform a series of experiments with increasing RH. Before the calibration experiments started, the CO_2 concentration in ambient air and inside the closed box was measured. The box was also coupled to a closed circuit (*fig. 1.7*) that could remove the CO_2 in the ambient air inside the box. That circuit contained a pump, a wash bottle with 80 mL of saturated $Ca(OH)_2$ solution plus 20 g of NaOH beads that removed CO_2 from the air according to reaction, an empty wash bottle to trap the water droplets present in air, and a U-shaped glass tube containing 10 g of silica gel to dry the air. The experiments showed that factory calibrations can be improved and differences between sensors can be compensated. The field campaigns demonstrated that the calibrated low-cost monitoring systems could be used as a low-cost scientific instrument.

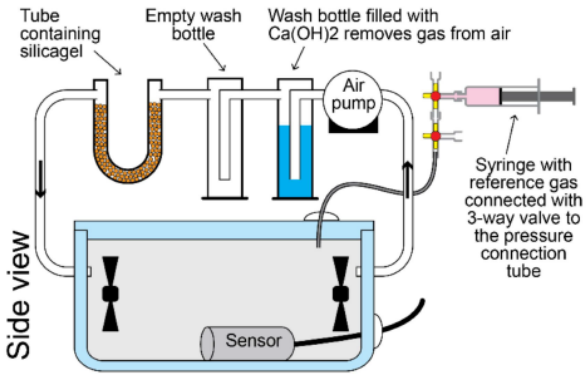

Fig. 1.7 - Setup to calibrate a CO2 sensor and the circuit to remove ambient CO2 inside the box

In another paper [12] the authors developed a state-of-the-art laboratory chamber (*fig. 1.8*) for evaluating the performance of low-cost air quality sensors under controlled conditions, a 1.3 m^3 stainless-steel outer chamber and a 0.11 m^3 Teflon-coated stainless-steel inner chamber are used to create controlled aerosol and gaseous atmospheres, respectively. In a typical fine particle (PM2.5) experiment testing a PM sensor, the outer chamber is first conditioned to a target temperature and RH. Aqueous potassium chloride solution (KCl, 17% by weight) is nebulized in the aerosol generator and is subsequently dried in the PALAS dryer. Dry particles are then directed into the chamber. The chamber operates in a dynamic mode at a flowrate of about 23 L/min. On average, 2-3 h are needed to achieve a stable particle concentration and size distribution in the chamber. This procedure is repeated multiple times over multiple pollutant concentrations spanning from very low (e.g., below typical ambient conditions or less than 50% of the U.S EPA National Ambient Air Quality Standards

(NAAQS)) to very high (e.g., well above ambient conditions or more than 200% of the U.S. EPA NAAQS), and under different temperature (e.g., 5-35 °C) and RH levels (e.g., 20-80%). As suggested by the specification, the chamber system can reach a broad range of temperature and relative humidity conditions (T range: -32 °C to +177 °C; RH range: 10-95%). The chamber's capability to create stable and reproducible temperature and relative humidity conditions was demonstrated in two separate experiments. In the first experiment, the temperature set point was increased by 5 °C every 30 min from 5 to 35 °C, while maintaining relative humidity constant at 40%.

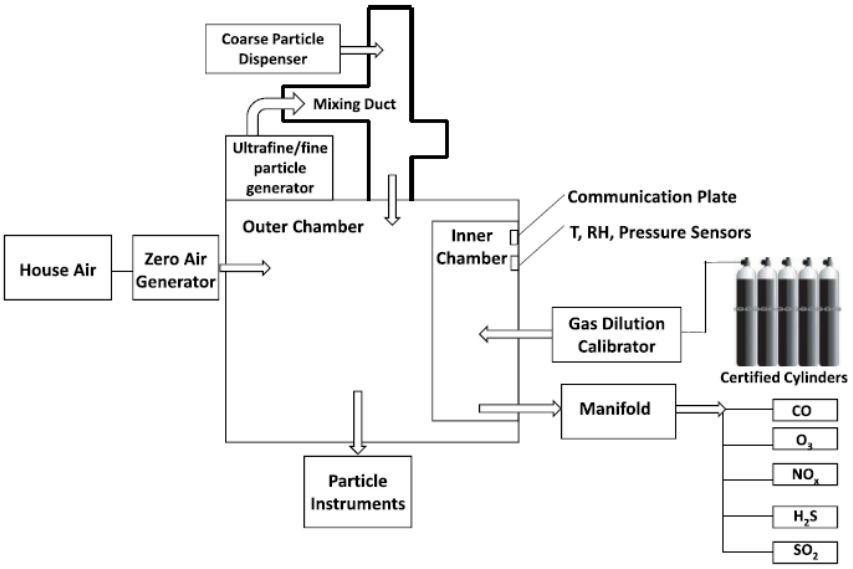

Fig. 1.8 – Schematic of the chamber system

The chamber could deliver accurate temperature control. At each step, the temperature reached the target value within 3 min from the previous set point and remained stable throughout the 30 minutes period. All experiments were carried at 20 °C. A series of experiments were conducted to demonstrate the system's capability of generating stable and reproducible aerosol and gaseous atmospheres, as well as its capability of simulating a wide range of temperature (T) and relative humidity (RH) conditions. Preliminary testing results have shown important information regarding sensor accuracy, precision, detection limit, correlation coefficient, interferents, and climate susceptibility. There some more papers that focus more on the experiments and the calibration side, treating the environmental conditions and the chamber in which the tests take place in a more marginal way, like for example [13], in which the authors compared the results obtained from six different sensors using different technologies: laser, infrared and mechanical. Three different types of aerosol generators were utilized to generate/airborne test particles. For test

particles in the sub-micrometer size range, a custom-made Collison atomizer was used to produce droplets of different mean sizes by nebulizing solutions in different salt concentrations. Three different solutions, Sodium Chloride; Methylene blue; and Fluorescein sodium, were also used to produce droplets of different composition. However, the readouts of all tested optical PM sensors are sensitive to the physical and chemical properties of test particles. The composition of test particles (i.e., refractive index) have significant effect on the readouts of optical PM sensors. The level of composition effect depends on the optical design of PM sensors. The study further demonstrated that the calibration of low-cost optical PM sensors should be performed under the steady particle concentration condition, instead of the transient particle mass concentration condition even at very slow transient conditions (i.e., under calm air condition). It might be due to the presence of dead space in the particle passages of low-cost optical PM sensors. Another article [14] proposed to find a new calibration system for low-cost Sensor Network in air pollution monitoring. The linearity of the gas sensors of SO_2, NO_2, CO, and O_3 was examined to evaluate the performance of the sensors before calibration. Three sensors for each type of gas were involved. The sensors were tested under steadily increased concentration, which was from 0 ppm to 40 ppm for CO sensors, 0 ppb to 400 ppb for NO_2, O_3, and SO_2 sensors with three more test points (at 20 ppb, 50 ppb, and 80 ppb) below 100 ppb. The gas sensors were tested under increasing temperature (from -10 °C to 40°C), which included 9 sensors for SO_2, NO_2, and CO and 8 sensors for O_3. CO and NO_2 sensors were not influenced by the increase of temperature from 15 °C to 21°C. CO sensors showed a positive relationship and NO_2 sensors showed a negative relationship with the increase of temperature from 15 °C to 21 °C. The gas sensors were also subject to the influence of relative humidity (RH). The four kinds of gas sensors were tested from 20% RH to 90% RH. In the lab test, the CO, NO_2, and O_3 sensors had good linearity before calibration, which had R2 values were very close to 0.99. Furthermore, at a constant concentration, the SO_2 and CO sensors showed a positive relationship with the increase of temperature, while the NO_2 and O_3 sensors showed a negative relationship as the temperature increased. The CO sensors were insusceptible to change of relative humidity, but the NO_2 and O_3 sensors had a positive relationship, and the SO_2 sensors had a negative relationship with the increase of relative humidity. The four types of gaseous sensors were calibrated after examining the influence of temperature and relative humidity on sensors. The CO sensors gave stable outputs when NO_2, SO_2, and O_3 standard gases were introduced to test the cross-interference effect. However, the NO_2, SO_2, and O_3 sensors experienced cross-interference in varying degrees. After the last stage of calibration in the laboratory, all four types of gaseous sensors were calibrated

again to eliminate those effects. Then, the four types of gas sensors and the PM10 and PM2.5 sensors were calibrated in the field through combined supervision calibration and transfer calibration. Except for the R2 value for CO sensors, all other sensors had their R2 values improved and slopes become much closer to 1 after calibration. Another article [15] deals with the creation of a mobile floating CO_2 monitoring system. Calibration was performed by creating a simple "leaky chamber", breathing into this chamber created an artificially raised CO_2 concentration which gradually fell, exposing the sensors to steadily lowering CO_2 concentrations. Applications for this sensor are highly varied and the base design can be augmented with additional components or code to suit the user's needs. They demonstrate application by measuring CO_2 evasion rates of peatland streams, confirming these areas as sources of CO_2 which would require extensive monitoring campaigns to capture the variability in evasion rates over space and time. The sensor design presented here is quick to assemble, accessible to novice users, and cheap. Use of a custom PCB creates a streamlined and sturdy device that is easily replicable. In another paper [16] the authors focused on the calibration of sensor network for outdoor measurement of PM2.5. Their idea is to calibrate PM sensors to use them to monitor PM values in Temuco City, Chile, since burning wood is widely valued as a primary energy source in the domestic sphere due to its low price compared to other heating alternatives. However, its massive use and inefficient combustion generate high emissions of air pollutants. In this experiment, an airtight chamber with forced airflow was prepared to generate the combustion of a specific vegetable mass inside, generating PM2.5 in the process. Regarding the calibration methods shown in this work, they proposed a nonlinear calibration model, which was compared with a polynomial calibration model. Furthermore, the nonlinear model includes five parameters instead of the nine parameters of the polynomial model. This smaller number of parameters may make the model easier to calibrate and less susceptible to overfitting. In another paper [17] the authors managed to create a wearable system for outdoor air quality monitoring. In the experiment, the calibrator generated three different concentrations (50, 110, 220 ppb) for both O_3 and NO_2 with an airflow of 350 mL/min, they have presented a portable sensing device that effectively outputs air quality data from commercially available sensors towards cloud storage. With the tests we conducted, the air quality analyzer was proven to have a consistent response when exposed to target air pollutants and a high-to-moderate correlation with referential measures in outdoor environments. The system can easily be extended to create urban networks that can monitor pollution events with improved spatial resolution. They have shown a successful wearable implementation of the new generation of MOS gas and PM sensors, with a low cost, low-

powered, and small-sized solution. The devices have been used for several months without any perceptible quality reduction, and the network's robustness allowed for easy deployment in urban environments.

All the experiments carried out on these papers display some problems, for example about the complex preparation of the cameras used, or that some instrumentation used is very expensive. Some papers deal more with outdoor air quality, others don't use PM as their pollutant. Sometimes the experimental procedure is not touched on, or the focus are the comparisons between different sensors that measure the same thing but using different technologies.

2. Materials and methods

In this chapter the topics discussed are the conceptualization of the experimental chamber, the experimental setup and procedure, the methodology used for the data analysis and the development of a mathematical model. For these reasons explained before, the aim of this work is to develop a simpler camera to calibrate low-cost sensors and looking to achieve values compatible with the ranges indicated by the Environmental Protection Agency as the standard for good quality air. Here's the breakdown of the final prototype of our experimental setup (*fig. 2.1*).

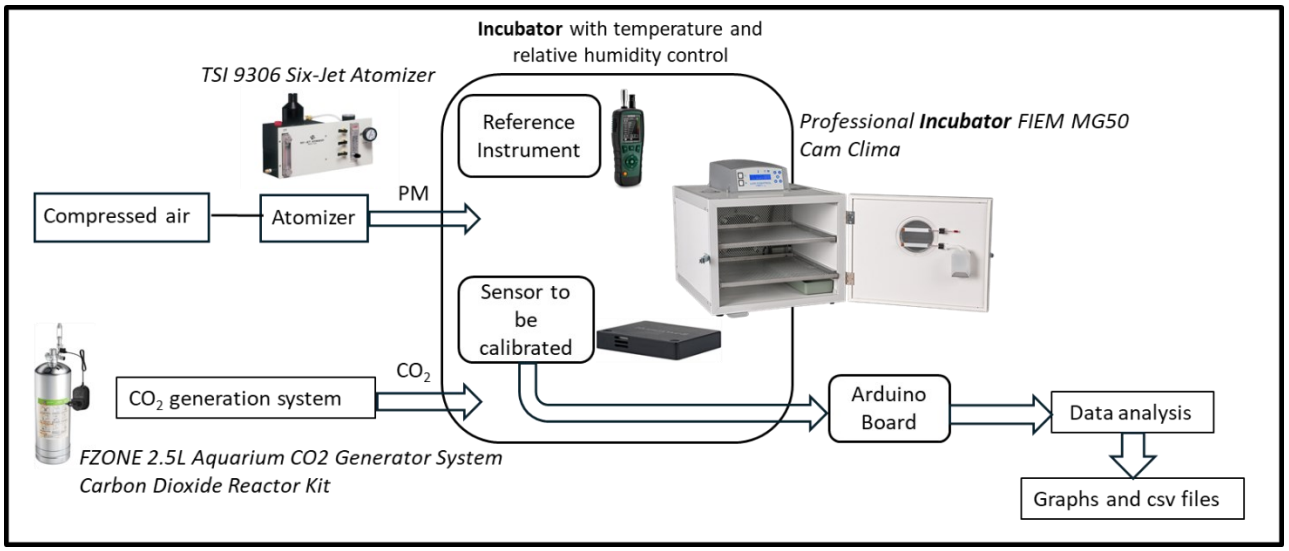

Fig. 2.1 – The flowchart of the final prototype

Table 2.1 – The reference instruments and the low-cost sensors selected

Reference Instruments	Low-cost sensors
Extech VPC260 Particle Counter	Honeywell HPMA115C0 Particle Sensor
Extech CO240 CO_2 Meter	E+E Elektronik EE895-M16HV3 CO_2 Sensor
Extech VFM200 VOC/Formaldehyde Meter	Seeed Studio 101020512 VOC and CO_2 Sensor

The atomizer, whose operation will be explained later, takes air from the compressed air system to inject particulate matter in the camera, in parallel there's a CO_2 generation system that exploits the reaction between citric acid, sodium bicarbonate and water to produce CO_2 and pumps it in the chamber. The chamber is a professional incubator that can control the temperature (up to

40 °C) and the relative humidity level inside. Inside the incubator there are reference instruments for PM, CO_2 and VOC (as shown in the table above) that act as a meter for the low cost sensors to calibrate them. The sensors are connected to an Arduino board that allows the reading of the data for the analysis. The entire system is set up on a cart for handling, thus making the system movable and transportable. Since every component had to be bought and that's a long process, we used a temporary system to start experimenting with.

2.1 Experimental setup

2.1.1 The Sensor

The selection of sensor to put in the camera was made having some fundamental needs: only LCS were considered and the capability of being programmed by the Arduino Mega 2560 board. The chosen sensor was: Plantower PMSA003I particle concentration sensor (minimum distinguishable particle diameter: 0.3 µm). The PMSA003I sensor (*fig. 2.2*) can both track and monitor particle concentration of PM1, PM2.5 and PM10. It comes with a compact format and an I2C interface so it can be used with Linux computers, such as Raspberry Pi or also with Arduino (which usually uses serial software). This breakout features two STEMMA QT connectors, no welding required. This sensor PMSA003I uses laser scattering to irradiate particles suspended in the air, after which it collects the scattered light to obtain the curve of the changes of the scattered light in time.

Table 2.2 - Data sheet for the PMSA003I sensor [18]

PARAMETER	VALUE
Particle Resolution	1 µg/m^3
Particle standard volume	0.1 L
Single response time	< 1 s
Total response time	1 s
Type of power supply	5.0 V: min 4.5 V / max 5.5 V
Working temperature range	-10 °C / 60 °C
Working humidity range	0-99 %
Storage temperature range	-40 °C / 80 °C

Dimensions	51.0 x 35.5 x 13.6 mm / 2.0” x 1.4” x 0.5”
Price	Approx. 50 euros

The diameter of the equivalent particles and the number of particles with different diameters per unit volume are then calculated by the microprocessor. The I2C data stream updates once per second providing the concentration of PM1.0, PM2.5 and PM10.0 in both standard and ambient units.

Fig. 2.2 – Plantower PMSA003I

2.1.2 The chamber

A polypropylene box (*fig. 2.3*) was picked as the chamber for the project and therefore it was subjected to some preparations.

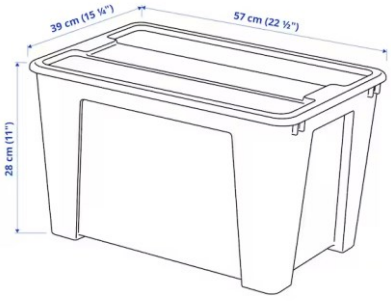

Fig. 2.3 – The box with its dimensions

Table 2.3 – Data sheet of the chamber

PARAMETER	VALUE
Material	Polypropylene
Dimensions (cm)	57 x 39 x 28
Volume	45 L
Price	Approx. 10 euros

The first preparation made was the drilling of two holes (*fig. 2.4*) on the shorter sides, a small one for the USB cable of the Arduino board, and a bigger one that was widen with a screwdriver using a tip made for countersinks. This hole was used to link the chamber with the atomizer using a rubber hose to allow the injection of particle matter into the box.

Fig. 2.4 – The position of the holes

- 1. Hole for the rubber hose that connects the atomizer and the box
- 2. Hole for the USB cable of the Arduino board

The sensor and the Arduino board were lifted, placed on a carton box, nearing them to the hole on the side, because we thought it was the best possible positioning.

2.1.3 The atomizer

For the injection of particulate matter an atomizer was used, the TSI 9306 Six-Jet Atomizer (*fig. 2.5*). An atomizer is a device that is used to emit liquid droplets as fine spray. “Atomize” means splitting up a large body into small, discrete particles. It works on Bernoulli's principle [19]. When high speed horizontal air passes over a vertical tube, it creates a low pressure and draws the air and liquid inside the vertical tube upward. An atomizer has a nozzle at the end of the horizontal tube which causes the liquid to break up into small drops and mixes it with the air.

Fig. 2.5 – TSI 9306 Six-Jet Atomizer

The TSI 9306 Six-Jet Atomizer [20] has a built-in pressure regulator and pressure gauge, as well as a self-contained dilution system. External controls allow one through six particle-generating atomizer jets to be selected, each producing particle concentrations greater than 10^7 particles/cm^3 at 6.5 L/min (nominal at 25 psig pressure), allowing the particle output to be adjusted. The system can be used to generate particles from virtually any liquid and can even produce solid particles from solutions or suspensions. Precise control of dilution air flow rate allows further control over the output concentration and helps dry the solid particles generated from a solution.

Table 2.4 – Data sheet for the TSI 9306 Six-Jet Atomizer

PARAMETER	VALUE
Number mean droplet diameter	0.35 µm for water
Particle concentration	$>10^6$ particles/cm^3
Number of jets	6
Flowrate per jet	6.5 L/min at 170 kPa (25 psig), 12 L/min at 380 kPa (55 psig)
Normal operating pressure	170 kPa (25 psig)
Maximum inlet pressure	550 kPa (80 psig)
Maximum outlet pressure	102 kPa (15 psig)
Outlet port dimensions	2.5 cm (1.0 inch)
Dimensions	32.0 x 12.5 x 26.0 mm / 12.6” x 4.9” x 10.2”
Weight	2.7 kg (1.2 lb)

The compressed air required for the operation of the atomizer was provided by the system installed in the “Mechanical and Thermal Measurement” laboratory.

2.1.4 The wiring of the sensor

The LCS sensor was programmed using an Arduino MEGA2560 [21] (*fig. 2.6*), a board based on the ATmega2560 microcontroller and equipped with 54 digital input/output pins (14 of which can be used as PWM signals), 16 analogue inputs, 4 UARTs, a 16MHz crystal, a USB connector, a power jack, a connector for ICSP programming and a button for resetting the board. The board also provides everything needed to support the operation of the microcontroller.

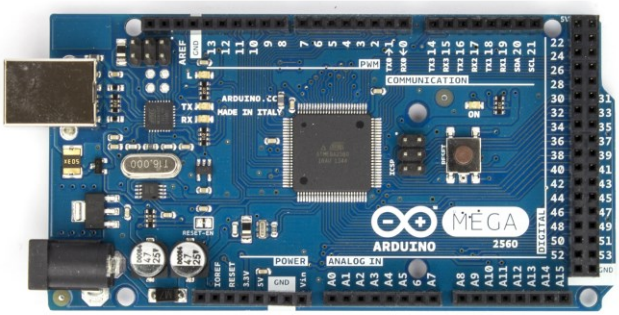

Fig. 2.6 – Arduino MEGA2560

The following steps were carried out to connect and program the PMSA300I sensor (*fig. 2.7*):

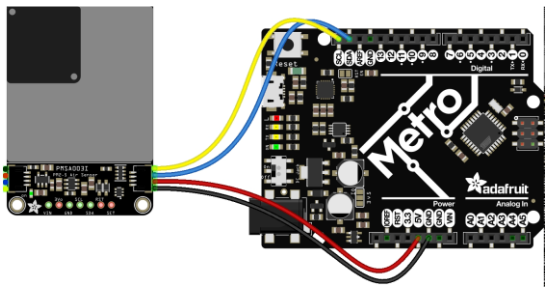

Fig. 2.7 – Connection of the PMSA300I sensor

- 1. Connect the VCC of the sensor to the 5V of the board (red wire)
- 2. Connect the GND of the sensor to the GND of the board (black wire)
- 3. Connect the SCL of the sensor to the SCL of the board (yellow wire)
- 4. Connect the SDA of the sensor to the SDA of the board (blue wire)

2.2 Experimental protocol

In this paragraph, the experiments carried out will be explained and the results achieved will be analyzed. The goal was to measure the particulate concentration inside the chamber using our low cost sensor. The atomizer was used for the injection of particulate matter, the results of different operating times of the atomizer were also compared in the tests. Once all the connections were made and the sensor was placed in position, various tests were carried out to study the behavior of the sensor.

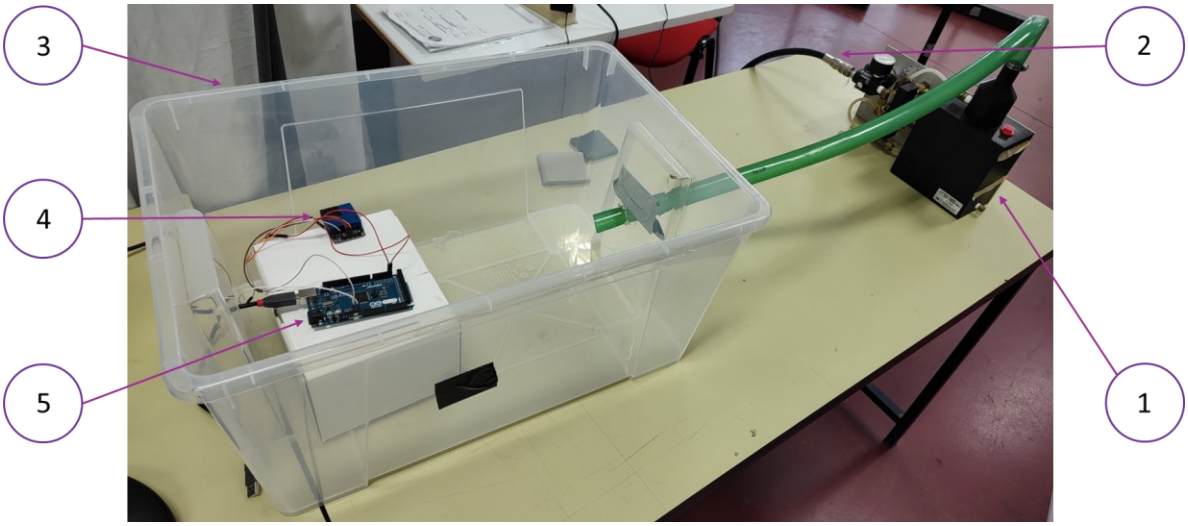

Fig. 2.8 – The experimental setup (without the closing lid of the box)

- 1. TSI 9306 Six-Jet Atomizer
- 2. Compressed air for the operation of the atomizer (system installed in the laboratory)
- 3. The polypropylene chamber
- 4. Plantower PMSA003I sensor
- 5. Arduino MEGA2560 board

The parameters of the experiments were chosen with a "trial and error" approach, basing only on the notions learned from the reference literature: the target was to get concentrations in the box compatible with the Environmental Protection Agency's range definable as the standard of good air quality (up to 54 µg/m^3 of PM10 average concentration for a day). To achieve this, a solution of water and sucrose was chosen as the working fluid of the atomizer, because the atomizer's manual suggested this to achieve particles with a size bigger than 1 µm, furthermore the possible alternative (solution of water and sodium chloride) had the disadvantage of being corrosive. The concentration of the solution was chosen after some attempts, finding out that 5% of sucrose was the sweet spot. The pressure of the atomizer was gradually lowered, during testing, from the initial 20 psi to 10 psi because the concentration levels were too high for our target. The number of atomizing jets turned on was set to 1 (of the possible 6), because it was more than adequate. The pressure of the compressed air system was already set to 4 bar, we left it at that level because there were already many variables into account, locking this one gave the opportunity to experiment more with other factors.

Table 2.5 – Parameters for the experiments

PARAMETER	VALUE
Working fluid for the atomizer	Solution of water and sucrose (5% solution)
Working pressure for the compressed air system	4 bar
Working pressure for the atomizer	10 psi (0.69 bar)
Number of atomizing jets turned on	1

2.3 Experimental procedure

The target of the experiments was to inject particulate matter into the chamber using the atomizer in order to measure the low cost sensor's response, visualized using a Python script ran by the PC connected to the Arduino board. The only variable of the experiments was the time frame of the particle injection, to point out the differences in concentration measured. Since there is no standard for these tests, the decision to make each experiment last 10 minutes was made, privileging a shorter time with the objective of making more tests, given that there was a need to wait after every experiment that the concentration dropped to the initial values. The same experimental pattern was always employed (*fig. 2.9*) using a stopwatch to make sure that the method was consistent: first minute in stationary conditions, with the atomizer turned off, therefore with a PM concentration in the box close or equal to zero. Then the atomizer was turned on for the time provided (1, 5, 10, 20 or 30 seconds), finally steady monitoring (with the atomizer switched off) to reach the end of the 10 minutes limit.

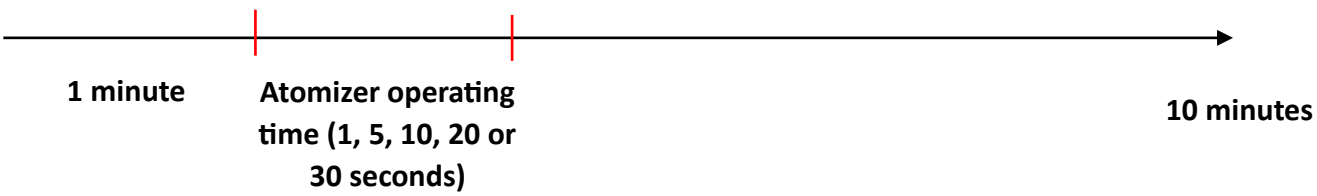

Fig. 2.9 – The experimental pattern

A Python script was prepared to plot the curves, the script read data from the Arduino board using the serial module, it plotted in real time the curves for both the PM10 and PM2.5

concentrations in the box using the matplotlib module, graphs with time on the abscissa (in seconds) and the particle concentration on the ordinate (in $\mu g/m^3$). After 10 minutes, the final graphs could be saved as an image. The script, at the end of every experiment, also generated a csv file (using the pandas module) with every concentration value registered by the sensor during the test and the time associated.

2.4 Data analysis

For the data analysis the parameters picked were:

- the maximum and the minimum values registered in the experiments to pick out the range of values recorded in a single test;
- the average particle concentration, but since in the first minute the atomizer is turned off, the average was calculated excluding the values recorded in that time period;
- the standard deviation [22], which is a measure of how dispersed the data is in relation to the mean. Low, or small, standard deviation indicates data are clustered tightly around the mean, and high, or large, standard deviation indicates data are more spread out. A standard deviation close to zero indicates that data points are very close to the mean, whereas a larger standard deviation indicates data points are spread further away from the mean.

2.5 The mathematical model

Mathematical modeling is a method that represents and explains real systems and occurrences using math formulas, descriptions and approaches. Mathematical models are used to examine, analyze and predict behavior and events [23]. They also utilized to solve complex problems and answer questions. We wanted to develop a mathematical model that could help us to analyze the results of our experiments, to estimate the differences between the expected data (generated by the mathematical model) and the experimental data (generated by our tests). In literature [24] we found out that the expected curves obtained from the experiments should have the exponential shape (*fig. 2.10*), and that this was the reference equation:

$$C = a * (1 - e^{-bt})$$

where:

- C is the particle concentration in the box [µg/m^3]
- a is the asymptotic value of the function if **t → +∞**
- b represents the growth speed of the curve when the particle injection occurs in the box
- t is the time [s]

The equation is based on the following hypothesis: **"there are no particulate leaks from inside the box to the outside"**.

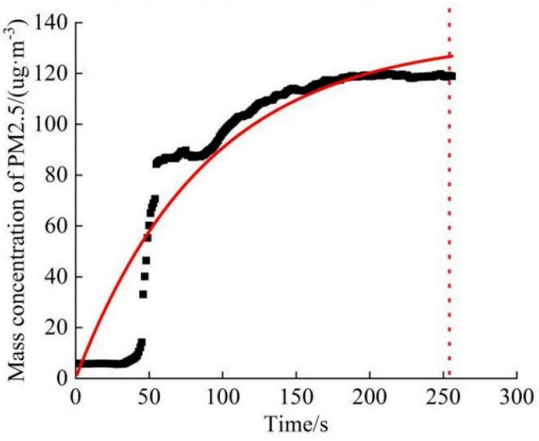

Fig. 2.10 – The curve (in red) that represents the equation

The idea was to take advantage of the curve fitting algorithms in Python to generate the curve (the fitted curve) based on our experimental data, calculating the variables a and b in the process, therefore giving the possibility to check how close our experiments' data come to the model. We tried to implement the experimental curves into the mathematical model developed using the free and open-source Python library SciPy, in particular the curve-fit function. Since the curve of our mathematical model was:

$$C = a * (1 - e^{-bt})$$

the function, to estimate the a and b variables needed initial guesses suggested by the user. The script starts from the initial guesses and then it adapts the curve suggested by the mathematical model to fit the experimental data the curve suggested by the mathematical model. The other graphs extracted from the analysis are more for the statistical significance of the tests, the second graph in the model results is the histogram of the residuals. The residual [25] is the difference between the observed value and the estimated value of the quantity of interest (for example, a sample mean). In our case the observed value is our experimental data, the estimated value is the fitted curve. The histogram of the residuals can be used to check whether the variance is normally distributed [26]. A symmetric bell-shaped histogram which is evenly distributed around zero indicates that the normality assumption is likely to be true. If the histogram indicates that random error is not normally distributed, it suggests that the model's underlying assumptions may have been violated. The third graph is a scatter plot (aka scatter chart, scatter graph), which uses dots to represent values for two different numeric variables. The position of each dot on the horizontal and vertical axis indicates values for an individual data point. Scatter plots are used to observe relationships between variables [27]. The dots in a scatter plot not only report the values of individual data points, but also patterns when the data are taken as a whole. Identification of correlational relationships are common with scatter plots. In these cases, we want to know, if we were given a particular horizontal value, what a good prediction would be for the vertical value. You will often see the variable on the horizontal axis denoted an independent variable, and the variable on the vertical axis the dependent variable. Relationships between variables can be described in many ways: positive or negative, strong or weak, linear or nonlinear. Scatter plots with a linear pattern have points that seem to generally fall along a line while nonlinear patterns seem to follow along some curve. If there is no clear pattern, then it means there is no clear association or relationship between the variables that we are studying. Linear patterns can be thought of as either positive or negative. In a positive pattern, as the value of the independent variable increases, so does the value of the other variable. This shows up in the scatterplot as a linear pattern that rises from left to right. In a negative pattern, as the independent variable increases, the value of the response decreases. This is seen as a linear pattern that falls from left to right. The strength of the relationship or association between two variables is shown by how close the points are to each other. This is true whether the pattern is linear, nonlinear, positive, or negative. A scatter plot can also be

useful for identifying other patterns in data. Data points can be divided into groups based on how closely sets of points cluster together. Scatter plots can also show if there are any unexpected gaps in the data and if there are any outlier points. This can be useful if we want to segment the data into different parts. In every scatter plot there's also the line in red that is originated from the Ordinary Least Squares (OLS) regression [28]. OLS provides a systematic approach to fitting a line that best represents the relationship between the independent variables and the dependent variable by minimizing the sum of the squared residuals, as stated before, the residuals represent the difference between the observed values and the predicted values. The goal of OLS is to find the line that minimizes the sum of the squared residuals, effectively finding the best fit for the data. By squaring the residuals, the impact of larger errors is emphasized and ensured a balanced consideration of positive and negative deviations from the predicted values. This sum is minimized by adjusting the coefficients of the linear regression model, ultimately finding the best-fit line that minimizes the overall error. The last graph is basically an enlargement of the third one, in which we tried to remove the outliers [29]. In statistics, an outlier is a data point that differs significantly from other observations. An outlier may be due to a variability in the measurement, an indication of novel data, or it may be the result of experimental error; the latter are sometimes excluded from the data set. An outlier can also cause serious problems in statistical analyses. The idea was to remove the outliers and then verify again the relationship between the independent variable and the dependent variable.

Another possibility given by the development of our model is that of being able to predict the time needed to reach a fixed concentration C inside the chamber, using the inverse equation of our mathematical model.

$$C = a * (1 - e^{-bt})$$

$$t = -\frac{1}{b} ln\left(1 - \frac{C}{a}\right)$$

3. Results and analysis

The Python script plotted these graphs as a result of the experiments, there are 4 curves in each graph, time on the x axis and the particle concentration on the y axis (graphs with 10 seconds as the injection time):

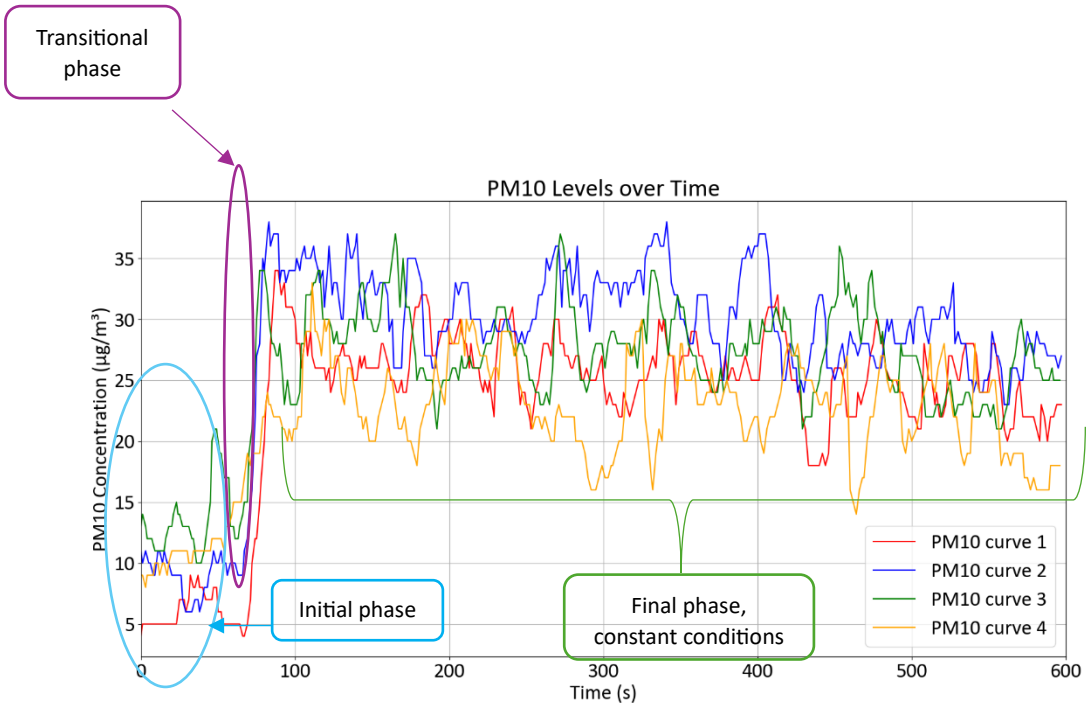

Fig. 3.1 – The PM10 graph (injection time: 10 seconds)

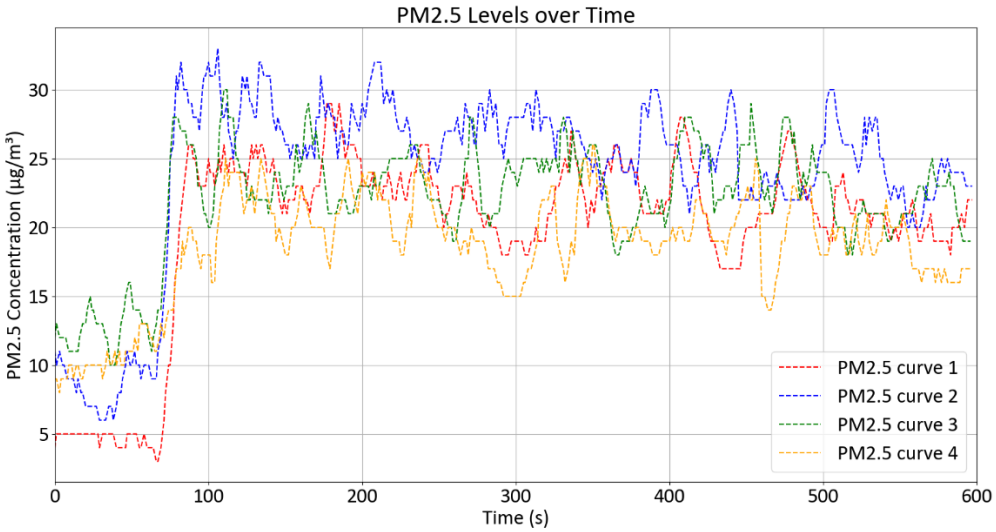

Fig. 3.2 – The PM2.5 graph (injection time: 10 seconds)

The PM2.5 graph uses a different line style only to differentiate them, it's immediately noticeable how the initial phase of the graphs isn't as close to zero as it should be, this is explainable as the sensor isn't really sensitive to the lower concentrations, the values registered rarely went lower than the 5-10 µg/m^3 zone, even for the first experiments performed in a day. Since the sensor's sampling rate was ~1 second, lots of values were registered and because of this the curves are very jagged, there are many peaks and valleys in a very short time. The Python script after every experiment also generated a .csv file with the data registered from the sensor (PM10 and PM2.5 concentrations) and the corresponding time associated with it.

Time	PM 10 (µg/m^3)	PM 2.5 (µg/m^3)
12:46:56	25.0	23.0
12:46:58	23.0	21.0
12:46:59	23.0	21.0
12:47:01	23.0	20.0
12:47:02	23.0	20.0
12:47:04	24.0	21.0
12:47:05	28.0	23.0
12:47:07	28.0	24.0
12:47:08	32.0	27.0
12:47:10	32.0	29.0
12:47:11	32.0	30.0
12:47:13	33.0	30.0
12:47:14	33.0	28.0
12:47:16	34.0	28.0

Fig. 3.3 – Excerpt of a csv file generated from an experiment

Before the analysis of the data, we tried to improve what we had gathered, since we had too many different concentrations in short time spans, we thought that a reduction of the considered data could simplify the analysis significantly, we opted to build "averaged" curves utilizing this pattern: every 10 consecutive points of the curve (or 10 consecutive concentration values in the csv file, same thing) were replaced by the average of those 10 values, basically 1 average point took the place of 10 consecutive points. Here's the graphs obtained after this operation:

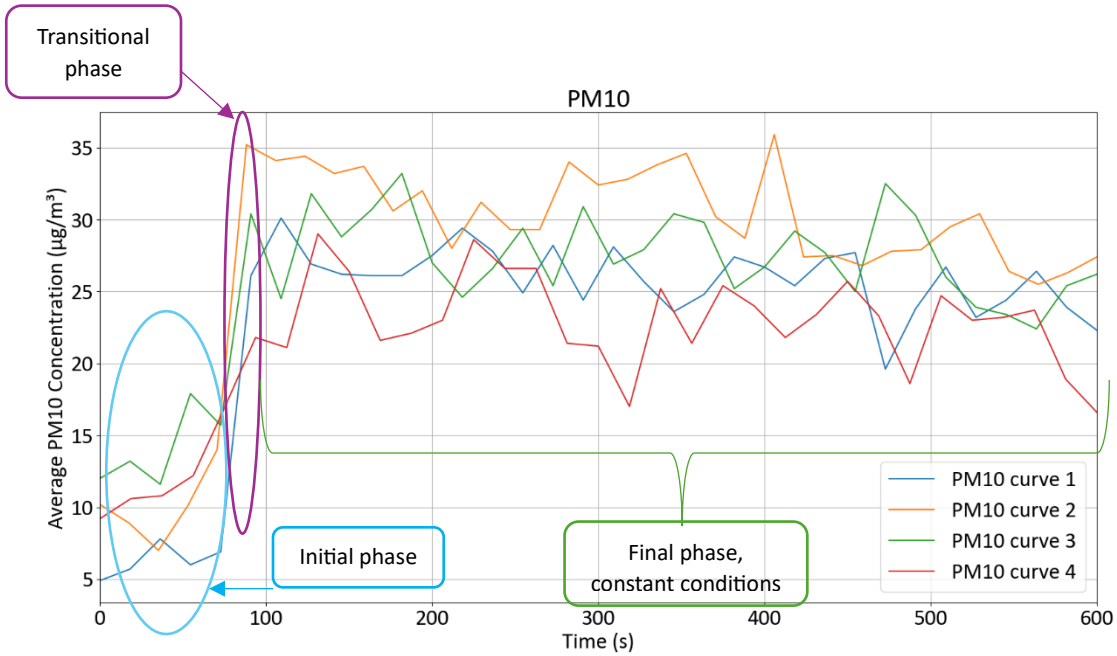

Fig. 3.4 – The averaged curve for PM10 (injection time: 10 seconds)

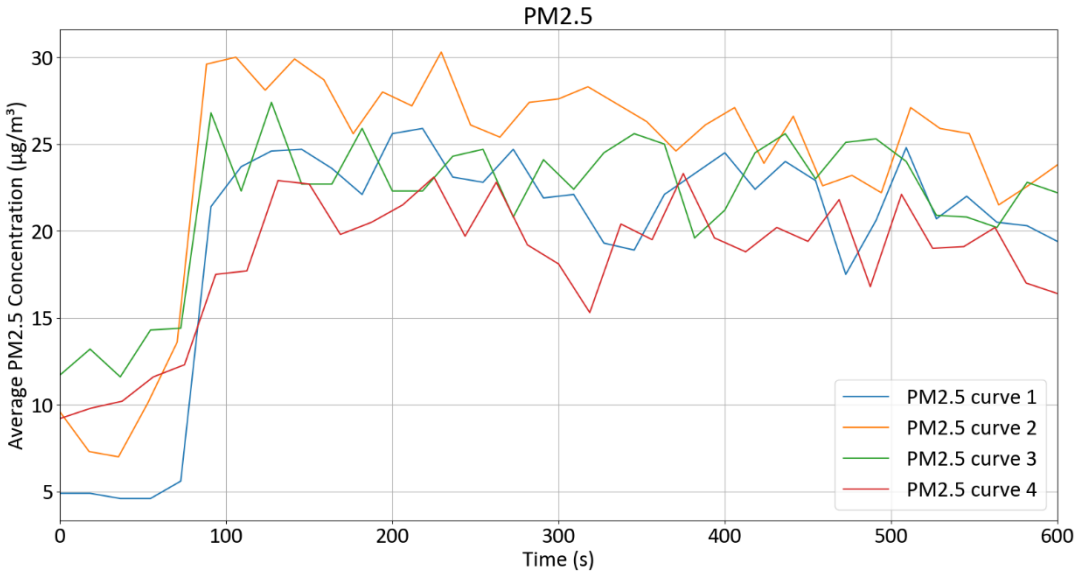

Fig. 3.5 – The averaged curve for PM2.5 (injection time: 10 seconds)

The reduction of the data improved the readability of the graphs, while maintaining the same trend of the "raw" curves (fig. 18 and 19), simplifying the analysis of the data, so we applied the same reduction algorithm to the other curves obtained from the experiments.

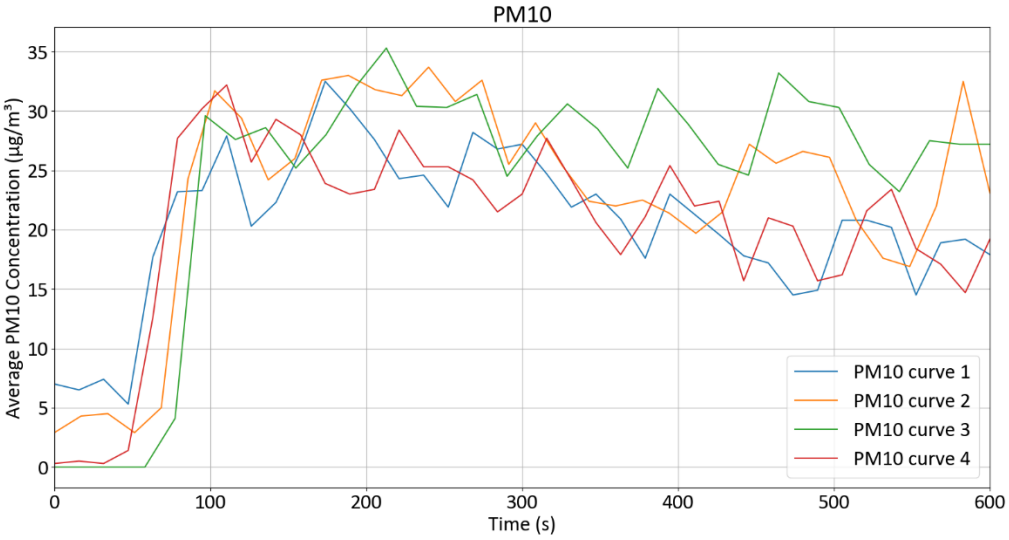

Fig. 3.6 – The averaged curve for PM10 (injection time: 20 seconds)

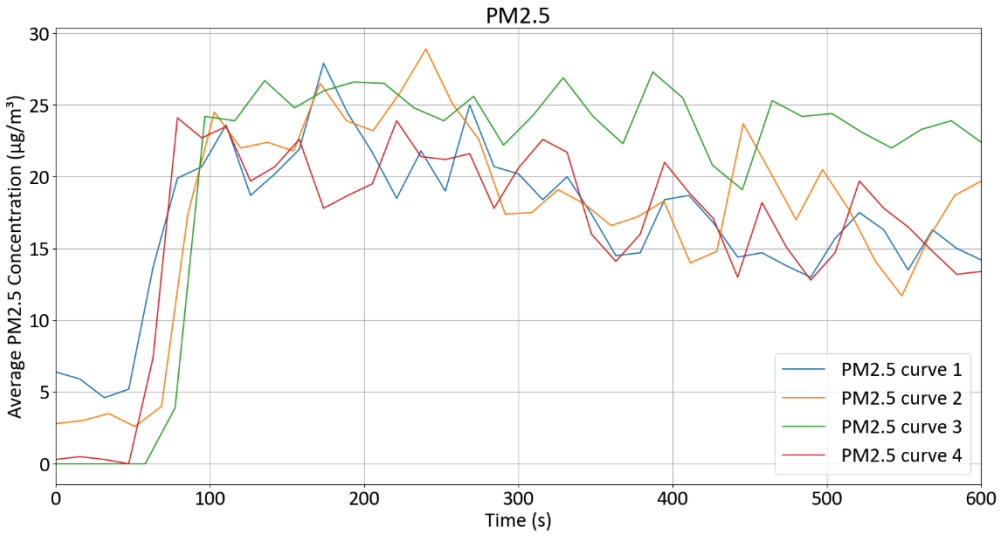

Fig. 3.7 – The averaged curve for PM2.5 (injection time: 20 seconds)

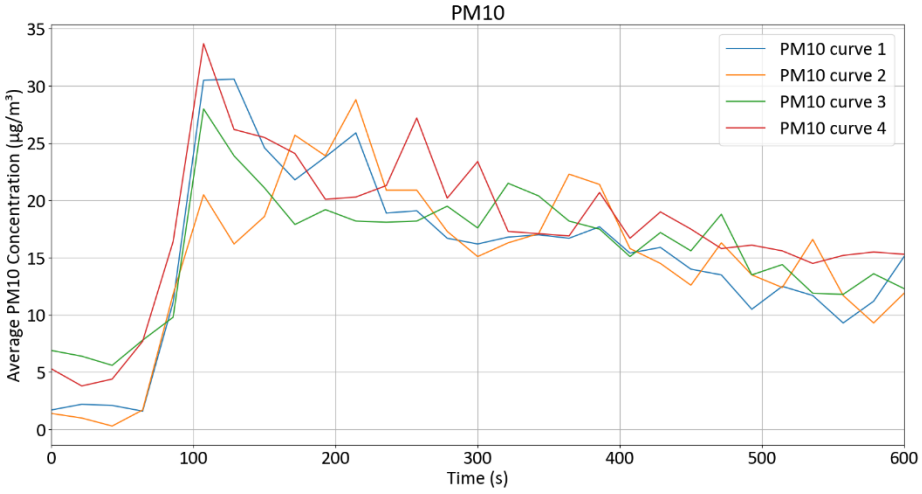

Fig. 3.8 – The averaged curve for PM10 (injection time: 30 seconds)

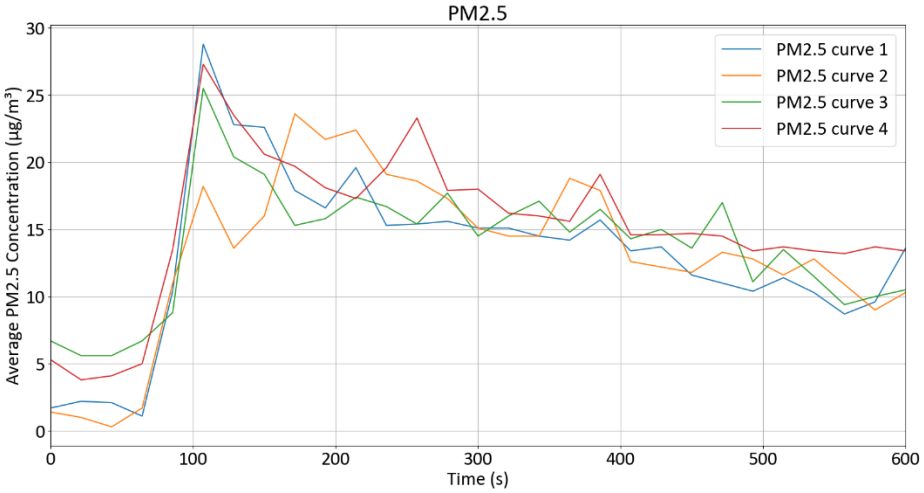

Fig. 3.9 – The averaged curve for PM2.5 (injection time: 30 seconds)

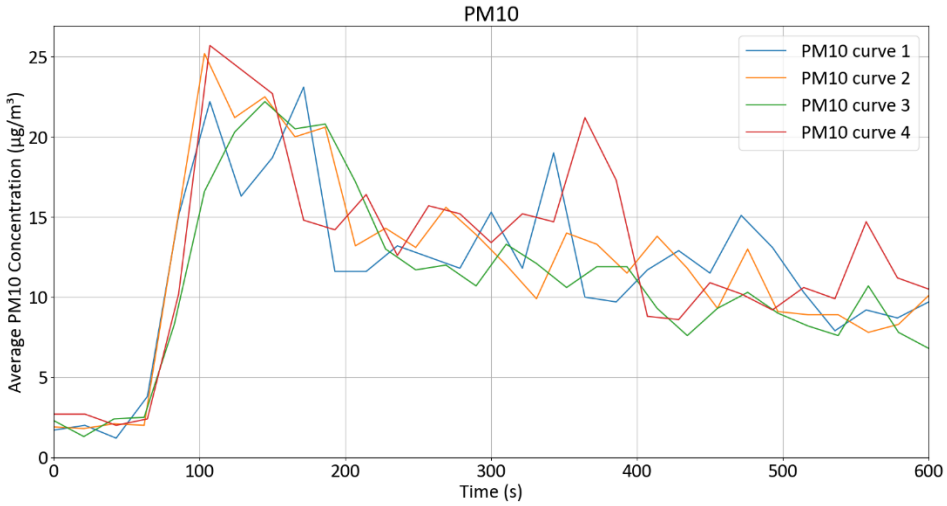

Fig. 3.10 – The averaged curve for PM10 (injection time: 5 seconds)

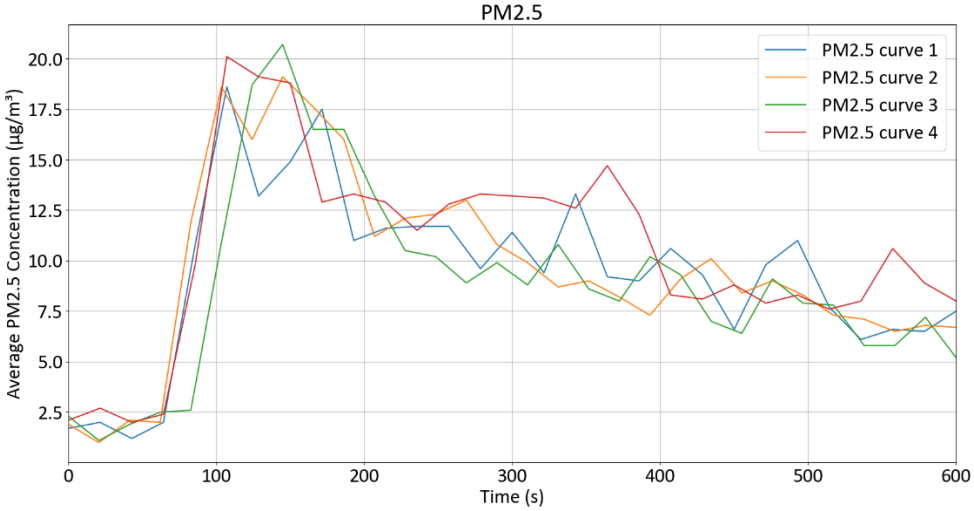

Fig. 3.11 – The averaged curve for PM2.5 (injection time: 5 seconds)

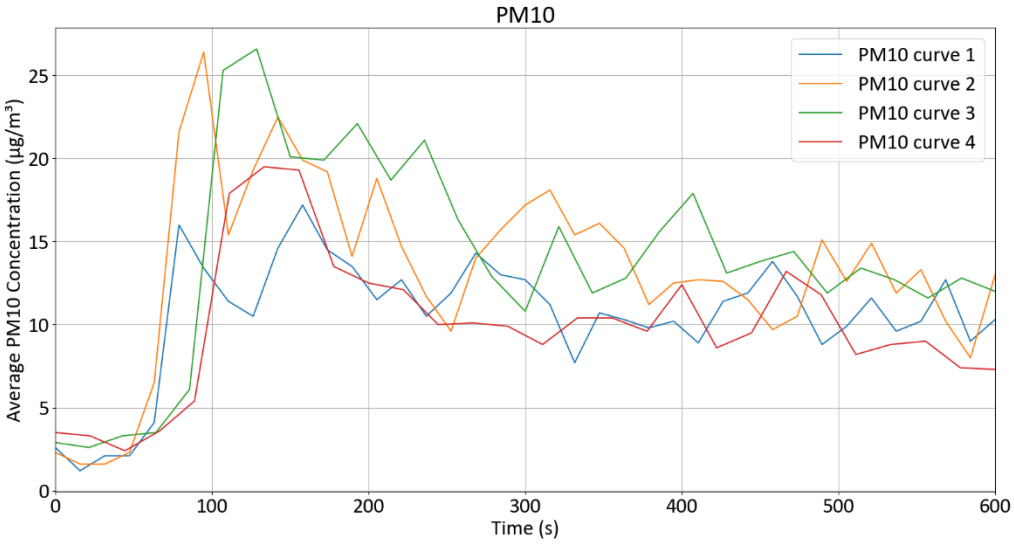

Fig. 3.12 – The averaged curve for PM10 (injection time: 1 second)

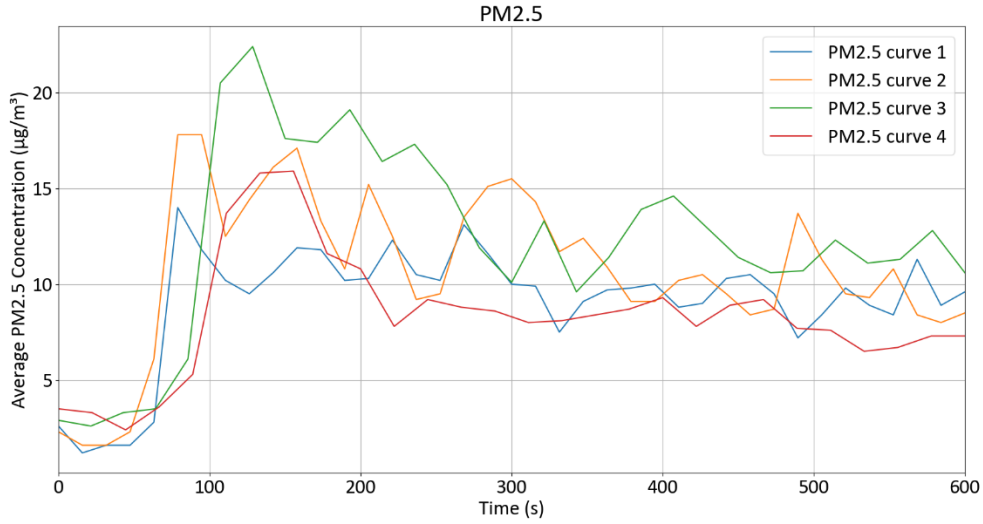

Fig. 3.13 – The averaged curve for PM2.5 (injection time: 1 second)

The shapes of the curves obtained are similar throughout the experiments, there's a drop in the concentrations registered in the last phase of the tests, ideally the trend should have stayed more stable, there are some spills of particulate matter that influence the tests, particularly the hose that links the atomizer and the chamber is a source of leakage in this regard.

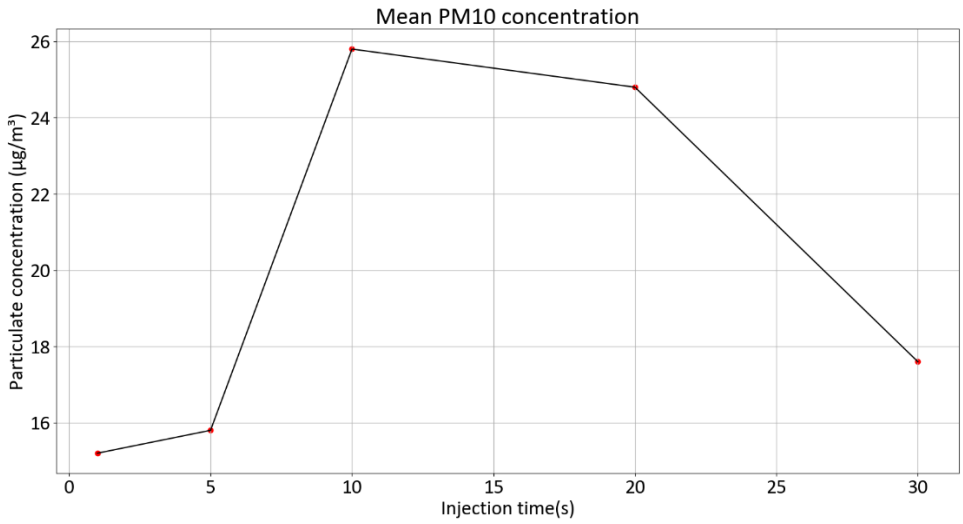

Fig. 3.14 – The curve that links every mean concentration (PM10) and the injection time used

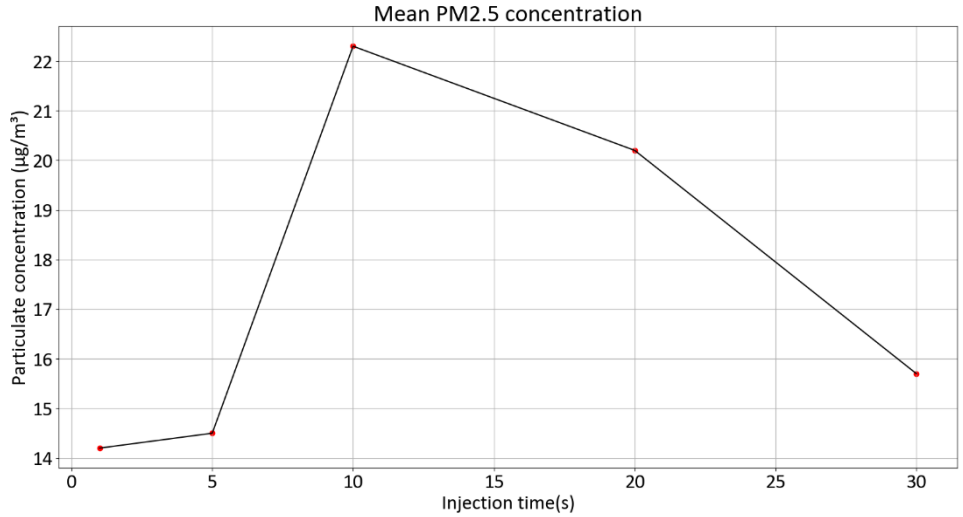

Fig. 3.15 – The curve that links every mean concentration (PM2.5) and the injection time used

In the last two graphs it's evident that the experiments with 10 and 20 seconds (the "intermediate" injection times) gave higher concentration values considering that the parameters were the same for every experiment, probably a possible explanation is that the experiments with 1, 5 and 30 seconds of injection time were made consecutively in a week when the other experiments were done a month earlier and, even if all the experiments took place in the same laboratory, parameters such as temperature and relative humidity slightly changed and the results are observable here.

After having showed the graphs resulted from the experiments, now it's time to analyze the numbers and the more statistical aspects derived from the tests. For every injection time there are

two tables, one for PM10 and one for PM2.5, every table has five columns: the first one identifies the number of the test, the second and the third one are the maximum and the minimum particle concentration registered, therefore picking out the range of values recorded in a single test. The fourth column is the average particle concentration, but since in the first minute the atomizer is turned off, we did the average excluding the values recorded in that time period. The fifth column is the standard deviation, which is a measure of how dispersed the data is in relation to the mean.

Table 3.1 - Results of the PM10 experiments (injection time: 1 second)

Test n°	Maximum concentration registered ($\mu g/m^3$)	Minimum concentration registered ($\mu g/m^3$)	Average concentration ($\mu g/m^3$)	Standard deviation ($\mu g/m^3$)
1	31	9	11.6	2.0
2	25	7	14.6	3.8
3	22	7	15.6	4.0
4	34	8	11.0	3.1

Table 3.2 - Results of the PM2.5 experiments (injection time: 1 second)

Test n°	Maximum concentration registered ($\mu g/m^3$)	Minimum concentration registered ($\mu g/m^3$)	Average concentration ($\mu g/m^3$)	Standard deviation ($\mu g/m^3$)
1	25	9	10.0	1.3
2	18	5	11.7	2.7
3	20	6	13.7	3.3
4	21	8	9.1	2.4

In the first two tables it's visible how, with one second of injection time, the standard deviations aren't very consistent throughout, there's some variability in the results, even if the conditions were always the same. The tests for PM2.5 are less variable than those for PM10, and the standard deviations are lower, this is expected, since PM10, by definition, contains in itself every particle that has a diameter of less than 10 µm, so PM2.5 particles are counted in the PM10 value, basically PM2.5 is a subset of PM10, therefore the maximum, the minimum and the average values should be equal or lower than PM10's corresponding values, and given this, there's a constriction towards the average values, hence lowering the standard deviation.

Table 3.3 - Results of the PM10 experiments (injection time: 5 seconds)

Test n°	Maximum concentration registered (µg/m^3)	Minimum concentration registered (µg/m^3)	Average concentration (µg/m^3)	Standard deviation (µg/m^3)
1	32	6	12.8	3.6
2	29	5	13.2	4.1
3	27	5	12.3	4.5
4	29	8	14.0	4.2

Table 3.4 - Results of the PM2.5 experiments (injection time: 5 seconds)

Test n°	Maximum concentration registered (µg/m³)	Minimum concentration registered (µg/m³)	Average concentration (µg/m³)	Standard deviation (µg/m³)
1	24	5	10.2	2.8
2	23	4	10.4	3.5
3	24	4	10.1	4.0
4	23	7	11.5	3.2

With 5 seconds of injection time the results are more stable, pointing out that the sensor has a more linear response compared to the 1 second tests, particularly in the PM2.5 tests in which the averages are condensed in a 1 µg/m^3 range, it's observable how a 4 seconds increase in injection time doesn't correspond in significantly higher concentrations.

Table 3.5 - Results of the PM10 experiments (injection time: 10 seconds)

Test n°	Maximum concentration registered (µg/m³)	Minimum concentration registered (µg/m³)	Average concentration (µg/m³)	Standard deviation (µg/m³)
1	34	18	25.9	2.2
2	38	23	30.4	3.0
3	37	21	27.6	2.9
4	33	13	23.1	3.1

Table 3.6 - Results of the PM2.5 experiments (injection time: 10 seconds)

Test n°	Maximum concentration registered (µg/m³)	Minimum concentration registered (µg/m³)	Average concentration (µg/m³)	Standard deviation (µg/m³)
1	29	17	22.4	2.1
2	33	20	26.2	2.3
3	30	18	23.3	1.9
4	26	14	19.9	2.1

These are probably the “best” results obtained in our tests, the standard deviations are fairly low and consistent, and the averages are quite higher than the 5 seconds experiments.

Table 3.7 - Results of the PM10 experiments (injection time: 20 seconds)

Test n°	Maximum concentration registered (µg/m³)	Minimum concentration registered (µg/m³)	Average concentration (µg/m³)	Standard deviation (µg/m³)
1	35	13	22.2	4.5
2	43	17	26.0	4.8
3	40	19	28.6	3.0
4	37	12	22.4	4.2

Table 3.8 - Results of the PM2.5 experiments (injection time: 20 seconds)

Test n°	Maximum concentration registered ($\mu g/m^3$)	Minimum concentration registered ($\mu g/m^3$)	Average concentration ($\mu g/m^3$)	Standard deviation ($\mu g/m^3$)
1	32	12	18.3	3.7
2	34	11	19.8	4.1
3	29	17	24.2	2.0
4	28	12	18.3	3.2

For the 20 seconds experiments we can see, comparing them to the 10 seconds experiments, that the concentrations recorded aren't much higher, there's only a slight difference, the standard deviations are bigger, and they have a larger range between them.

Table 3.9 - Results of the PM10 experiments (injection time: 30 seconds)

Test n°	Maximum concentration registered ($\mu g/m^3$)	Minimum concentration registered ($\mu g/m^3$)	Average concentration ($\mu g/m^3$)	Standard deviation ($\mu g/m^3$)
1	33	8	17.2	5.1
2	34	8	17.4	4.7
3	26	10	17.2	3.2
4	37	10	19.2	3.8

Table 3.10 - Results of the PM2.5 experiments (injection time: 30 seconds)

Test n°	Maximum concentration registered (µg/m³)	Minimum concentration registered (µg/m³)	Average concentration (µg/m³)	Standard deviation (µg/m³)
1	30	8	14.5	3.6
2	25	8	15.2	3.9
3	29	8	14.9	2.8
4	30	12	16.7	3.1

The last experiments, with 30 seconds of injection time, showed again that an increase in the operating time of the atomizer doesn't correspond in a rise of the concentration recorded, on the contrary there's a big decrease in the averages recorded even if the standard deviations are quite online with the ones of the 20 seconds tests.

3.1 Model results

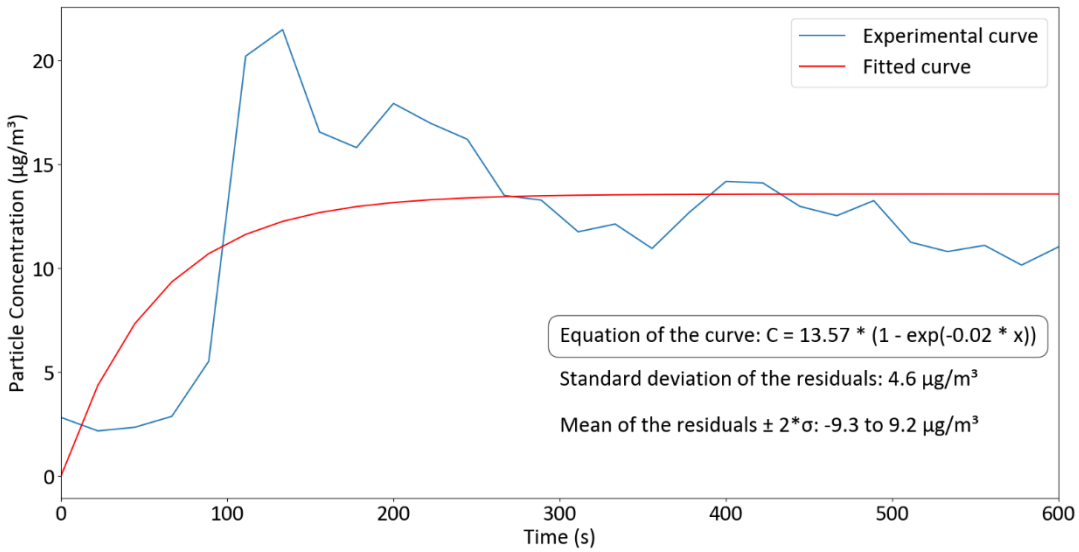

Fig. 3.16 – The PM10 fitted curve generated from the experiments with 1 second of injection time

In red there's the fitted curve that Python suggested as the best fit, in blue there's the curve that represents the experimental data. The four curves shown before were condensed in one curve that represents them all. The equation of the fitted curve suggested is shown in the rounded box, the numerical values of the variables a and b are included. The standard deviation of the residuals is reported under the box and was used to estimate the uncertainty of our model using the confidence level of 95%. Confidence intervals use the standard error to derive a range in which we think the true value is likely to lie [30]. A confidence interval gives an indication of the degree of uncertainty of an estimate and helps to decide how precise a sample estimate is. It specifies a range of values likely to contain the unknown population value. These values are defined by lower and upper limits. The width of the interval depends on the precision of the estimate and the confidence level used. A greater standard error will result in a broader interval; if the interval is wider, the estimate is less precise.

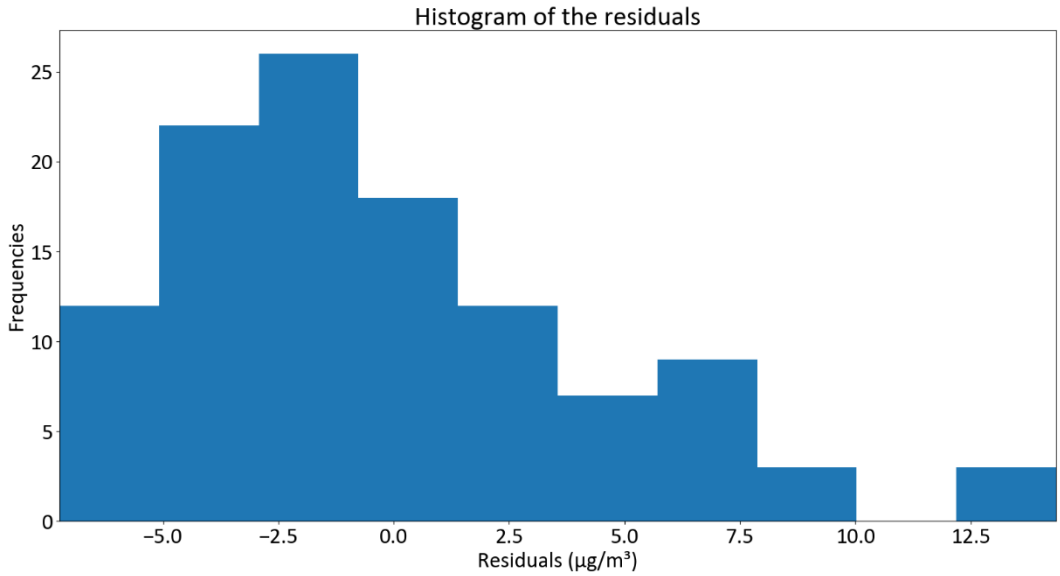

Fig. 3.17 – The histogram of the residuals (PM10) generated from the experiments with 1 second of injection time

The histogram of the residuals suggests that there are some outlier points, and since the graph is not centered there's an imbalance issue, the random error is not normally distributed, random error is a chance difference between the observed and true values of something.

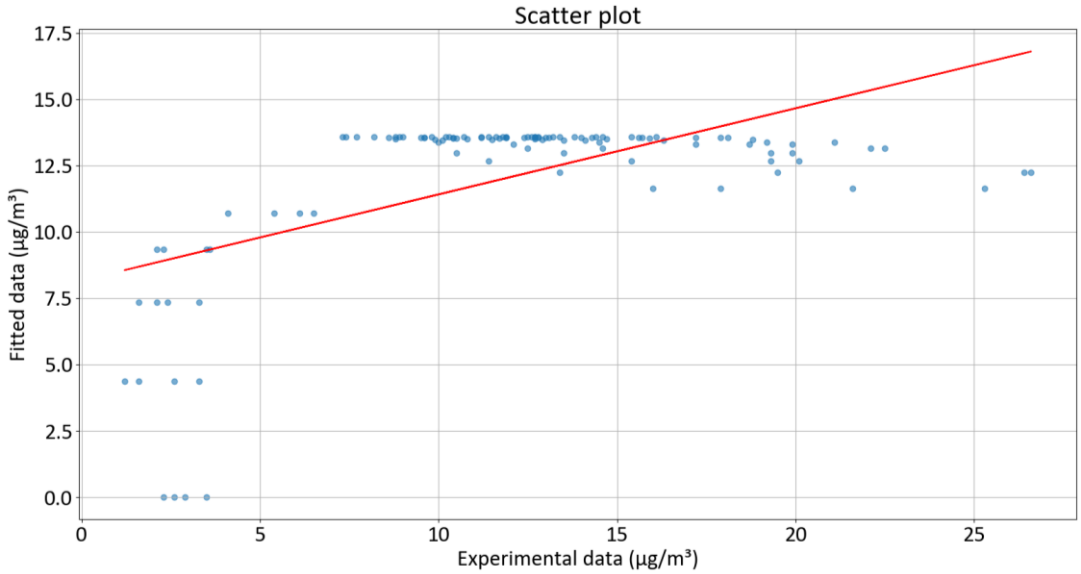

Fig. 3.18 – The scatter plot (PM10) generated from the experiments with 1 second of injection time

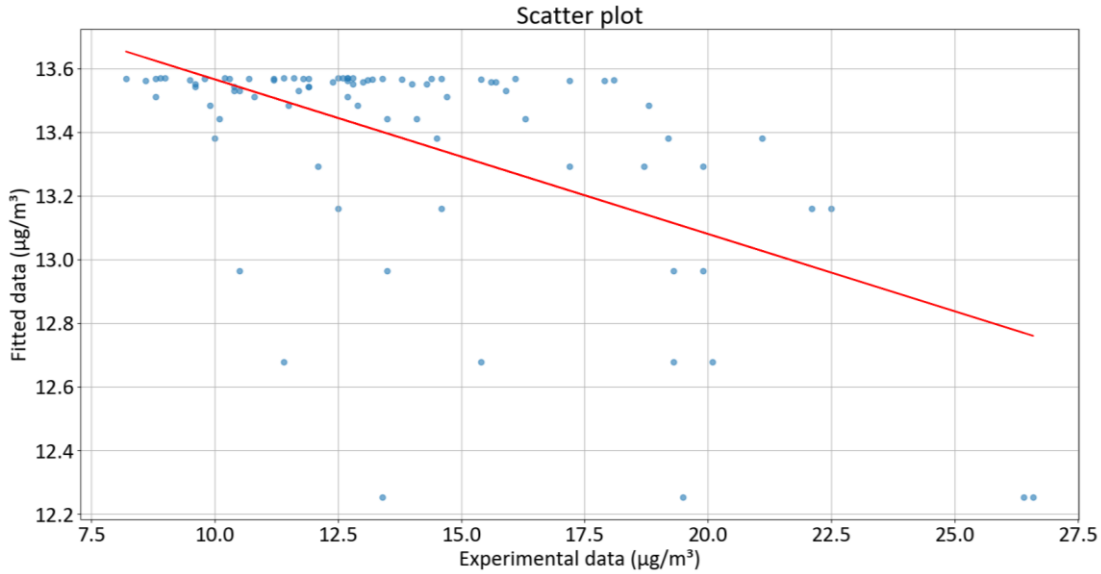

Fig. 3.19 – Enlargement of the scatter plot (PM10) generated from the experiments with 1 second of injection time

The scatter plots highlight that there's a relationship between the experimental and the fitted data, even more considering the enlargement in which some outlier points were excluded, the red line's angle isn't perfect (45°) but not very distant.

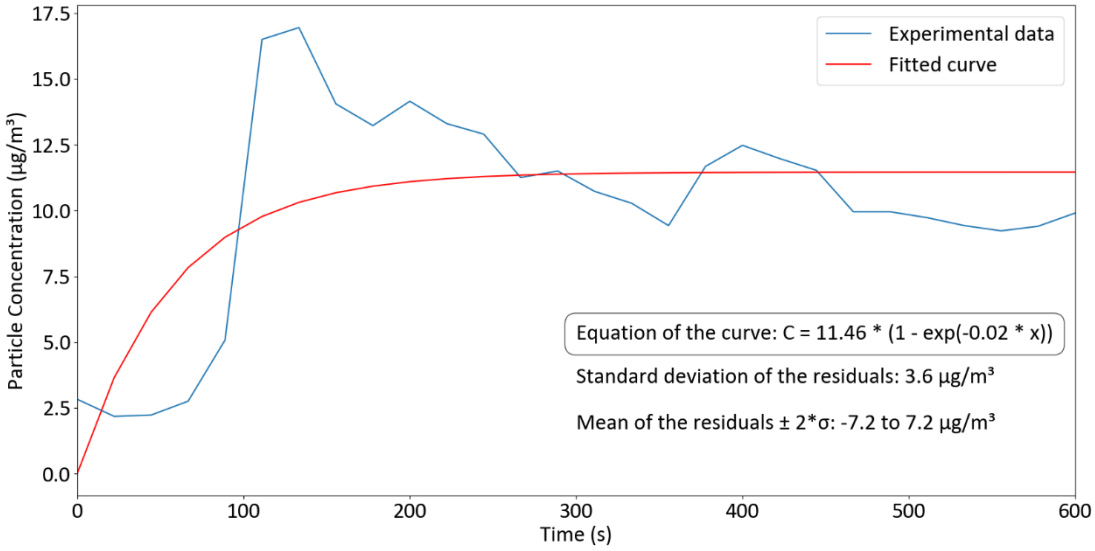

Fig. 3.20 – The PM2.5 fitted curve generated from the experiments with 1 second of injection time

Similarly to what was discussed before, the PM2.5 graph shows that the curve fitting algorithm manages to near the experimental curve better than before, the standard deviation is lower, and the uncertainty range is shorter.

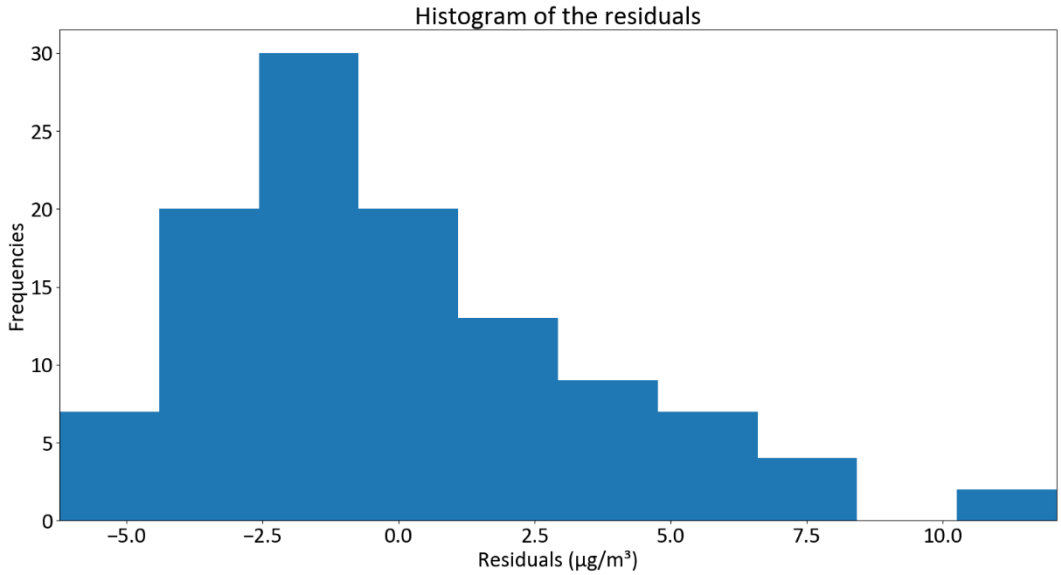

Fig. 3.21 – The histogram of the residuals (PM2.5) generated from the experiments with 1 second of injection time

The histogram of the residuals for PM2.5 is better than the one before, it's more regular but still highlights some outlier points and a random error not normally distributed.

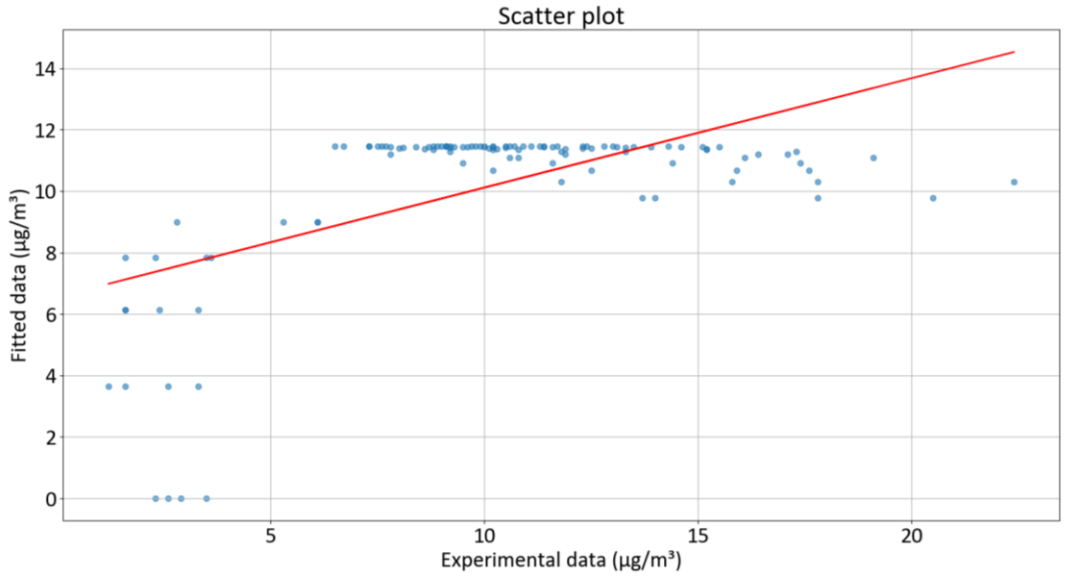

Fig. 3.22 – The scatter plot (PM2.5) generated from the experiments with 1 second of injection time

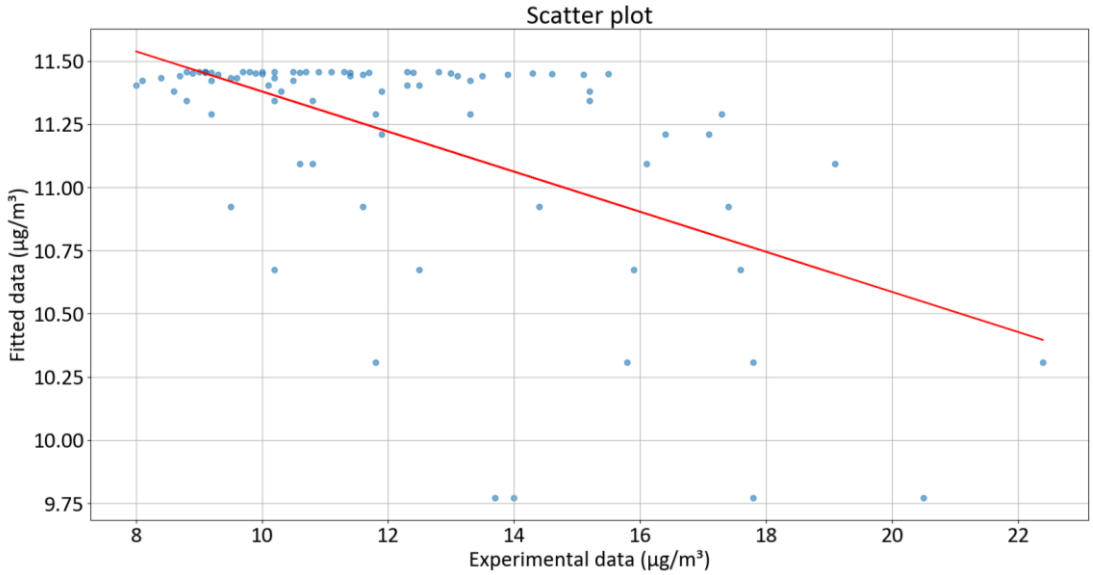

Fig. 3.23 – Enlargement of the scatter plot (PM2.5) generated from the experiments with 1 second of injection time

The scatter plots show there's a stronger bond between the experimental and the fitted data compared to the PM10 tests as expected from the previous two graphs.

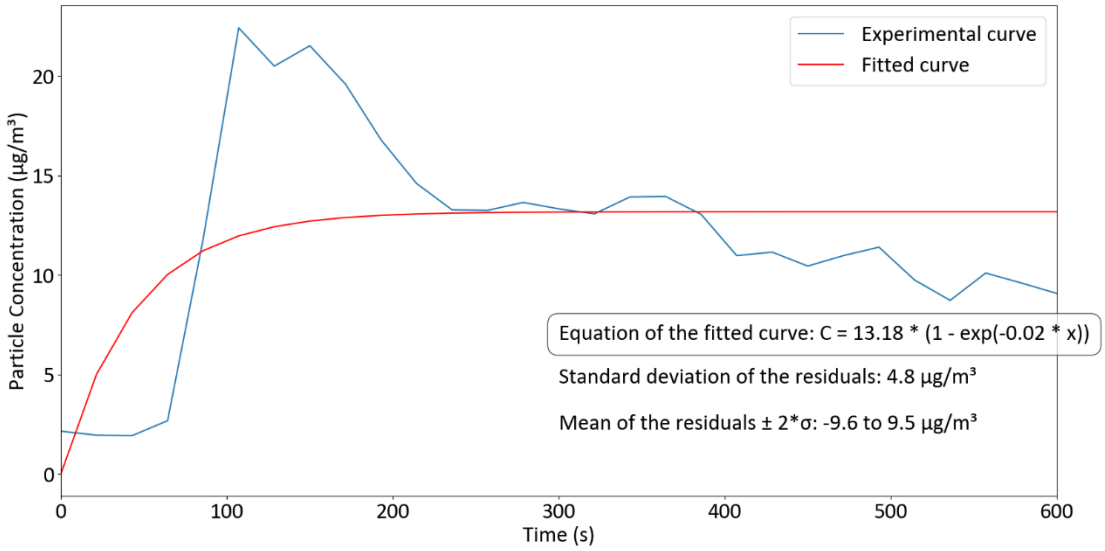

Fig. 3.24 – The PM10 fitted curve generated from the experiments with 5 seconds of injection time

The 5 seconds fitted curve graph is fairly like the 1 second one, since the standard deviation is very similar, the gap in injection time does not really show differences here.

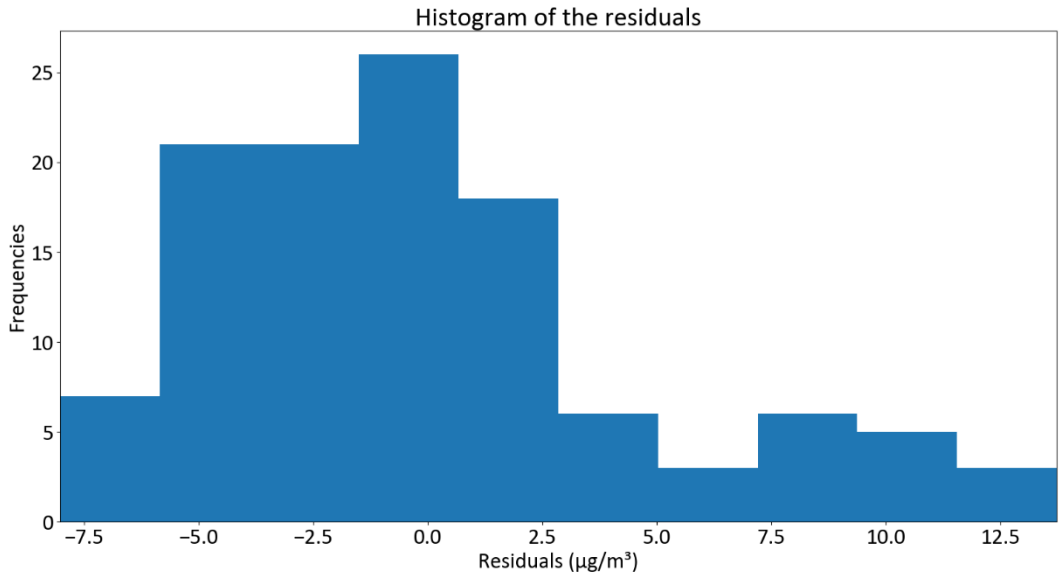

Fig. 3.25 – The histogram of the residuals (PM10) generated from the experiments with 5 seconds of injection time

The histogram of the residuals is more centered than the one second one, but it's more irregular even if it's distributed better.

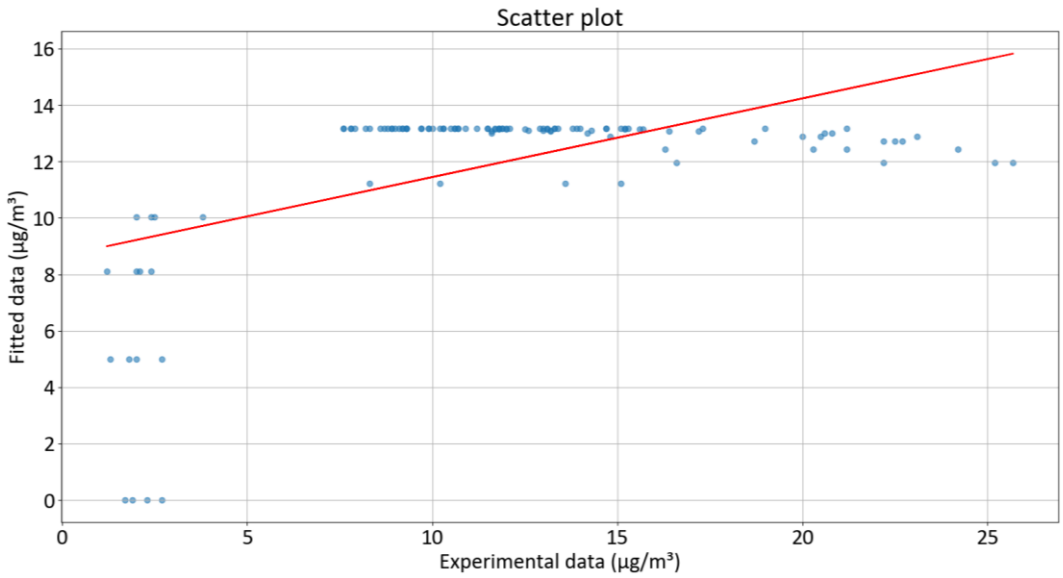

Fig. 3.26 – The scatter plot (PM10) generated from the experiments with 5 seconds of injection time

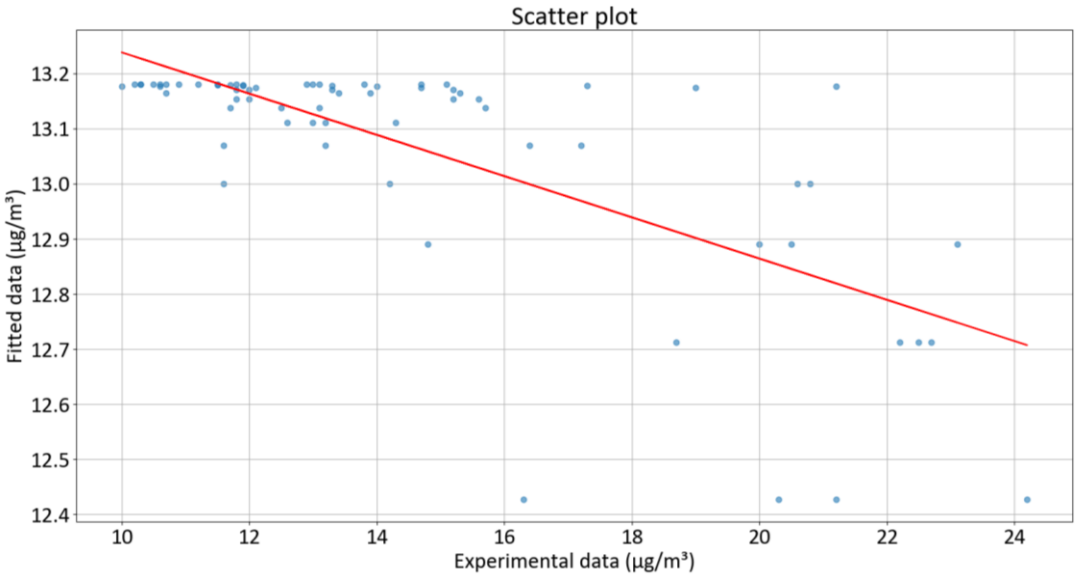

Fig. 3.27 – Enlargement of the scatter plot (PM10) generated from the experiments with 5 seconds of injection time

The scatter plots follow the trends observed before, without big differences.

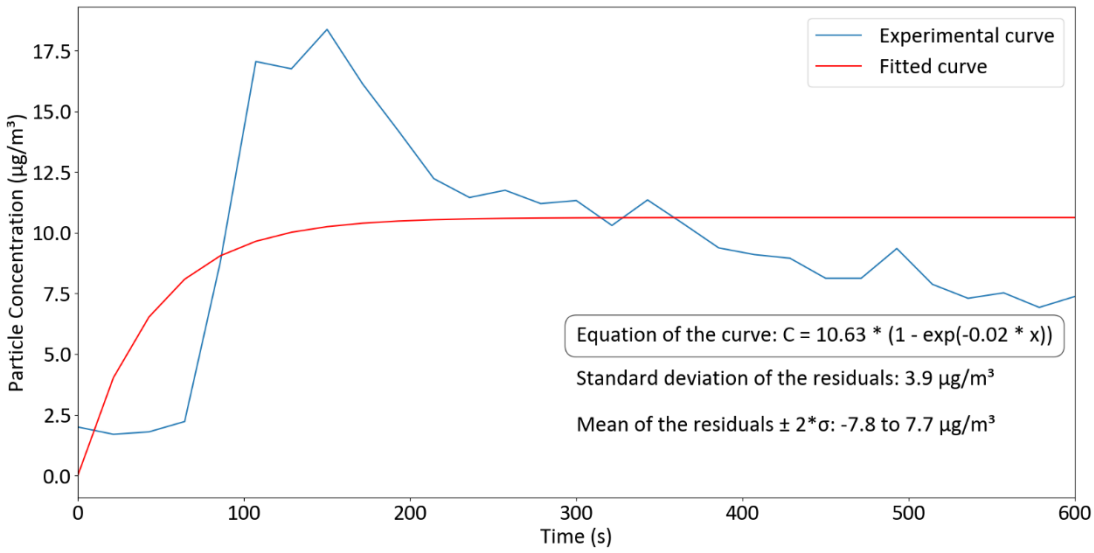

Fig. 3.28 – The PM2.5 fitted curve generated from the experiments with 5 seconds of injection time

This graph is in line with the PM10 one, a little more centered on the average and with a smaller uncertainty range.

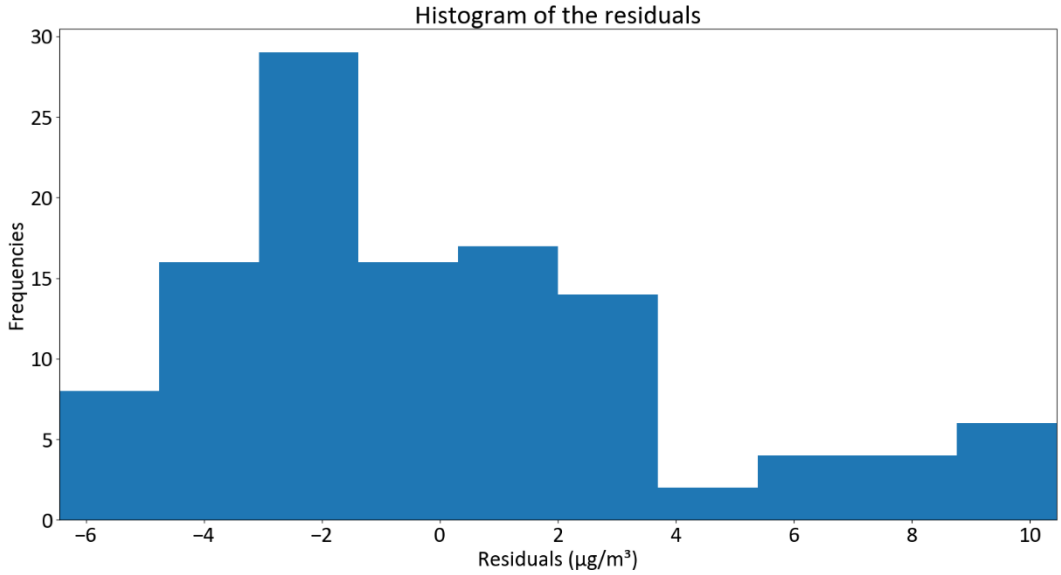

Fig. 3.29 – The histogram of the residuals (PM2.5) generated from the experiments with 5 seconds of injection time

The histogram is better distributed than the PM10 one, but the observations are similar.

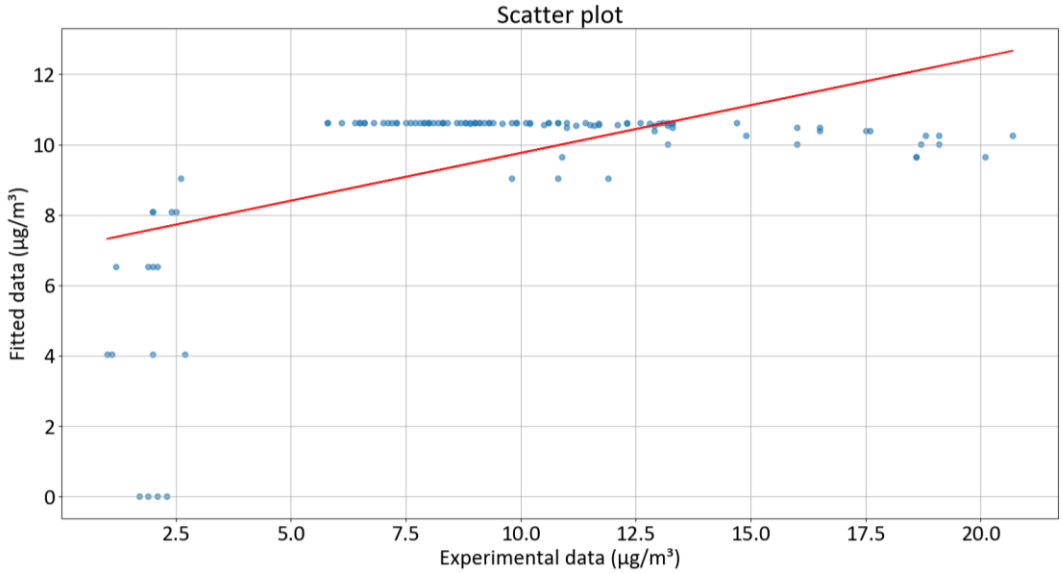

Fig. 3.30 – The scatter plot (PM2.5) generated from the experiments with 5 seconds of injection time

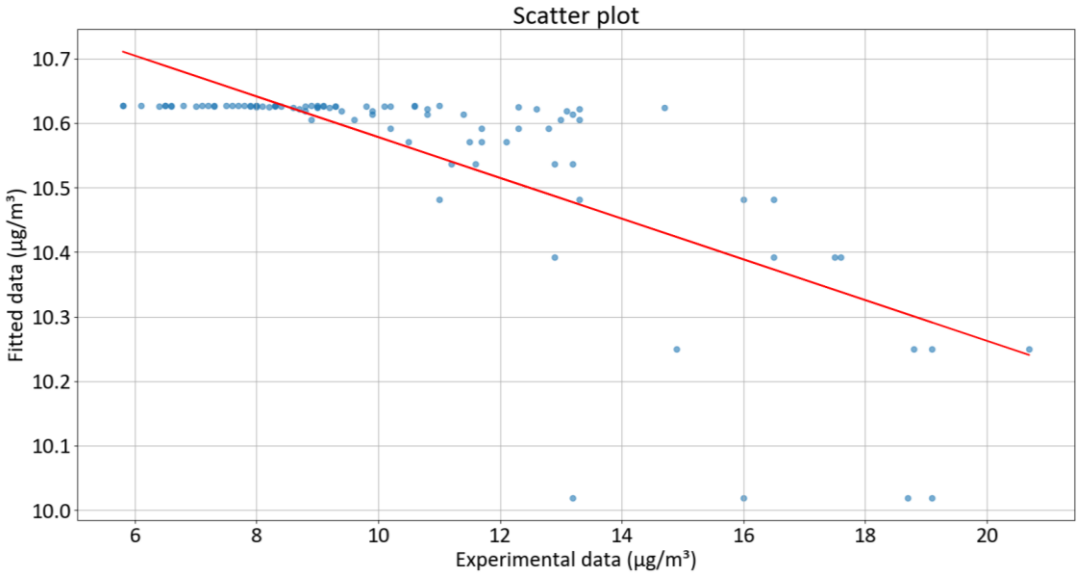

Fig. 3.31 – Enlargement of the scatter plot (PM2.5) generated from the experiments with 5 seconds of injection time

The scatter plots don't show many differences, the alteration between one and five seconds is very subtle.

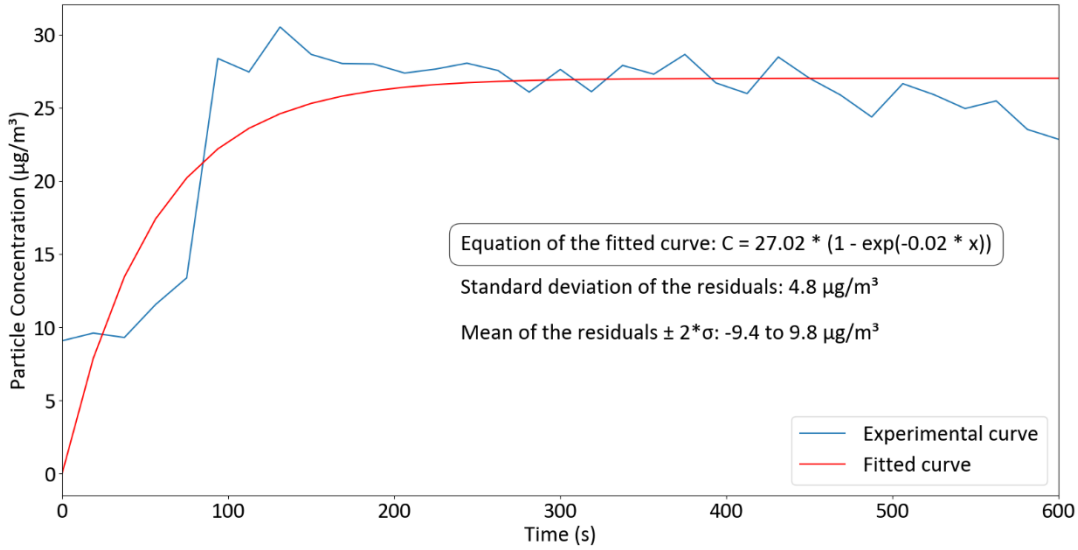

Fig. 3.32 – The PM10 fitted curve generated from the experiments with 10 seconds of injection time

The graph for the 10 seconds tests seems to follow better the trend of our experimental curve, even if the statistical measurements are close to the values registered before.

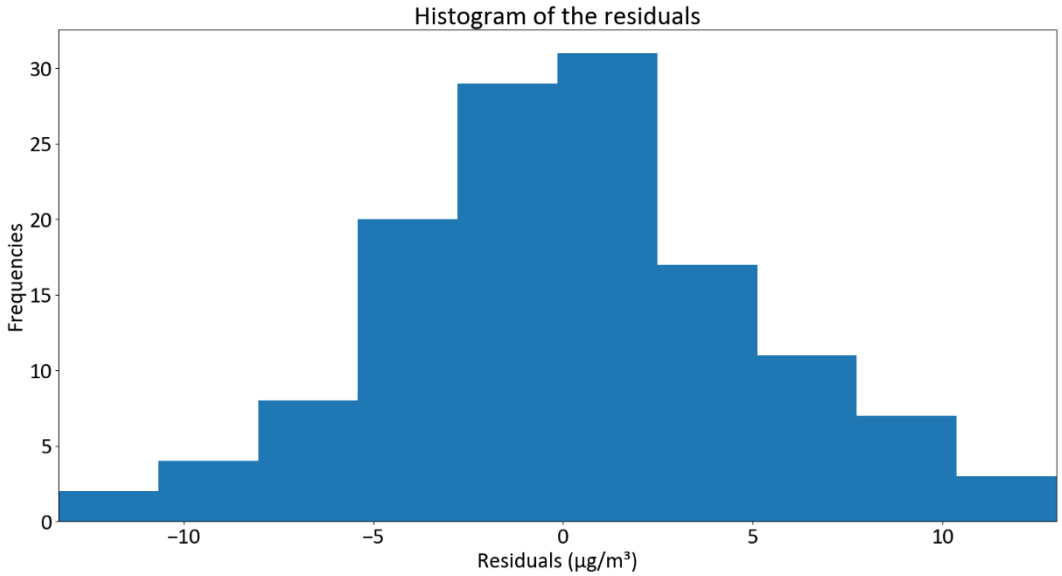

Fig. 3.33 – The histogram of the residuals (PM10) generated from the experiments with 10 seconds of injection time

The histogram of the residuals has a better shape than the ones before it, and it's perfectly centered on the x axis, suggesting that the best measurements statistically wise came with this time period of the atomizer's operating time.

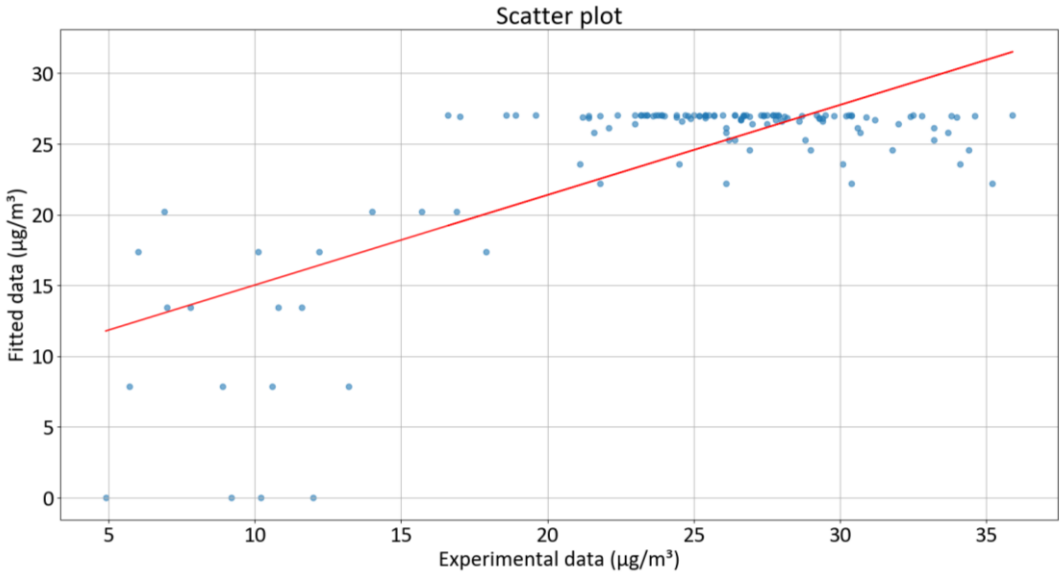

Fig. 3.34 – The scatter plot (PM10) generated from the experiments with 10 seconds of injection time

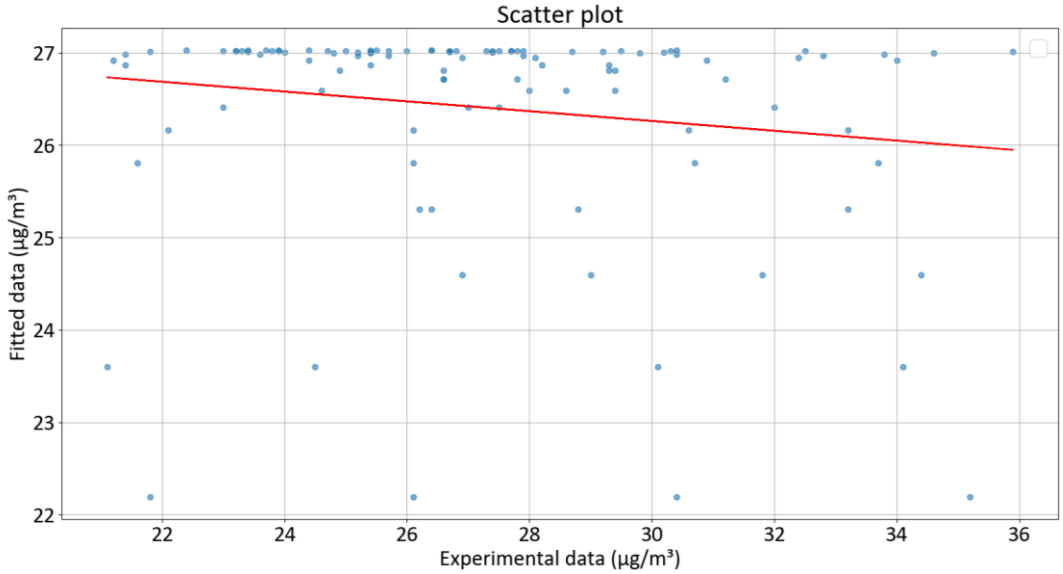

Fig. 3.35 – Enlargement of the scatter plot (PM10) generated from the experiments with 10 seconds of injection time

The scatter plots show that there's a good bond between the independent variable and the dependent variable.

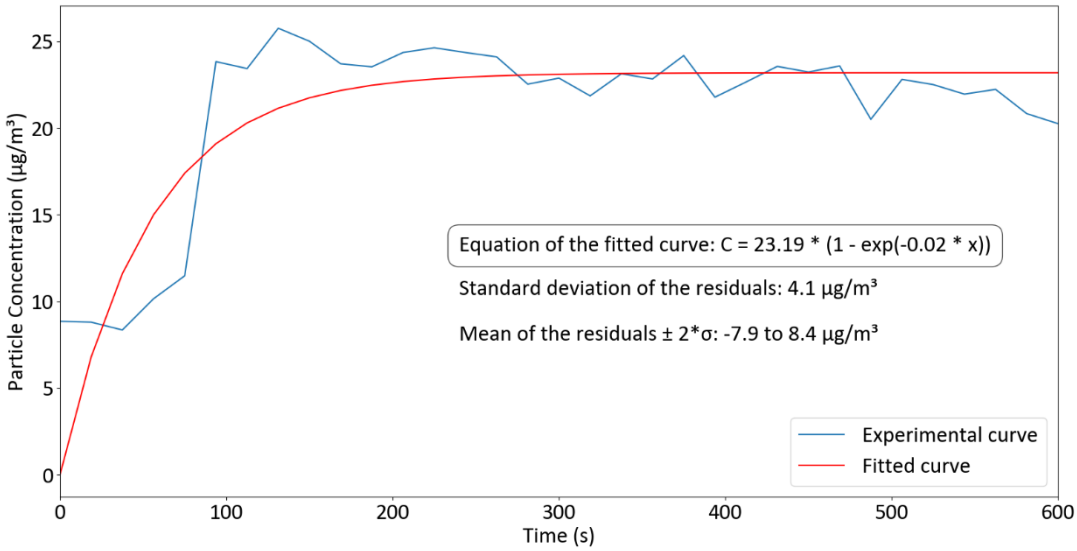

Fig. 3.36 – The PM2.5 fitted curve generated from the experiments with 10 seconds of injection time

The fitted curve generated from PM2.5 measurements is very similar to the PM10 one.

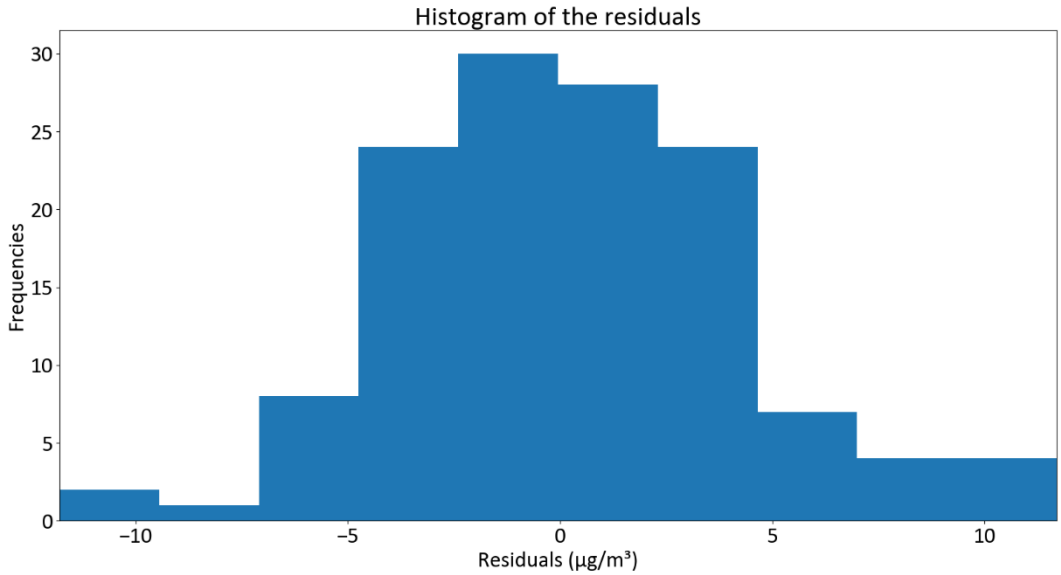

Fig. 3.37 – The histogram of the residuals (PM2.5) generated from the experiments with 10 seconds of injection time

The second graph is more irregular than the PM10 one, but it's still centered on the abscissa, this chart is in countertrend with the previous PM2.5 showings.

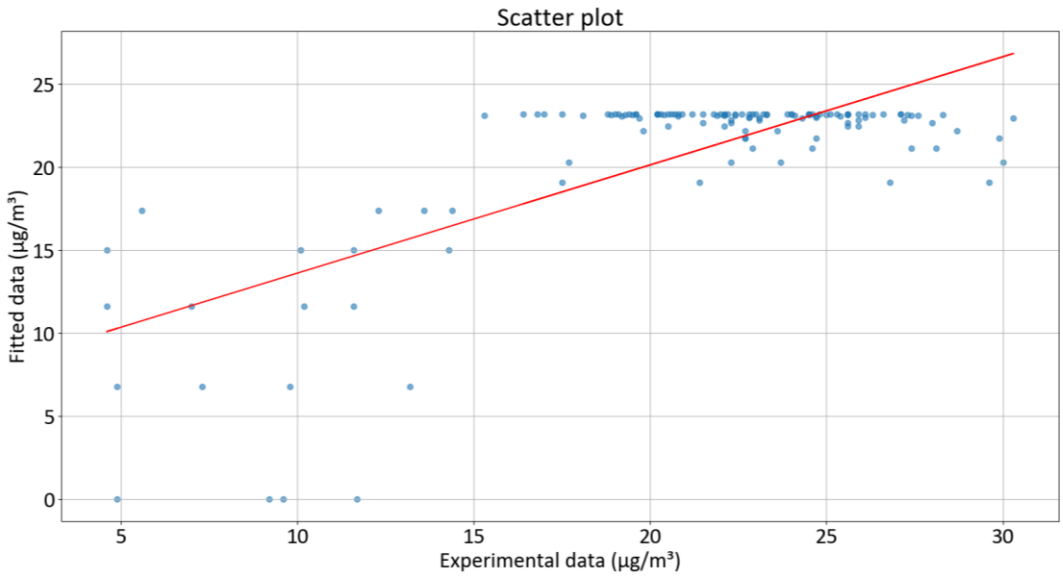

Fig. 3.38 – The scatter plot (PM2.5) generated from the experiments with 10 seconds of injection time

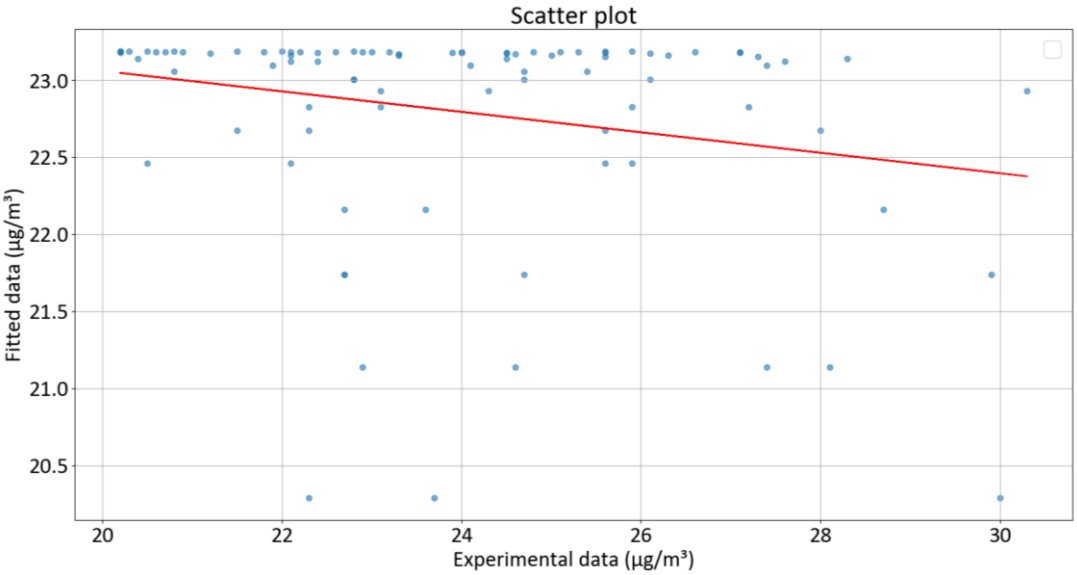

Fig. 3.39 – Enlargement of the scatter plot (PM2.5) generated from the experiments with 10 seconds of injection time

The scatter plots follow closely the PM10 ones.

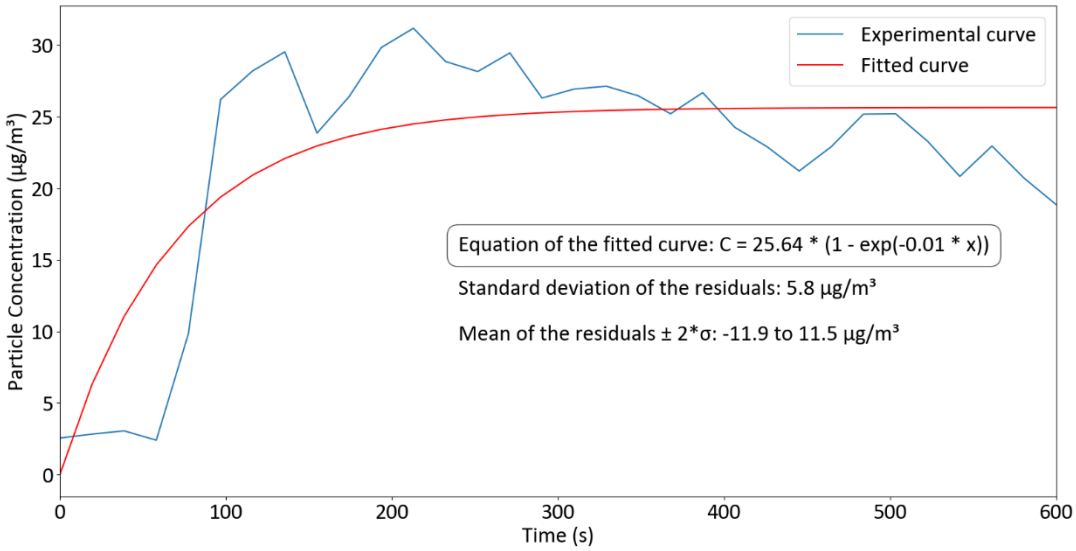

Fig. 3.40 – The PM10 fitted curve generated from the experiments with 20 seconds of injection time

For the experiments with 20 seconds the standard deviation and the uncertainty are higher than the PM10 ones, especially the uncertainty.

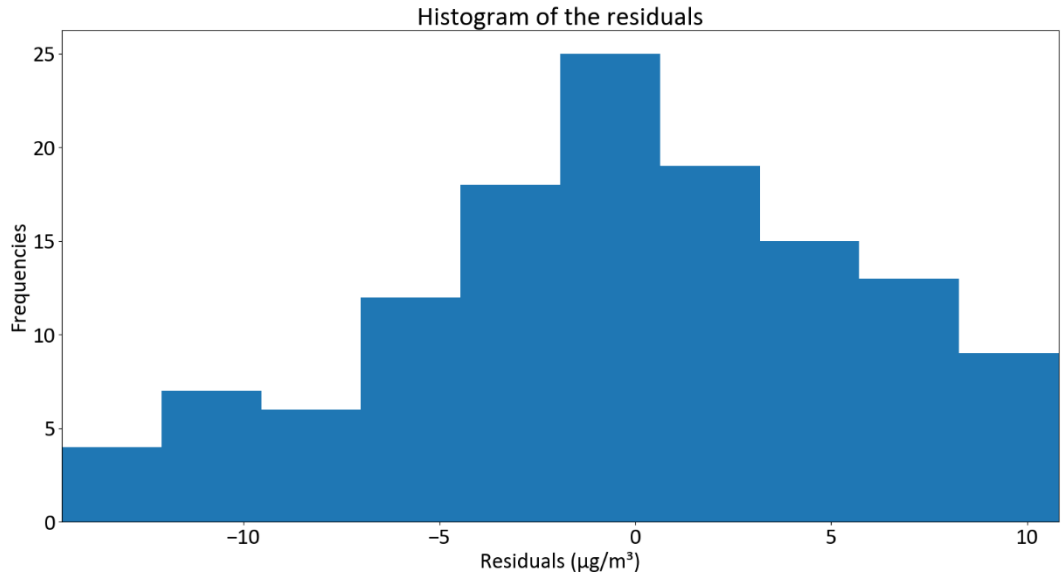

Fig. 3.41 – The histogram of the residuals (PM10) generated from the experiments with 20 seconds of injection time

The shape of the histogram of the residuals is quite regular but it's slightly off centered.

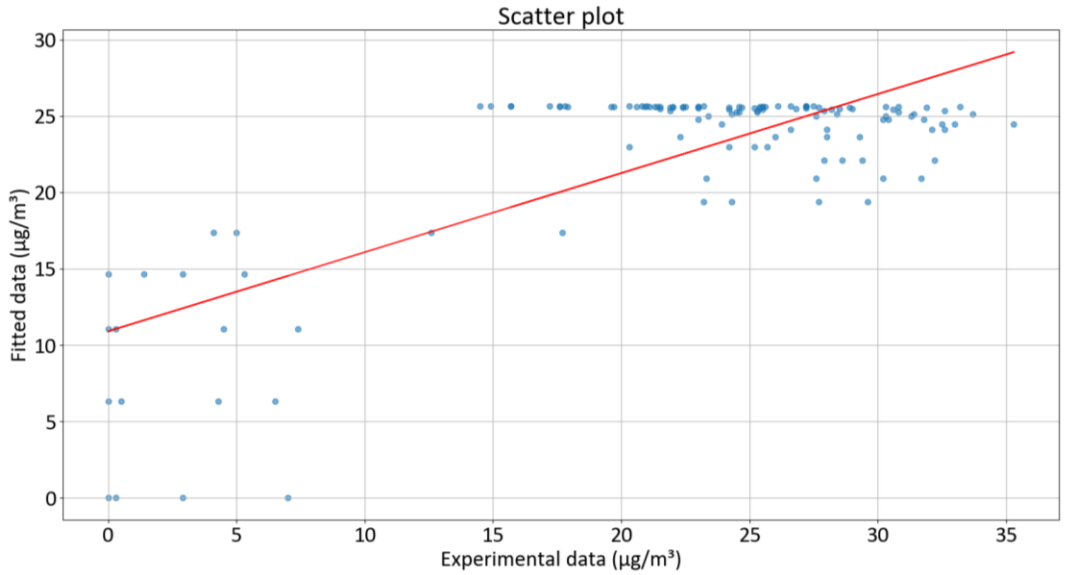

Fig. 3.42 – The scatter plot (PM10) generated from the experiments with 20 seconds of injection time

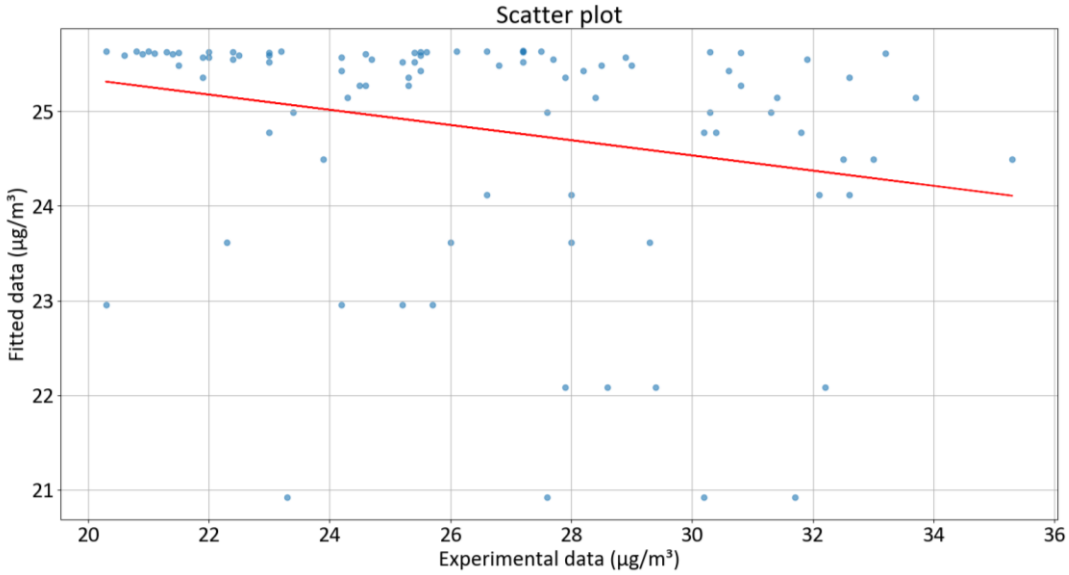

Fig. 3.43 – Enlargement of the scatter plot (PM10) generated from the experiments with 20 seconds of injection time

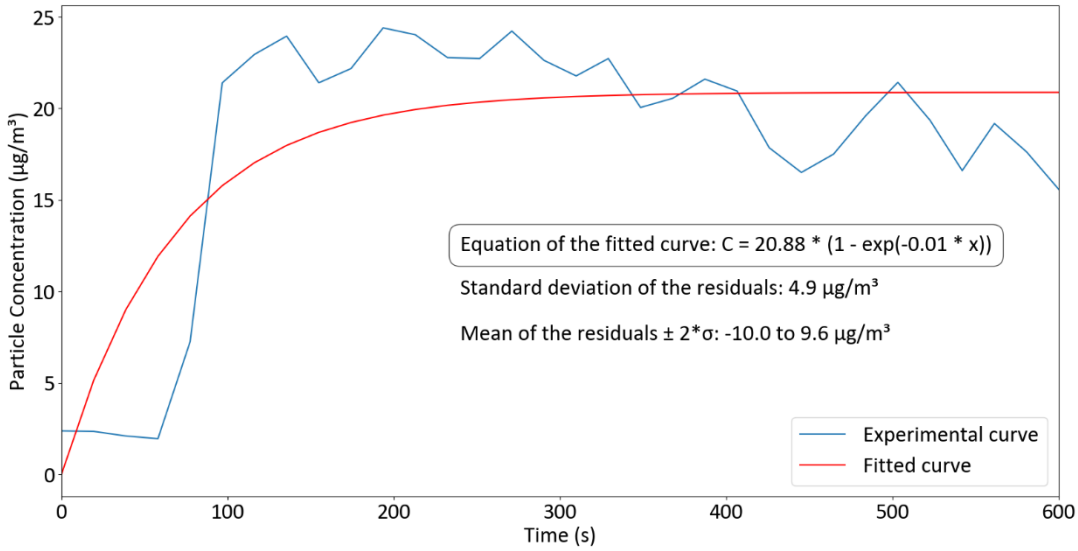

Fig. 3.44 – The PM2.5 fitted curve generated from the experiments with 20 seconds of injection time

As expected, lower standard deviation than the PM10 one.

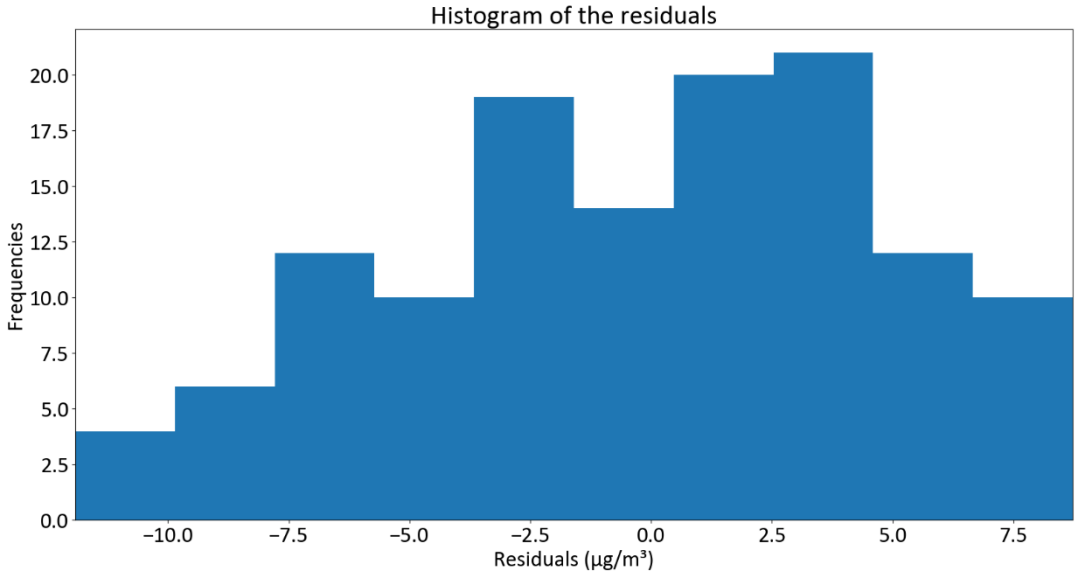

Fig. 3.45 – The histogram of the residuals (PM2.5) generated from the experiments with 20 seconds of injection time

Compared to the PM10 one it's much more irregular, indicating that there were some random errors in these experiments.

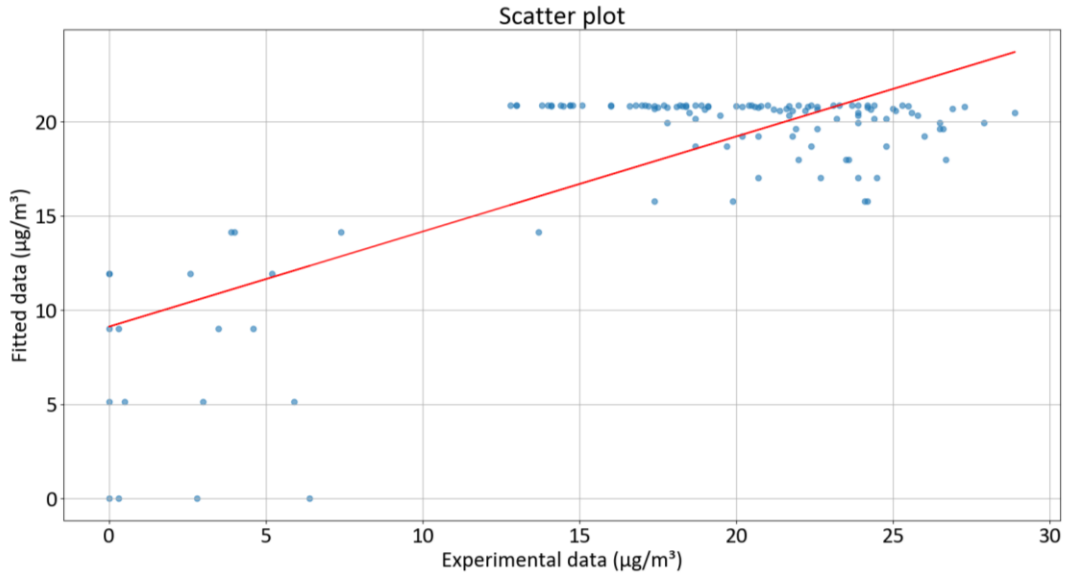

Fig. 3.46 – The scatter plot (PM2.5) generated from the experiments with 20 seconds of injection time

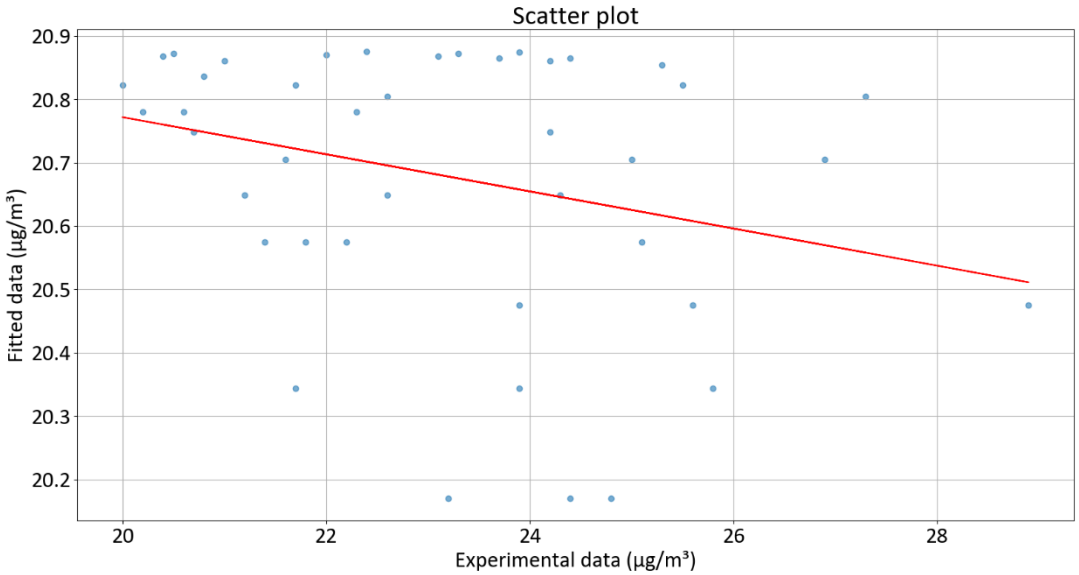

Fig. 3.47 – Enlargement of the scatter plot (PM2.5) generated from the experiments with 20 seconds of injection time

Tighter relationship between the variables compared to the PM10 scatter plots.

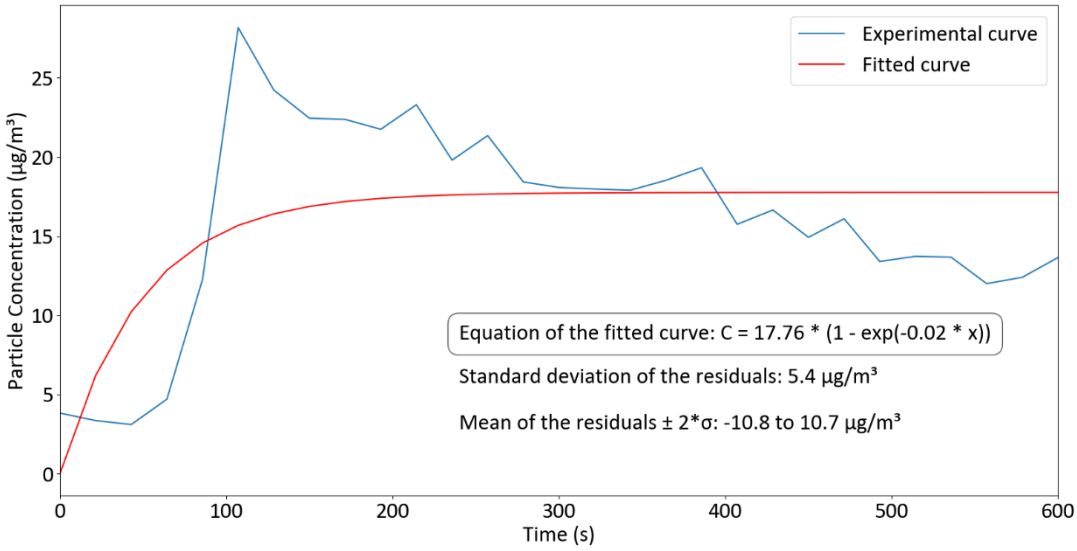

Fig. 3.48 – The PM10 fitted curve generated from the experiments with 30 seconds of injection time

The standard deviation is shorter than the 20 seconds graph, and the uncertainty range is similar but slightly better here.

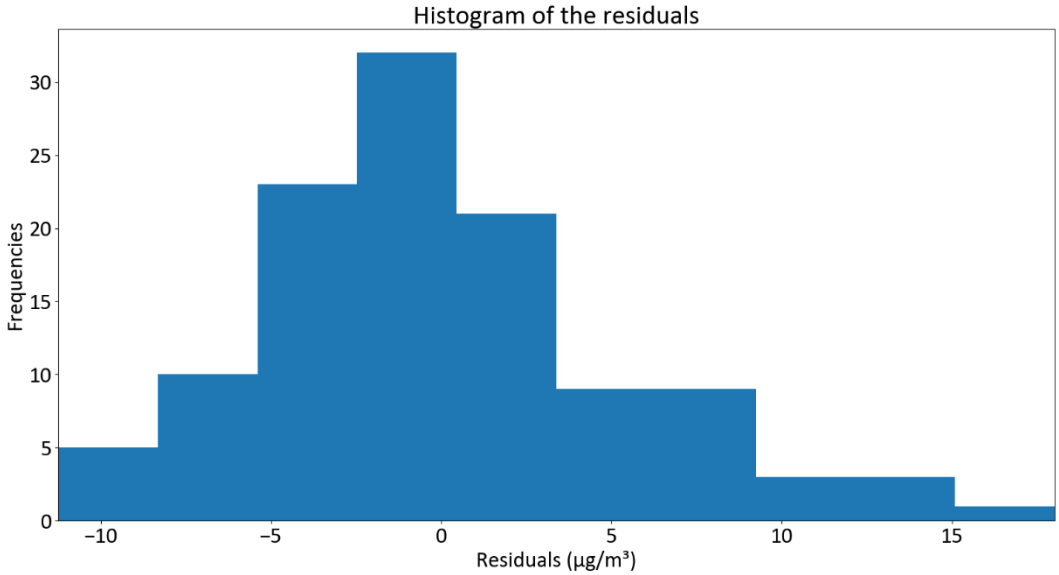

Fig. 3.49 – The histogram of the residuals (PM10) generated from the experiments with 30 seconds of injection time

The histogram of the residuals is very regular in shape, but it's slightly shifted, it's still one of the better ones in this regard.

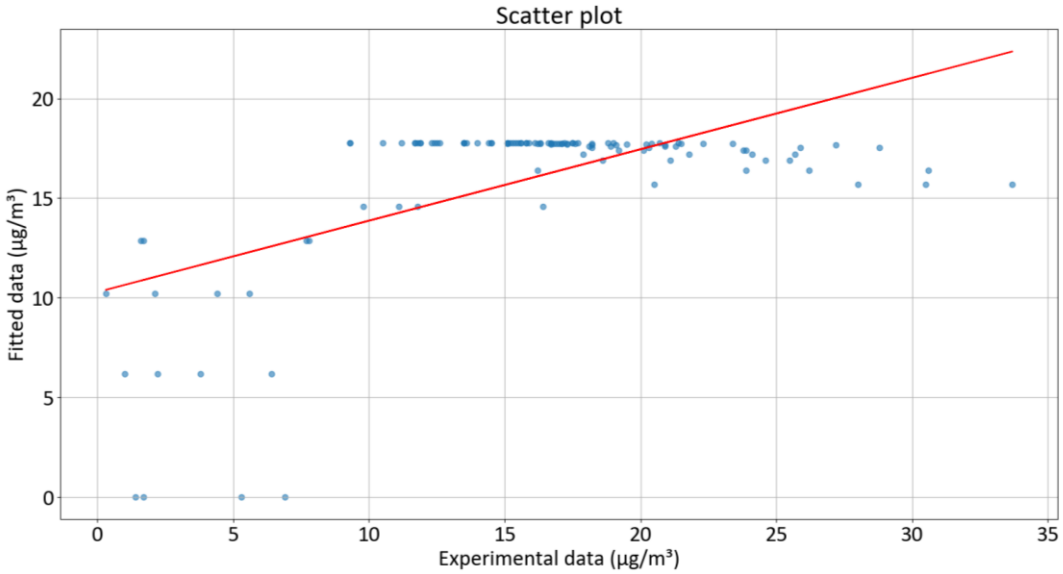

Fig. 3.50 – The scatter plot (PM10) generated from the experiments with 30 seconds of injection time

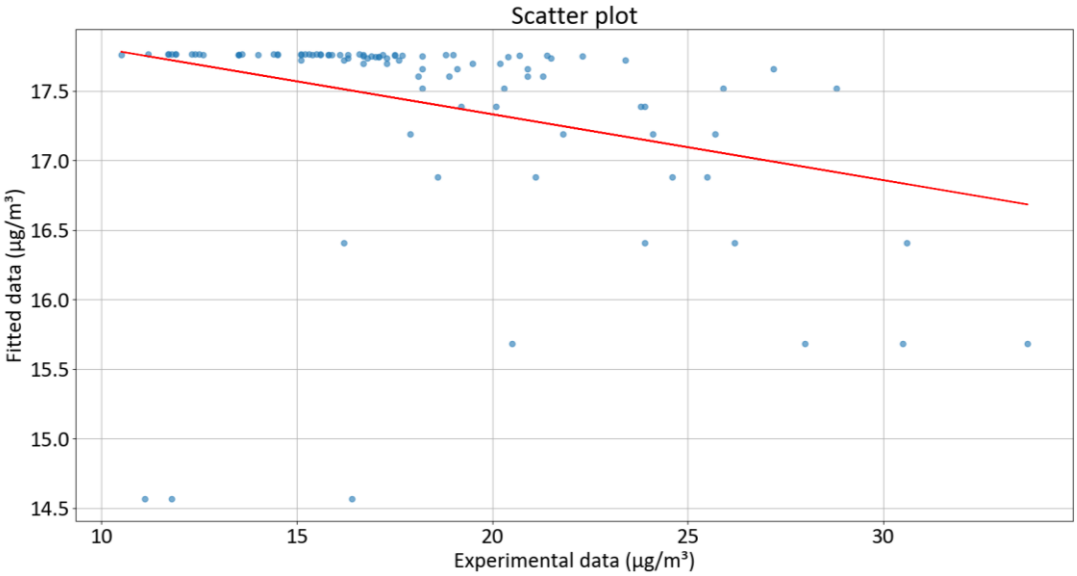

Fig. 3.51 – Enlargement of the scatter plot (PM10) generated from the experiments with 30 seconds of injection time

Plots very similar to the ones of the 20 seconds experiments.

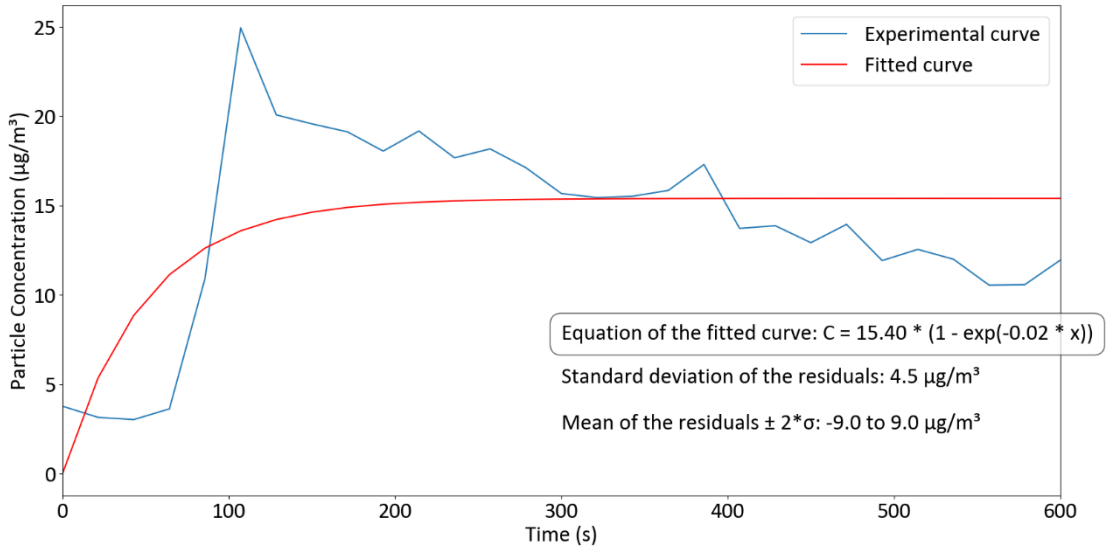

Fig. 3.52 – The PM2.5 fitted curve generated from the experiments with 30 seconds of injection time

Trend very similar to the PM10 one, lower standard deviation.

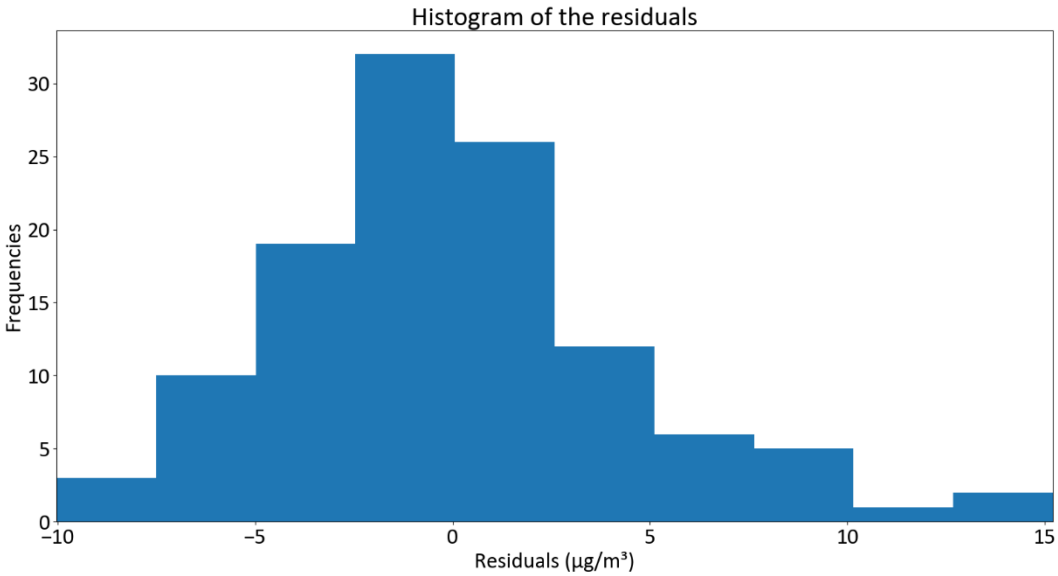

Fig. 3.53 – The histogram of the residuals (PM2.5) generated from the experiments with 30 seconds of injection time

The histogram isn't quite as regular as the PM10 one, and it's off centered, another plot in countertrend with what was shown before.

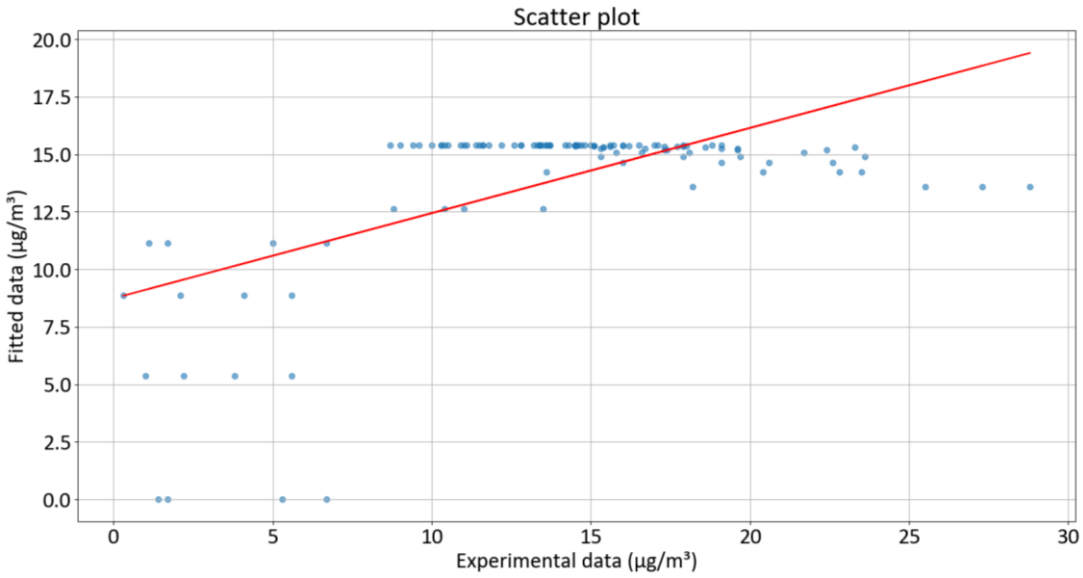

Fig. 3.54 – The scatter plot (PM2.5) generated from the experiments with 30 seconds of injection time

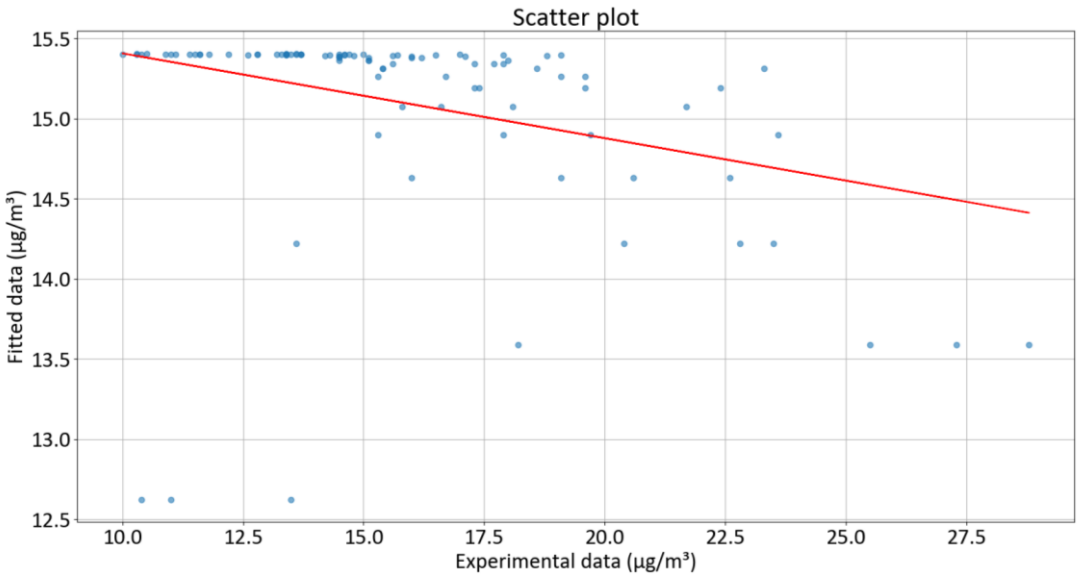

Fig. 3.55 – Enlargement of the scatter plot (PM2.5) generated from the experiments with 30 seconds of injection time

Plots in line with the PM10 trend.

Table 3.11 - Results for PM10 tests using the curve fitting in Python

Injection time (s)	Equation of the fitted curve	Standard deviation of the residuals (µg/m³)	Uncertainty (mean of the residuals ± 2*σ) (µg/m³)
1	$C = 13.57 * (1 - e^{-0.02t})$	4.6	-9.3 to 9.2
5	$C = 13.18 * (1 - e^{-0.02t})$	4.8	-9.6 to 9.5
10	$C = 27.02 * (1 - e^{-0.02t})$	4.8	-9.4 to 9.8
20	$C = 25.64 * (1 - e^{-0.01t})$	5.8	-11.9 to 11.5
30	$C = 17.76 * (1 - e^{-0.02t})$	5.4	-10.8 to 10.7

Table 3.12 - Results for PM2.5 tests using the curve fitting in Python

Injection time (s)	Equation of the fitted curve	Standard deviation of the residuals (µg/m³)	Uncertainty (mean of the residuals ± 2*σ) (µg/m³)
1	$C = 11.46 * (1 - e^{-0.02t})$	3.6	-7.2 to 7.2
5	$C = 10.63 * (1 - e^{-0.02t})$	3.9	-7.8 to 7.7

10	$C = 23.19 * (1 - e^{-0.02t})$	4.1	-7.9 to 8.4
20	$C = 20.88 * (1 - e^{-0.01t})$	4.9	-10.0 to 9.6
30	$C = 15.40 * (1 - e^{-0.02t})$	4.5	-9.0 to 9.0

Another possibility given by the implementation of our mathematical model is that we can use the inverse equation:

$$C = a * (1 - e^{-bt})$$

$$t = -\frac{1}{b} ln\left(1 - \frac{C}{a}\right)$$

to extract the time needed to reach a certain concentration, obviously the time extracted will have the same uncertainty of the fitted curve, but it could be useful to understand the matter from a different facet, considering that the curve of our mathematical model has an asymptotic trend the time past to reach that zone of the graph is a peculiarity of every experiment. Python estimates the a and b variables basing on the experimental data, a possible development is that predictive algorithms could estimate a and b before the tests basing on real-life data, giving another possibility to the tester.

4. Conclusion

The goal of this project is to develop a calibration bench for low cost particulate matter sensors for air quality measurement, and to develop a mathematical model that could help to measure the particle concentration. The experiments made, after having developed a dedicated chamber and an experimental procedure, represent the core of this work. The tests done suggested that longer operating times of the atomizer don't lead to significant higher concentrations registered in the chamber, on the contrary the tests with the intermediate operating times (10 and 20 seconds) revealed as those with the highest concentrations registered, 38 and 43 $\mu g/m^3$ respectively and with the lowest standard deviations recorded, between 2.2 and 3.1 $\mu g/m^3$ for PM10 and among 1.9 and 2.3 $\mu g/m^3$ for PM2.5 10 seconds experiments. The highest concentrations registered in the chamber, 40 and 43 $\mu g/m^3$, were compatible with the range indicated by the Environmental Protection Agency as the standard for good air quality (daily average maximum concentration of 54 $\mu g/m^3$). The mathematical model developed allowed us to use curve fitting in Python and it was used to estimate how much our experimental data came closer to the theoretical curve expected from the model. The developed mathematical model is a simplification of reality since the hypothesis of no particulate leakage from the chamber is not respected because the tube that links the atomizer and the chamber is a source of leakage. The analysis of the results explained that the model, although not super accurate with standard deviations of the residuals between 3.6 and 5.8 $\mu g/m^3$, and uncertainties ranging from a lowest of -7.2 to 7.2 $\mu g/m^3$ to a highest of -11.9 to 11.5 $\mu g/m^3$, showed good premises. Unlike what was found before, the tests that recorded the lowest standard deviations and the lowest uncertainties are those done with the shortest injection time of the atomizer, 1 second, with standard deviations of the residuals of 4.6 $\mu g/m^3$ for PM10 tests and 3.6 $\mu g/m^3$ for PM2.5 tests, with uncertainties ranging from -9.3 to 9.2 $\mu g/m^3$ for PM10 experiments and -7.2 to 7.2 $\mu g/m^3$ for PM2.5 tests, the curve fitting algorithm works better with a shorter operating time of the atomizer. This work tried to evaluate the performances of particulate matter low cost sensors using an experimental chamber, which is relatively new concept, and the results showed that good performances are already obtainable. Considering that the technology is growing really fast, probably we aren't very far away from the propagation of reliable air quality measurement instruments that exploit low cost sensors, and it could be a revolution.

A future development on the entire project is the evaluation of the influence of parameters such as temperature and relative humidity on the concentrations registered in the tests, the work

done highly suggests that they could play a big role, another possible future evolution could be the implementation of predictive algorithms in order to forecast the values of the two variables that define the equation of the model without any real-life testing.

Summary in Italian

Lo scopo di questo progetto era quello di stimare le prestazioni di sensori a basso costo sviluppando un banco di calibrazione per sensori di particolato, e l'obiettivo principale era la prototipazione e lo sviluppo di un modello matematico che potesse aiutarci a misurare le concentrazioni di particelle al fine di ottenere curve matematiche ottenute dai test condotti con l'obiettivo di stimare le prestazioni dei sensori testati, trattandosi di un argomento relativamente recente, non ci sono molti standard stabiliti quindi abbiamo cercato di elaborare una procedura sperimentale che potesse aiutarci in questo percorso. L'obiettivo degli esperimenti era iniettare materiale particolato nella camera utilizzando l'atomizzatore per misurare la risposta del sensore a basso costo, visualizzata utilizzando uno script Python eseguito dal PC collegato alla scheda Arduino. L'unica variabile degli esperimenti è stata l'intervallo temporale dell'iniezione delle particelle, per evidenziare le differenze nella concentrazione misurata. Non esistendo uno standard per queste prove, si è deciso di far durare ogni esperimento dieci minuti, privilegiando un tempo più breve con l'obiettivo di fare più prove, dato che era necessario attendere dopo ogni esperimento che la concentrazione scendesse al livello valori iniziali. È stato sempre utilizzato lo stesso schema sperimentale utilizzando un cronometro per assicurarsi che il metodo fosse coerente: primo minuto in condizioni stazionarie, con l'atomizzatore spento, quindi con una concentrazione di PM nella scatola prossima o uguale a zero. Successivamente è stato acceso l'atomizzatore per il tempo previsto (1, 5, 10, 20 o 30 secondi), infine monitoraggio costante (ad atomizzatore spento) per raggiungere la fine del limite di 10 minuti. È stato preparato uno script Python per tracciare le curve, lo script legge i dati dalla scheda Arduino utilizzando il modulo seriale, ha tracciato in tempo reale le curve per entrambe le concentrazioni di PM10 e PM2.5 nella scatola utilizzando il modulo matplotlib, grafici con il tempo sull'ascissa (in secondi) e la concentrazione delle particelle sull'ordinata (in $\mu g/m^3$). Dopo 10 minuti, i grafici finali potrebbero essere salvati come immagine. Lo script, al termine di ogni esperimento, generava anche un file csv (utilizzando il modulo pandas) con ogni valore di concentrazione registrato dal sensore durante il test e il tempo associato. Per l'analisi dei dati i parametri scelti sono stati: i valori massimi e minimi registrati negli esperimenti per individuare l'intervallo di valori registrati in una singola prova; la concentrazione media delle particelle, ma poiché nel primo minuto l'atomizzatore resta spento, la media è stata calcolata escludendo i valori registrati in quel periodo di tempo; la deviazione standard. Volevamo sviluppare un modello matematico che potesse aiutarci ad analizzare i risultati dei nostri esperimenti, a stimare le differenze tra i dati attesi (generati dal modello

matematico) e i dati sperimentali (generati dai nostri test). L'idea era di sfruttare gli algoritmi di fitting in Python per generare la curva (la curva adattata) basata sui nostri dati sperimentali, calcolando le variabili a e b nel processo, dando quindi la possibilità di verificare quanto fossero vicini i dati dei nostri esperimenti rispetto al modello. Abbiamo provato a implementare le curve sperimentali del modello matematico sviluppato utilizzando la libreria Python gratuita e open source SciPy. Un'altra possibilità data dallo sviluppo del nostro modello è quella di poter calcolare il tempo necessario per raggiungere una concentrazione fissata C all'interno della camera, utilizzando l'equazione inversa del nostro modello matematico. Prima dell'analisi dei dati abbiamo cercato di migliorare quanto avevamo raccolto, poiché avevamo troppe concentrazioni diverse in tempi brevi, abbiamo pensato che una riduzione dei dati considerati potesse semplificare notevolmente l'analisi, abbiamo optato per costruire delle curve "mediate" che utilizzano questo schema: ogni 10 punti consecutivi della curva (o 10 valori di concentrazione consecutivi nel file csv, stessa cosa) sono stati sostituiti dalla media di quei 10 valori, praticamente 1 punto medio ha preso il posto di 10 punti consecutivi. L'andamento delle curve ottenute nel corso degli esperimenti è simile, si nota un calo delle concentrazioni registrato nell'ultima fase delle prove, idealmente l'andamento sarebbe dovuto rimanere più stabile, si notano alcune fuoriuscite di particolato che influenzano le prove, in particolare il tubo flessibile che collega l'atomizzatore e la camera è una fonte di perdite in questo senso. I test effettuati hanno suggerito che tempi di funzionamento più lunghi dell'atomizzatore non portano a concentrazioni significativamente più elevate registrate nella camera, al contrario i test con tempi di funzionamento intermedi (10 e 20 secondi) si sono rivelati quelli con le concentrazioni più elevate registrate, 38 e 43 $\mu g/m^3$ rispettivamente e con le deviazioni standard più basse registrate, tra 2.2 e 3.1 $\mu g/m^3$ per PM10 e tra 1.9 e 2.3 $\mu g/m^3$ per PM2.5 negli esperimenti da 10 secondi. Le concentrazioni più alte registrate nella camera, 40 e 43 $\mu g/m^3$, sono compatibili con l'intervallo indicato dall'Environmental Protection Agency come standard per una buona qualità dell'aria (concentrazione massima media giornaliera di 54 $\mu g/m^3$). Il modello matematico sviluppato ci ha permesso di utilizzare l'adattamento della curva in Python ed è stato utilizzato per stimare quanto i nostri dati sperimentali si avvicinassero alla curva teorica prevista dal modello. Il modello matematico sviluppato è una semplificazione della realtà in quanto l'ipotesi di assenza di fuoriuscite di particolato dalla camera non viene rispettata poiché il tubo che collega l'atomizzatore e la camera è fonte di perdita. L'analisi dei risultati ha spiegato che il modello, sebbene non estremamente accurato con deviazioni standard dei residui tra 3.6 e 5.8 $\mu g/m^3$ e incertezze che vanno da un minimo di -7,2 a 7,2 $\mu g/m^3$ a un massimo di -11,9 a 11,5 $\mu g/m^3$, ha mostrato buone premesse. A differenza

di quanto riscontrato prima, i test che hanno registrato le deviazioni standard più basse e le incertezze più basse sono quelli effettuati con il tempo di iniezione dell'atomizzatore più breve, 1 secondo, con deviazioni standard dei residui di 4.6 $\mu g/m^3$ per i test PM10 e 3.6 $\mu g/m^3$ per test PM2.5, con incertezze che vanno da -9.3 a 9.2 $\mu g/m^3$ per esperimenti PM10 e da -7.2 a 7.2 $\mu g/m^3$ per test PM2.5, l'algoritmo di fitting della curva funziona meglio con un tempo di funzionamento più breve dell'atomizzatore.

References

[1] European Environment Agency, *https://www.eea.europa.eu/publications/environmental-burden-of-cancer/air-pollution.*

[2] European Commission, *https://environment.ec.europa.eu/topics/air/air-quality/eu-air-quality-standards_en.*

[3] European Parliament*, https://www.europarl.europa.eu/legislative-train/theme-a-european-green-deal/file-revision-of-eu-ambient-air-quality-legislation.*

[4] Environment Protection Agency*, https://www.epa.gov/criteria-air-pollutants/naaqs-table.*

[5] Wikipedia, *https://en.wikipedia.org/wiki/Air_quality_index#United_States.*

[6] VITALITY, *https://fondazionevitality.it/.*

[7] *Huang* Zhang, Renhui Ruan, Shruti Choudhary, Houzhang Tan, Pratim Biswas, *Numerical and experimental investigation on the performance of a ventilated chamber for low-cost PM sensor calibration*, Journal of Aerosol Science, 2020.

[8] T. Sayahi, D. Kaufman, T. Becnel, K. Kaur, A.E. Butterfield, S. Collingwood, Y. Zhang, P.E. Gaillardon, K.E. Kelly, *Development of a calibration chamber to evaluate the performance of low-cost particulate matter sensors,* Environmental Pollution, 2019.

[9] Vanessa Schwarstzhaupt Gamboa, Eder Julio Kinast, Marçal Pires, *System for performance evaluation and calibration of low-cost gas sensors applied to air quality monitoring*, Atmospheric Pollution Research, 2022.

[10] Nam H. Nguyen, Huy X. Nguyen, Thuan T. B. Le, Chinh D. Vu, *Evaluating low-cost commercially available sensors for air quality monitoring and application of sensor calibration methods for improving accuracy*, Open Journal of Air Pollution, 2021.

[11] Rosa Amalia González Rivero, Luis Ernesto Morera Hernández, Olivier Schalm, Erik Hernández Rodríguez, Daniellys Alejo Sánchez, Mayra C. Morales Pérez, Vladimir Nuñez Caraballo, Werner Jacobs and Alain Martinez Laguardia, *A low-cost calibration method for temperature, relative humidity, and carbon dioxide sensors used in air quality monitoring systems*, MDPI, 2023.

[12] Vasileios Papapostolou, Hang Zhang, Brandon J. Feenstra, Andrea Polidori, *Development of an environmental chamber for evaluating the performance of low-cost air quality sensors under controlled conditions*, Atmospheric Environment, 2017

[13] Di Liua, Qiang Zhangb, Jingkun Jiangb, Da-Ren Chena, *Performance calibration of low-cost and portable particulate matter (PM) sensors*, Journal of Aerosol Science, 2017.

[14] Houxin Cui, Ling Zhang, Wanxin Li, Ziyang Yuan, Mengxian Wu, Chunying Wang, Jingjin Ma, Yi Li, A new calibration system for low-cost Sensor Network in air pollution monitoring, Atmospheric Pollution Research, 2021.

[15] S.L. Brown, C.S. Goulsbra, M.G. Evans, T. Heath, E. Shuttleworth, *Low cost CO_2 sensing: a simple microcontroller approach with calibration and field use,* HardwareX, 2020.

[16] Carlos Muñoz, Juan Huircan, Francisco Jaramillo, Álex Boso, *Calibration of Sensor Network for Outdoor Measurement of PM2.5 on High Wood-Heating Smoke in Temuco City,* MDPI, 2023.

[17] Sergio Palomeque-Mangut, Félix Meléndez, Jaime Gomez-Suarez, Samuel Frutos-Puerto, Patricia Arroyo, Eduardo Pinilla-Gil, Jesús Lozano, *Wearable system for outdoor air quality monitoring in a WSN with cloud computing: Design, validation and deployment,* Chemosphere, 2022.

[18] Adafruit, https://www.adafruit.com/product/4632.

[19] Toppr, https://www.toppr.com/ask/question/what-is-an-atomiser/.

[20] TSI, https://tsi.com/products/aerosol-generators-dispersers/polydisperse-generators/six-jet-atomizer-9306/.

[21] Arduino, https://store.arduino.cc/products/arduino-mega-2560-rev3.

[22] National Library of Medicine, https://www.nlm.nih.gov/oet/ed/stats/02-900.html.

[23] Indeed, https://www.indeed.com/career-advice/career-development/mathematical-modeling.

[24] Li, Jie, et al. *Emission characteristics and assessment of potential health risks on PM2.5-bound organics from incense burning*, Atmospheric Pollution Research, 2022.

[25] Wikipedia, https://en.wikipedia.org/wiki/Errors_and_residuals.

[26] OriginLab, https://www.originlab.com/doc/Origin-Help/Residual-Plot-Analysis.

[27] Atlassian, https://www.atlassian.com/data/charts/what-is-a-scatter-plot.

[28] Medium, https://medium.com/@VitorCSampaio/understanding-ordinary-least-squares-ols-the-foundation-of-linear-regression-1d79bfc3ca35.

[29] Wikipedia, https://en.wikipedia.org/wiki/Outlier.

[30] Office for National Statistics, https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/uncertaintyandhowwemeasureit.