



DEVELOPMENT OF A TOMOGRAPHIC ATMOSPHERIC MONITORING SYSTEM BASED ON DIFFERENTIAL OPTICAL ABSORPTION SPECTROSCOPY

RUI FILIPE CANTARINO VALENTE DE ALMEIDA

Master in Biomedical Engineering

BIOMEDICAL ENGINEERING

NOVA University Lisbon
September, 2021



NOVA

NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTMENT
OF PHYSICS

DEVELOPMENT OF A TOMOGRAPHIC ATMOSPHERIC MONITORING SYSTEM BASED ON DIFFERENTIAL OPTICAL ABSORPTION SPECTROSCOPY

RUI FILIPE CANTARINO VALENTE DE ALMEIDA

Master in Biomedical Engineering

Adviser: Prof. Dr. Pedro Vieira
Auxilliary Professor, NOVA University Lisbon

BIOMEDICAL ENGINEERING

NOVA University Lisbon
September, 2021

Development of a Tomographic Atmospheric Monitoring System based on Differential Optical Absorption Spectroscopy

Copyright © Rui Filipe Cantarino Valente de Almeida, NOVA School of Science and Technology, NOVA University Lisbon.

The NOVA School of Science and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

To my parents.

ACKNOWLEDGEMENTS

First and foremost, I need to thank FCT NOVA and FutureCompta, for having me and letting me work with them and develop this thesis. In addition, I want to acknowledge the contribution from the Portuguese Foundation for Science and Technology, who funded this work through the NOVA I4H doctoral programme, grant PDE/B-DE/114549/2016.

I would also like to thank my advisor, Prof. Pedro Vieira, for all the support he has given me, on many levels and throughout the years.

To my friends, old and new, I extend my utmost gratitude for helping me tilt my head back and laugh once in a while.

Finally, to my family, without whom I would have never been able to do anything, much less something as significant as what I present in this dissertation.

“I know the pieces fit, 'cause I watched them fall away. ”
(TOOL - Schism)

“Silenzio, Bruno! ” (Luca Paguro)

ABSTRACT

The aim of this thesis is to describe the design and development of a proof of concept for a commercially viable large area atmospheric analysis tool, for use in trace gas concentration mapping and quantification.

Atmospheric monitoring is a very well researched field, with dozens of available analytical systems and subsystems. However, current systems require a very important compromise between spatial and operational complexity. We address this issue asking how we could integrate the [Differential Optical Absorption Spectroscopy \(DOAS\)](#) atmospheric analysis technique in a [Unmanned Aerial Vehicle \(UAV\)](#) with tomographic capabilities.

Using a two-part methodology, I proposed two hypothesis for proving the possibility of a miniaturised tomographic system, both related to how the spectroscopic data is acquired. The first hypothesis addresses the projection forming aspect of the acquisition, its matrix assembly and the resolution of the consequent equations. This hypothesis was confirmed theoretically by the development of a simulation platform for the reconstruction of a trace gas concentration mapping.

The second hypothesis deals with the way in which data is collected in spectroscopic terms. I proposed that with currently available equipment, it should be possible to leverage a consequence of the Beer-Lambert law to produce molecular density fields for trace gases using passive [DOAS](#). This hypothesis was partially confirmed, with definite conclusions being possible only through the use of complex autonomous systems for improved accuracy.

This work has been a very important first step in the establishment of [DOAS](#) tomography as a commercially viable solution for atmospheric monitoring, although further studies are required for definite results. Moreover, this thesis has conducted to the development of a [DOAS](#) software library for Python that is currently being used in a production environment. Finally, it is important to mention that two journal articles were published from pursuing this work, both in important journals with Impact Factors over 3.0.

Keywords: [DOAS](#), tomography, [UAV](#), drones

RESUMO

Era o objectivo deste trabalho descrever o processo de desenho e implementação de uma prova de conceito para um sistema de avaliação atmosférica comercialmente viável, para uso no mapeamento das concentrações de compostos traço na atmosfera.

A avaliação atmosférica é um campo muito estudado, estando no presente momento disponíveis para instalação diversos sistemas e subsistemas com estas capacidades. No entanto, é marcante o compromisso que se verifica entre a resolução espacial e a complexidade operacional destes equipamentos. Nesta tese, desafio este problema e levanto a questão sobre como se poderia desenvolver um sistema com os mesmos fins, mas sem este premente compromisso.

Usando uma metodologia a duas partes, proponho duas hipóteses para comprovar a exequibilidade deste sistema. A primeira diz respeito à formação da matriz tomográfica e à resolução das equações que dela derivam e que formam a imagem que se pretende. Confirmei esta hipótese teoricamente através do desenvolvimento de uma plataforma de simulação para a reconstrução tomográfica de um campo de concentrações fantoma.

A segunda é dirigida a aquisição de dados espectroscópicos. Proponho que com o material presentemente disponível comercialmente, deverá ser possível aproveitar uma consequência da lei de Beer-Lambert para retirar os valores de concentração molecular de gases traço na atmosfera. Foi apenas possível validar esta hipótese parcialmente, sendo que resultados mais conclusivos necessitariam de equipamentos automatizados dos quais não foi possível dispôr.

No final, este trabalho constitui um importante primeiro passo no estabelecimento da técnica de [DOAS](#) tomográfico como uma alternativa comercialmente viável para a análise atmosférica. Ademais, o desenvolvimento desta tese levou à escrita de uma biblioteca em Python para análise de dados [DOAS](#) actualmente usada em ambiente de produção. Por fim, importa realçar que dos trabalhos realizados no decorrer da tese foram publicados dois artigos em revistas científicas com *Impact Factor* acima de 3.

Palavras-chave: [DOAS](#), tomografia, drones

CONTENTS

List of Figures	xi
List of Tables	xv
List of Listings	xvi
Acronyms	xvii
1 Background, Motivation and Introduction	1
1.1 Starting Points	1
1.2 Research Questions	2
1.3 Air Pollution	3
1.3.1 Criteria Pollutants	4
1.3.2 Air Pollution Effects on Human Health	7
1.3.3 Air Pollution effects on ecosystems	13
1.3.4 Air Pollution Sources	15
1.4 Air Pollution Monitoring Techniques	23
2 Literature Review	25
2.1 DOAS Tomography	26
3 Theoretical Background	30
3.1 Unmanned Aerial Vehicle	31
3.2 Optical Systems	32
3.3 DOAS	34
3.3.1 Types of DOAS experiments	37
3.4 Important Notes on DOAS in practice	41
3.4.1 Cross section conditioning	41
3.4.2 Spectral Calibration	42
3.5 Tomographic algorithms and reconstruction techniques	44
4 Methods	51
4.1 First Hypothesis	53
4.1.1 Unmanned Aerial Vehicle	53

CONTENTS

4.1.2	Collection System	54
4.1.3	Ground Station	58
4.1.4	Tomosim	59
4.2	Second Hypothesis	67
4.2.1	The Lambertian Hypothesis	67
4.2.2	Testing the Lambertian Hypothesis	69
5	Discussion	72
5.1	First Hypothesis	72
5.1.1	The UAV	72
5.1.2	Collection System	73
5.1.3	TomoSim	73
5.2	Second Hypothesis	73
5.2.1	Experiment - First Run	73
5.2.2	Experiment - Second Run	73
	Bibliography	74
	Appendices	
	Annexes	

LIST OF FIGURES

1.1	Bimodal distribution for particle volume and the phenomena that lead to their formation [Vallero2014]	5
1.2	Ozone mean and peaks in European rural sites. Peaks are plotted blue, mean is plotted green. Note the decrease in peak concentration and the steadiness of the mean signal.	6
1.3	Annotated anatomy of the respiratory system [Vallero2014].	8
1.4	Possible abnormalities caused by Air Pollution (AP) exposure <i>in utero</i> . Notice that time of exposure is of critical importance [Vallero2014].	11
1.5	Developmental stages of the lung throughout life vs the risks of AP exposure in each stage [Vallero2014].	12
1.6	Dramatic photograph depicting the Capelinhos' lighthouse, half a kilometer from the eruption site, surrounded by a cloud of ash Particulate Matter (PM), volcanic gas and water vapor with more than 1km in height[TSF2017].	16
1.7	House almost completely buried by the Black Sunday dust storm. Several houses were entirely swallowed during this storm, trapping people inside, as if a big blizzard had hit them. Unlike a blizzard though, there was nothing anyone could do to keep the dirt outside, and all surfaces were covered black [Reis2008].	17
1.8	Carbon Dioxide atmospheric concentrations since the year 1000. Note the seemingly exponential increase since the 1800s. Plotted and published by the 2 Degrees Institute [co2levels2020] with data from ice cores [Etheridge] and in situ monitors [Tans].	18
1.9	Schematic presentation on the sources of anthropogenic pollution and its categorisation according to the IPCC. Adapted from [CABI2019]	19
1.10	Chinese energy production, Gross Domestic Product (GDP) and Carbon Dioxide (CO ₂) emissions. Data collected from the World Bank and International Energy Agency websites [WorldBank, IEA]	20
1.11	General trends for European emissions. Data presented in % emissions of year 2000. Note the downward global trend in pollutant emissions, and its decoupling with the European GDP [EEA2019].	22
1.12	European emissions divided by activity sector. The global decreasing trend is confirmed, as industries all around are producing less and less AP with the passing years [EEA2019].	22

LIST OF FIGURES

1.13	A tree diagram providing a general overview of the various types of atmospheric pollution monitoring systems.	24
2.1	Simplified diagram of the Systematic Mapping Study (SMS) method. This type of study is focused on registering every step of the search and the reading of the retrieved material, in an effort to make the process completely repeatable. Adapted from [Kitchenham2007].	27
2.2	BABII assembly geometry. In this experiment campaign, the telescopes illuminated retroreflecting targets that were positioned in two steel towers on both sides of a busy motorway in Germany, connecting Heidelberg to Mannheim [Pundt2005a].	28
2.3	Erna Frins's paper [Frins2006] proposes a very relevant passive DOAS application that can conceptually be employed in a DOAS tomography scenario. In this 2006 paper, the authors use multi-axis measurements of sun-illuminated targets to estimate absorption paths without using radiative transfer models.	29
3.1	Simplified schematic flowchart of the DOAS algorithm, including the non-linear part.	38
3.2	Active DOAS schematic.	38
3.3	Passive DOAS schematic.	39
3.4	Several examples of possible passive DOAS experiment geometries. All these examples use only the light from an astronomical source to detect many atmospheric trace gases. All these examples were taken from [Platt2007].	39
3.5	Occultation measurement schematic representation [Platt2007]	40
3.6	Schematic representation of the limb satellite measurement geometry.	40
3.7	Schematic representation of the nadir satellite measurement geometry.	41
3.8	Before being used in DOAS calculations, literature trace gas cross section spectra must undergo a conditioning process to make sure they are compatible with the current experiment's instruments. Figure adapted from [Beekman2010]	43
3.9	This figure, taken from [Beekman2010], shows the difference between NO ₂ cross sections. The red line comes from the literature, and has been conditioned to the experiment spectrometer. The blue line was actually collected with that device. The slight discrepancies that can be observed are calibration defects.	43
3.10	High resolution solar spectrum commonly used to calibrate the various spectral signals in DOAS experiments [Chance2010].	45
3.11	Schematic representation for coordinate setting. The image depicts a parallel projection setting [Kak2001a].	46
3.12	The Fourier Slice Theorem (FST), a schematic representation [Asl2013a].	48
3.13	Schematic representation of an equiangular fan-beam projection, taken from [Kak2001].	50
4.1	General diagram of the atmospheric monitoring system, in an overview style.	52

4.2	Schematic representation of the intended pollution measurement geometry. The second hypothesis of the Research Question (RQ) tells us that to get the Region Of Interest (ROI) we can subtract the concentrations obtained on the left of the ROI from the ones obtained to the right of the ROI.	52
4.3	A schematic for the custom-designed drone system on which the pollution monitoring system would be mounted.	53
4.4	Conducted experiment to evaluate advantages of using a direct light connection instead of fibre optics. The spectrometer support (that can be seen in red) was designed in Autodesk Inventor and 3D printed.	55
4.5	Plots obtained through the collection of light managed by the assemblies depicted in Figure 4.4	55
4.6	Telescope - Spectrometer coupling. 3D design rendered through Autodesk Inventor.	56
4.7	The spectrum captured on a cloudy but luminous day using the ET90 telescope and the Avantes Mini spectrometer.	57
4.8	UML diagram for the DOAS library. The Object Oriented Programming (OOP) approach that was followed allows for an instrument-oriented experiment parametrisation, which is not available in any other software	59
4.9	A flowchart describing the beginning stages of the Tomosim simulation.	60
4.10	Flowchart representing the Tomosim operation during the reconstruction phase.	60
4.11	Illustration of the projection gathering algorithm based on the assembly of groups of acquisitions as fanbeam projections. The lines presented on the drawing come from the right to the left. Each is a ray in fan, corresponding to one stop of the drone. In the second measurement moment, the drone moves to the exit point on the left and repeats the first moment collection. Here, both fans ans rays are separated by 5 degrees and there are 6 rays within each fan. A "real life"acquisition would feature a lot more rays, but that would be graphically complicated for the reader. Both ray and fan angular distance are customisable at runtime. Map taken from Google Maps in 2019. ©Google.	61
4.12	P_2 Calculation. The software uses the position of the drone, as defined by the vertical angle, β , and the distance between the drone and the centre of the trajectory, D , to determine P_2 through the solution of a second degree equation.	62
4.13	A graphical representation of the new spectral phantom, custom built for the TomoSim application.	64
4.14	Error estimation graphical representation. Note errors are extremely exaggerated for visualisation purposes.	65
4.15	First draft of the Unified Modelling Language (UML) diagram for the new architectural model of the Tomosim library. Note the usage of several classes from the DOASlibrary described in Section 4.1.3.	67
4.16	Location of observer points for the physical experiment.	68
4.17	Published spectrum of the CREE XHP50.2 LED light. The LED that was used in this experiment corresponds to the blue line [CREE2021].	69

LIST OF FIGURES

LIST OF TABLES

1.1	Main research question.	3
1.2	Secondary research questions.	3
1.3	Main avenues for the formation of ozone in the troposphere.	5
1.4	Global energy production, divided according to the fuel used to obtain it and the production country.	20
4.1	Specifications table for the Omegon MightyMak 60 telescope [Omegon2021].	57
4.2	Table summarising the new phantom's construction details, as a sum of 5 Gaussian profiles and an ellipse designed using TomoPhantom. In this table, Co is the object's amplitude, Xo and Yo are its center coordinates, and a and b are the objects half-widths. The table is constructed using TomoPhantom's particular syntax and more information can be obtained at [Kazantsev2018].	63
4.3	Summary table for the two experiment assemblies. Note the difference in terms of material, due to the two different roles both assemblies play during the experiment. This is translated into not having the need of an artificial light source in the West Bank's assembly.	70
4.4	Actions are the indivisible unit upon which each capture is built. The prescribed actions for this experiment are described in this table.	70
4.5	Captures are particular sets of actions that are conducted according to a specific order, depending on the time of day on which the capture is run. This table describes the prescribed captures on which this experiment consisted.	70
5.1	Component list for the hexacopter that was designed as a part of this thesis.	72

LIST OF LISTINGS

ACRONYMS

ALRI	Acute Lower Respiratory Infections 9
ANS	Autonomic Nervous System 10
AP	Air Pollution ix, xi, 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23
API	Application Programming Interface 63
ASCII	American Standard Code for Information Interchange 70
BAEP	Brainstem Auditory-Evoked Potentials 12
BTU	British Thermal Unit 18
CNS	Central Nervous System 12, 13
CO	Carbon Monoxide 4, 21, 40
CO₂	Carbon Dioxide xi, 18, 20, 40
COPD	Chronic Obstructive Pulmonary Disease 9, 12
CT	Computed Tomography 45
CVD	Cardiovascular Disease 10
CVM	Cardiovascular Mortality 10
DOAS	Differential Optical Absorption Spectroscopy vii, viii, ix, xii, xiii, 1, 2, 3, 4, 7, 26, 27, 28, 29, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 58, 59, 66, 67, 68
EEA	European Environmental Agency 4, 5, 7, 21
EPA	Environmental Protection Agency (United States) 3, 4
ESCAPE	European Study of Cohorts for Air Pollution Effects 9
FBP	Filtered BackProjection 47, 48, 49, 63
FFF	Forest Fire Finder 1, 2, 56
FP₇	European Union's Seventh Framework Programme 9
FST	Fourier Slice Theorem xii, 47, 48
FT	Fourier Transform 47, 48
FWHM	Full Width at Half Maximum 42

ACRONYMS

GDP	Gross Domestic Product xi , 20 , 21
H₂O	Water 40
H₂S	Hydrogen Sulfide 15
ICE	Internal Combustion Engine 7 , 20 , 21
IEP	InfraEstruturas de Portugal 69
IFT	Inverse Fourier Transform 48 , 49
IOT	Internet Of Things 2
LPG	Liquefied Petroleum Gas 21
MAX-DOAS	MultiAxis-DOAS 39
MLEM	Maximized Likelihood Expectation Maximization 63 , 64
MSE	Mean Squared Error 66
NH₃	Ammonia 21
NO	Nitrogen Oxide 7
NO₂	Nitrogen Dioxide 4 , 6 , 7 , 9 , 11 , 14
NO_X	Nitrogen Oxides 7 , 21
O₂	Oxygen 7 , 24
O₃	Ozone 4 , 5 , 7 , 11 , 14
OOP	Object Oriented Programming xiii , 58 , 59 , 66
PAH	Polycyclic Aromatic Hydrocarbons 12
PM	Particulate Matter xi , 4 , 7 , 9 , 10 , 11 , 15 , 16 , 17
Project ATMOS	ATmosphere MOnitoring System Project 1 , 2 , 53
PT2020	Portugal2020 1 , 2
ROI	Region Of Interest xiii , 51 , 52 , 61 , 62 , 66
RPi	Raspberry Pi 54
RQ	Research Question xiii , 2 , 51 , 52
RTK GPS	Real Time Kinematic Global Positioning System 65
RTK-GPS	Real Time Kinematic GPS 54 , 58
SART	Simultaneous Algebraic Reconstruction Technique 63
SGA	Small for Gestation Age 11
SMA	SubMiniature version A 54
SMS	Systematic Mapping Study xii , 27
SO₂	Sulfur Dioxide 4 , 6 , 11 , 14 , 15
UAV	Unmanned Aerial Vehicle vii , ix , x , 31 , 53 , 58 , 72
UML	Unified Modelling Language xiii , 58 , 67
USB	Universal Serial Bus 54 , 70

VOC	Volatile Organic Compound 5 , 17
WFM-DOAS	Weighting Function Modified Differential Optical Absorption Spectroscopy 40
WHO	World Health Organization 5

TODO LIST

link to theory	63
Link to theory	63
link to theory	63
link to theory	64
link to theory	66
Figure: This is actually a table	72

BACKGROUND, MOTIVATION AND INTRODUCTION

1.1 Starting Points

This thesis describes the work that I have done in the past 4 years on the design and development of a miniaturized system for atmospheric monitoring based on optical spectroscopy. The project itself was the major part of the [ATmosphere MOnitoring System Project \(Project ATMOS\)](#), an initiative that was contemplated with European funding through a [Portugal2020 \(PT2020\)](#) initiative and came as a response to the growing weight that [AP](#) has in the whole Western World.

The potential impact of [AP](#) on human health is amply documented. Numerous papers have, for decades, established many links between air quality and several common ailments like respiratory syndromes and cardiovascular diseases. Similar connections have also been found regarding the probability of gestational malformations and several types of cancer. On a different level, and of perhaps less immediate concern, are the effects that have been observed on ecosystems. Frequently these effects are difficult to predict (and timely mitigate) and in some cases have been known to interfere with people's livelihood. If not addressed, these interferences will certainly hinder economies and limit the quality of life of populations globally. The severity of this problem makes it clear that we need to tackle it intelligently, and this approach requires that we can measure, trace and track [AP](#) effectively, which beckons engineers and scientists to create more technology for this specific purpose.

The idea behind this thesis was born in 2015, at NGNS-IS (a Portuguese tech startup). At the time, the company's flagship product was the [Forest Fire Finder \(FFF\)](#). The [FFF](#) was a forest fire detection system, capable of mostly autonomous and automatic operation. The system was the first application of [DOAS](#) for fire detection, and for that it was patented in 2007 (see [[Vieira2007, Application2008](#)]). The [FFF](#) is a remote sensing device that scans the horizon for the presence of a smoke column, sequentially performing a chemical analysis of each azimuth, using the Sun as a light source for its spectroscopic operations [[Valente de Almeida2017](#)].

The [FFF](#) was deployed in several "habitats", both nationally (Parque Nacional da Peneda-Gerês and Ourém) and internationally (Spain and Brazil). One of the company's clients at the time was interested in a pollution monitoring solution, and asked if the spectroscopic system would be capable of performing such a task. The challenge

resonated through the company's structure and the team then started reading about the concept of AP and how both populations and entities were concerned about it. It became clear that, while there were already several methods to measure AP, there was a clear market drive for the development of a system that could leverage the large area capabilities of a DOAS device while being able to provide a more spatially resolved "picture" of the atmospheric status. With this in mind, the company managed to have the investigation financed through a PT2020 funding opportunity. This achievement was a clear validation of the project's goals and of the need there was for a system with the proposed capabilities. It was, however, not enough. FFF was a very good starting point, but there was still a lot of continuous research work needed before any of the goals that had been set were achieved. This led to the publication of this PhD project, in a tripartite consortium between FCT-NOVA, NGNS-IS and the Portuguese Foundation for Science and Technology. Its main goal was to develop an atmospheric monitoring system prototype that would be able to spectroscopically map pollutant concentrations in a two-dimensional way.

In April 2017, NGNS-IS was integrated in the Compta group, one of the oldest IT groups operating in Portugal. Despite its age, this company is one of the main presences in some of the most modern industrial fields, like Internet Of Things (IOT) applications. Project ATMOS's pollutant tracing capabilities made it an almost perfect fit in one of IOT's most resounding niches, the *Smart Cities* trend. Unfortunately, the transition between one company and the other, regardless of the project's adequacy, was anything but smooth. Almost two years later, in the beginning of 2019, engulfed in a sea of endless bureaucracy and ill intent on behalf of the managing governmental authorities (who seemed always more interested in seeing the project fail than anything else), Project ATMOS was terminated and financing was cut.

Obviously, this was a very hard blow to recover from, especially from the perspective of this thesis. After all, most of the work that was planned implied the use of very expensive equipment that was on the verge of being purchased. Since this was not going to be possible, we had to reimagine it. How could the RQ presented in Section 1.2 still be answered when we could not use drones or buy any more telescopes or spectrometers? The point of the thesis thus became to present a modeled proof of concept for the initially proposed system, instead of a minimum viable product.

1.2 Research Questions

This thesis had the defining primary objective of modelling a proof of concept for a spectroscopy-based pollution monitoring system that can be fitted onto a highly mobile platform, such as a drone. The other defining characteristic of the system is that it needs to be able to cover large areas (which could be remote and / or inaccessible) and have a good spatial resolution.

Early in the project's life, the team reached a conceptual milestone: what if it was possible to perform an atmospheric tomographic scan? Preliminary research into the literature indicated this had already been done (focused literature review available in Chapter 2). This started to systematise our objectives.

- To use a tomographic approach for the mapping procedure;

- To ensure the designed system would be small and highly mobile;
- To use a single light collection point, minimizing material costs.

The main research question was thus formed, and is presented in Table 1.1. From it, four secondary research questions are derived, which are presented in Table 1.2.

Table 1.1: Main research question.

RQ1	<i>How to design a miniaturized tomographic atmosphere monitoring system based on DOAS?</i>
------------	---

Table 1.2: Secondary research questions.

RQ1.1	<i>What would be the best strategy for the system to cover a small geographic region?</i>
RQ1.2	<i>What would be the necessary components for such a system?</i>
RQ1.3	<i>How will the system acquire the data?</i>
RQ1.4	<i>What should the tomographic reconstruction look like and how to perform it?</i>

1.3 Air Pollution

Daniel Vallero, in his book "Fundamentals of Air Pollution" [Vallero2014] makes a very important observation: Air Pollution has no universal definition. Its meaning is intertwined with

the context with which it is measured and observed, with the ecosystem in which it is perceived and even with the pollutant concentration (not every toxic compound is toxic at every concentration). The Environmental Protection Agency (United States) (EPA) defines Air Pollution as the following:

Air Pollution is the presence of contaminants or pollutant substances in the air that interfere with human health or welfare, or produce other harmful environmental effects.

He then analyzes this definition through two possible lenses, the one that comes with the interference produced by air contaminants; and the one that comes from the harm they may cause. He notes that both points of view come with a heavy burden of ambiguity, incompatible with a scientific definition. We can thus observe that preferable to address the issue through its measurable effects and consequences. These are well-established and well known, and scientists all around the world have been publishing extensively about them for some decades now. The correlation between Air Pollution and an increased mortality in heavily industrialized areas was first established in Europe, in the 19th century, but the first time it was taken seriously was

during the 1952 killer-smog incidents, in London [Platt2007]. At the time, a combination of very cold weather, an anticyclone and fireplace emissions caused a thick smog to fall over London, directly causing thousands of deaths and indirectly many more [Bell2008, Office2019]. The disastrous consequences of this incident had a huge impact in the civil society, resulting in a series of policies and laws, among which the Clean Air Acts of 1956 and 1968, which are broadly considered to be some of the first actions to decrease pollution in human societies. Much work has been done, and it has resulted in remarkable progress since the definition of those two policies. We are in fact in a much better place than we were some years or decades ago, but pollution is still a part of everyday reality for the whole of civilization. In the current day and age, both European and American regulatory and surveillance bodies (the European Environmental Agency (EEA) and the EPA, respectively) have identified a group of six *criteria pollutants* that need to be monitored effectively.

1.3.1 Criteria Pollutants

Criteria pollutants are a group of six chemical species, commonly found mixed in air, that constitute a serious hazard for human, animal, and general environment health. These pollutants were defined in 1970, through the Clean Air Act in the United States, but have since been widely adopted. These six pollutants are:

1. Particulate Matter;
2. Ozone;
3. Carbon Monoxide;
4. Sulfur Dioxide;
5. Nitrogen Dioxide;
6. Lead (Pb);

They are collectively addressed as criteria pollutants, solely because environmental agencies around the world have since been using them, and their atmospheric concentrations as criteria for setting standards. Of the 6 criteria pollutants, 4 are of direct importance for this dissertation, as DOAS can be used to measure and quantify them. They are PM, Ozone (O_3), Sulfur Dioxide (SO_2) and Nitrogen Dioxide (NO_2). This section aims to provide a brief presentation of these four species.

1.3.1.1 Particulate Matter

Particles are aggregates of many molecules, that can be similar or different. There are several processes in which particles can be formed, and they can be chemically active or, for instance, act as surfaces on which trace gases may condense. The chemical composition of a particle is immensely variable, and it can even be difficult to determine, since one particle sample (usually collected over a period of some hours) can have thousands of species, namely hydrocarbons. One interesting aspect of PM is that its volume distribution is bimodal, having two clear maxima below $1\mu m$ and another

around $10\mu\text{m}$. The reason for this separation is the way in which the particles were formed. Coagulations and condensation tend to form the smaller particles, while the larger peak is comprised mostly of crustal (as in from the Earth's crust) and sea solids, together with some smaller aggregated particles – smaller particles that have adhered onto one another. Figure 1.1 illustrates this bimodal distribution and the responsible formation phenomena.

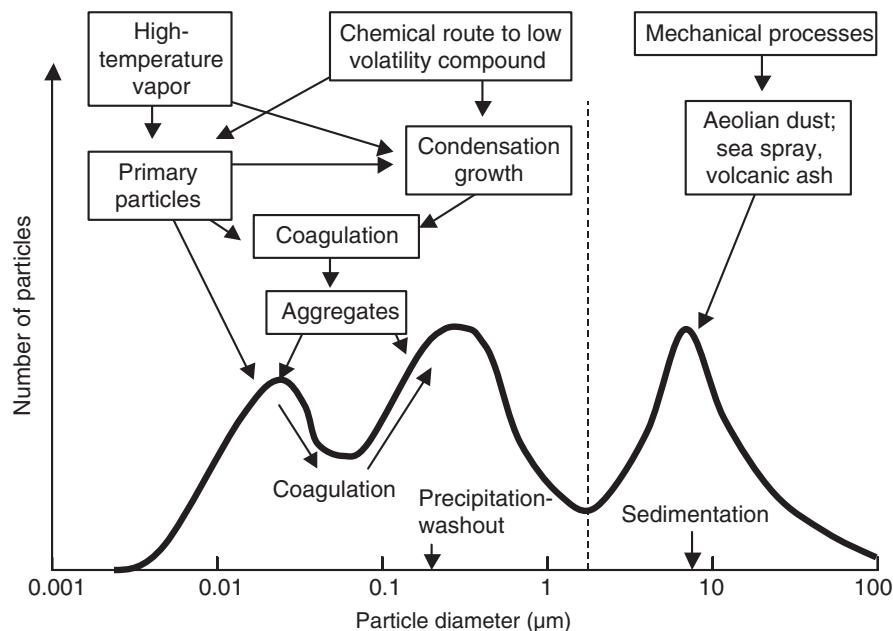


Figure 1.1: Bimodal distribution for particle volume and the phenomena that lead to their formation [Vallero2014]

1.3.1.2 Tropospheric Ozone

Ozone is one of the most important trace gases in the atmosphere, in functional terms. The O_3 concentration in lower troposphere has had a sharp rise in recent decades, which indicates that this rise is anthropogenic in nature [Bourdeau]. In the last EEA reports, European concentrations of O_3 have remained approximately stable and just above the World Health Organization (WHO)-set limit for the protection of human life [EEA2016, Guerreiro2019]. In both reports, the EEA states that although ozone precursor concentrations have been steadily declining, concentrations for O_3 remain the same, although the amplitude of peak events is shown to be decreasing (see Figure 1.2)

Ozone formation in the lower atmosphere is also interesting. While other pollutants are directly emitted by its sources, O_3 is a secondary pollutant, which means that it is formed through chemical reactions that its precursors endure. With the exception of Volatile Organic Compound (VOC), these precursors are results from human activity. There innumerable pathways for ozone formation. Table 1.3 summarizes the main avenues.

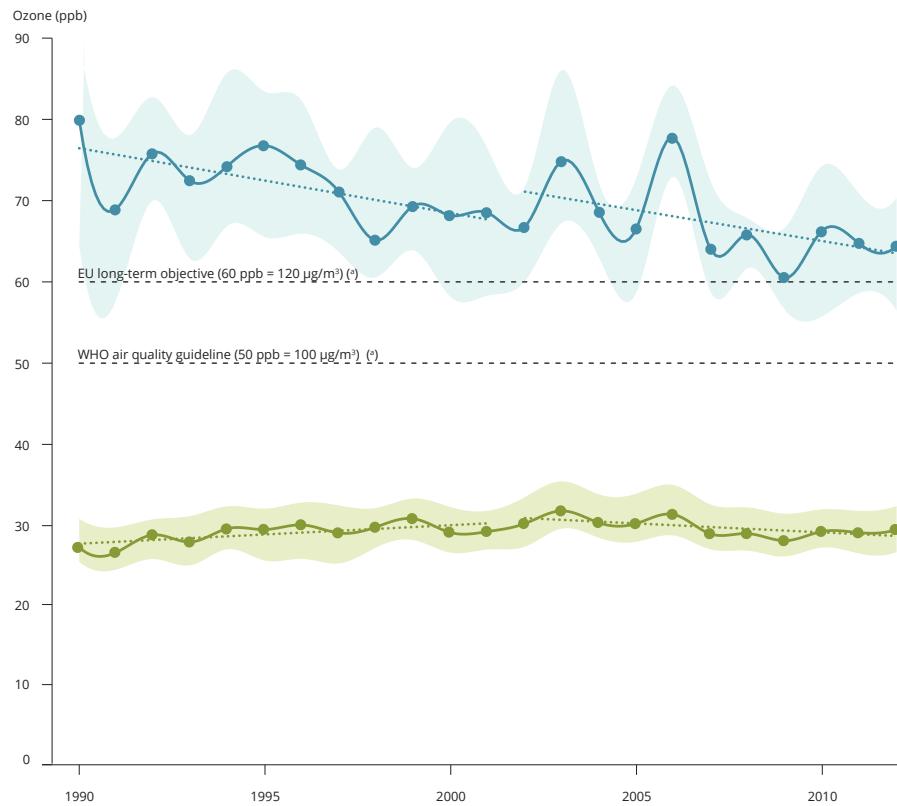


Figure 1.2: Ozone mean and peaks in European rural sites. Peaks are plotted blue, mean is plotted green. Note the decrease in peak concentration and the steadiness of the mean signal.

Table 1.3: Main avenues for the formation of ozone in the troposphere.

1.3.1.3 Sulfur Dioxide

Nitrogen, Sulphur, Phosphorous and Potassium are very familiar to those accustomed to dealing with plant life. For plants, these 4 chemicals are. Their life depends on them. However, from an AP point of view, sulphur and nitrogen compounds are common issues and have to be addressed separately. We will deal with NO_2 in Section 1.3.1.4. Sulphur compounds are released into the atmosphere primarily through anthropogenic activities (for instance, diesel-powered engines), although marine biogenic gases represent a significant proportion of the emitted sulphuric species.

These compounds can be released in reduced and oxidised forms, although reduced forms are generally oxidised onto SO_2 . SO_2 has many complex atmospheric chemical reactions, but proceeds to the sulphate ion through at least three pathways. If given enough time, sulphur oxidises into SO_4^{2-} , usually forming sulphuric acid. This reacts readily with ammonia in the atmosphere, and produces ammonium sulphates that, together with their precursor sulphuric acid, are readily dissolved in water [Bourdeau].

1.3.1.4 Nitrogen Dioxide

NO_2 is the most important atmospheric trace gas as far as this thesis is concerned, for reasons that will become clear in Section ???. Nitrogen oxides are toxic, but that is not the only reason why they are problematic. In fact, they are one of the main precursors of tropospheric ozone, namely through photolysis, as described in Equation 1.1.



M , in the case of Equation 1.1, is any third body that absorbs excess vibration energy and can stabilise the newly formed molecule. But O_3 can go on to react with Nitrogen Oxide (NO) and regenerate NO_2 and Oxygen (O_2), as in Equation 1.2.



Now, these reactions and several others that involve nitrogen species occur whether there is pollution or not, which results in an equilibrium O_3 concentration of 30 ppb. In a polluted atmosphere, say by Internal Combustion Engine (ICE), this balance is destroyed. Fossil fuel combustion increases Nitrogen Oxides (NO_x) concentration, but also concentrations of hydrocarbons through incomplete combustion. This means that in addition to the balanced atmospheric reactions of nitrogen species, we now have to consider the oxidation of these long hydrocarbon chains, which produce NO_2 without destroying O_3 , thus upsetting the initial balance.

1.3.2 Air Pollution Effects on Human Health

Arguably, there is no medium in which it is more important to consider AP by its effects than in the human body. However, even this has its caveats. The body's response to any given substance changes with the dose that is administered to it, something which has been known to us for centuries:

What is it that is not poison? All things are poison and nothing is without poison. It is the dose alone that makes a thing not poison.

– Paracelsus

This quote, originally in the writings of one of the fathers of modern medicine, the Swiss Paracelsus, was taken from Patricia Frank's book called *The Dose Makes The Poison* [Frank2011] and is one of the core tenets of toxicology even today. There are, however, some substances which do not need anything close to a high dose to cause harm to human health, and in general, atmospheric pollutants fall in that category. According to the EEA, heart disease and stroke are the most common causes of premature death due to Air Pollution. The same organization states that the most prominent atmospheric pollutants in terms of the effects they have on human health are PM, NO_2 and O_3 [EEA2016, EEA2007]. In this thesis, I will focus mostly on them, not only because of their health importance, but also because of their spectral nature, which allows us to detect them using DOAS [Platt2007]. Of course, a complete description of how AP affects the human body is a colossal task which is well beyond the scope of this thesis. Therefore, I will focus my attention on the more prominent

symptoms that are results of these chemicals: respiratory syndromes, cardiovascular diseases, problems during gestation and finally, neurologic consequences of AP.

1.3.2.1 Respiratory effects of Air Pollution

The respiratory system's main functions are the delivery of oxygen into the blood stream and the removal of carbon dioxide from the body. Air enters the body from the upper airways and flows to the alveolar region, where oxygen diffuses across the lung wall into the blood stream, from which it is transported to the tissues where it diffuses yet again and is made available to the mitochondria in the cells, that use it for cellular respiration [Nilsson2010]. The whole system is in permanent interaction with the atmosphere, and is therefore exposed to all kinds of air pollutants and trace gases, and therefore it is only natural that respiratory effects are among the most direct health complications originating in AP [Vallero2014].

The region in which a given pollutant is, within the respiratory system (see Figure 1.3), is of great importance. After the air is inhaled through the nose, the air is heated or cooled to body temperature, as well as humidified, in the upper airways. The trachea leads the air into the bronchi, where flow is divided several times before reaching the alveoli, where oxygen is supposed to enter circulation. Since air flows within the different regions of the pulmonary system are completely different, AP is also handled differently among them. Moreover, it is also important to consider that pollutants also vary according to their own physical properties, and pollutant absorption is also a function of this. Particles' absorption depends on their aerodynamic characteristics, as well as soluble fraction and density. Gaseous pollutants are dependent exclusively on their vapor pressure, solubility and density [Nilsson2010, Vallero2014].

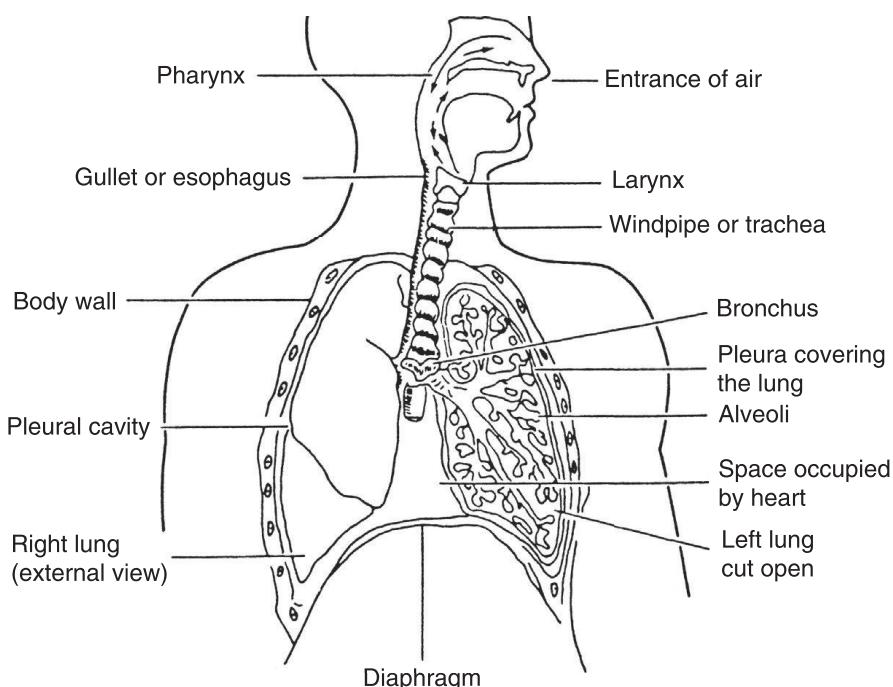


Figure 1.3: Annotated anatomy of the respiratory system [Vallero2014].

The respiratory system has several (imperfect) mechanisms in place to prevent particles from reaching the blood stream. Larger particles are deposited in the nose, by impaction on the hairs and bends of the nose. Smaller particles are immune to this first barrier, and manage to get to the trachea and bronchi, where they are filtered also by impaction, this time on the walls of the innumerable bifurcations of the bronchial tree. The smallest particles are removed through Brownian motion, which ends up pushing them against the alveolar membrane. Deposited substances are then removed through the action of cilia in the pulmonary system's walls or by coughing, sneezing or blowing one's nose [Vallero2014].

While the body is quite efficient at filtering out particles from the respiration process, the same cannot be said about gaseous pollutants. Removal of these compounds can only be achieved through absorption, which depends almost exclusively in the gases' solubility. High solubility compounds are absorbed directly in the upper airways (SO_2 , for instance), while less soluble gases (such as O_3 and NO_2) are absorbed in the lungs themselves. Irritant gases trigger a variety of responses, in which one can include sneezing, coughing or bronchoconstriction. These gaseous compounds are then diffused through to the bloodstream or the lungs themselves try to convert them into other substances via biochemical processes. In some cases, this attempt to detoxify a pollutant can lead to much more problematic circumstances. For instance, the lung is known to activate procarcinogens, substances that are only carcinogenic after being metabolized in a certain way [Vallero2014].

Acute symptoms of AP exposure are very varied, and range from mild irritation to complete respiratory failure, depending mostly on level of exposure and individual sensitivity to the chemical compound. One of the most important acute manifestations of AP exposure are encompassed within the Acute Lower Respiratory Infections (ALRI) group. There are several studies in which the relationship between this issue and AP is deducted and explained, mostly in developing countries, and it remains as one of the major causes for infantile death [CABI2019, UNICEF2013]. Children are one of the most affected demographics by AP [EEA2016], and one of the chief reasons for this is that the human respiratory system is still developing in this stage of life.

In a 2016 review [Goldizen2016], the authors searched the literature for childhood adverse effects of AP, with a particular focus on respiratory problems. They have found evidence for a number of respiratory complications and diseases that were previously reported in the literature caused or exacerbated by AP. Effects are many, and vary immensely in nature, severity and affected populations. Short term effects, like coughing and wheezing were found for the three types of major pollutant and several others; several papers mention an association between the occurrence of respiratory infections and exposure to AP, namely concerning PM and NO_2 . The same review found reports of decreased lung function in children and asthma exacerbation in children due to Air Pollution. Moreover, a person exposed to high levels of AP during childhood are also more likely to develop syndromes like Chronic Obstructive Pulmonary Disease (COPD), and to have exacerbated symptoms of this disease. Finally, and perhaps more concerning, the carcinogenic nature of several of the constituents of AP leads to findings relating the appearance of respiratory cancers to exposure levels during childhood. Many of the conclusions of this review come from a large-scale European effort called European Study of Cohorts for Air Pollution Effects (ESCAPE), that intended

to investigate long-term health effects of AP in Europe. ESCAPE was an European Union's Seventh Framework Programme (FP7) initiative that ended in 2014.

1.3.2.2 Air Pollution and cardiovascular issues

After being absorbed by the respiratory system, oxygen is distributed to all cells of the body through the cardiovascular system. Air pollutants, like particles and trace gases, are also capable of penetrating the lung barrier and therefore share the same fate. There are several pathways with which AP and negatively affect the cardiovascular system. The most immediate of which is probably an imbalance in the Autonomic Nervous System (ANS) caused by direct inflammation and oxidative stress in the respiratory system. The second most immediate pathway is systemic inflammation caused by Air Pollution. Finally, soluble AP compounds in the bloodstream also contribute to Cardiovascular Disease (CVD) by increasing inflammation and oxidative stress in the cardiovascular system [Brook2008, Vallero2014].

The link between Air Pollution and cardiovascular effects started being made during the twentieth century, given a series of incidents (like London's 1952 killer-smog) that happened in the urban areas of industrialized countries. Nowadays, Cardiovascular Mortality (CVM) has been shown to be intricately connected to AP. In fact, in a 2013 review indicated that an annual increase of $10\mu\text{g}/\text{m}^3$ in fine PM and NO₂ led to an increase of 11% and 13% respectively in terms of CVM and premature atherosclerosis, in spite of absolute AP concentrations were maintained below the European policy-recommended thresholds. Road traffic exposure studies have reported similar findings, with subjects having increased coronary calcium scores [Bourdrel2017].

Arrhythmia is one of the other cardiovascular issues that might be caused by AP. There is still some debate regarding whether or not there is a causal relationship between the two, but there have been several studies in which increased levels of Air Pollution were correlated with arrhythmia-related hospital admissions. Moreover, there seems to be a correlation between low heart rate variability and AP, which is considered a marker for ANS imbalance and an important risk factor for CVM[Bourdrel2017].

The risk of stroke is also clearly exacerbated by the presence of increased levels of AP. In fact, it is currently thought that AP is responsible for about 29% of the burden of stroke, globally. Studies have shown that an increase of $5\mu\text{g}/\text{m}^3$ in the annual PM_{2.5} concentration leads to a remarkable 19% increase in the risk of stroke, which was found to be more significant in non-smokers. A positive correlation was also found between gaseous pollutants (NO₂, CO and SO₂) concentration and the risk of stroke or stroke mortality.

Short term effects of AP on the cardiovascular system seem to be predominantly the triggering acute coronary incidents. For instance, a positive correlation was found between short term increases in AP and non-fatal myocardial infarctions.

1.3.2.3 Gestational and developmental complications

Mammals are in their life's most vulnerable stage while they are still developing inside their mother's womb. This is the time when there is a greater rate of tissue expansion and creation, creating an enormous need for nutrients. These are supplied by the

mother's blood, crossing the placenta and reaching the fetus through its umbilical cord. High rates of tissue formation and proliferation render the forming being unstable and therefore more susceptible to the appearance of some kind of morphological abnormality. At this time, there is no separation between the mother's blood and the fetus, meaning that whatever chemical reaches the progenitor's bloodstream also reaches the growing fetus. If the mother is exposed, so is the fetus [Vallero2014].

There are numerous chemicals that can affect the female reproductive system, of which some are habitual components of AP. They have been associated to several highly adverse affects, and interfere with such things as the processes by which the body is able to produce eggs, or other processes that enable the formation of a single cell by the union of the sperm and the egg (the zygote). After conception, AP has been known to reduce uterine nurturing capabilities, and hinder the new being's development. Some of them are even teratogens, meaning that they induce birth defects. Figure 1.4 illustrates the kind of defects that come with exposure, according to the time at which the mother was exposed.

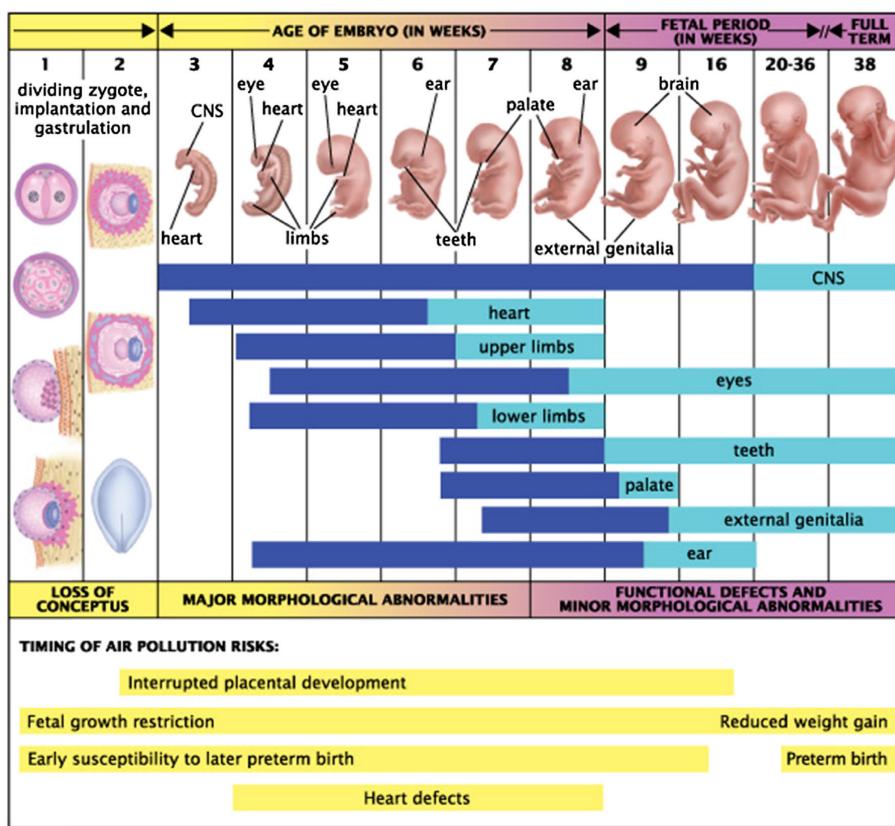


Figure 1.4: Possible abnormalities caused by AP exposure *in utero*. Notice that time of exposure is of critical importance [Vallero2014].

There are already several studies that correlate higher AP exposure levels to birth defects or the probability of negative outcomes. For instance, in [Li2019], researchers have studied the association between AP exposure levels (for the mother) and the appearance of premature Small for Gestation Age (SGA) by collecting more than 40000

births in Changzhou Maternity (China) and studying the mother's typical environment. This study has found a positive association between **SGA** and exposure to **PM_{2.5}** in two or three pollutants models of **AP** (with **NO₂** and **SO₂**), during the third trimester of gestation. Another, perhaps more comprehensive study, was performed using Swedish data from 1997 to 2007, and found that there was a positive association between **O₃** exposure and the appearance of pre-eclampsia (a potentially deadly complication of pregnancy), estimating that about 1 in 20 pre-eclampsia cases were caused by **AP** [Olsson2013].

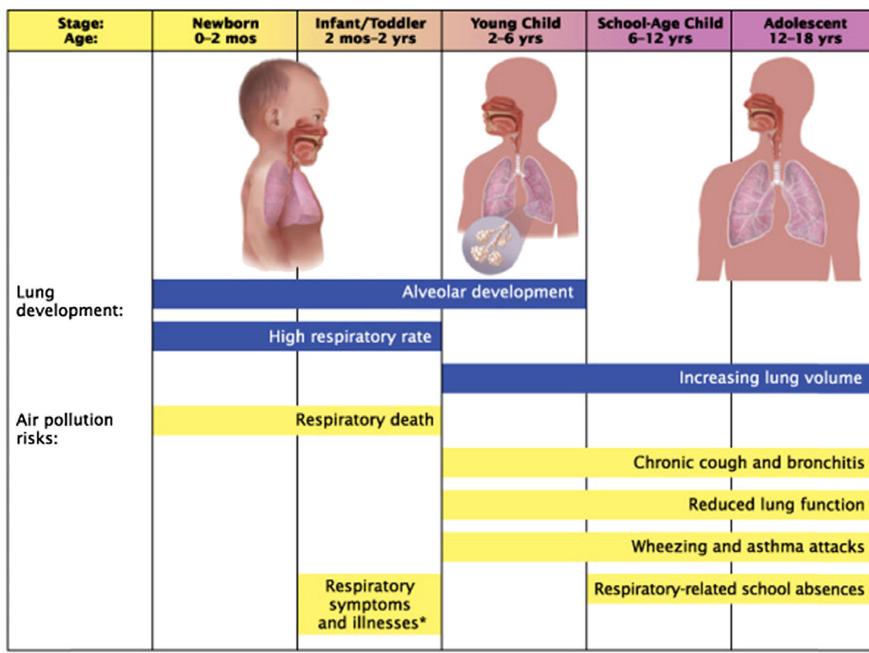
Besides uterine development compromises, birth defects and reproductive difficulties, **Air Pollution** has also been associated with hindrances to the child's neurodevelopment. In a New York study was able to associate lower levels of mental development at age 3, in African-American children with valid prenatal **Polycyclic Aromatic Hydrocarbons (PAH)** exposure data. In another study from the neighboring Boston, **AP** was associated with generally lower cognitive test scores, even when correcting for several influencing factors. On a different level, **AP** was shown to produce significant delays in the central conduction times of **Brainstem Auditory-Evoked Potentials (BAEP)** tests in children, indicating that there might be important repercussions of **AP** to vestibular and auditory development.

Although most other systems are affected by **AP**, it does have a particularly heavy toll on the respiratory development. This is because the lungs are not completely developed at birth, and are only finished in the late teens. The level to which **AP** affects the respiratory system development varies greatly with the stage of life in which the effect is produced, and severity is also very varied. Acute negative effects range from respiratory death to chronic cough [Vallero2014]. Moreover, childhood (and prenatal) exposure to **AP** has been associated with the emergence of conditions such as **COPD** and asthma.

1.3.2.4 Neurological disorders

The brain and the **Central Nervous System (CNS)** were one of the last to be included in the range of organs that are affected by **AP**. While the effects of **AP** on the respiratory and cardiovascular systems are quite broad and include some "surprises", the fact that these systems were affected by **Air Pollution** was evident and expectable, given the type of exposure these systems endure. The **CNS**, on the other hand, has a more difficult to express relationship with **AP**, and has required more sophisticated methods to detect [Vallero2014, Genc2012]. It was in the beginning of this century that the first connections between **AP** and the emergence of neurological disorders started to be made, and from then on, we have progressed into thinking that not only are they related, but also that **AP** might be one of the key driving forces in the onset of certain neurological diseases, including the most dreaded of them all, Alzheimer's and Parkinson's [Genc2012, DePradoBert2018, Calderon-Garciduenas2014].

The reason why **AP** is able to reach and damage the **CNS** is a continuation (or even an extension) of the ways in which it affects the cardiovascular system. By crossing the alveolar barrier into the bloodstream, **AP** acts as an oxidative stress source. As it can also do in lung tissue, **Air Pollution** creates some local proinflammatory effects in the cardiovascular system, affecting the vascular endothelium cells. This can lead



*Air pollution exposure has also been more recently linked to respiratory symptoms and illnesses in early life including cough, bronchitis, wheeze and ear infections

Figure 1.5: Developmental stages of the lung throughout life vs the risks of AP exposure in each stage [Vallero2014].

to a systemic inflammatory status, which is accompanied by the production of proinflammatory cytokines (a type of message-protein that is used by organisms to trigger certain types of response, like inflammation [Zhang2007]). Now, since blood vessels in the brain are extremely responsive to this kind of message, their presence can activate cerebral endothelial cells and disrupt the blood-brain barrier [Genc2012].

In 2018, a consortium of several Spanish universities and researchers wrote a review detailing the until-then-published articles dealing with the neurological implications of AP [DePradoBert2018]. This review identifies several articles that connect the long-term exposure to Air Pollution with adverse impacts on the brain and brain structures. *In vitro* and *in vivo* studies, focusing on traffic related emissions and their effect on gray matter cells, have found that these display significant alterations. On other studies identified by the review, it was shown that white matter, the myelinated part of the brain, is particularly sensitive to AP and its volume is significantly decreased both in the elderly and children, as consequence of prolonged exposure to it.

There are also several articles that show that there is an association between exposure to air pollutants and impairments on brain function. In Section 1.3.2.3, I have already mentioned a study that was conducted in New York, and that found that the children that they were using as subjects were found to have measurable cognitive deficits in comparison with children of the same age living in less polluted areas which are compatible with the affected areas of the brain that were detected through neuroimaging studies [DePradoBert2018].

1.3.3 Air Pollution effects on ecosystems

The Earth is home to an almost unbelievable number of different ecosystems. The ubiquitousness of AP means that all of them are in some way or another affected by this problem. In general terms, the threat posed by AP to any given habitat is a function of its biodiversity, defined as the number of different living beings that inhabit a certain environment (in all biological kingdoms) [OxfordPress2020]. Living beings within an ecosystem are like nodes in a graph, with many connections to any particular node. More biodiversity corresponds to a greater number of nodes and an even larger number of links, which means that there is a greater probability that some of those links become disrupted by Air Pollution in some way.

Water based environments are greatly affected by AP. Material deposition on the surface of the water can have serious consequences in terms of habitat conditions for holding life. In this regard, the most important air pollutants are NO₂ and SO₂, which significantly decreases the water's pH. On its own, this represents a major problem. The acidifying effects of nitrogen and sulfur deposition became very pronounced in Scandinavia (among other places). Thousands of this territory's lakes, once teeming with wildlife, became effectively lifeless. Those that did not reach this point, have seen the number of fish living on their waters dwindle to numbers from which there may be no return [DeWit2015]. Sulfur and nitrogen depositions also enrich surface waters, altering the solubility and other physical aspects on the surface of the water, which in turn inevitably leads to disruptions in species abundance and diversity. Moreover, indirect effects may also take their tolls. For instance, O₃ does not play any significant role in the chemical behavior of a water body, but it can influence the number of predators around this habitat, which will compromise the predator-prey balance of the aquatic environment [Vallero2014, Lovett2009].

In terrestrial ecosystems, AP effects are not smaller in importance or complexity, and they are different for each type of being. To the Flora, AP can have a subtle to deadly effect, depending on variables like pollutant chemical species, exposure time, or plant life stage in which exposure happens. For instance, O₃ is especially poisonous to plants. Even small concentrations of this gas will cause plant growth to decrease significantly. It enters the plant through the stomata and reduces photosynthesis through increased oxidative stress. Many times, although concentrations are not enough to outright kill the plant, they are enough to make them more susceptible to other attacks like pathogens, insects or environmental conditions. Ozone is commonly responsible for huge financial losses that come from the diminished agricultural yields. And while it is true that due to several policies, AP is in a clear downward trend since the 1980s in urban regions, it is also true that in many rural areas, these changes have been smaller or non-existent, making these losses even more relevant [Lovett2009, Vallero2014].

Forests are among the most susceptible environments to AP. They suffer from the previously described mechanisms of AP damage, like acidic deposition, but also suffer from other, less direct pollution risks. Emission of greenhouse gases can induce changes in humidity, temperature, and general climate profile of a forest. The combination of direct and indirect risks result in an exacerbation of both, leading to more and more forest losses due to Air Pollution. The damage done to forests all around

the world is especially problematic given the biodiversity that these ecosystems contain within themselves. Rainforests in particular are thought to contain more than half of the world's terrestrial species. These species have many times adapted to a particular kind of microhabitat which only exists in the specific rainforest in which it lives. Changes in these specific conditions, whether caused by [Air Pollution](#) or any other cause, are leading to alarming extinction rates in forests and rainforests globally [[Vallero2014](#), [Lovett2009](#)].

Of course it is not only the flora that suffers with [Air Pollution](#). Direct implications of [AP](#) on animals approximate those that fall upon humans. We are an animal species, after all. Our main difference is the adaptation capabilities that our superior intellect grants us, which allows us to escape more or less unscathed for a longer period of time, and to combat what we cannot escape from in ways which are simply unaccessible to other animal species. So, although [AP](#) has direct effects on all animals that are exposed to it, ecosystem damage and eventual destruction remains the most perilous factor for this biological realm [[Vallero2014](#), [Lovett2009](#)].

1.3.4 Air Pollution Sources

There are almost as many [AP](#) sources as there are pollutants. The first major division between these sources is whether they are natural or anthropogenic. However, this separation is not always clear, as one source can lead to another and boundaries become fuzzy within their own context. The most prominent example of such is the case of accidental fires. While they are most of the times classified as a natural source of [AP](#), their origin lies most of the times in human activities. In this section, I will present a selection of the most important naturally occurring air pollutants and examples of how they have affected human lives throughout the times. The selection itself does not intend to be complete description of pollution sources, but rather paint a general picture of the subject.

1.3.4.1 Natural Sources of Air Pollution

Although people, governments and institutions tend to speak far more seldomly of them than of their man-made counterparts, natural sources of air pollutant are not only abundant, but also important. One of the main natural sources of [AP](#) are volcanic eruptions. These phenomena are responsible for the emission of immense quantities of [PM](#) and gases such as [SO₂](#), [Hydrogen Sulfide \(H₂S\)](#) and methane. Depending on the type of volcanic eruption, the emitted cloud of gas and [PM](#) can remain airborne for long periods of time, even disrupting modern life at times, namely in what concerns air travel. The last eruption to happen in Portuguese soil took place in the remote Azorian island of Faial. In September 1957, the Earth shook almost continuously for around two weeks. Finally, on the 27th, 100 m Northeast the Capelinhos islands, the sea was seen to boil and project vapor and volcanic material hundreds of meters into the air. In the following hours, the underwater volcano finally exploded, emitting large quantities of volcanic ash and gases into the atmosphere. The phenomenon lasted for more than a year, and the final ejection of lava took place in October 1958 (see Figure 1.6). The eruption had a significant social impact, in addition to its ecologic importance. In the end, 40% of Faial's population left the island as a result [[Vallero2014](#), [TSF2017](#)].

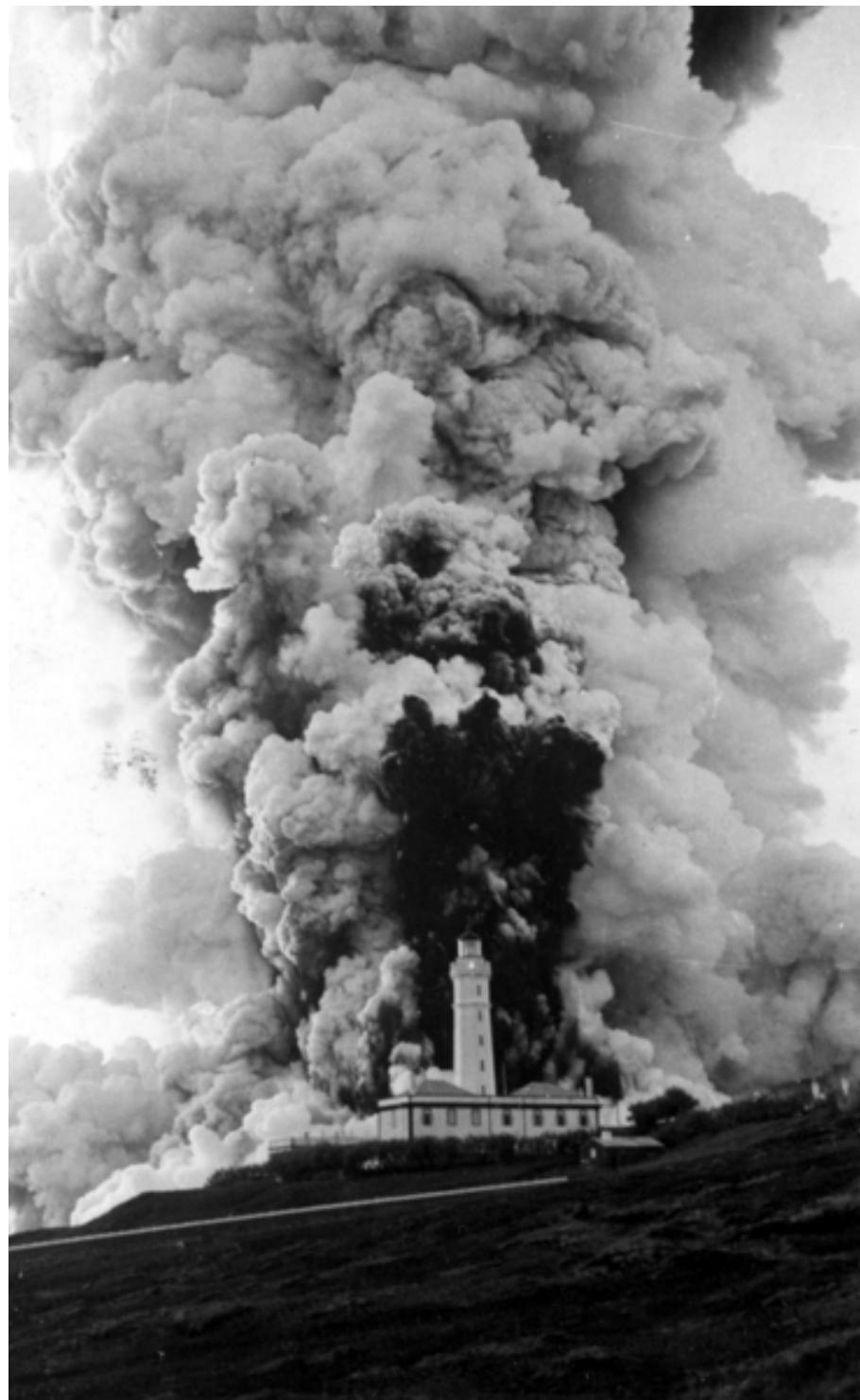


Figure 1.6: Dramatic photograph depicting the Capelinhos' lighthouse, half a kilometer from the eruption site, surrounded by a cloud of ash PM, volcanic gas and water vapor with more than 1km in height[TSF 2017].

Oceans are also a significant source of AP. Aerosol particles of salt are continuously emitted from these large masses of salt water, which damage many human created structures, namely metallic constructions. In certain parts of the world, another important source of Particulate Matter (especially because of its consequences in the inhabitants' daily life) are dust storms. The most famous of these events, and one of the most deadly storms in the recorded history of the US territory was the infamous *Black Sunday* dust storm. Starting on Palm Sunday, 14 April, this sky-blackening dust storm punished the peoples from the panhandles of Texas and Oklahoma, burying entire houses (see Figure 1.7) under the dust and destroying the livelihoods of thousands of Americans. Dust storms were an important part of the US history during the 1930s and led to the creation of the Soil Conservation Service, a branch of the US Department of Agriculture [Vallero2014, Agriculture2012, Reis2008].



Figure 1.7: House almost completely buried by the Black Sunday dust storm. Several houses were entirely swallowed during this storm, trapping people inside, as if a big blizzard had hit them. Unlike a blizzard though, there was nothing anyone could do to keep the dirt outside, and all surfaces were covered black [Reis2008].

Fires are also one of the largest sources of natural air pollution in the world. The uncontrolled burning of organic matter that is a large forest fire creates a large quantity of air pollutants that range from smoke to unburned (or partially burned) hydrocarbons, nitrogen and carbon oxides, and ash particles. Besides the obvious dangers of this kind of burnings for human life and activities, forest fires can also cause indirect damages, such as disruptions in supplies and travel due to reduced visibility [Vallero2014].

Trees and forests in themselves are also responsible for a certain quantity of air pollution. Although they have the main part in the carbon dioxide conversion into oxygen, through photosynthesis, plants and trees are still the largest emitters of hydrocarbons in the planet, as attested by the blue haze that is visible on top heavily forested areas, resulting from chemical reactions between VOCs produced by the trees.

This counter-intuitive fact was in the origin of the infamous Ronald Reagan speech in which he "blamed" trees for much of AP, in a time when anthropogenic AP was at its apogee in the US and Europe. Plants are also the emitters of another kind of PM, which is of particular importance both to themselves and humans, which are the pollens. This is a bio-aerosol - a type of aerosol that is or was part of a living being - associated with a number of diseases [Vallero2014].

Finally, I will discuss Radon gas. This is a natural occurring radioactive gas that is part of the radiative decay of Uranium present in all rocks. Although chemically inert, Radon is radioactive and, as all radioactive substances, emits particles when it decays. Although present virtually everywhere, outdoor concentrations of Radon are typically too small to cause any problems. The problem with this gas comes essentially from indoor concentrations, namely at home. Being a gas, Radon is able to enter people's houses, exposing the inhabitants. Prolonged exposure to Radon gas is the second biggest cause of lung cancer and authorities estimate that between 3 and 14% of lung cancer cases are caused by this gas. In Portugal, Radon concentrations were found to be below the European prescribed limit in two thirds of the houses in a 2001 study, but in 17% of the cases, concentrations were not only above this limit, but also over the highest tolerable limit [Vallero2014, WorldHealthOrganization2016, ProTeste2003].

1.3.4.2 Anthropogenic Sources of Air Pollution

Air Pollution that originates from human activities is called anthropogenic. Since the first industrial revolution, mankind has been using more and more resources to fuel our progress and continuously improving way of life. Of course, the consumption of natural resources has some unpleasant and sometimes dangerous consequences. The most important of which, looking from the lens of this thesis, is the incredible increase in the levels of AP. If one had any doubts whatsoever, all it would take would be a look into the atmospheric CO₂ concentration chart (Figure 1.8) from a few centuries back to the current day to completely dissipate them.

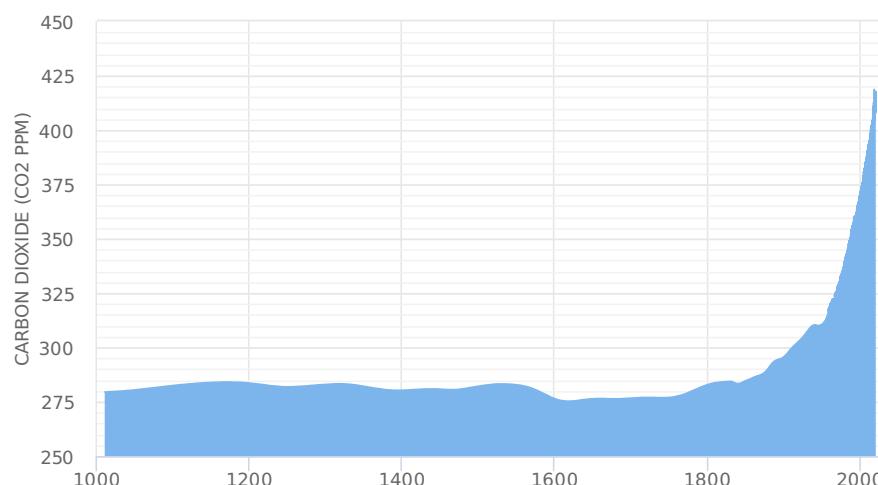


Figure 1.8: Carbon Dioxide atmospheric concentrations since the year 1000. Note the seemingly exponential increase since the 1800s. Plotted and published by the 2 Degrees Institute [co2levels2020] with data from ice cores [Etheridge] and in situ monitors [Tans].

There are literally hundreds of sources of AP, but it is possible to categorize them into 4 main *families*: industrial processes, energy (includes transportation), agriculture and forestry, and waste. Of these 4 broad categories, as displayed in Figure 1.9. The most prominent is without a doubt the energy sector, although we also have to bear in mind that any and all combustion used in the other sectors is counted as energy production [International Agency for Research on Cancer 2016, CABI 2019].

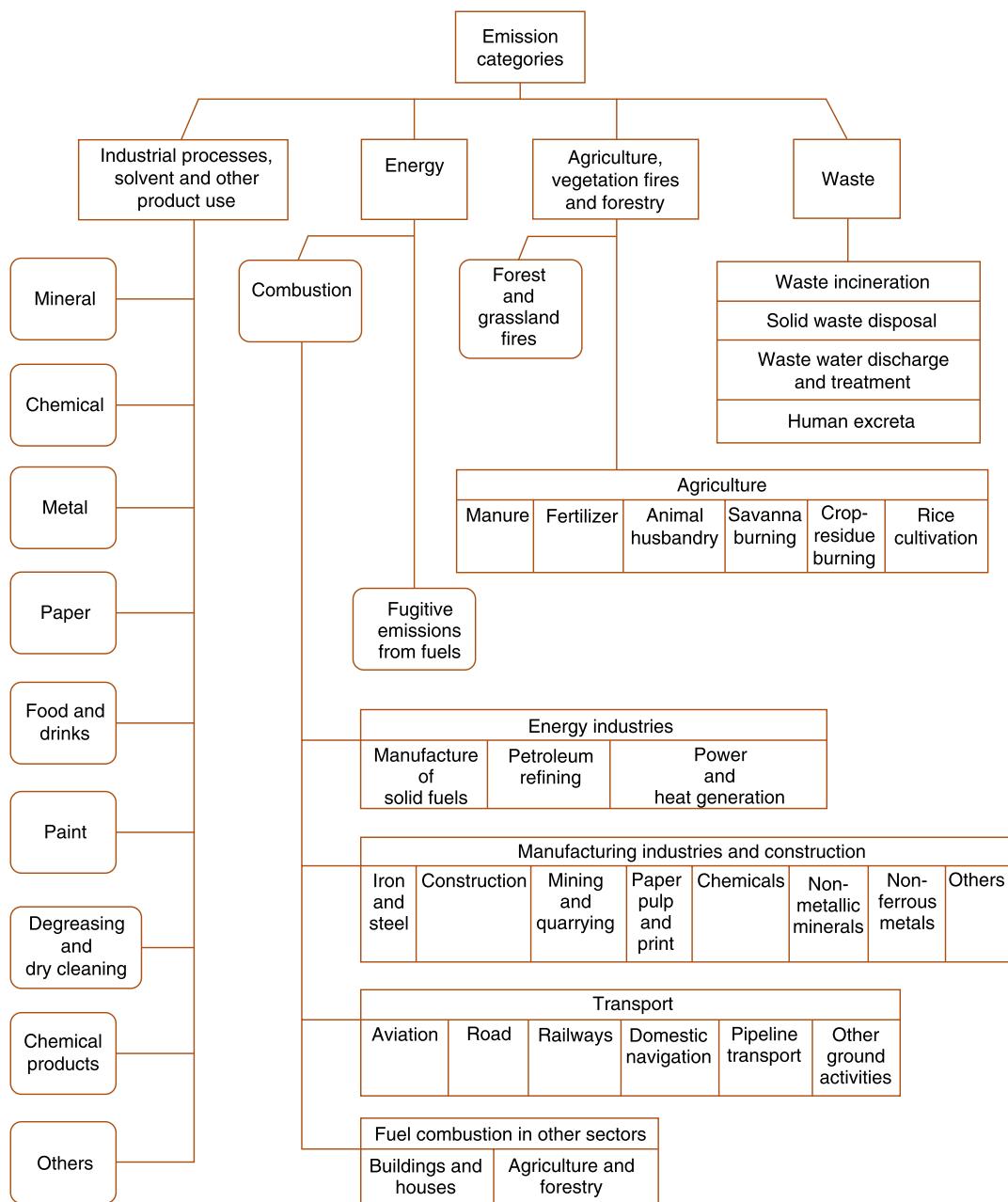


Figure 1.9: Schematic presentation on the sources of anthropogenic pollution and its categorisation according to the IPCC. Adapted from [CABI 2019]

From 2002 to 2011, fossil fuel combustion has been responsible for an average of

8.3 petagrams of carbon per year. This truly gigantic carbon footprint is in its majority explained by the world's energy needs, which are ever increasing up to now. In 1990, total energy demand was situated at 356 quadrillion **British Thermal Unit (BTU)**, having grown to 410 quadrillion **BTU** in 2010. In 2020, energy demand estimates are located at 600 quadrillion **BTU**, of which almost a quarter was expended by China [CABI2019].

It is important that we focus a little bit more on the Chinese case. It is now somewhat near commonsense to regard China as the factory of the world, and this of course is tied to Chinese energy consumption and production. On the same line of reasoning, this must mean that in some way, the country's energy expenditure is connected to the amount of financial resources that it produces, the **GDP**. Looking at the plots in Figure 1.10, one can see that all these numbers are highly correlated. When we ponder on the case of Chinese **AP**, and wonder why has this problem not been addressed previously, given its imposing dimensions and growing importance, one must take into account that, given the indirect importance of **AP** on Chinese people's gains, it is highly likely that the country's governments will be reluctant to decrease it in any expedient form [CABI2019, IEA, WorldBank].

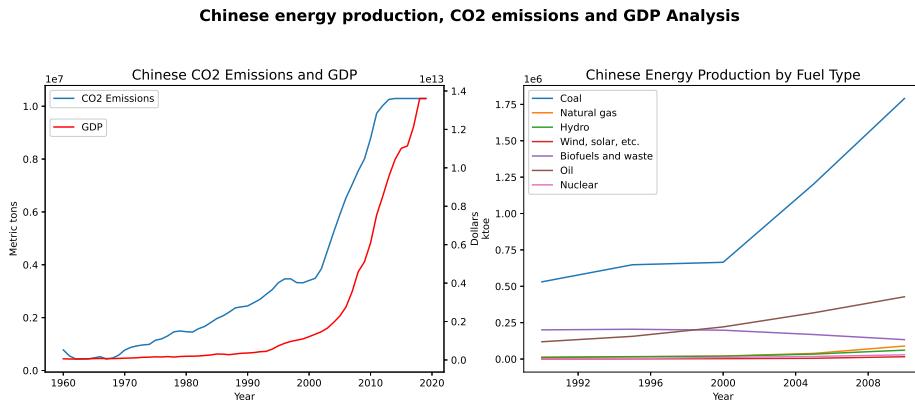


Figure 1.10: Chinese energy production, **GDP** and **CO₂** emissions. Data collected from the World Bank and International Energy Agency websites [WorldBank, IEA]

Another important conclusion that we can take from the plot in Figure 1.10 is that China has a large and historical dependence on the use of coal as fuel for energy production. This adds to the problem described in the above paragraphs, as coal is the single most damaging fossil fuel available. Not only does China get most of its energy from coal burning, but is also responsible for more than half of the world's production and consumption of this substance (see Table 1.4).

ICE are the single most important means for powering human transportation. Almost every vehicle in the world uses a kind of **ICE**. These motors operation is an application of the Otto cycle, in which the chemical energy in the fuel is converted to mechanical energy. These engines are as ubiquitous as the fossil fuels that have powered them since the beginning of the automobile revolution, in the early 20th century. Fossil fuels have several features that make them ideal to power our vehicles. Their energy density is high, they are incredibly safe to manipulate and use, and fossil fuel infrastructure can be found in almost every far corner of the Earth. However,

Table 1.4: Global energy production, divided according to the fuel used to obtain it and the production country.

Country	Liq. Fuel (M barrels / day)	Coal (BTU)	Nat. Gas (T cu. ft)	Renew. (BTU)	Nuc. (b kWh)
China	10,6	80,6	5,1	10,6	93
USA	18,5	17,3	25,5	7,8	769
Europe	14,4	12,5	17,9	11,7	837
Middle East	16	0,1	14,8	0,2	1
India	3,6	12,6	2,1	3,5	30
Russia	3,4	4,5	15,7	1,7	166
Africa	3,5	4,3	2,7	4,7	12
Brazil	3,3	0,5	1,1	6,8	15
World	91,4	153,9	120,8	63,7	2345

using them releases a number of gaseous and particle-condensed side products into the atmosphere, and this makes traffic one of the most important sources of AP. For instance, traffic pollution is the main responsible for human exposure to NO_x gases. Without countermeasures, gasoline ICE equipping passenger vehicles emit around 1.8 g/km of these gases, while diesel emits 2.8 g/km and Liquefied Petroleum Gas (LPG) around 2.1 g/km. On heavy duty engines, like on trucks and tractors, these figures skyrocket to 14.7 g/km for diesel engines and around 5.1 g/km for LPG [CABI2019].

Energy production (including transportation) is clearly the single largest contributor to global AP. This does not mean that other human activities do not pollute or produce air pollutants. Pollutant contributions from the industry, the agricultural activities and waste disposal are also non-negligible. In fact, industries around the world are responsible for the production and emission of all the criteria pollutants. It is important to single out one particular activity, which is the burning of forest for land-use changes. Carbon Monoxide (CO) emissions for this purpose are very high due to the nature of the burning material, which emits more than 50 times more CO than fuel or coal [CABI2019].

1.3.4.3 The European Case

Europe has for long been on the forefront of the fight against AP emissions. The European Union has put in place a number of policies aiming at cutting (or even eliminating) emissions of human health compromising pollutant components. Few places in the world have been so demanding regarding their environmental practices, and numbers are a clear reflection of these adaptation efforts. In their 2019 report, the EEA state that European emissions have globally declined, and have been declining since at least the year 2000. Moreover, and in contrast with China's case, the GDP does not seem to be connected to AP emissions. As can be seen in Figure 1.11, emissions are decoupled from economic growth, as there are now less emissions per GDP unit than before [EEA2019].

If one extends this analysis further, and separates emissions by using their origin, the trends are approximately the same: except for Ammonia (NH₃) (a side product of agricultural activities) a clear reduction is present in all sectors. These results can be seen in Figure 1.12 and were presented in [EEA2019].

CHAPTER 1. BACKGROUND, MOTIVATION AND INTRODUCTION

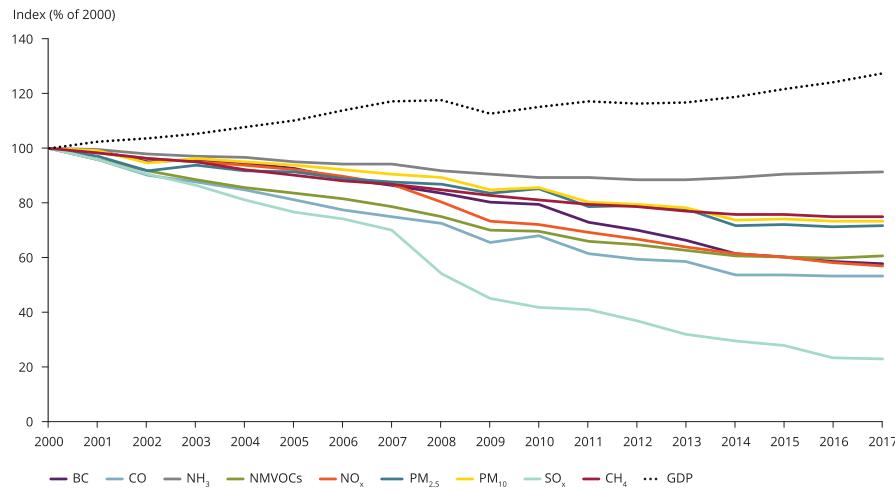


Figure 1.11: General trends for European emissions. Data presented in % emissions of year 2000. Note the downward global trend in pollutant emissions, and its decoupling with the European GDP [EEA2019].

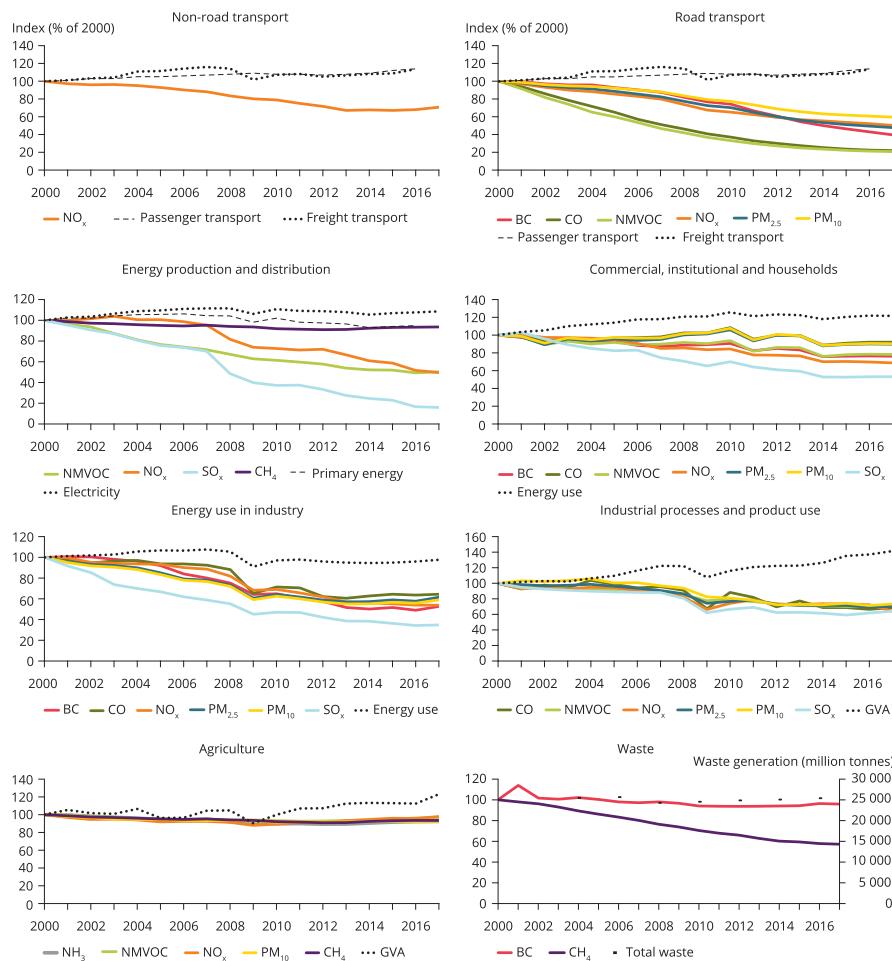


Figure 1.12: European emissions divided by activity sector. The global decreasing trend is confirmed, as industries all around are producing less and less AP with the passing years [EEA2019].

1.4 Air Pollution Monitoring Techniques

There is no doubt that AP is a global threat that affects everyone, both in personal terms (through the degradation of their health) and in societal terms, through the investments and limitations that we as a whole have to commit to in order to prevent larger, unmanageable problems. Reducing AP is a priority and a requirement for today's modern societies. This demands immediate and effective actions, which in turn imply that we have a solid and profound understanding of how pollutants are created, transported and transformed in the atmosphere. The scale on which these interventions must be conducted requires them to be made on a concerted and collaborative manner, and always leveraged by technological development [EEA2019]. Many of the air pollutants cannot be detected solely by our senses, or even if they can is at already dangerous concentrations. Technology is therefore a prerequisite to our fighting the problem of Air Pollution [Vallero2014].

Pollution monitoring is itself based on the ability of a given measurement method to determine concentrations for trace gases, aerosols or radiation quantities. As with many other test techniques, in various fields, pollution monitoring techniques have three very important aspects to verify. The first of which is sensitivity, and also the most demanding. Important trace gases in atmospheric chemistry have sometimes vestigial concentrations, and the ability to correctly detect them is many times a technical challenge. The second most important is specificity, which is the ability of an atmospheric measurement to measure each compound independently, without a component influencing another component's measurement either positively or negatively. Finally, any usable monitoring technique must be sufficiently precise as to provide valid measurements.

Air Pollution monitoring techniques and devices are too many to address them all in this document. However, a small introduction to the topic is in order. A valid point of entry in this discussion would be a broad categorisation of monitoring techniques with regard to the proximity with which they are designed to interact with the target pollutant. As with the case of fire detecting systems, reported in [Valente de Almeida2017], atmospheric pollution monitoring systems can be divided into three broad categories, which are further divided as in 1.13:

In-situ systems localised sensors that are influenced by gases in their immediate vicinity;

Large area (open path) systems able to evaluate the atmosphere around them (normally through somekind of optical measurement) up to a few km;

Remote systems placed in satellites, which were designed to perform a more global analysis of the atmosphere and identify larger-scale trends;

Sampling studies are still regarded as the gold standard for atmospheric pollution monitoring. They are fundamentally different from the other methods, because there is an important time gap between collection and analysis. Sampling methods cannot provide real time information on atmospheric composition. The gaseous analyte is moved into a collection medium before being transported to the laboratory, where a

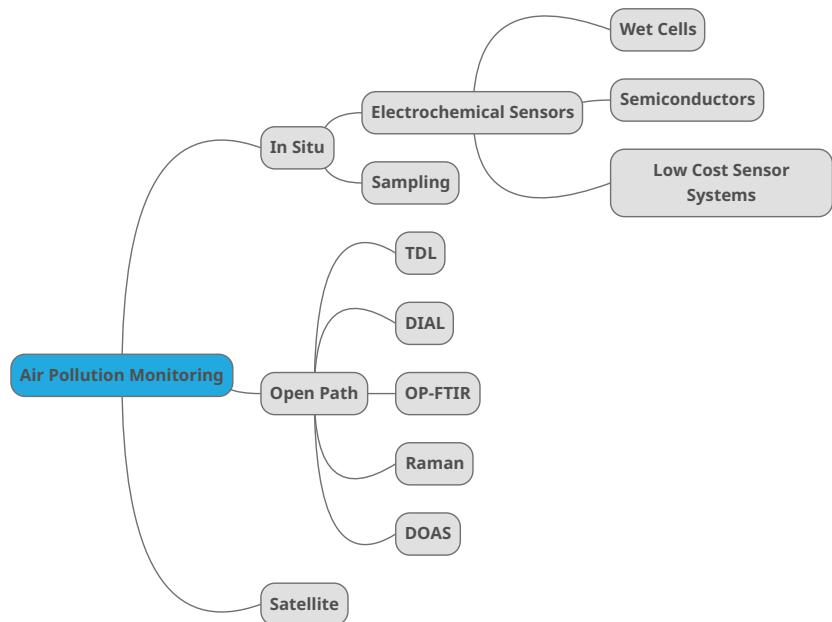


Figure 1.13: A tree diagram providing a general overview of the various types of atmospheric pollution monitoring systems.

very powerful analytical method such as mass spectroscopy or chromatography is used to retrieve the chemical composition of the sample. The use of this kind of technique determines the importance of sampling methods in the study of the atmosphere's composition, but they also make them quite inconvenient for day to day use [Vallero2014].

A much more convenient solution comes in the shape of electrochemical sensors. They were highly popularised in recent decades through their usage in the field of industrial hygiene [Clark1997]. The first generation of these sensors were called wet cells. The basic design for an O_2 sensor of this type is presented in Figure ???. When an electromagnetic force is applied to the cell, hydrogen appears at the cathode, and oxygen at the anode. Conversely, when the cathode is subject to an oxygen concentration and an appropriate electrolyte is used (for instance, KOH), an electronic charged develops through the oxidation of the anode into a hydroxide. If a load is presented to the system, the current flowing through it is proportional to the partial pressure of oxygen in the mixture [Dhall2021, Clark1997].

A similarly applicable technology is the use of semiconductor sensors. These usually operate by the detection

LITERATURE REVIEW

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.²

2.1 DOAS Tomography

In previous sections (Section 3.3) we have established that **DOAS** is a widespread spectroscopic technique with particular usage in the atmospheric research community. This technique uses either artificial or natural light to retrieve slant column densities, the number of molecules that interact with the said light in a given optical path. These values are described by Equation 3.4, and are indeed line integrals.

Also in a previous but different section (Section 3.5), we have seen that with enough of these integrals (which we can call projections), we can gather enough information to depict the interior structure of an object that we make traverse with radiation of some kind, as long as it is measurable and its behaviour is well known.

Joining the two notions, we can get to the idea that it would be possible to use **DOAS** in a tomographic manner, so long as we use sufficient light collection points. This has been proposed by Byer and Shepp [**Byer1979**] in 1979. At the time the authors did not mention specifically the **DOAS** technique, but the assembly they propose is essentially equivalent to one that would use this method. **DOAS** tomography was also mentioned in one of the main pieces of literature in the field, [**Platt2007**], as one of the most promising avenues available for **DOAS** progress in the future.

As a requirement for one of the courses that I chose to pursue during the curricular part of my doctorate, I had to complete an assignment that consisted in writing a

Systematic Mapping Study on the subject of **DOAS** tomography. The study is included in full in Appendix ???. It aims at characterising the literary panorama regarding the subject and tries to arrive at a standard application with respect to software, algorithms and instrumentation.

This type of study has a particular methodology focused on repeatability. Its complete explanation is out of scope of this section and even this document, but a simplified schematic is included in Figure 2.1. A more detailed and extensive description of this method can be found in several works available in the literature, namely [**Kitchenham2007**, **Kitchenham2009**, **Souza2019**].

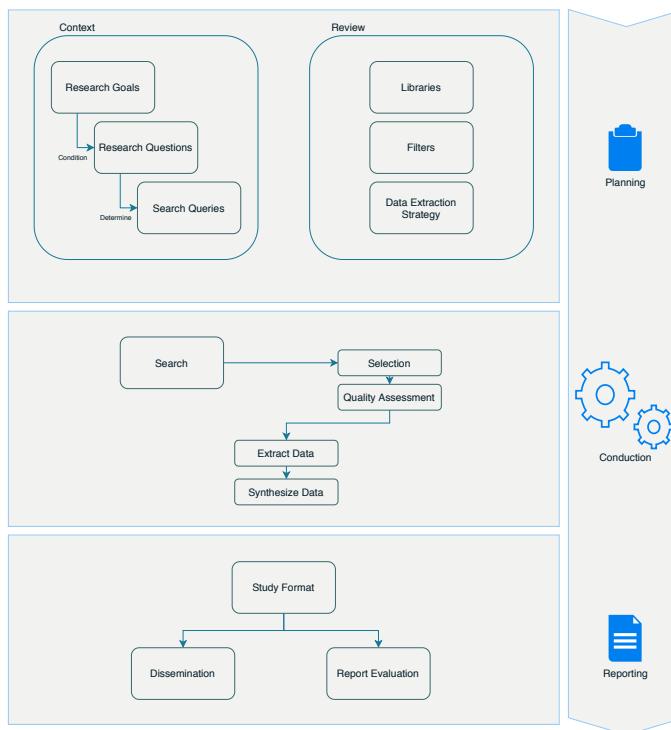


Figure 2.1: Simplified diagram of the **SMS** method. This type of study is focused on registering every step of the search and the reading of the retrieved material, in an effort to make the process completely repeatable. Adapted from [**Kitchenham2007**].

After defining search terms and conducting the actual search, we arrived at a total of 9 relevant papers. These were analysed under the light of the study's main goal, which was to identify a standard **DOAS** tomography device and application.

The first identifiable pattern is the marked prevalence of active **DOAS** systems. 11 out of the 13 retrieved papers describe or consider an active **DOAS** system of some kind. As stated in Section 3.3, active **DOAS** systems do have better analytical capabilities than their passive counterparts, although that comes at the cost of increased instrument complexity and operational costs.

Another immediate conclusion is that there is a "dominant" study. Almost half of the papers found originated from the BABII campaign, in which a group of researchers set out to quantify pollution through **DOAS** tomography along a busy German motorway, in the beginning of the 21st century [**Hartl2005**, **Hartl2006**, **Laepple2004**, **Pundt2005**, **Pundt2006**, **Pundt2005a**, **Mettendorf2005**].

All of the active **DOAS** systems were purposely built for their corresponding experiment (or group of experiments). BABII researchers used two telescopes with around 200mm diameter and 1m focal length to simultaneously illuminate 8 retroreflectors that were assembled onto two towers located on each side of the road. In one of the papers associated with this initiative, the same telescope instrumentation was used to validate the 2D reconstruction technique that was going to be used in the other papers. The campaign's main assembly is illustrated in Figure 2.2.

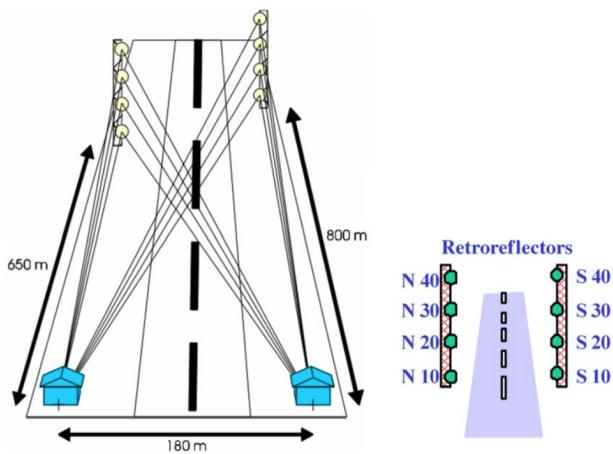


Figure 2.2: BABII assembly geometry. In this experiment campaign, the telescopes illuminated retroreflecting targets that were positioned in two steel towers on both sides of a busy motorway in Germany, connecting Heidelberg to Mannheim [Pundt2005a].

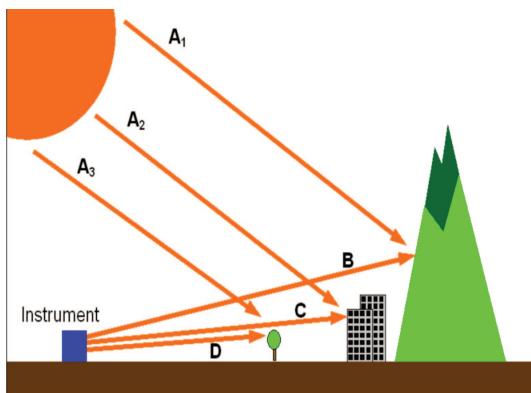
Another important initiative with respect to **DOAS** tomography was the study conducted in 2016 by Stutz et al [Stutz2016]. The approach in this case was to use a similar telescope to detect the light emitted by a narrow interval UV LED light source (290nm) to create a fence line monitoring system for Benzenes, Toluene and Xylene. The team managed to apply this system in a successful manner in refineries in Los Angeles and Houston. One of the most interesting aspects of this study is that it details a tomographic system that could easily be commercially deployed.

Another type of **DOAS** tomography system was proposed by researchers in the Cork Institute of Technology [ODriscoll2003, ODriscoll2003a, Murphy2003]. In their three papers, the authors describe 1) a new multipath instrument that significantly increases the amount of projection information in this kind of application; 2) a tomographic reconstruction algorithm based on evolutionary algorithms; and 3) the application of **DOAS** tomography to a simulated urban canyon scenario. Although all three papers present technological innovation, it would not be fair not to say that from a strictly literary point of view, these were among the weakest retrieved by the search process.

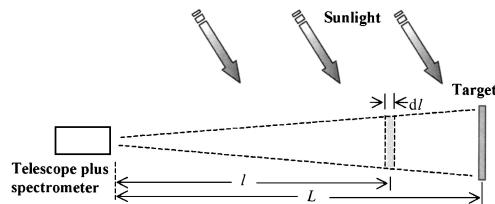
Regarding passive DOAS applications, the two papers we have found come with two completely different paradigms. The first article [Johansson2009] was written in 2009 and details the application of a tomographic inversion algorithm to a scanning DOAS application, designed to work with trace gas plumes like the ones above volcanoes or power stations. The team present a system composed of two DOAS devices, with sufficient distance with themselves as to allow tomographic reconstruction, but sufficiently small to allow the light path to be considered a straight line from the point

of last scattering to the detector. The authors applied an adapted version of the Lower Third Derivative (LTD) algorithm to the projections obtained by pointing the set of fixed DOAS apparatus towards the plume in different angles. Besides simulations for their proposed method, the authors have also conducted practical experiments, both over a power plant in Spain and a volcano in Italy. Results from these experiments display a good agreement between reality and simulation results, proving the technique's validity.

The second Passive DOAS application is a paper published by Frins et al. [Frins2006]. In this study, the researchers detail a particular application in which they measure light coming from bright and nonreflecting sun-illuminated objects in their field of view. They use this light to retrieve column density values for a number of trace gases. The proposed method also includes a way with which to remove the stratospheric contribution that appears in the measured light besides the target column. The authors discuss how radiative transfer can influence measurements, but they also present a number of approaches to mitigate this problem, ensuring the validity of their approach. Besides presenting the method, the authors also describe an experiment they conducted by assembling and manoeuvring a DOAS system on top of a building in Heidelberg, Germany.



(a) Schematic representation of Frins's assembly [Frins2006].



(b) The physical principle behind Frins's paper [Frins2006].

Figure 2.3: Erna Frins's paper [Frins2006] proposes a very relevant passive DOAS application that can conceptually be employed in a DOAS tomography scenario. In this 2006 paper, the authors use multi-axis measurements of sun-illuminated targets to estimate absorption paths without using radiative transfer models.

In summary, the search has found that active tomographic DOAS is far more common than the passive counterpart (11 out of 13 articles discussed this method). This preference can be explained by the fact that the results produced by this kind of system are generally superior to those obtained by passive methods. However, passive applications are normally much less demanding on a technical level, and are simpler to run and assemble. Much as a result of this, we have also identified that the systems used in the literature were not mobile or had a very low mobility level which in turn caused that all the systems were working with low projection numbers. This should be taken into account in future research on the topic. As a final note, we would also like to point out that there is no commercially available systems for this kind of application,

CHAPTER 2. LITERATURE REVIEW

although some of the articles, like the one by Stutz in 2016 [**Stutz2016**] detail systems which could easily be adapted to that end.

THEORETICAL BACKGROUND

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellen-tesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.²

3.1 Unmanned Aerial Vehicle

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellenesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellenesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellenesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellenesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.³

3.2 Optical Systems

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu

neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellenesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.⁴

3.3 DOAS

Differential Optical Absorption Spectroscopy is a well established absorption spectroscopy technique that is widely used in the field of atmospheric studies [Platt2007]. In this section, I present a short introduction to the field. The first part of this introduction is based on [Valente de Almeida 2017], an article we have published in 2017, marking the conclusion of the initial studies for this PhD thesis. The second and last part refers to additional research that I have conducted since then.

DOAS itself is based on Lambert–Beer’s law, which can be written as [Platt2007]

$$I(\lambda) = I_0(\lambda) \cdot \exp(-\sigma(\lambda) \cdot c \cdot L), \quad (3.1)$$

Where λ is the wavelength of the emitted light; $I(\lambda)$ is the light intensity as measured by the system; $I_0(\lambda)$ is the intensity of the light as emitted by the source; and $\sigma(\lambda)$ is the absorption cross section of absorber, which is wavelength dependent; c is the concentration of the absorber we want to measure.

This law allows the definition of optical thickness (τ) [Platt2007]:

$$\tau(\lambda) = \ln \left(\frac{I_0(\lambda)}{I(\lambda)} \right) = \sigma(\lambda) \cdot c \cdot L. \quad (3.2)$$

In a laboratory setting, Eq. (3.1) or (3.2) can be used to directly calculate an absorber’s concentration, provided there is knowledge of its cross section. In the open atmosphere, however, absorption spectroscopy techniques are far more complex. On one hand, $I_0(\lambda)$ is not accessible since we measure from inside the medium we want to measure. On the other hand, there are several environmental and instrumental effects that influence measurement results. These effects include the following [Platt2007].

- Rayleigh scattering is due to small molecules present in the atmosphere and is heavily influenced by wavelength (hence the blue colour of the sky).
- Mie scattering is caused by particles and larger molecules suspended in the atmosphere and is not very dependent on the wavelength (hence the white colour of clouds).
- Instrumental and turbulence effects are the instrument’s transmissivity and atmospheric turbulence in the optical path also limit light intensity.

In addition, we also have to take into account that, in the atmosphere, there are a number of trace gases that interfere with passing light. Another aspect worth mentioning is that our device is never pointed directly at the light source (the Sun) but always processes light that has been scattered at some unknown point in the optical path. This means that the light that reaches our detector is only the scattered fraction of the sunlight, depending on the system's position and geometry, as well as wavelength. The expansion of Lambert–Beer's equation to include all these effects results in Eq. (3.3).

$$\begin{aligned} \epsilon_{\text{TG}}(\lambda, s) &= \sum_i \sigma_i(\lambda, s) \cdot c_i(s) \\ I(\lambda) &= I_0(\lambda) \cdot A(\lambda, \dots) \cdot S(\lambda) \cdot \\ &\quad \cdot \exp \left[- \int [\epsilon_{\text{TG}}(\lambda, s) + \epsilon_{\text{M}}(\lambda, s) + \epsilon_{\text{R}}(\lambda, s)] ds \right] \end{aligned} \quad (3.3)$$

Where $A(\lambda, \dots)$ is the fraction of scattered light that reaches the device, $S(\lambda)$ represents instrumental and turbulence effects, $\sigma_i(\lambda, s)$ is the absorption cross section of absorber i , c_i is the concentration of absorber i . $\epsilon_{\text{TG}}(\lambda, s)$ is the absorption by the i trace gases, $\epsilon_{\text{R}}(\lambda, s)$ represents Rayleigh's extinction coefficient and $\epsilon_{\text{M}}(\lambda, s)$ represents Mie's extinction coefficient. The interest of this equation lies within the retrieval of c_i , a given absorber's concentration. Since the integral is taken along the total atmospheric path of the measured photons, and considering that their cross sections do not vary significantly in atmospheric conditions, it is possible to define the concept of slant column, which is of great importance [Merlaud2013].

$$\text{SC}_i = \int c_i(s) ds \quad (3.4)$$

This quantity, as Eq. (3.4) shows, equals the integral of an individual absorber's concentration along the atmospheric optical path of relevance. Now, without knowledge of $I_0(\lambda)$, these equations cannot give us absolute concentration values. We can, however, use another scattered light spectrum as reference in Eq. (3.2). Instead of absolute densities, this will yield relative changes in the atmosphere. We thus arrive at Eq. (3.5).

$$\begin{aligned} \ln \left(\frac{I_{\text{ref}}}{I} (\lambda) \right) &= \ln \left(\frac{A_{\text{ref}}}{A} (\lambda, \dots) \right) + \ln \left(\frac{S_{\text{ref}}}{S} (\lambda) \right) \\ &\quad + \sum_i (\sigma_i(\lambda) \cdot \Delta \text{SC}_i(\lambda)) + \Delta \tau_{\text{M}}(\lambda) \end{aligned}$$

Where ΔSC_i is the relative slant column of absorber i ; $\Delta \tau_{\text{M}}$ is the relative Mie scattering term, integrated to its optical thickness; and $\Delta \tau_{\text{R}}$ is the relative Rayleigh scattering term, integrated to its optical thickness. This is where the principle of DOAS is applied. Instrument features, scattering and other atmospheric effects have broad absorption spectral profiles, which vary slowly with wavelength. Several trace absorbers have narrow and rapidly varying spectral signatures in at least a small section of the spectrum. By using Eq. (3.5), we can separate these contributions [Danckaert2015].

$$\sigma(\lambda) = \sigma'(\lambda) + \sigma_0(\lambda) \quad (3.5)$$

Here, the broad part of the optical thickness ($\sigma_0(\lambda)$) can be separated from the narrow part ($\sigma'(\lambda)$ – differential) by approximating it by a low-order polynomial, resulting in Eq. (3.6).

$$\ln\left(\frac{I_{\text{ref}}}{I}(\lambda)\right) = \sum_{i=1}^n \sigma'_i(\lambda) \cdot \Delta SC_i + \sum_{j=0}^m a_j \cdot \lambda^j \quad (3.6)$$

Where $\sum_{i=1}^n \sigma'_i(\lambda) \cdot \Delta SC_i$ is the differential part (narrowband, rapidly varying with wavelength) and $\sum_{j=0}^m a_j \cdot \lambda^j$ is a low-order polynomial, used to remove the broadband spectral features resulting from atmospheric and instrumental phenomena.

In practice, the mathematical solving of Eq. (3.6) is not enough since it does not account for the Ring effect or the non-linearities that result from stray light and wavelength shift in measured and cross-section spectra.

The Ring effect is a consequence of rotational Raman scattering: molecules in the atmosphere do not absorb photons in a purely elastic (Rayleigh scattering) fashion. A small portion of the light–matter interaction is in fact inelastic [**Brinkmann1968**, **Merlaud2013**]. This changes the light source frequencies as seen from the detector. This phenomenon was first noticed by Grainger and Ring in 1962. At the time, they noticed that the well-known Fraunhofer lines would slightly change when one observed them by using moonlight instead of scattered daylight [**GRAINGER1962**]. Mathematically, the Ring effect is introduced into the DOAS expressions as a synthetically produced pseudo-absorber.

Up until this point, the DOAS problem can be solved by a system of linear equations of the form displayed in Equation 3.7.

$$\tau = A \cdot X \quad (3.7)$$

A is an $m \times n$ matrix. Its columns are the differential cross-sections for the measurement target gas, $\sigma'_i(\lambda)$, and the wavelength powers ($\lambda, \lambda^2, \lambda^3, \dots$) according to the polynomial used for the broadband extraction described above, $P(\lambda) = \sum_{j=0}^m a_j \lambda^j$. The lines of matrix A are the wavelength window of study, as seen by the spectrometer used in the experiment. This leads to there being many more lines than columns in A . The system is thus overdetermined. Solving it requires the use of numerical approximations, and the most common approach is a least-squares approximation, in which the best solution minimises $\chi^2 = [\tau - A \cdot X] [\tau - A \cdot X]^T$. The upper script T used in this expression denotes the transpose of the preceding matrix [**Merlaud2013**, **Platt2007**].

A crude DOAS algorithm might not go any forward. In effect, this is precisely the approach followed in [**Valente de Almeida2017**] for smoke detection using machine learning techniques. A more refined measurement, aimed at quantification of the target trace gases and not at determining the presence of a particular type of atmospheric event requires an additional consideration, and some more algorithmic steps. There are in fact some non-linearities in the complete retrieval process. These non-linearities are not retrievable through linear algorithms alone. These non-linear effects present

themselves as shifts, stretches and offsets in the measured signal. They do not directly change the dependent variable values (i.e., the radiance), but instead "move" the scale of the independent variable (i.e., the wavelength). Therefore, linear approximations such as least-squares minimisation are not sensitive to them. A non-linear approximation algorithm such as Levenberg–Marquardt [Press2007] can be used, and the expression that is being solved is presented in Equation 3.8.

$$\ln \left(\frac{I_{\text{ref}}(\lambda)}{I(\lambda + \text{shift}) + \text{o_set}} \right) = \sum_{i=1}^n \sigma_i'(\lambda) \cdot \Delta \text{SC}_i + \sum_{j=0}^m a_j \cdot \lambda^j \quad (3.8)$$

Programming-wise, solving these equations is an iterative two-stage process, which runs until one of the following typical stop criteria are met.

Maximum iteration number: This is a self-explanatory criterion. It limits the number of times the algorithm's cycle can run;

Minimum improvement threshold: As the algorithm proceeds, this criterion ensures that it progresses in the correct direction, i.e., minimising χ^2 ;

Minimisation target: If χ^2 becomes lower than this given threshold, the cycle is terminated.

The cycle begins by the determination of the concentration values of the target trace gases and the χ^2 score is calculated. The wavelength window is shifted, stretched and offset according to the initial conditions. The χ^2 is re-calculated and compared to the previous value. This cycle is iteratively repeated in the best minimisation direction until any of the stop criteria are met and the cycle is terminated. The algorithm is illustrated (simplified) in Figure 3.1.

3.3.1 Types of DOAS experiments

There are two main families of DOAS assemblies, with different goals and capabilities:

- Active systems, of which a simple illustration is presented in Fig. 3.2, are characterized by relying on an artificial light source for their measurements. A spectrometer at the end of the light path performs spectroscopic detection. Active DOAS techniques are very similar to traditional in-lab absorption spectroscopy techniques [Platt2007];
- Passive DOAS techniques, illustrated in Fig. 3.3, use natural light sources, such as the Sun and the moon, in their measurement process. An optical system is pointed in certain elevation and azimuth angles and sends the captured light into a spectrometer, connected to a computer. The system returns the total value of the light absorption in its path [Platt2007, Merlaud2013].

Within the two main DOAS families, there are several types of possible experiment. Differences in the design of these assemblies originate from a number of different target requirements. Active or passive applications can differ with relation to their intended spectral range, light throughput or resolution, among others.

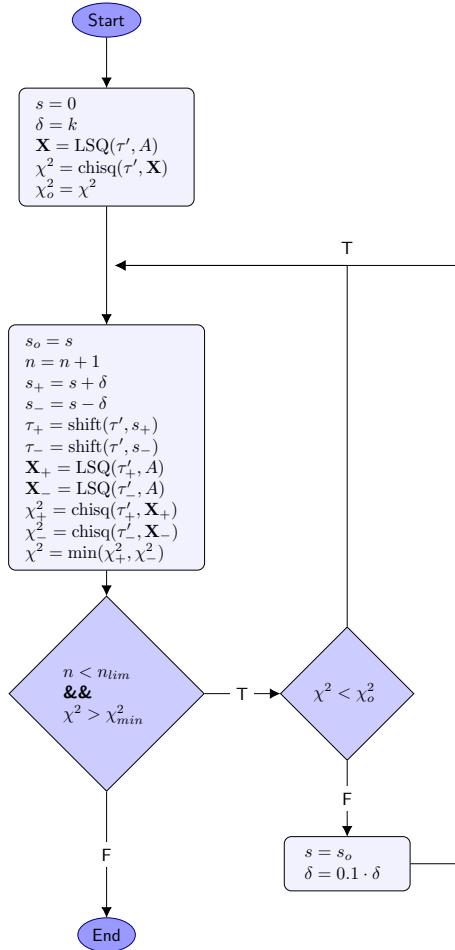


Figure 3.1: Simplified schematic flowchart of the DOAS algorithm, including the non-linear part.

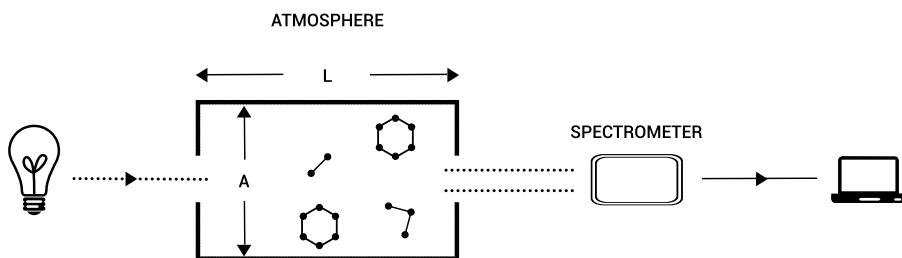


Figure 3.2: Active DOAS schematic.

In active experiments, the choice of the light source is the most critical aspect of the whole experimental design. Active DOAS light sources must be stable, have a very high throughput (these experiments are often conducted over long optical paths) and must have an adequate cost to purchase, maintain and operate. This is especially true in long-running experiments, which must remain working for months or even years. Moreover, the spectral range of the emitted light is also of central importance, because most trace gases have very particular spectral cross sections. The spectral structure of the emitted light is also an important feature to consider, for similar reasons.

The sun and the moon are the two most important light sources when it comes to

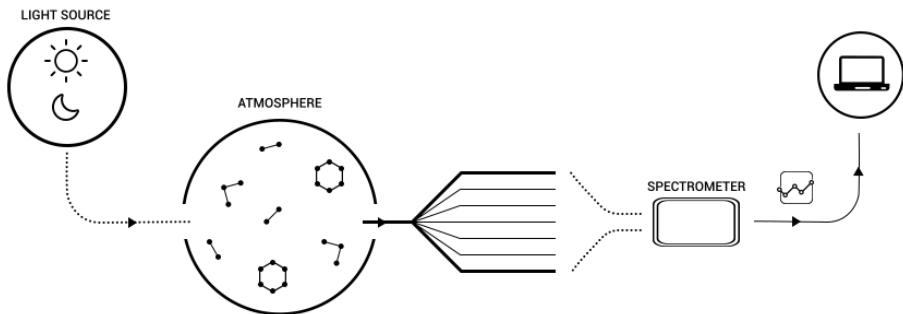


Figure 3.3: Passive DOAS schematic.

passive DOAS applications. Sunlight can be used directly or after a scattering event, the latter being the more common. In these experiments, instead of pointing directly at the sun, the collector is pointed at a certain point in the atmosphere, entering the system after it has been scattered. There are many possible geometries to a scattered sunlight DOAS experiment. Some of them are schematically represented in Figure 3.4.

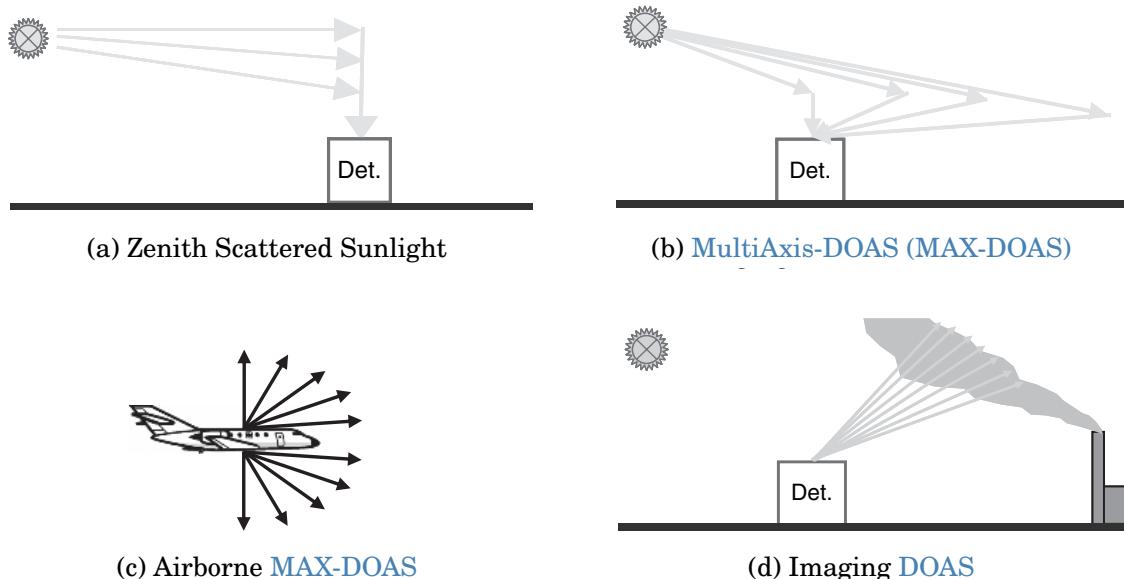


Figure 3.4: Several examples of possible passive DOAS experiment geometries. All these examples use only the light from an astronomical source to detect many atmospheric trace gases. All these examples were taken from [Platt2007].

Although instrumentally simpler than their active counterpart, passive DOAS applications require more care in the retrieval process. In this kind of application, light sources are extremely far away. Additionally, they are normally highly structured. This means that one has to be extremely mindful when using it for the retrieval of small concentration changes. In addition to this, there is always the need to convert the system's direct measurement, a column density, into vertical densities. Since in scattered sunlight measurements, the optical path is impossible to calculate in a precise manner, this requires the use of complex radiative transfer models [Platt2007, Frins2006].

3.3.1.1 Satellite Measurements

One particularly interesting use of passive DOAS are satellite measurements. There are three types satellite DOAS experiments:

Occultation measurements: this is a direct sunlight measurement. Light comes from the sun and traverses the Earth's atmosphere in a tangential manner before entering the satellite's light collector;



Figure 3.5: Occultation measurement schematic representation [Platt2007]

Limb: a scattered sunlight measurement, in which the collector is pointed towards the Earth, at an angle. Light reaches the detector after being scattered in the atmosphere, the ground, or both;

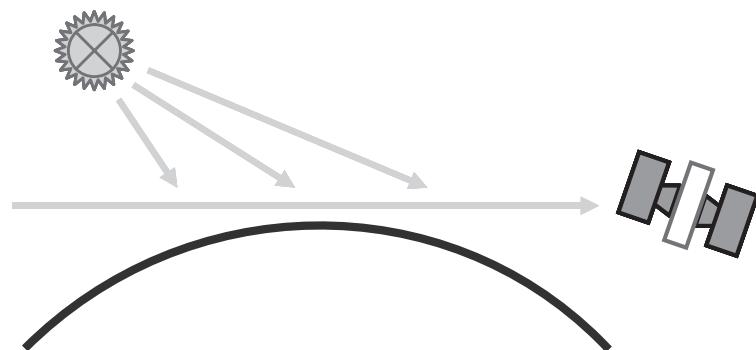


Figure 3.6: Schematic representation of the limb satellite measurement geometry.

Nadir: this is the most common measurement geometry for satellite experiments. In this mode, light that gets reflected off the Earth's surface is captured by the collecting device, while it is pointing directly down.

Satellite based DOAS measurements have also been important because they have given rise to new trace gas retrieval techniques. Through them, new trace gases, previously unreachable through DOAS have been quantified on a global level, such as carbon monoxide through Weighting Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) [Buchwitz2000, Buchwitz2004].

WFM-DOAS is a trace gas column retrieval algorithm, developed in the beginning of the 21st century, with the main goal of determining the total columns of gases such as CO, Water (H_2O) or CO_2 from satellite nadir data [Buchwitz2004]. Instead of

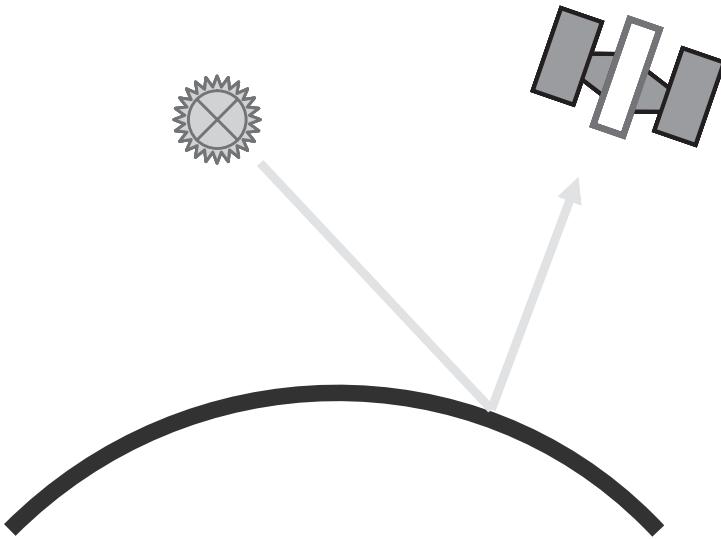


Figure 3.7: Schematic representation of the nadir satellite measurement geometry.

using literature-obtained cross-section data for each target trace gas, the [WFM-DOAS](#) approach uses weighing functions, calculated through the application of a radiative transfer model, such as SCIATRAN [[Rozanov2002](#)]. Still, the retrieval is based on a fitting process. Equation 3.9, adapted from [[Buchwitz2000](#), [Buchwitz2004](#)] can be called the [WFM-DOAS](#) equation.

$$\left\| \ln I_i^{obs}(\mathbf{V}^t) - \left[\ln I_i^{mod}(\bar{\mathbf{V}}) + \sum_{j=1}^J \frac{\partial \ln I_i^{mod}}{\partial V_j} \Bigg|_{\bar{V}_j} (\hat{V}_j - \bar{V}_j) + P_i(a_m) \right] \right\|^2 = \|RES\|^2 \rightarrow min \quad (3.9)$$

In Equation 3.9, I_i^{obs} is the observed sun-normalised radiance (the ratio between a nadir radiance measurement and solar irradiance) for the center wavelength λ_i of detector pixel number i . \mathbf{V} are the vector that have vertical columns as their components. These can be true, V^t , modelled, \bar{V}_j , or approximated, \hat{V}_j . The true columns are unknown as we only have the radiance value and not the things on which it depends, modelled vertical columns are taken from the literature (they are climatological values) and \hat{V}_j are one of the fitting parameters. The other being a_m , the low order polynomial (P) coefficients. RES is the fit residuum, which is minimised for the fitting.

3.4 Important Notes on [DOAS](#) in practice

Theoretically, the procedure described in Section 3.3 would be enough to obtain the target trace gas concentrations. In practice, this is an over-simplification. There are several additional required steps, most of them concerning a certain conditioning that one has to apply to both the collected and literature spectral signals.

3.4.1 Cross section conditioning

Unless conditions are absolutely stable, and one is able to record all data with the same device, which is impractical to the point of infeasibility, external literary sources are required for the DOAS analysis. These data are obtained with standardised trace gas sampling, extremely high resolution spectrometers and in carefully designed laboratory experimental setups.

To use external cross sections in one's own experiments, these data must be adapted to one's equipments. The very high resolution spectra coming from the literature are convolved with the instrument function of the real experiment's spectrometer. This function can be thought of as the device's impulse response. Since the spectrometer is inevitably imperfect, this response is not nearly as sharp as the impulse itself, and is normally modelled as a Gaussian curve fitted to a known source's well defined, impulse-like narrow structures. An Hg-Cd lamp was used in Merlaud's 2013 work, for instance [Merlaud2013]. For most spectrometers, this step is not actually required, as manufacturers already provide a very accurate spectral resolution value.

After calculating or fetching this value from the device's manual, it is a matter of creating a Gaussian kernel that can be used to filter the high resolution spectra through convolution. The **Full Width at Half Maximum (FWHM)** of the Gaussian kernel is equal to the spectrometer's resolution. One can ensure this by using the formula in Equation 3.10 which relates the width of the kernel with the standard deviation used to create it. The signal that results from the convolution of the literature cross sections and the gaussian kernel can be thought of as how the cross section would look if it had been acquired using the experiment's spectrometer.

$$FWHM = 2\sqrt{2\sigma^2 \ln 2} \rightarrow FWHM \approx 2.355 \cdot \sigma \quad (3.10)$$

The final step in this process of conditioning high resolution literature cross sections to the experiment that one is conducting is the discretisation of this signal. Physically, spectra are continuous signals. In practice, they were all captured using finite-resolution devices, and therefore are a digital signal. However, the resolution of the literature cross sections is so high in comparison to the usual resolution of a DOAS experiment spectrometer that one can think of this signal as being *quasi*-continuous. In order for them to be used in DOAS calculations, they have to be "re-discretised" onto the experiment's spectrometer sensor resolution. Since it is seldom the case (if ever) that the literature and the low resolution pixels coincide, a mathematical routine is used to interpolate the former onto the latter. The most commonly used routine is the cubic spline interpolation, widely regarded as the best compromise between computational expense and accuracy [Beekman2010]. Overall, cross sections used in DOAS undergo a process illustrated in Figure 3.8 [Danckaert2015, Beekman2010].

3.4.2 Spectral Calibration

One other small but important correction that can bias results is the lack of proper calibration between cross sections and collected spectra. Figure 3.9 shows how a conditioned cross section present slight wavelength related discrepancies.

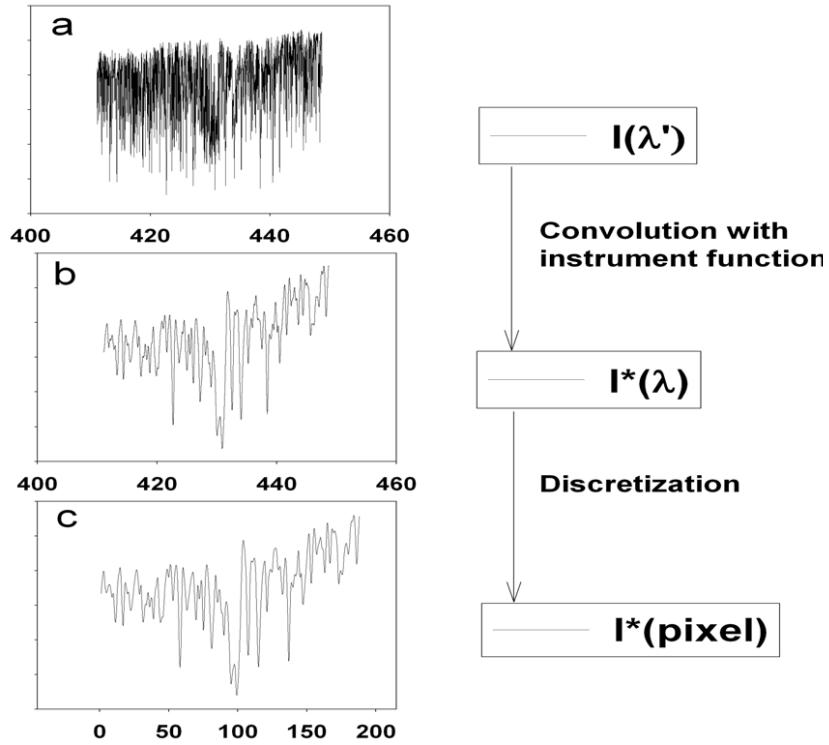


Figure 3.8: Before being used in DOAS calculations, literature trace gas cross section spectra must undergo a conditioning process to make sure they are compatible with the current experiment's instruments. Figure adapted from [Beekman2010]

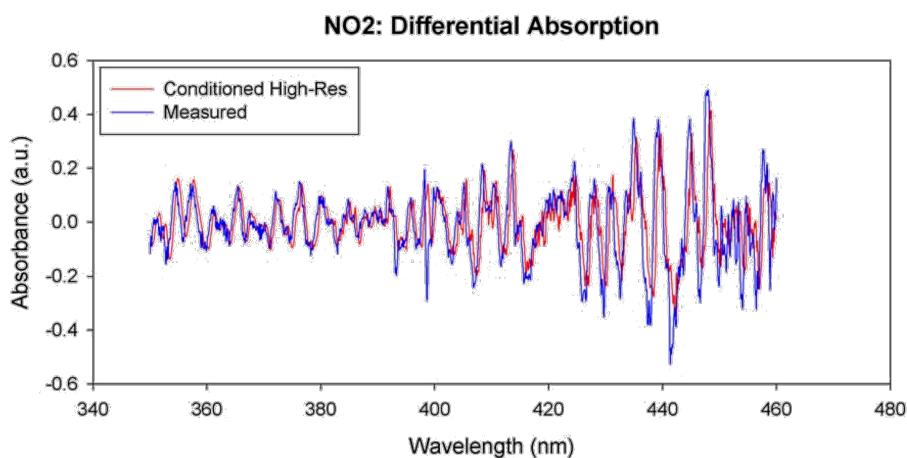


Figure 3.9: This figure, taken from [Beekman2010], shows the difference between NO₂ cross sections. The red line comes from the literature, and has been conditioned to the experiment spectrometer. The blue line was actually collected with that device. The slight discrepancies that can be observed are calibration defects.

These small nonconformities are caused not only because both types of data were captured by different spectrometers, but also because conditions most certainly changed. This means that situations in which calibration must be handled with care are the norm. This is done by running an iterative process that makes slight adjustments to the function mapping wavelength to pixel number, as explained in the next few paragraphs.

The light that enters the spectrometer is scattered by the diffraction grating before reaching the device's sensor, which is divided into pixels. These pixels have a well defined and finite width. At the centre of each one of them, the signal's intensity can be described by Equation 3.11, in which I is the intensity, i the pixel number and λ the wavelength.

$$I(i) = \int_{\lambda(i)}^{\lambda(i+1)} I(\lambda) d\lambda \quad (3.11)$$

The expression in Equation 3.11 assumes that the signal has been conditioned properly, as described in Section 3.4.1, through the convolution of the cross section data with the instrument function and adequate discretisation. The smaller $\Delta\lambda = \lambda(i+1) - \lambda(i)$, the finer the instrument's resolution. This is, of course, inherent to the system and cannot be changed. What can and should be changed during analysis is the central wavelength assigned to each pixel. This assignment is in fact what constitutes the instrument's calibration.

The most common way of conducting said calibration is by introducing a polynomial describing wavelength to pixel relationship. This polynomial is the wavelength-pixel mapping function, and can be written as in Equation 3.12.

$$\lambda(i) = \sum_{k=0}^q \gamma_k \cdot i^k \quad (3.12)$$

In Equation 3.12, γ_k determines how the pixels are mapped to the wavelength (λ), and the type of mapping effect depends on k . Changing γ_0 shifts the signal left or right; changing γ_1 introduces linear distortions to the pixel mapping, i.e., stretching or squeezing of the signal. One can fit this polynomial up towards $k = \infty$, in theory, but normal spectrometer calibrations are only run up to the second or third degree. The Avantes spectrometers that were used in this dissertation are factory-calibrated to a fourth order polynomial.

The non-linear portion of the DOAS algorithm, described in Section 3.3, is indeed a wavelength calibration. However, it is important to previously calibrate every spectral signal before actually running the DOAS fitting process. This is because the types of optimisation algorithm used in this technique, such as Levenberg–Marquardt, tend to find local minima or to not converge if there is significant misalignment between the spectra.

For passive applications using the sun as the light source, one can use a high resolution solar spectrum (see Figure 3.10), and the Fraunhofer bands in it (which have a very well defined wavelength that can be used as ground truth) to align the various spectra. For active applications, one can use the light source's own spectrum, either

provided by the manufacturer or previously collected in the absence of atmospheric or other effects.

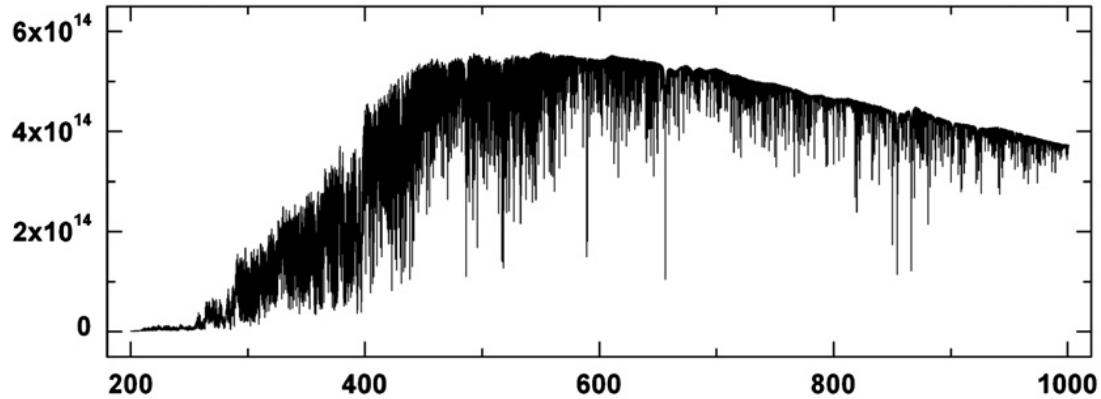


Figure 3.10: High resolution solar spectrum commonly used to calibrate the various spectral signals in DOAS experiments [Chance2010].

3.5 Tomographic algorithms and reconstruction techniques

Tomography is the cross-sectional imaging of an object through the use of transmitted or reflected waves, captured by the object exposure to the waves from a set of known angles. It has many different applications in science, industry, and most prominently, medicine. Since the invention of the Computed Tomography ([Computed Tomography \(CT\)](#)) machine in 1972, by Hounsfield [[Gunderman2006](#)], tomographic imaging techniques have had a revolutionary impact, allowing doctors to see inside their patients, without having to subject them to more invasive procedures [[Kak2001](#)].

The central thought around tomographic image reconstruction is to recreate the information contained inside a target physical body, without having to cut it or open it in any way. The theory is based on Radon's idea that it is possible to "*represent a function written in \mathbb{R} in the space of straight lines (\mathbb{L}) through its line integrals*" [[Radon1986](#)].

Say we have a body that we want to fully characterise without cutting open or destroying in any way. Now imagine we can traverse it with some kind of radiation, ray by ray, and that we are able to measure the rays after they traverse the target. What we would capture would be relative to the emitted radiation, of course, but it would also contain information on how that ray had interacted with the target body's matter. In the case of the ubiquitously used X-Ray radiation, the measurement would be one of the total attenuation "imprinted" onto the ray by the target body's molecules, in the ray's particular direction. If said body is heterogeneous, the total attenuation can be derived by the infinitesimal sum of all different attenuation phenomena caused by the object's several different constituents (the same can be said of a homogeneous object, but in that case there is only one type of attenuation present). This means that each one of the rays contains information regarding the constitution of said body.

The question that arises is thus "*how we can use this information to create a spatially accurate representation of this target's interior composition?*". The answer to this

3.5. TOMOGRAPHIC ALGORITHMS AND RECONSTRUCTION TECHNIQUES

question lies on many factors, but the most prominent of which are surely choosing the quantity that we are trying to find (that characterises the object) and assembling the projections (that is what we call the line integrals in tomographic imaging) in a way that allows solving an equation system for the aforementioned quantity. This assembly, a matrix of projections organised by their angles and position within the detector, is called sinogram. All tomography methods revolve around finding the relationship between it and the system's geometrical description [**Bruyant2002**, **Kak2001**, **Herman1973**, **Herman1995**, **Herman2009**, **Defrise2003**].

Let's consider the case in which we deal with a single ray of solar light entering the atmosphere at a given point. Since the atmosphere contains numerous absorbers and comparable atmospheric effects, the ray changes from the point where it enters the atmosphere to the point at which it is measured by a detector. Total absorption will depend on the pollutant species, their cross-section and their concentration, since it obeys Lambert-Beer's law. Looking from another angle, this absorption is also the line integral that we will use to reconstruct our image. With **DOAS**, it is possible to measure several pollutants at the same time, but for simplicity (and since it is one of the most studied compounds in the field), let's consider that the single pollutant in our atmospheric mixture is NO₂.

The problem of tomographic reconstruction can be approached in a number of ways, depending mostly on the authors. In my literary search, I have found that Kak and Slaney [**Kak2001**] have certainly explained this problem in one of the clearer ways available. Therefore, I shall base the rest of my presentation in their writings, and complement with other authors' notes wherever necessary.

Considering the coordinate system displayed in Figure 3.11. In this schematic, the object is represented by the function $f(x, y)$. The (θ, t) parameters can be used to define any line in this schematic. Line AB in particular can be written:

$$x \cdot \cos(\theta) + y \cdot \sin(\theta) = t \quad (3.13)$$

And if we were to write a line integral along this line, it would look like Equation 3.14, the Radon transform of function $f(x, y)$:

$$P_\theta(t) = \int_{-\infty}^{\infty} f(x, y) \cdot \delta(x \cdot \cos(\theta) + y \cdot \sin(\theta) - t) dx dy \quad (3.14)$$

Where δ , the delta function, is defined in Equation 3.15.

$$\delta(\phi) = \begin{cases} 1, & \phi = 0 \\ 0, & \text{otherwise} \end{cases} \quad (3.15)$$

As I have mentioned previously, a projection is a set of line integrals such as $P_\theta(t)$. Geometry plays a very important role in how the integrals are written and solved for reconstruction. The simplest case is the one where the set is acquired in a row, describing what is called a parallel geometry. Another more complex case is when a single point source is used as origin for all rays, forming a fan. This is called a fan-beam array. There are other possible geometries, but they fall out of the scope of this work and will therefore not be addressed any further.

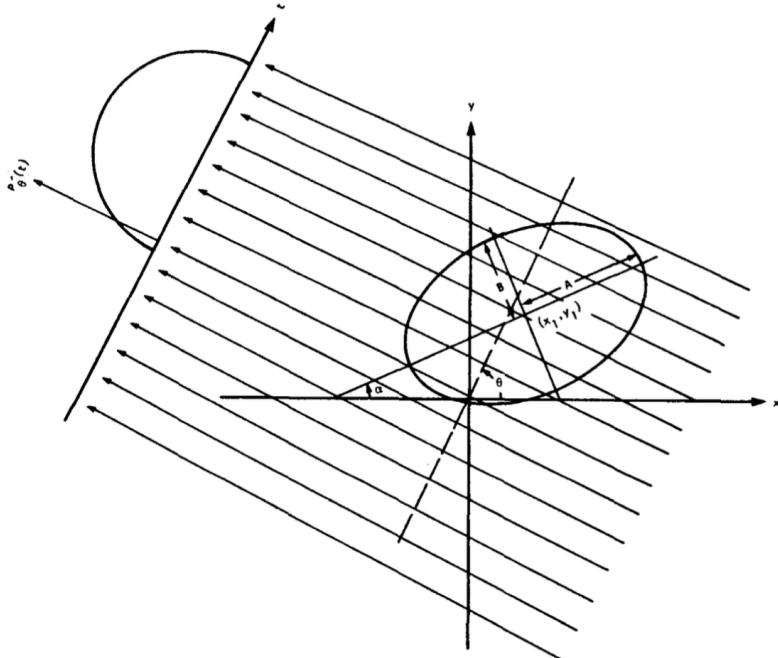


Figure 3.11: Schematic representation for coordinate setting. The image depicts a parallel projection setting [Kak2001a].

The Fourier Slice Theorem (**FST**) is the most important component of the most important algorithm in tomographic inversion, the Filtered BackProjection algorithm (**Filtered BackProjection (FBP)**). **FST** is based on the equality relation between the two-dimensional Fourier Transform (**Fourier Transform (FT)**) of the object function and the one-dimensional **FT** of the object's projection at an angle θ . Let's start by writing the 2D **FT** for the object function, Equation 3.16, and the 1D **FT** of projection P_θ , in Equation 3.17.

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cdot \exp [-j2\pi(ux + vy)] dx dy \quad (3.16)$$

$$S_\theta(\omega) = \int_{-\infty}^{\infty} P_\theta \cdot \exp [-j2\pi\omega t] \quad (3.17)$$

For simplicity, let's consider the 2D **FT** at the line defined by $v = 0$ in the frequency domain. We rewrite the 2D **FT** integral as:

$$F(u, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cdot \exp [-j2\pi\omega ux] dx dy \quad (3.18)$$

Notice that y is not present in the phase factor of the **FT** expression anymore, and this means we can rearrange the integral as:

$$F(u, 0) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \mathbf{f}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right] \cdot \exp [-j2\pi\omega ux] dx \quad (3.19)$$

Now, the **bold** part of Equation 3.19 is similar to Equation 3.14. It is precisely that equation, considering $\theta = 0$ and a constant value of x , as in Equation 3.20.

3.5. TOMOGRAPHIC ALGORITHMS AND RECONSTRUCTION TECHNIQUES

$$P_{\theta=0}(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (3.20)$$

This in turn can be substituted in Equation 3.19, finally arriving at:

$$F(u, 0) = \int_{-\infty}^{\infty} P_{\theta=0}(x) \cdot \exp [-j2\pi ux] dx \quad (3.21)$$

And this is the one-dimensional **FT** for the projection at angle $\theta = 0$. Finally, the enunciation of the Fourier Slice Theorem:

The Fourier Transform of a parallel projection of an image $f(x, y)$ taken at angle θ gives a slice of the two-dimensional Fourier Transform, $F(u, v)$, subtending an angle θ with the u -axis (see Figure 3.12)

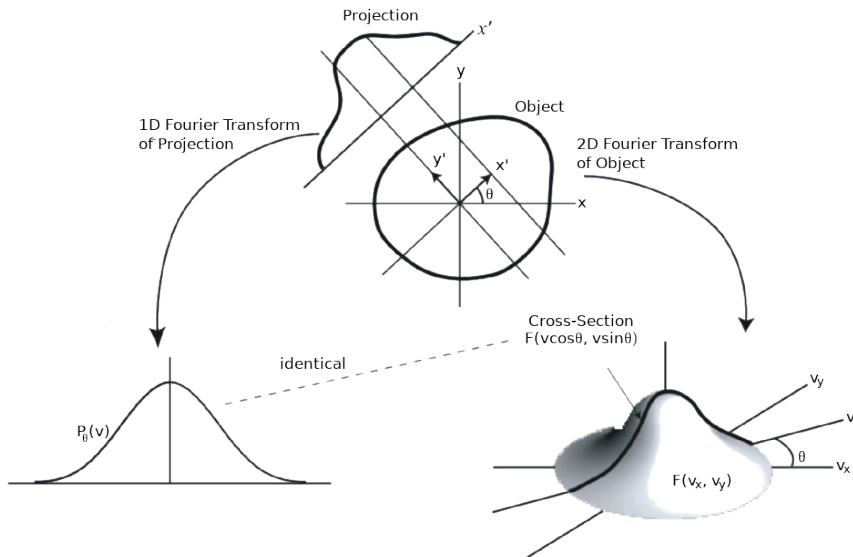


Figure 3.12: The **FST**, a schematic representation [Asl2013a].

If one takes the **FST** into account, the idea behind the **FBP** seems to appear almost naturally. Say one has a single projection and its Fourier transform. From the **FST**, this projection is the same as the object's two-dimensional **FT** in a single line. A crude reconstruction of the original object would result if someone were to place this projection in its right place in the Fourier domain and then perform a two-dimensional **Inverse Fourier Transform (IFT)**, while assuming every other projection to be 0. The result, in the image space, would be as if someone had smeared the object in the projections direction.

What is really needed for a correct reconstruction is to do this many times, with many projections. This brings a problem with the method: smearing the object in all directions will clearly produce a wrong *accumulation* in the center of the image, since every projection passes through the middle (remember we are still talking about parallel geometry projections) and are summed on top of each other, but on the outer edges, this does not occur. If one does not address this, the image intensity levels

in the reconstructed image will be severely overestimated in the center and underestimated in the edges (due to normalization). The solution is conceptually easy: we multiply the Fourier transform by a weighting filter proportional to its frequency (ω) and that encompasses its relevance in the global scheme of projections. If there are K projections, then it is adequate for this value to be $\frac{2\pi|\omega|}{K}$. As an algorithm, **FBP** can be written as in Algorithm 1.

Algorithm 1: The Filtered BackProjection Algorithm

Result: A reconstructed image of the projected object.

```

for  $\theta \leftarrow 0$  to  $180$  by  $\frac{180}{K}$  do
    measure projection  $P_{\text{theta}}(t)$ ;
    FT( $P_{\theta}(t)$ ), rendering  $S_{\theta}(\omega)$ ;
    Multiply by  $\frac{2\pi|\omega|}{K}$ ;
    Sum the IFT of the result in the image space;
end
```

Parallel projections, in which the object is scanned linearly from multiple directions, have the advantage of having a relatively simple reconstruction scheme. However, they usually result in acquisition times which are in the order of minutes. A faster way of collecting the data is one where all radiation emanates from a single point-source, which rotates around the target object (as well as the detectors). There are two types of fan-beam projections: equiangular and equally spaced. In this project, I have only worked with equiangular processes, so I will not include an explanation for equally spaced fan-beam projections. The reader may find this well described (much better than I would be able to) in [Kak2001] and [Herman1973].

Consider Figure 3.13. If our projection data were acquired through a parallel ray geometry, we would be able to say that ray SA belonged to a projection $P_{\theta}(t)$, in which θ and t would be written:

$$\theta = \beta + \gamma \quad \text{and} \quad t = D \cdot \sin \gamma \quad (3.22)$$

In Equation 3.22, D is the distance between the source S and the origin O ; γ is the angle of a ray within a fan and β is the angle that the source S makes with a reference axis. Through these relationships one can *translate* the parallel projection's FBP algorithm to the fan-beam case, which involves several complex geometric transformations, although the overall rationale is exactly the same.

Another particularity of fan-beam projection data is the fact that they can be sorted into a parallel projection. For that, one starts with the premise that if one were to substitute the fan geometry for parallel beams, most of the fan-beam rays would also appear in some projection of the parallel setup. This re-sorting algorithm starts with Equation 3.22. Now, if we call a fan-beam projection taken at angle β $R_{\beta}(\gamma)$, and a parallel projection taken at angle θ $P_{\theta}(t)$, one could thus write Equation 3.23, which can already be used to re-sort any fan-beam projection into parallel beam geometry.

$$R_{\beta}(\gamma) = P_{\beta+\gamma}(D \cdot \sin \gamma) \quad (3.23)$$

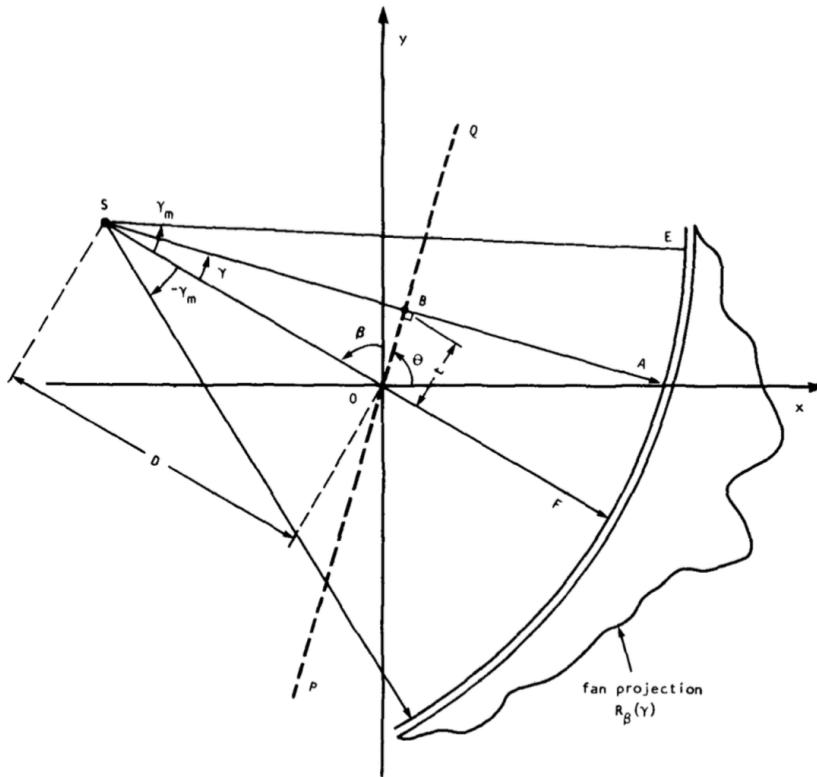


Figure 3.13: Schematic representation of an equiangular fan-beam projection, taken from [Kak2001].

Let's call the angular interval between fan-beam projections can be written $\delta\beta$, and the angular interval of rays within each fan is written $\delta\gamma$. In the case that they are the same ($\beta = \gamma = \alpha$), then it is the case that they can both be replaced by multiples of that interval in Equation 3.23, which becomes Equation 3.24.

$$R_{m\cdot\alpha}(n\cdot\alpha) = P_{m\cdot\alpha+n\cdot\alpha}(D \cdot \sin n\cdot\alpha) \quad (3.24)$$

Or, in non-mathematical notation, the n^{th} ray of the m^{th} radial projection (R) is the same as the n^{th} ray in the $(m+n)^{\text{th}}$ parallel projection. Although being much simpler than directly applying the [FBP](#) algorithm to the fan-beam projection data, this method has a limitation, which is the non-uniformity of the generated parallel projections. This can usually be corrected through interpolation [Kak2001a].

4

METHODS

Macroscopically, the approach to the **RQ** was conducted by working with two hypothesis:

First Hypothesis: The definition of a particular set of algorithmically defined projections in such a manner that they might be used for tomographic reconstruction of column densities of trace gases in the atmosphere, in a given **ROI**;

Second Hypothesis: We can retrieve the column density for a given trace gas (or set of trace gases) between two points by performing a spectral measurement in both of these points in the same direction and subtracting them one from the other.

Testing the first hypothesis implies the design of a whole system and requires a very diverse and multidisciplinary approach. I have already covered that the system should be drone-mounted. But which drone? And what is the collection system? And then again, what trajectory shall this drone adopt? Do we know it works?

Figure 4.1 is a general diagram of the envisaged solution. Section 4.1 aims to provide a complete description of how I have arrived at the system that I propose, and of the subsystems that comprise it. Subsection 4.1.1 is dedicated to the description of the drone and its subsystems, the blue box in the figure, Subsection 4.1.2 is about the particularities of the collection system and corresponds to the red box in the figure. The ground station is addressed in Subsection 4.1.3, which is the orange box in the figure. Finally, the trajectory simulation platform (green box) is the matter of Subsection 4.1.4.

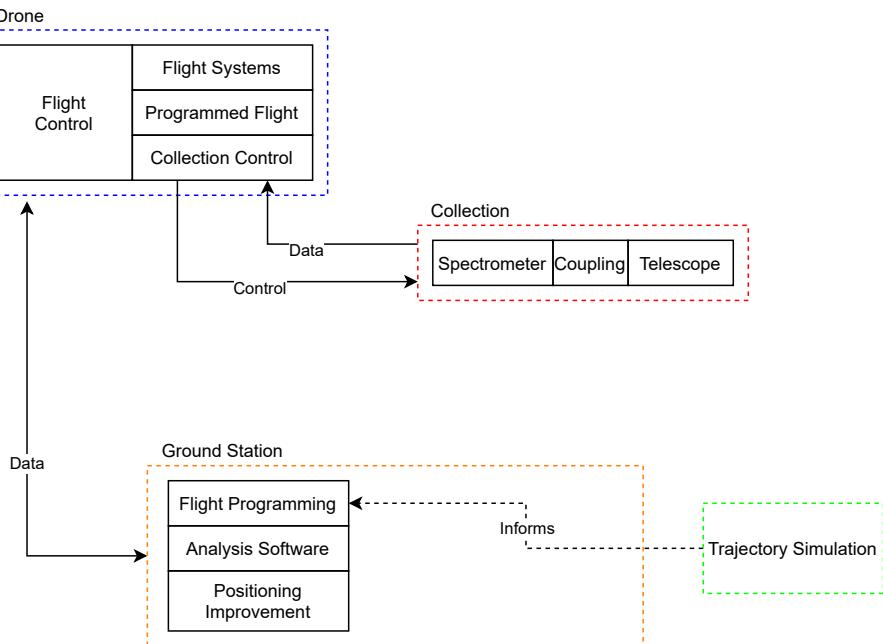


Figure 4.1: General diagram of the atmospheric monitoring system, in an overview style.

On Section 4.2, this chapter moves to a different subject. It is aimed at the testing of the second hypothesis. Figure 4.2 is a pictorial representation of what the proposed pollution monitoring system is designed to (ideally) do. Our second hypothesis is based on Lambertian theory, and tells us that one way of getting the density values in the ROI is to subtract whatever density is obtained to the left of the ROI from the density obtained by measuring on the right of the ROI. It is clear that, from a mathematical point of view, this is the case. But can I measure it with current-day off-the-shelf equipment? Subsection 4.2.1 addresses the theoretical perspective and reasoning for this idea. Subsection 4.2.2 presents the experiment that I have designed to test it.

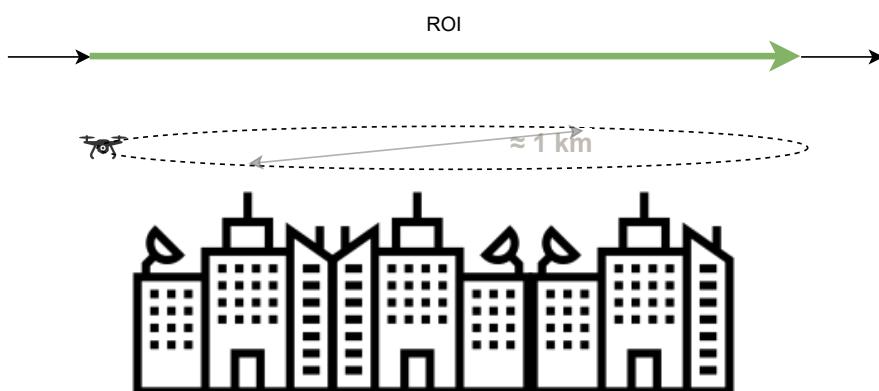


Figure 4.2: Schematic representation of the intended pollution measurement geometry. The second hypothesis of the RQ tells us that to get the ROI we can subtract the concentrations obtained on the left of the ROI from the ones obtained to the right of the ROI.

4.1 First Hypothesis

4.1.1 Unmanned Aerial Vehicle

Although it was not possible to assemble the final physical system onto an actual drone, I have managed to specify it completely in a custom design for which Figure 4.3 is a basic schematic ¹.

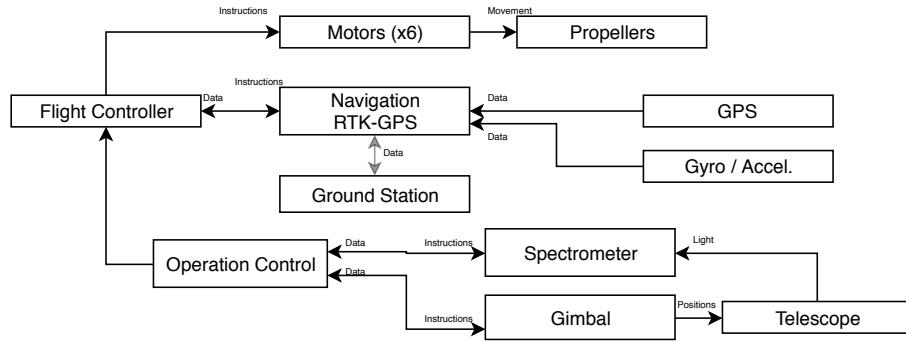


Figure 4.3: A schematic for the custom-designed drone system on which the pollution monitoring system would be mounted.

Our flight time requirements are between 30 to 40 minutes for each measurement cycle. During this time, the drone would have to carry itself and the collection equipment while still being able to land in a safe position before recharging or replacing the batteries. Sufficiently capable commercial systems were prohibitively expensive, thus I had to design my own drone.

I began by selecting the type of drone I would use. Hexacopter models are known around enthusiasts to be more efficient and stable, especially if there is a high-payload requirement. They have several advantages over the more "traditional" quadcopter designs. The most important is certainly the fact that they can fly and land safely in case of one or even two motors malfunctioning. Moreover, they are much stabler than their 4 motor counterpart, and they are for the most part immune to (up to) moderate winds. Finally, they are able to climb higher, since they are less affected by thin air. They also have disadvantages, such as their steeper cost, their size and their inherently higher complexity. These drawbacks were easily manageable though, through the European funds of project [Project ATMOS](#) and the technical abilities of the team in which the project was being developed.

I have therefore settled on a DJI S900 frame. This was, at the time of selection, being discontinued by the Chinese brand. However, this model remained popular among custom builders because of the familiarity people already had with it and the fact that it was still commonly available in specialised retailers. I was advised by two specialists to increase the device's breadth through the addition of custom made 368mm arms. This inclusion served two purposes: on the one hand, it reduced weight, since the original arms were made of aluminium and the new ones were carbon fibre; on the other hand, they granted the room needed for wider propellers, with 17" blades. To

¹To the reader: although this is still a fairly advanced design, recent advances in drone technology might render it a bit obsolete in some aspects. Please bear in mind that this design was finished in March 2019

accommodate the change in propeller size, it was also advisable to replace the original motors by more powerful units. In this case, I chose the DJI E1200. Larger propellers, lighter body and more powerful motors led to significant improvements in both flight time and carrying capacity. According to the manufacturer, this configuration is able to handle a maximum weight per rotor of 3900g, or almost 24 kg in total. Assuming our drone to weight 6 kg, this gives us a very comfortable weight margin for our payload, which in any case should be slightly below 2 kg [DJI2015].

A Pixhawk flight controller takes care of the aerial dynamics, movement and positioning. This flight controller comes with all the needed sensors (gyroscopes, barometers, magnetometers, accelerometers, etc.), requiring only an external navigation unit. The unit natively supports [Real Time Kinematic GPS \(RTK-GPS\)](#), a combination of inertial sensors and traditional satellite navigation data that achieves positioning precisions of a few tens of centimetres [Zimmermann2017, Triantafyllou2011].

4.1.2 Collection System

The flight controller unit is in permanent communication with the collection system, through the spectrometer controlling computer. This is an [Raspberry Pi \(RPi\)](#)-based single board device. This equipment controls the spectral acquisition process, which is conducted by an Avantes Mini spectrometer with 2048 channels and a [Universal Serial Bus \(USB\) 3.0](#) connection. The Avantes spectrometer can be interfaced with both Unix-based and Microsoft Windows operating systems through a set of software libraries that are available for download at this manufacturer's website. There are several physical interfaces available for both types of operating system. Nowadays, the most common and expedient way to run this connection is through the [USB](#) port. This is also the connection that allows a higher data throughput (allowing smaller integration times to be used). In the end, the proposed system will only run in Unix-based computers, but the Windows version was very important for initial experiments, which at the time were programmed in C# (code included in Appendix ??). The spectrometer control flow is very different in the two approaches. The Unix version works by continuously polling the spectrometer for the spectral data, which is the somewhat obvious strategy. The Windows version uses a sophisticated technique: Windows messages. This is a set of fixed-value event flags that are fired at the Operating System level and can be intercepted by running programs.

This type of spectrometer is usually shipped with an [SubMiniature version A \(SMA\)](#) connector, for direct connection of a fibre optics cable. While this is ideal as a bench-top solution and when size and weight restrictions are looser, we found it not to be the best fit for our particular case. We tend to think of fibre optics as an almost lossless medium for data transmission. This is, of course, true, but it does not hold for its connectors. They represent one of the most significant sources for signal loss in an optics fibre line. Avantes does not mention any value for [SMA](#) connector losses, but the traditional figure that appears in most manufacturer's catalogues is around 1dB [diamond2021, Optics, amphenol2002]. To test how a direct connection fares when compared to a fibre connection, I designed a spectrometer support that can be attached to the telescope via its red-dot rail and fabricated it using a 3D printer. Technical drawings available in Appendix ??.

CHAPTER 4. METHODS

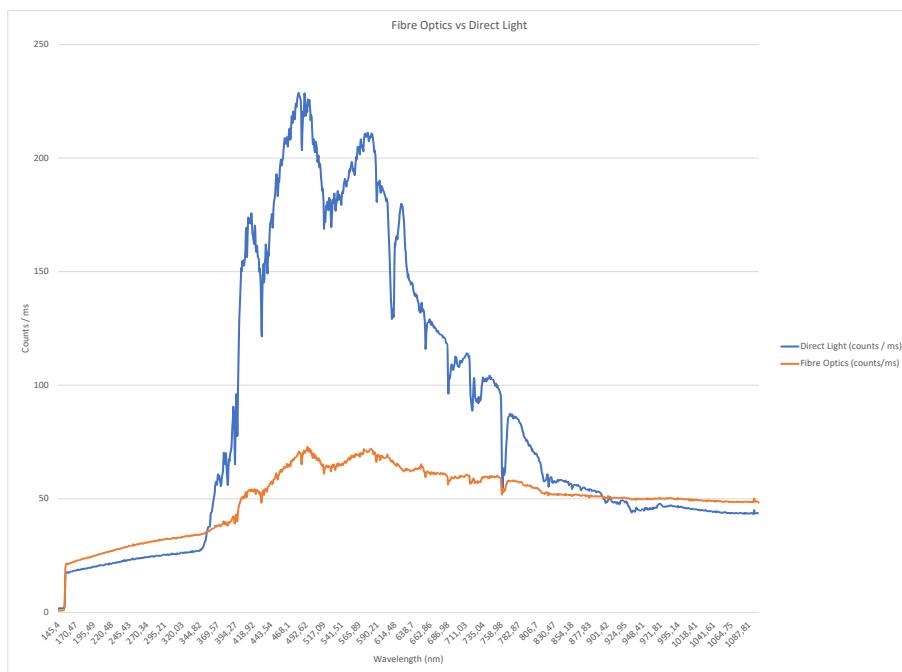
The assembly is displayed in Figure 4.4. In this experiment, I have found 75% energy difference between the fibre optics cable and a direct measurement in which the light goes directly from the telescope into the spectrometer's connector. These results are plotted in Figure 4.5.



(a) This assembly uses a traditional fibre optics connection to transport light between the telescope and the spectrometer.

(b) This assembly allows light to pass directly from the telescope to the spectrometer.

Figure 4.4: Conducted experiment to evaluate advantages of using a direct light connection instead of fibre optics. The spectrometer support (that can be seen in red) was designed in Autodesk Inventor and 3D printed.



spectrometer's entrance. This design was made with Autodesk's Inventor CAD software suite ² and some renderings of the designs can be seen in Figure 4.6. Technical drawings are also available in Appendix ??.

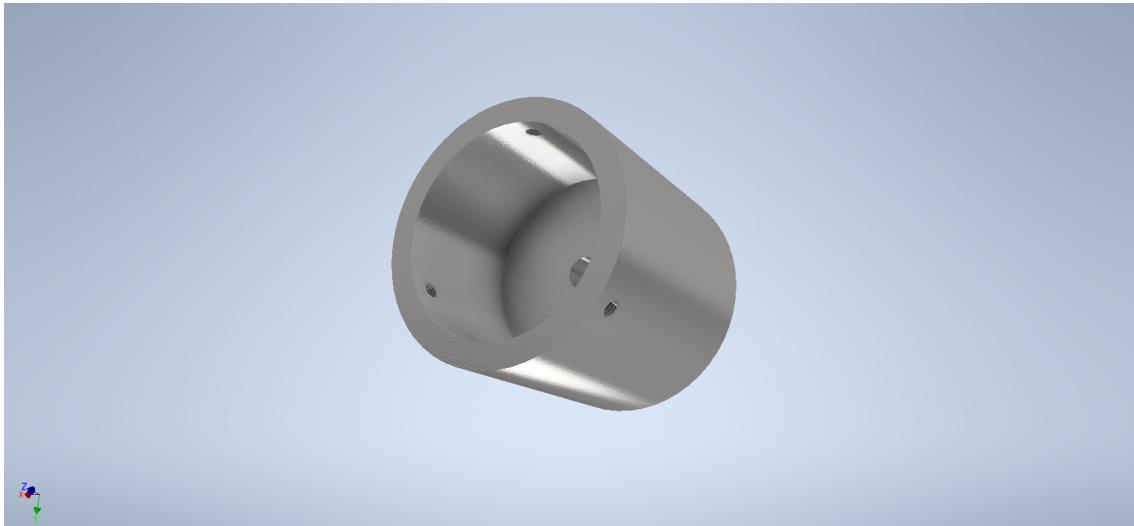


Figure 4.6: Telescope - Spectrometer coupling. 3D design rendered through Autodesk Inventor.

The centrepiece of this collection system is of course the telescope. From the beginning of the project, I understood that this was going to be one of the most important aspects of the endeavour. It has gone through some iterations before reaching the final design solution. The telescopes with which I began my experiments were some of the ones used in the FFF project, which were normally Meade ET90 or similar (90mm diameter Maksutov-Cassegrain telescopes). These seemed too bulky to mount on a drone. They would fit, of course, but movement would become a very probable problem. Looking for alternatives at the time rendered no results. The 90mm diameter telescope market has a lot of competition, but smaller tubed reflecting telescopes are somewhat rare. On the refractors side of things, the problem was otherwise. Telescopes are usually built to look at the skies. Manufacturers try to offer the best balance between magnification, image quality and ease of use in each price range. Therefore, commercial refractor telescopes were too long to include in our drones, as they reached for larger focus lengths to get more magnification power. Our needs were in fact quite different. We wanted the light collection capabilities of a telescope tube, but since our idea was to connect it to a spectrometer, optical quality was not at all a priority, and neither was magnification. With this in mind, I started working on a custom made tube design.

I started by choosing the appropriate tube diameter. On a cloudy but luminous day, I used an ET90 to capture the spectrum that I show in Figure 4.7. This spectrum was collected with the Avantes Mini spectrometer, with an integration time of 50 ms. The Poissonian nature of light entering a telescope allows me to assume that its relationship with time is linear. Moreover, we know that the quantity of light that

²Thanks to a very generous protocol between Autodesk and many universities around the world that makes this software free for all FCT NOVA's students.

enters the telescope varies with the square of its diameter. Thus we can write the calculations in Equation 4.1.

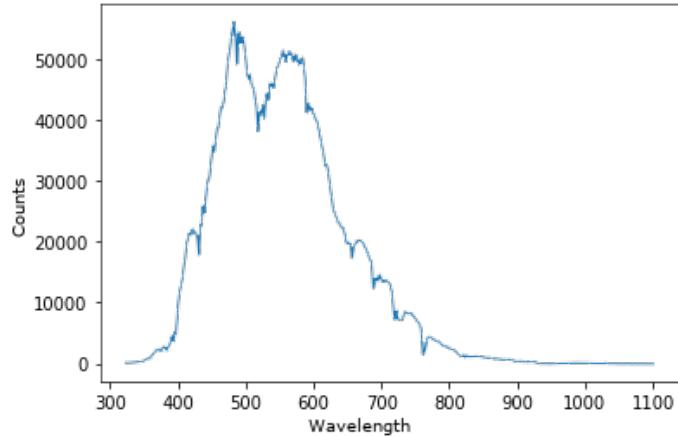


Figure 4.7: The spectrum captured on a cloudy but luminous day using the ET90 telescope and the Avantes Mini spectrometer.

$$E \propto D^2 \Leftrightarrow \frac{E_{90}}{E_{25}} = \frac{90^2}{25^2} \approx 13 \quad (4.1)$$

The same spectrum of Figure 4.7 would take 13 times longer to collect with a telescope that had 25mm instead of 90mm diameter. In other words, 650 ms. This is a reasonable compromise given our application.

The first design attempt was a Maksutov-Cassegrain telescope. However, this was quickly abandoned due to 1) calculations indicated an impossible length for this kind of telescope and a 25mm diameter; and 2) the inherent complexity of these telescopes made assembling one prohibitively costly in terms of time.

The second design was much simpler. It was a refracting tube with the same diameter (25mm) and a focal length of 300mm. Designs for this telescope and its optical simulation were produced using OSLO-EDU software and are included in Appendix ???. This ended up not being the telescope used in the final design, because as I was working in this part of the project, Omegon released their MightyMak 60, a 60mm diameter Maksutov-Cassegrain telescope that is small enough to be fitted onto a drone assembly. This was a perfect timing because the commercial availability of a suitable telescope allowed significant time savings on the whole ordering, assembling and testing process. The specifications of the new telescope are included in Table 4.1 [[Omegon 2021](#)].

Table 4.1: Specifications table for the Omegon MightyMak 60 telescope [Omegon2021].

Feature	Value	Unit
Aperture	60	mm
Focal Length	700	mm
f/	11,7	N/A
Maximum Useful Magnification	120	x
Weight	0,6	kg

4.1.3 Ground Station

As illustrated by Figure 4.1, the ground station is the device responsible for:

1. Improving positioning precision;
2. Programming the UAV's trajectory;
3. Analysing the incoming data.

Now, the first point is handled in automatic fashion by the RTK-GPS devices that are included in the assembly, so there is no intervention required on my behalf; the second point, flight programming, is achieved via Arducopter's programming API's. These are software libraries that allows one to program a drone's flight pattern using regular programming languages, such as Python. Unfortunately, there was not enough time for me to learn how this is done. Finally, the ground station is supposed to process the data that come from the drone.

The DOAS software library is a Python package developed specifically for this thesis data processing operations. It was, as other components, designed using an OOP approach and following the SOLID principles of Object Oriented Programming. This piece of software was written in response to the initial research that I undertook and that returned no usable results in terms of modular, compact Python libraries for DOAS applications, that I could use in my work. It is, as far as I know, the only DOAS solving application with this kind of structure.

The library (UML diagram presented in Figure 4.8) models a DOAS application through the instrumentation lens. A DOAS application is always parametrised through its spectrometer's physical features and limitations, which in turn determine the structure of the analysed spectral data, and even the differential cross sections of the trace gases that are to be studied. Of course, this library is much more limited in its capabilities than some specific programs that have become commonplace in this kind of application, such as QDOAS [Danckaert2015], but the fact that it can be operated through a Python program and that one can manipulate the data through such tools as Pandas DataFrames more than make up for this lack. Moreover, since it is in effect a software library, it is also as flexible as one is willing to expand it.

An important side note that attests to this library's relevance is that it has been fully integrated in FutureCompta's Bee2Fire software, with further developments being conducted through this team's efforts. This is the first commercially applicable result provided by the work in this thesis.

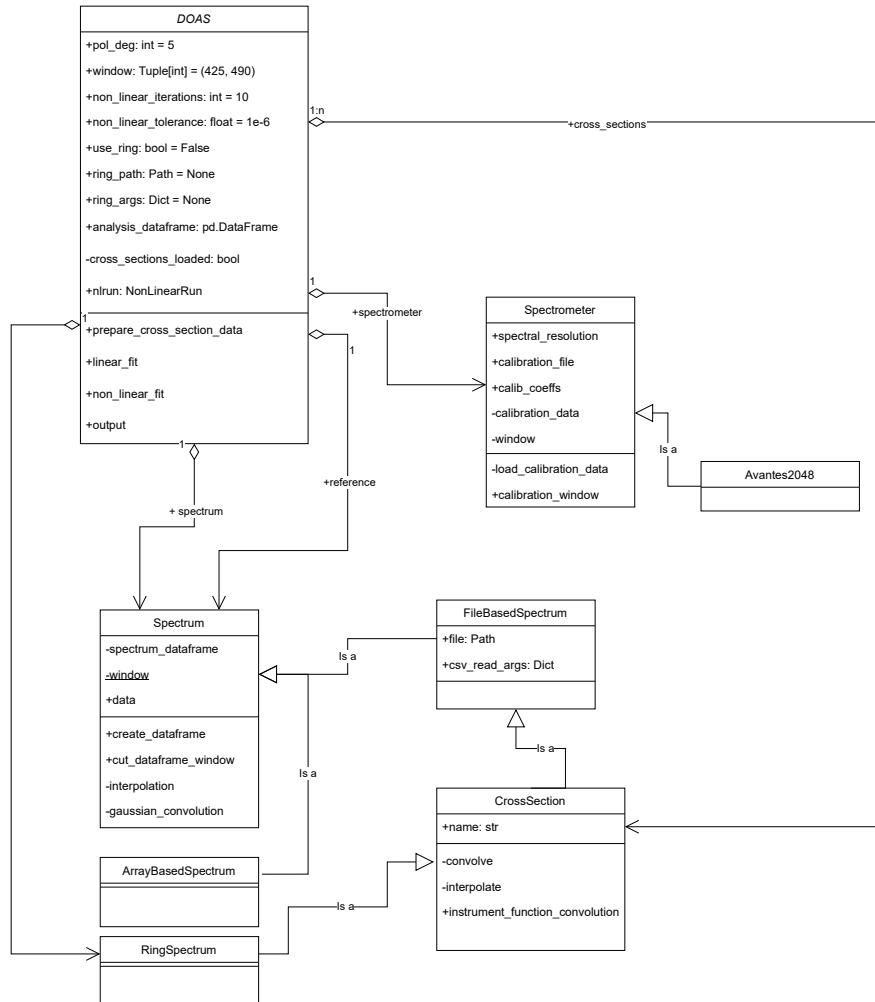


Figure 4.8: UML diagram for the **DOAS** library. The **OOP** approach that was followed allows for an instrument-oriented experiment parametrisation, which is not available in any other software

4.1.4 Tomosim

Tomosim is the software program designed for the simulation of the image reconstruction process in the proposed atmospheric monitoring system of this dissertation. It corresponds to the last box of the schematic presented in Figure 4.1. The general workflow of this piece of software can be viewed in Figure 4.9 and Figure 4.10.

One of the main goals of this piece of software is the evaluation and characterisation of possible drone trajectories that allow the capture of sufficient projection data to perform the tomographic inversion. The trajectory is a key element of the system, as it determines basically everything in the experiment, and most prominently, its tomographic geometry. Currently, Tomosim can only use one type of geometry. This is the fanbeam geometry, as described in Section 3.5. It was chosen because it seemed to be more promising taking into account the several important restrictions that this system imposes, namely in terms of flight time reduction. In essence, the drone's trajectory (illustrated in Figure 4.11) is a horizontal circle which is parametrised to be at a certain height and to have a certain diameter. Both of these dimensions are

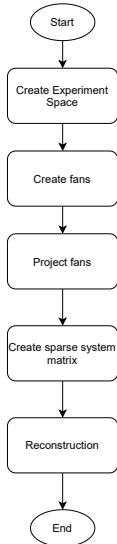


Figure 4.9: A flowchart describing the beginning stages of the Tomosim simulation.

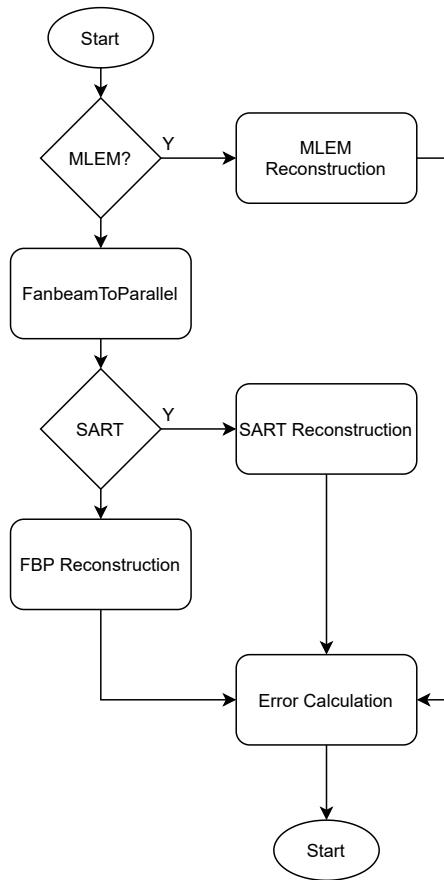


Figure 4.10: Flowchart representing the Tomosim operation during the reconstruction phase.

set at experiment / measurement time. The drone stops on this circle at regular angular intervals, say α degrees. Each one of these stops ($360 / \alpha$ stops) will generate a fanbeam projection, by pointing the optical system inwards (with respect to the circular macro-trajectory) and performing a series of spectroscopic measurements in different directions and also at regular intervals, say γ degrees. The particular case in which $\alpha = \gamma$ is very interesting, because it then opens the possibility for resorting the fanbeams into parallel virtual-projections that are much easier to reconstruct tomographically, as introduced in Section 3.5.

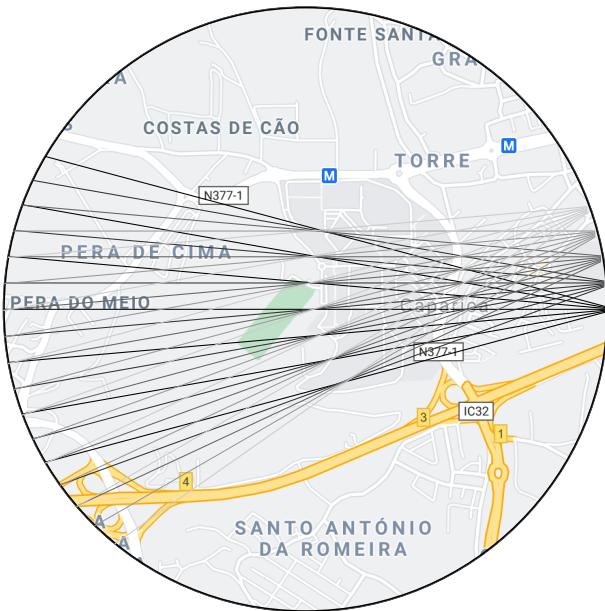


Figure 4.11: Illustration of the projection gathering algorithm based on the assembly of groups of acquisitions as fanbeam projections. The lines presented on the drawing come from the right to the left. Each is a ray in fan, corresponding to one stop of the drone. In the second measurement moment, the drone moves to the exit point on the left and repeats the first moment collection. Here, both fans and rays are separated by 5 degrees and there are 6 rays within each fan. A "real life" acquisition would feature a lot more rays, but that would be graphically complicated for the reader. Both ray and fan angular distance are customisable at runtime. Map taken from Google Maps in 2019. ©Google.

Of course, since we want to calculate molecular density in a give geographic region, we need to delimit our measurements to this space. The system is able to calculate the point where a given ray of light exits our ROI from the point of entry and its direction. Figure 4.12 is a schematic snapshot of a point in which the drone is taking a spectrum in one of its stops. Here, the drone's position (P_1) is given by the distance D and the angle β . The gimbal is pointing at a direction at an angular distance of γ form line OP_1 . Point P_2 , which is not known, is at the intersection between the trajectory's circumference and line P_1P_2 . Now, any point on this line can be expressed parametrically, with the sum of a point and a vector; while to say a point is on a circumference is the same as saying the distance between that point and the centre of this circumference is equal to its radius. The situation can be described by the expressions in Equation 4.2.

$$X = P_1 + t \cdot (P_2 - P_1) \quad (4.2)$$

$$|P_2| = D^2$$

Unravelling these expressions, and making use of the algebraic property that says that $|A|^2 = A \cdot A$, the expression becomes a second degree equation, as stated in Equation 4.3, writing $P_2 - P_1$ as V.

$$t^2V^2 + 2 \cdot V \cdot P_1 \cdot t + P_1^2 - D^2 = 0 \quad (4.3)$$

If line P_1P_2 non-tangentially intersects the circumference, solving Equation 4.3 renders two values for t (which correspond to P_1 and P_2). Selection is made by determining the returned value of t which maximises the euclidean distance between the produced point and P_1 .

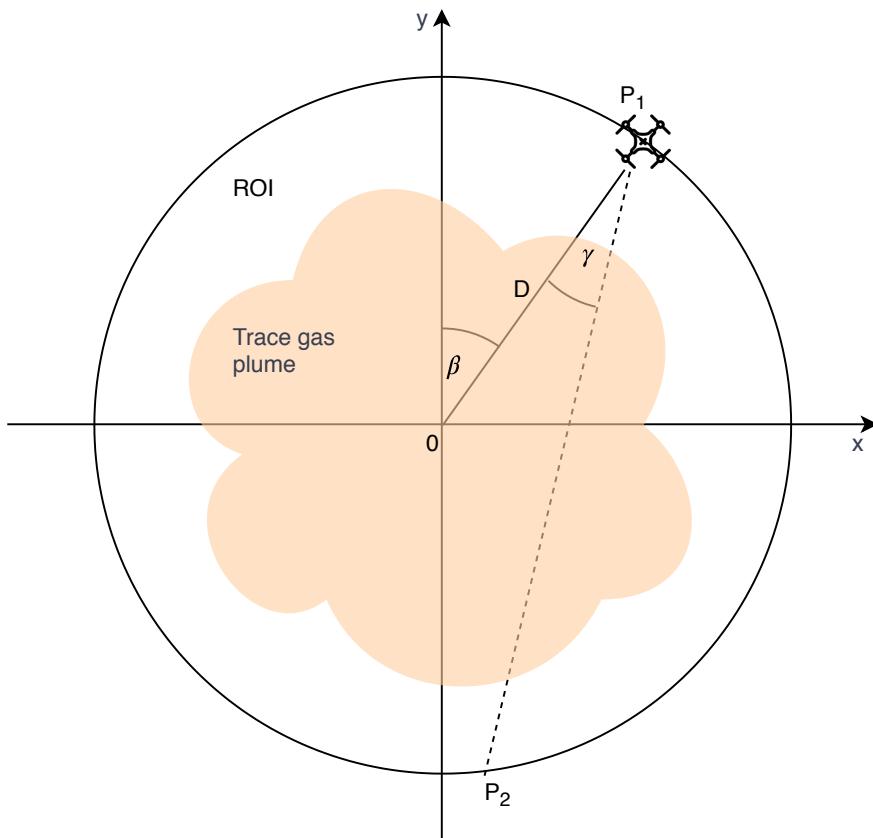


Figure 4.12: P_2 Calculation. The software uses the position of the drone, as defined by the vertical angle, β , and the distance between the drone and the centre of the trajectory, D , to determine P_2 through the solution of a second degree equation.

At this point, I have described the drone's trajectory and how the simulated drone calculates where the light enters and exits the ROI. But what exactly is in it? What will the simulation reconstruct?

A phantom is a device that represents the human body or some of its parts. They have been used in medical physics since the beginning of the field. In medical imaging, for instance, phantoms started being used in the late nineteenth century and early

twentieth century. At the time, it was very difficult to find volunteers for any kind of experiment that involved radiation, due to the common effects that were rapidly reported by the first people subject to this kind of intervention [Dewerd2014]. In spite of this difficulty, scientists and researchers still had to determine the dosimetry properties and physical limitations of their radiative devices, so medical physicists had to develop their own test models, or phantoms, for this effect. More recently, phantoms have been designed to develop computed tomography applications and algorithms. These phantoms mimic the body's attenuation properties in the X-Ray section of the electromagnetic spectrum, for instance.

Although the system that I propose does not aim at measuring or using the human body (or any other animal's), the concept still stands. To evaluate our reconstruction methods and the validity of our data gathering strategies, I needed an atmospheric phantom.

The distribution of gases in the atmosphere is completely different from biological tissue. Therefore, medical imaging phantoms were not adequate. The design that I have created is based on the premise that a two-dimensional Gaussian peak is more appropriate to describe the smoother nature of gaseous distribution [Stachniss2009]. This in contrast with the crisply defined edges of a medical tomography phantom such as Shepp-Logan's head phantom [Shepp1974].

To design the phantom itself, I used a library called TomoPhantom [Kazantsev2018], a tomographic phantom generator that provides a Python Application Programming Interface (API), making it trivial to include in the Tomosim simulator. The new phantom is comprised of 5 Gaussian profiles, depicting a static gas mixture. An ellipse is also in the phantom, near one of the corners. This serves mainly as a reference point for reconstruction, given its more solid and crisp nature. The new phantom can be seen in Figure 4.13 and its features are stated in Table 4.2.

Table 4.2: Table summarising the new phantom's construction details, as a sum of 5 Gaussian profiles and an ellipse designed using TomoPhantom. In this table, Co is the object's amplitude, Xo and Yo are its center coordinates, and a and b are the objects half-widths. The table is constructed using TomoPhantom's particular syntax and more information can be obtained at [Kazantsev2018].

Type	Co	Xo	Yo	a	b	Angle
Gaussian	1	-0,1	-0,1	0,25	0,5	-45
Gaussian	1	0,6	0	0,65	0,45	-45
Gaussian	1	-0,6	-0,4	0,8	0,8	0
Gaussian	1	-0,4	0,8	0,7	0,7	0
Ellipse	1	0,4	-0,8	0,3	0,15	0

The phantom is then discretised, using Siddon's algorithm before being reconstructed using one of the algorithms described in Section ?? : FBP, Maximized Likelihood Expectation Maximization (MLEM) or Simultaneous Algebraic Reconstruction Technique (SART). Here, there is an additional step if opting for FBP or SART. Both algorithms come directly from a dedicated tomography library and expect to receive parallel projection data as their input. Tomosim resorts the collected fanbeam projections onto a virtual parallel geometry by using the procedure described in Section ??.

[link to theory](#)

[Link to theory](#)

[link to theory](#)

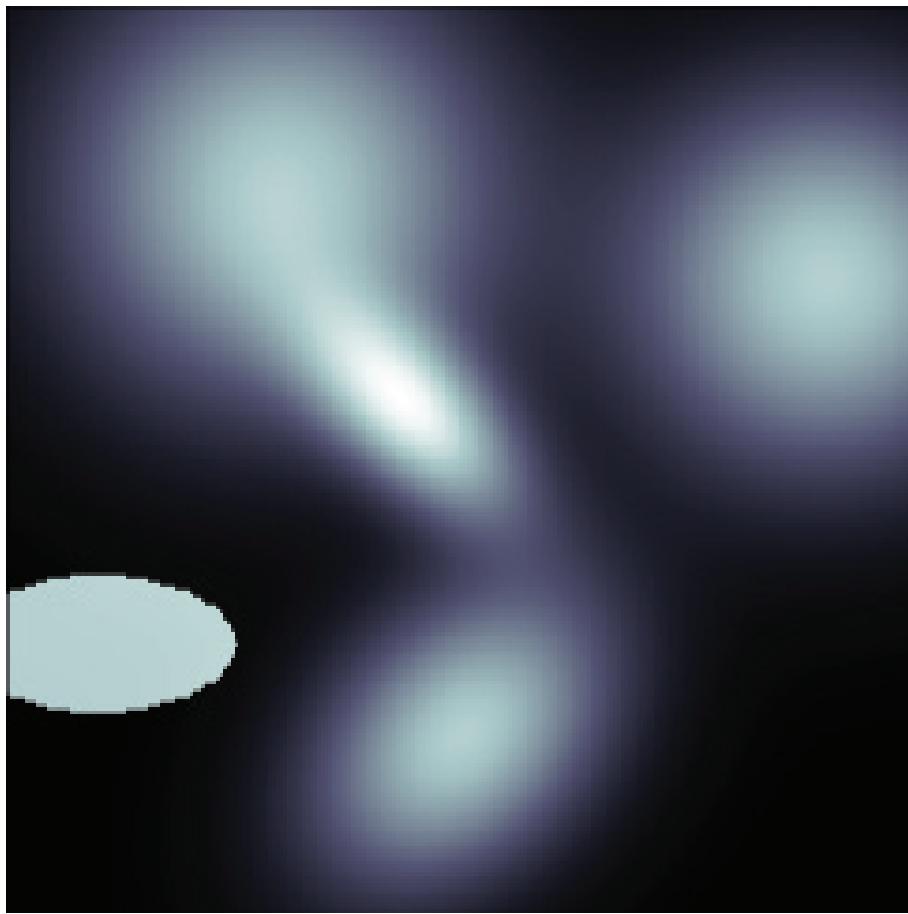


Figure 4.13: A graphical representation of the new spectral phantom, custom built for the TomoSim application.

I could not find any library that already applied this resorting algorithm, so I wrote my own, inspired by Matlab's `fan2para` routine.

In contrast, the `MLEM` algorithm was hand coded especially for Tomosim, and relies heavily on the fact that (as described in Section ??), this algorithm is a series of algebraic operations between matrices.

Error estimation is an important part of every tomography method and especially in simulations, as it allows one to approach the results with much more confidence of their similarity to the real world. Error sources for the Tomosim simulator come in four different natures: time errors, geometric errors, spectroscopic errors and reconstruction errors.

Time errors come from the fact that there are two moments of measurement. In a dynamic system, the time that passes between the two is enough for concentrations to change significantly. Tomosim does not address these errors, because they can be eliminated by the introduction of a second drone carrying the same type of equipment, which would eliminate said time difference.

Geometric errors exist due to the drone not being able to situate itself perfectly. There is always a positioning error, no matter how sophisticated the onboard equipment is. This type of error is addressed in the simulation through a Monte Carlo like approach. Positioning and pointing errors are assumed to have normal distributions.

[link to theory](#)

Each time a point is calculated by the drone, it generates a normally distributed random number, with a mean of 0 and a standard deviation equal to the nominal error of the positioning system. This number is then added to the theoretical point. Figure 4.14 is a graphical representation of the reasoning behind the calculation of the geometric error. The image deals with two types of error. One comes from the [Real Time Kinematic Global Positioning System \(RTK GPS\)](#) positioning system (the positioning error, ϵ_p); and the other that comes from the gimbal (the pointing error, ϵ_γ). The two ϵ values are the nominal error for the positioning and the pointing devices. The error is introduced in the simulation through the values of β and D (see Figure 4.12) while calculating P_2 . Given the very low nominal error for the gimbal, the small angle approximation is valid ($\sin \theta = \theta$). This is used to determine the theoretical value of P_2 , located on the device's circular trajectory. Finally, the software adds the positioning error, using the same process as in P_1 's case. The error depiction in Figure 4.14 is extremely exaggerated for visibility.

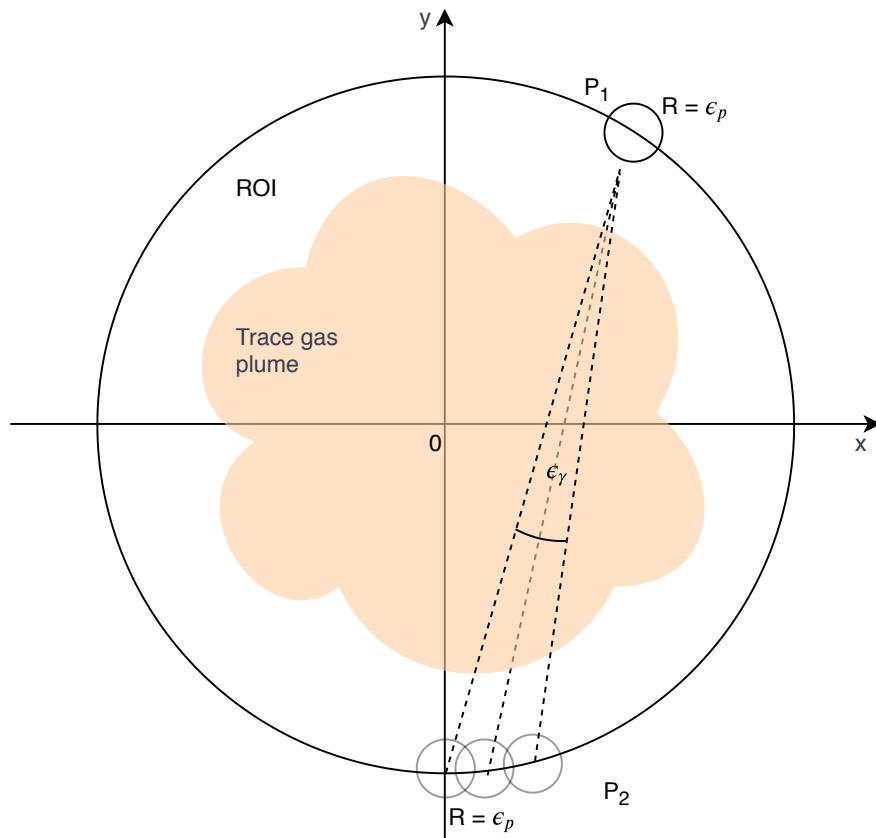


Figure 4.14: Error estimation graphical representation. Note errors are extremely exaggerated for visualisation purposes.

The third type of error are the spectroscopic errors. These come from the spectroscopic equipment that is used to gather projections. To take this noise into account, the simulator adds a Gaussian noise spectrum to each measurement, which is configurable through its standard deviation, as was previously done in [[Stutz1996](#)]. This is a valid approach, insofar as the captured spectra are perfectly calibrated regarding spectral shift and squeeze. Since this is a simulation software, this is an acceptable

assumption.

The final type of error that the system needs to contend with is the reconstruction error. In tomographic inversion problems, it is common to use techniques such as the **Mean Squared Error (MSE)** as a metric for an algorithm's performance. Tomosim was also evaluated in this light and in two separate ways. The first was to calculate the **MSE** for each pixel of the whole image. This information can still be viewed as an image (it is a two-dimensional grid of values) and paints an immediate picture of the general behaviour of the reconstruction algorithm. Moreover, it can tell the viewer if there are any types of shapes or areas in which the algorithm has more difficulties. The second way of using **MSE** to evaluate the reconstruction is to calculate a score through Equation 4.4. In this equation, and with respect to this simulator, f is the original image and g the one reconstructed from projections.

$$E = \sqrt{\frac{\sum |g(x, y) - f(x, y)|^2}{\sum |f(x, y)|^2}} \quad (4.4)$$

Just as the **DOAS** library, which is described in Section 4.1.3, the Tomosim simulation was programmed using the **OOP** paradigm. And also like the **DOAS** library, the SOLID principles of **OOP** were generally observed. The global flowcharts of the software's operation can be viewed in Figure 4.9 and Figure 4.10. As the application grew and I continued to work on it, it became quite clear that the original architecture was compromising the tool's performance and the code would need to be heavily overhauled. Although the time-frame of this work did not allow this, I was able to devise a basic new architecture for this software.

Ideally, Tomosim would be an extension of the more generic **DOAS** library. The only reason why it is not is purely historical: Tomosim's development was started before the former library library. Therefore, the first part of any refactoring exercise should be to mend this error. The new architecture, represented in Figure 4.15, is characterised by an increase in modularity. This is achieved through the introduction of the new classes "Reconstructor" and "RecEvaluator", which are composed into an also new class called "Experiment". The introduction of the two first classes greatly increases Tomosim's flexibility, as one would now be allowed to introduce any new reconstruction technique or evaluation method, without ever having to touch code already in place. To ensure this functionality, it is highly advisable that these two classes are built using the Template Design Pattern, with structural enforcing in the form of key abstract methods. This new design also favours the implementation of a mix between the factory design pattern and the strategy design pattern. This would result in factory methods being used to create Experiment objects based on some input parameters that would (for instance) determine the types of reconstruction and / or evaluation performed by said object.

Improving the system's architecture is a necessary and very important development, but the refactoring effort must also comprehend the performance side of the application. Currently, Tomosim's spends most of its running time discretising the **ROI**. This is expected. However, since Siddon's algorithm is not implemented using arrays, but instead makes heavy use of Python lists and for-loop-based iterations, it is absolutely imperative that this component is optimised

link to theory

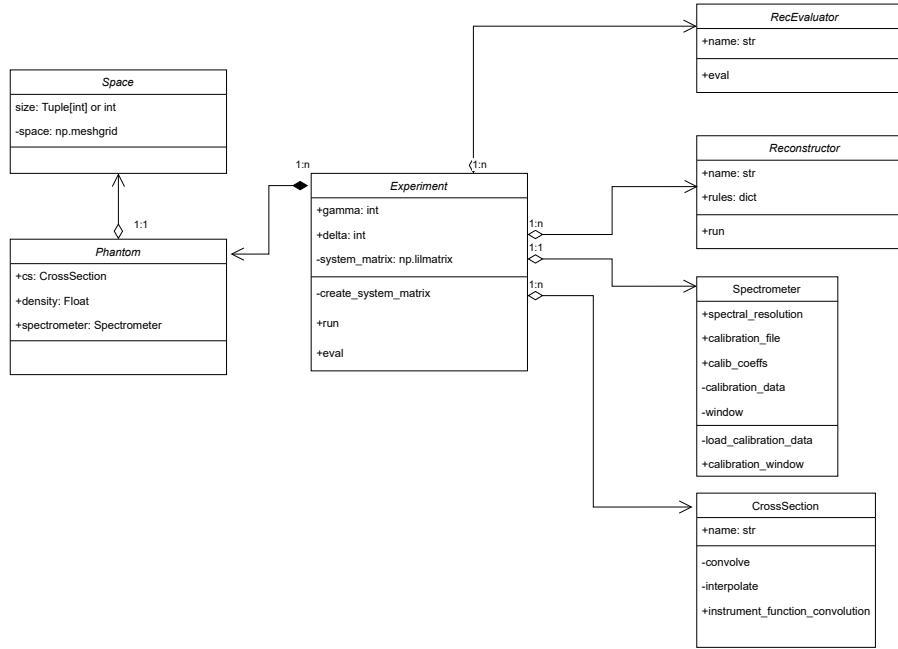


Figure 4.15: First draft of the UML diagram for the new architectural model of the Tomosim library. Note the usage of several classes from the DOASlibrary described in Section 4.1.3.

4.2 Second Hypothesis

4.2.1 The Lambertian Hypothesis

As stated in this chapter's introductory notes, this thesis main body of work revolves around two base assumptions, our hypothesis. The first one, about the information capturing by the idealised system, was addressed in Subsection 4.1.2. The second, more physical in nature, is the subject matter of this section. Our hypothesis states that the light absorption between points A and B (let's call it A_{AB}) should be equal to the difference of the absorptions in A and B . We can write this, in a *Lambertian* manner as in Equation 4.5.

$$I_B = I_A \cdot \exp \left[-AB \cdot \sum_i \sigma_{ABi} \cdot c_{ABi} \right] \quad (4.5)$$

This is to say that the light intensity reaching point B is given by the intensity reaching A , exponentially decreased by the absorbers at interval AB . The intensities at A and B are written as in Equation 4.6.

$$\begin{aligned} I_B &= I_0 \cdot \exp \left[-L_B \cdot \sum_i \sigma_{Bi} \cdot c_{Bi} \right] \\ I_A &= I_0 \cdot \exp \left[-L_A \cdot \sum_i \sigma_{Ai} \cdot c_{Ai} \right] \end{aligned} \quad (4.6)$$

If we join all this information in the same expression, the equation is transformed into its final form, presented in Equation 4.7.

$$I_0 \cdot \exp \left[-L_B \cdot \sum_i \sigma_{Bi} \cdot c_{Bi} \right] = I_0 \cdot \exp \left[-L_A \cdot \sum_i \sigma_{Ai} \cdot c_{Ai} \right] \cdot \left[-AB \cdot \sum_i \sigma_{ABi} \cdot c_{ABi} \right] \quad (4.7)$$

Equation 4.7 can be greatly simplified: we take the natural logarithm of both sides and we state that $\sum_i \sigma_{Xi} \cdot c_{Xi} = S_i$. These operations result in the simplified form of Equation 4.8.

$$L_B \cdot S_B = L_A \cdot S_A + L_{AB} \cdot S_{AB} \quad (4.8)$$

Now, $L_X \cdot S_X$ can be thought of as the wavelength dependent light absorption in path X . In this case, the wavelength interval is always the same. We can therefore conclude that, theoretically, our hypothesis is valid: light absorption between points A and B can be expressed in terms of the absorption on both these points and corresponds to their difference.

Although mathematically this seems clear-cut, in the real world things can become more problematic, since we have to deal with the imperfections that characterise a real physical system. Noise, instrumental limitations, adverse environmental effects, etc.. The experiment we describe in the next few paragraphs aimed at determining target trace gas concentration in a set analysis field. This field is dimension-wise compatible with those that would be employed in the final working system. This experiment is represented in Figure 4.16.

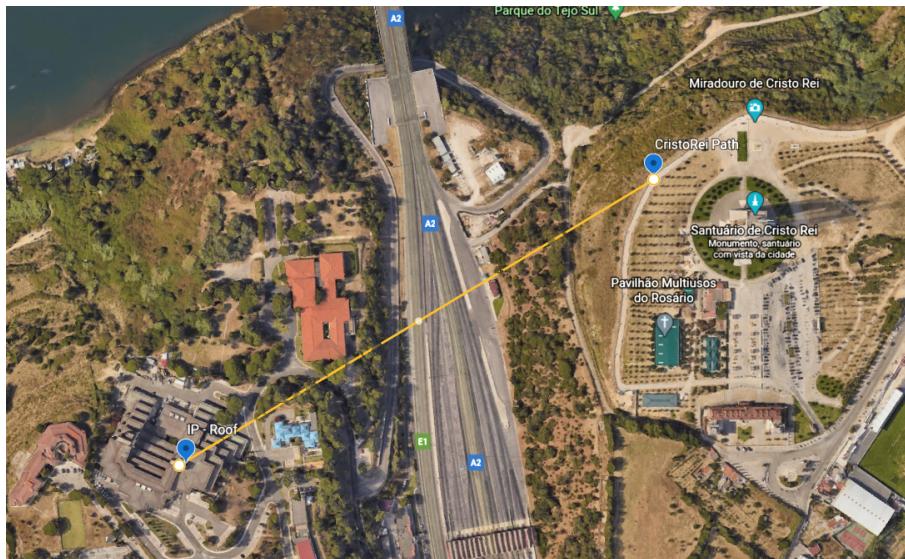


Figure 4.16: Location of observer points for the physical experiment.

The goal of the experiment was to compare passive and active DOAS measurements performed with a very short time difference between them. The passive measurement would employ the same acquisition strategy as the drone is expected to use. This comparison will be used to test our second hypothesis.

Finding two appropriate experiment sites proved to be the first difficulty: both telescopes should see sky on the back of the other telescope. Otherwise, the contribution from the terrain's reflection would have to be taken into account and the experiment conditions would be very different from the ones the drone will have. There are not

many site pairs that provide this, and most of the ones that exist are private and authorisations are not easy to obtain. In the end, we managed to run the experiment in the facilities of *InfraEstruturas de Portugal (IEP)* and the *Cristo-Rei* sanctuary, near our own base.

4.2.2 Testing the Lambertian Hypothesis

The experiment involved two different optical assemblies, which are summarised in Table 4.3. Both assemblies play two roles, which reflect the comparison between active and passive that is the entire aim of the test. To simplify, we will address the two as West Bank and East Bank. The West Bank is, as the name implies, the assembly that is placed further West, i.e., the one that is installed on IEP's roof. By exclusion, the East Bank assembly is the one placed on the sanctuary. The West Bank assembly is comprised of a telescope and tripod, a spectrometer (with the necessary fibre optics attached) and a laptop.

The East Bank assembly has exactly the same parts, but in addition to them, it features a hand-held torch that sports an XHP50.2 CREE LED. The manufacturer states that this torch is capable of illuminating by itself up to a distance of 300 m and produces luminous flux of at least 1500 lm. By fitting this torch on the telescope's eyepiece channel, we are able to further collimate the light that it produces, making it reach much further distances than originally stated, and being easily picked up by the other telescope. This is plain to see in Figure ???. The light spectrum that pertains to the CREE LED in use is published in this device's datasheet, and presented in Figure 4.17, which largely corroborates the spectrum in Figure 4.18, taken by the same spectrometers that were used in the experiment, at a distance of approximately 50 m from the torchlight.

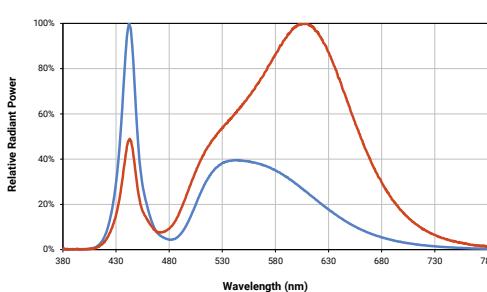


Figure 4.17: Published spectrum of the CREE XHP50.2 LED light. The LED that was used in this experiment corresponds to the blue line [CREE2021].

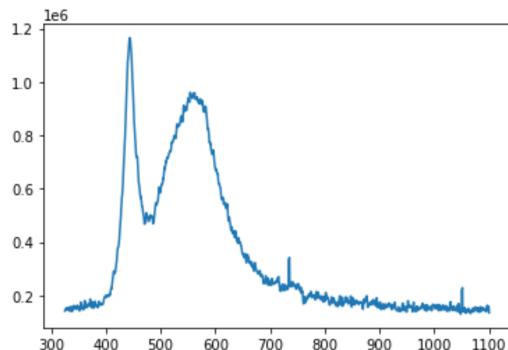


Figure 4.18: Spectrum measured with the experiment spectrometers, at a distance of approximately 50 m from the torchlight.

The experiment itself is scheduled to start at around 06:00 A.M.. It consists in capturing spectral measurements in both modes (active and passive) periodically, with the least amount of time possible between measurements in the same capture. In this case, I am calling capture to a particular group of actions that are defined in Table 4.4. Captures are defined according to the time at which they are run, and are summarised in Table 4.5. Closing time for this experiment was set on 11:00 A.M.. This time window

Table 4.3: Summary table for the two experiment assemblies. Note the difference in terms of material, due to the two different roles both assemblies play during the experiment. This is translated into not having the need of an artificial light source in the West Bank's assembly.

	West Bank	East Bank
Spectrometer	Avantes USB 2048 channels	Avantes USB 2048 channels
Telescope	Meade ET90	Meade ET90
Artificial Light Source	N/A	Goobay CREE XHP50.2 torch
Laptop	Windows 10 laptop	Windows 10 laptop
Software	AvaSoft 8.11	AvaSoft 8.11

ensures measurements are taken during sunrise and until after the morning rush hour is over.

Table 4.4: Actions are the indivisible unit upon which each capture is built. The prescribed actions for this experiment are described in this table.

Action ID	Action	Description
A	Active trace gas concentration determination	With the two telescopes facing each other, we collect spectra for two minutes with the light source turned off and another 2 min with the light source turned on.
B	Passive trace gas concentration determination	With the two telescopes aligned and approximately facing West, we collect spectra for 2 minutes.
C	Passive reference collection	The West telescope points upwards and collects data for 2 minutes.

Table 4.5: Captures are particular sets of actions that are conducted according to a specific order, depending on the time of day on which the capture is run. This table describes the prescribed captures on which this experiment consisted.

Time Frame	Period	Action
05:00 - Sunrise	15 minutes	A
Sunrise	Once	C
Sunrise - 11:00	15 minutes	A and B

As displayed in Table 4.3, the spectrometers are both the same model, manufactured by Avantes and with 2048 channels, powered through the same [USB](#) cable that is used for data transfer. The spectra are acquired through Avantes' own collection software, AvaSoft 8. The spectrometer are configured to have an integration time of 20ms and immediately store every measurement on an [American Standard Code for Information Interchange \(ASCII\)](#) file. With the kind of lighting conditions that we are dealing with, this integration time allows us not to worry about saturation. However, to build usable spectra we need to sum the collected files. This is valid because given

CHAPTER 4. METHODS

the very little time it takes to make a measurement (2 minutes), the sun can be considered a constant light source, and therefore we can consider the photons to have a Poissonian statistic distribution [**Fox2006**].

DISCUSSION

In Chapter 4, I introduced the methods with which I conducted the research and work in this thesis. In addition, I also introduced the notion that this project is divided naturally in two main parts, according to the research questions that I aim to answer.

This chapter maintains this division. In Section 5.1, I present and discuss what has come from the work conducted with reference to the technological problem of obtaining sufficient tomographic data for the retrieval of an atmospheric trace gas column density map for a given geographic area, through optical spectroscopy techniques.

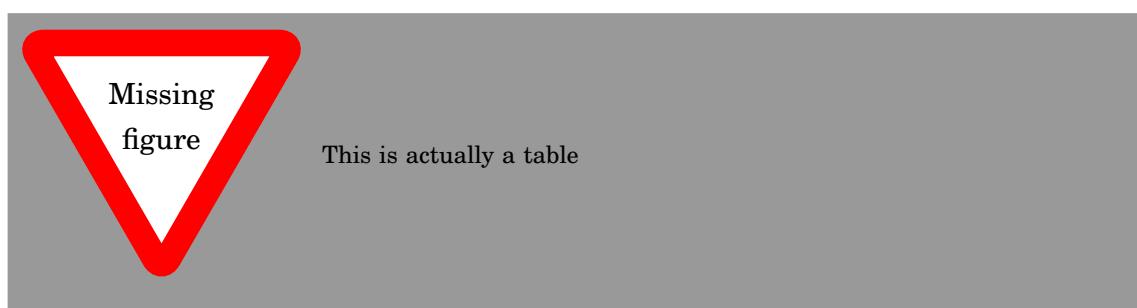
Section ?? is, naturally, dedicated to the results coming from the second hypothesis presented in Chapter 4. This is to say the *Lambertian* hypothesis that was tested through a practical experiment designed especially to this end.

5.1 First Hypothesis

5.1.1 The UAV

The component list presented in Table ?? follows the drone concept explained in Section 4.1.1. The idea was to build a hexacopter based on the very flexible and commonly used DJI S900 frame. This drone had to be capable of maintaining flight for at least 30 minutes while carrying a payload of 2kg (which is actually more than the collection system weighs). Note the inclusion of custom-made carbon fibre arms and 17" propellers.

Table 5.1: Component list for the hexacopter that was designed as a part of this thesis.



5.1.2 Collection System

The collection system is composed by the parts listed in Table ???. The collection system involved the design of several parts using specialised CAD software.

5.1.3 TomoSim

5.2 Second Hypothesis

5.2.1 Experiment - First Run

5.2.2 Experiment - Second Run

