POLITECNICO DI TORINO
ISAE-SUPAERO

Master's Degree in Aerospace Engineering

Master's Degree Thesis

Data analysis and acquisition using spectrometer EM27/SUN to validate satellite missions for greenhouse gas monitoring

Supervisors

Prof. Manuela BATTIPEDE

Eng. Christel GUY

Eng. Denis JOUGLET

Candidate

Giovanni LODIA

October 2024

Summary

The accurate monitoring of greenhouse gas (GHG) emissions, particularly carbon dioxide (CO_2), is essential in the global effort to address climate change. My internship focused on the development of tools and methodologies to enhance ground-based observations, particularly using the EM27/SUN Fourier Transform Spectrometer, within the framework of CNES's satellite validation missions. The primary objective was the creation of the "PROFFAST dashboard", a tool designed for real-time data treatment and visualization during GHG measurement campaigns. This dashboard enables immediate analysis and dynamic adjustment of field operations, improving data collection efficiency. Additionally, the thesis involved participating in a summer observation campaign, which aimed to validate the performance of multiple EM27/SUN spectrometers and to refine models of CO2 distribution. The results demonstrated consistency with expected CO2 trends and provided insights into the vertical distribution of CO2, influenced by factors such as vegetation. These contributions are crucial for advancing the precision of GHG monitoring and for supporting the validation of satellite-based observation systems, ultimately aiding in the global effort to mitigate climate change.

Acknowledgements

Durante questo percorso di doppia laurea tante cose sono cambiate, io in primis. In tutti questi cambiamenti una costante è stata il supporto e l'affetto della mia famiglia, a cui devo tutto e a cui vanno i miei più sentiti ringraziamenti. Se sono quello che sono è solo grazie a voi. Ci tengo a ringraziare le persone con le quali ho condiviso gli ultimi due anni, in particolare Pietro, Vincenzo, Adrian, Giuseppe, Davide, Pier, Feriel, Ângela. Grazie per avermi ascoltato, aiutato e consigliato quando ne avevo bisogno. Mai avrei pensato che un legame così forte potesse formarsi in così poco tempo. Un grazie speciale va a tutti gli amici con cui ho smesso di condividere la mia quotidianità trasferendomi prima a Torino e poi a Tolosa : Gianni, Alberto, Andrea, Mirko, Annasiria. Siete la prova che strade diverse non possono scalfire i veri legami. Dovunque sarete nei prossimi anni avrete sempre un posto nel mio cuore e nella mia vita. A mio malgrado, ringrazio anche il Politecnico di Torino per avermi dato l'opportunità di intraprendere questo percorso di doppia laurea e comprendere che lo studio, e per estensione il lavoro, non è la cosa più importante al mondo e che non dovrebbe soffocare le vere priorità della vita. Per finire, vorrei ringraziare tutte le altre incredibili persone che ho conosciuto e a cui ho voluto bene in questi anni universitari. C'è ognuno di voi in ogni pezzo.

Gianluca

Table of Contents

List of Tables

List of Figures

Acronyms

AERONET
AErosol RObotic NETwork

CAMS
Copernicus Atmosphere Monitoring Service

CNES
Centre Nationale d'Etudes Spatiales

COCCON
COllaborative Carbon Column Observing Network

CRA
Centre de Recherche Atmosphérique

ECMWF
European Center for Medium-range Weather Forecast

FTIR
fourier transform spectrometer

GaAs
Gallium arsenide

GHG
greenhouse gases

GOSAT
Greenhouse Gases Observing Satellite

IPDA
Integrated Path Differential Absorption

KIT
Karlsruhe Institute of Technology

LERMA
: Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères

LIDAR
: Light Detecting And Ranging

LOA
: Laboratoire d'Optique Atmosphérique

LSCE
: Laboratoire des Sciences du Climat et de l'Environnement

NCEP/NCAR
: National Centers for Environmental Prediction/National Center for Atmospheric Research

NIR
: Near InfraRed

NOAA
: National Oceanic and Atmospheric Administration

OCO2
: Orbiting Carbon Observatory 2

OPD
: Optical path difference

ppb
: parts per billion

ppm
: parts per million

S5P
: Copernicus Sentinel-5 Precursor

SFTP
: Secure file transfert protocol

SNR
: Signal to Noise Ratio

SWIR
: Short Wave InfraRed

TCCON
: Total Carbon Column Observing Network

UAVs
: unmanned aerial vehicles

Chapter 1

Introduction

Anthropogenic greenhouse gases (GHG) emissions play a crucial role in global warming and thus in the environmental crisis that humanity is facing in the last decades. This is because GHG alter the radiation balance of the Earth. This balance is depicted in fig. 1.1: the only source of energy of the Earth is the radiation coming from the Sun. Our star can be considered as a black body at 5772K, therefore it emits mostly in the visible spectrum (direct application of Planck's Law). Of the incoming Sun radiation only a fraction reaches the surface, since part of it is absorbed by the atmosphere and another big part is reflected both by the atmosphere and the surface. The Earth being a gray body, it emits all the radiation that it absorbs. Since the average temperature of the Earth is 288K, the radiation that it emits peaks in the infrared. GHG are sensitive to radiation in this spectral domain, while transparent to the radiation coming from the Sun. They absorb the radiation coming from the Earth, which is re-emitted towards the surface. The increase in GHG in the atmophere increases thus the net radiation at the top of atmosphere. This phenomen is called radiative forcing [1].

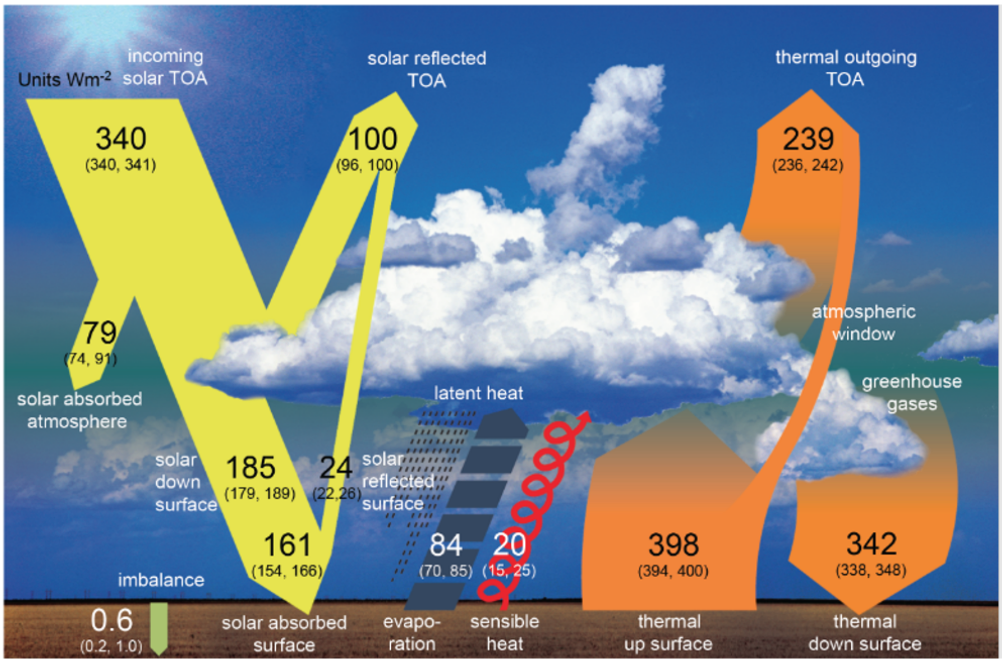


Figure 1.1: Radiative balance of Earth [2]

While water vapor is the most powerful greenhouse gas in the atmosphere, human activity does not significally impact its concentration. On the other hand, it affects the concentration of numerous others. Of these, carbone dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) are the most powerful and long lived. Figures 1.2 show the trend in these gases' concentration during the past decades. These concentrations show seasonal variations, mostly due to vegetation, as well as a constant increase during the years.

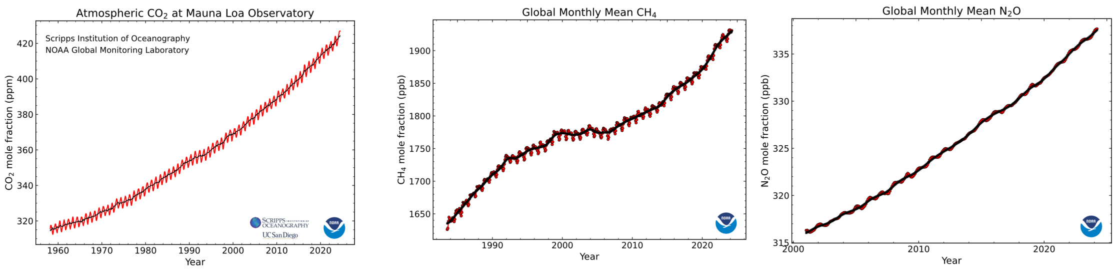


Figure 1.2: Concentrations of major GHG [3]

1.1 Monitoring greenhouse gases emissions

In order to face the impellent and globally shared problem of climate change, most countries aknowledged the problem and agreed on limiting their emissions in summits like the **Paris Agreement** (2015) and the **Glasgow Climate Pact** (2021). Accurate and precise measurement

and monitoring of GHG emission is needed to and establish trends over time, characterize sources and sinks and understand the impact of countries' actions to mitigate their emissions.
There are several ways to carry out this monitoring. The most important are aerial observations, satellite-based observations and ground-based observations [4].

1.1.1 Satellite-based observations

Satellite measurements are extremely useful because they provide global coverage of Earth's surface and atmosphere and, since a single mission can last many years in orbit, a coherent evolution of emissions overtime. The downside of this observation method is the limited spatial resolution and revisit time. Nevertheless, satellites can detect larger releases of methane and carbon dioxide.

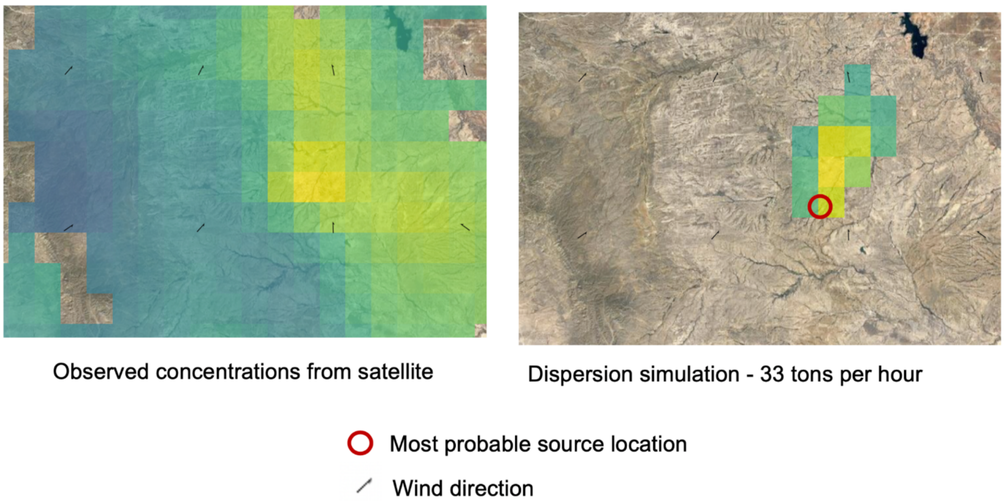


Figure 1.3: Methane plume over the Permian Basin in the US detected using data from the Copernicus Sentinel-5P satellite [5]

Numerous satellite missions and constellations have been deployed in the past years. Some of them are shortly described below:

- **Orbiting Carbon Observatory 2** (OCO-2): launched in 2014 by NASA, it carries three NIR/SWIR (Near Infrared/ Short Wave Infrared) spectrometers to analyse atmospheric carbon dioxide and oxygen distribution [6].

- **Copernicus Sentinel-5 Precursor** (S5P): launched in 2017 by ESA, perform atmospheric measurements with high spatio-temporal resolution, to be used for air quality, ozone, UV radiation, and climate monitoring and forecasting [7]. It carries the TROPOspheric Monitoring Instrument (TROPOMI)

The Centre Nationale d'Etudes Spatiales (CNES), or the french space agency, is planning to launch two satellite missions for GHG monitoring: MicroCarb and MERLIN. These missions will be explained in detail in the following chapters.

1.1.2 Aerial observations

Aerial observations can be carried using planes, unmanned aerial vehicles (UAVs) or atmospheric balloons and they allow to measure regional gas plumes or to spot small emission sources. Compared to satellites, these measurements have better resolution while mantaining a fairly broad coverage. A method of aerial emission measurement which is worth mentioning is the AirCore Atmospheric Sampling System, invented by the National Oceanic and Atmospheric Administration (NOAA). It consists in a tube that samples the atmospheric column below it that can be later analyzed to trace GHG profiles from the middle stratosphere to the ground.

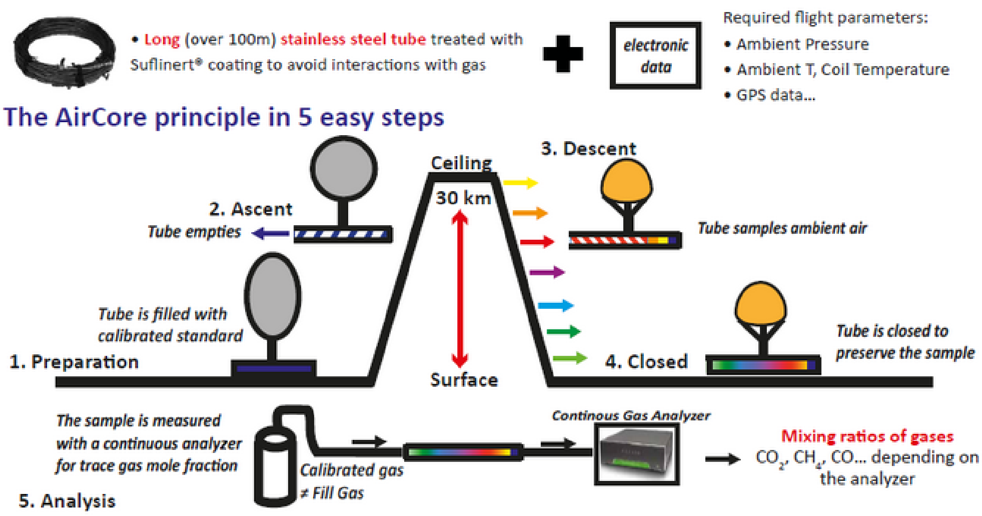


Figure 1.4: Overview of AirCore functioning [8]

AirCore technology provides outstanding precision, with a bias of 0.07 ppm for CO2 and 0.4 ppb for CH4 [9], making it a reliable reference to compare with other measurement sources. currently, a CNES base in Aire-sur-l'Adour conducts atmospheric balloon releases containing AirCore sensors on a monthly basis.

1.1.3 Ground-based observations

Lastly, surface observations can provide precise and localised measurements, able to detect smaller sourcer of GHG emissions and to do it continuously. At the surface, the interaction between the atmosphere and the ground often makes it difficult to correctly interpret data. For this reasons column measurements can be employed: instruments like the fourier transform spectrometer (FTIR) can provide column-averaged dry-air mole fractions, which are insensitive to variations in atmospheric water vapour and surface pressure. In 2004 a ground-based network of FTIRs, called Total Carbon Column Observing Network (TCCON), has been created to retrieve precise and accurate (precision of 0.25% for carbon dioxide) column abundances of CO_2, CH_4, N_2O and CO from solar absorption spectra. These measurements provide the primary validation dataset for retrievals of XCO_2 and XCH_4 from space-based instruments, and are a powerful tool to improve the understanding of the carbon cycle [10].

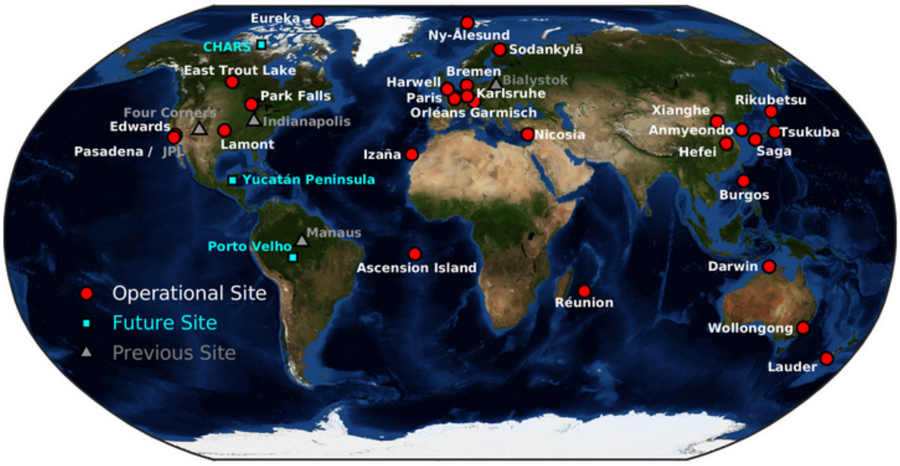


Figure 1.5: TCCON network[11]

As it can be seen from fig. 1.6, there are wide areas in the world that are not covered by the TCCON network, given that logistical and cost issues limit its expansion. In 2011 the Karlsruhe Institute of Technology (KIT) and Bruker Optics developed a new type of of portable FTIR spectrometer for the measurement of the main greenhouse gases called EM27/SUN. In the following years many of these devices started to be operated around the world and they proved to have a very high level of performance and stability overtime [12]. For this reason KIT started a new infrastructure for greenhouse gas measurements using the EM27/SUN spectrometer. This network, called COllaborative Carbon Column Observing Network, currently operates supporting TCCON and providing inter-calibrations and measurements in regions that are not suitable for TCCON spectrometers) [13]. Currently, more than 200 EM27/SUN are part of the COCCON network.

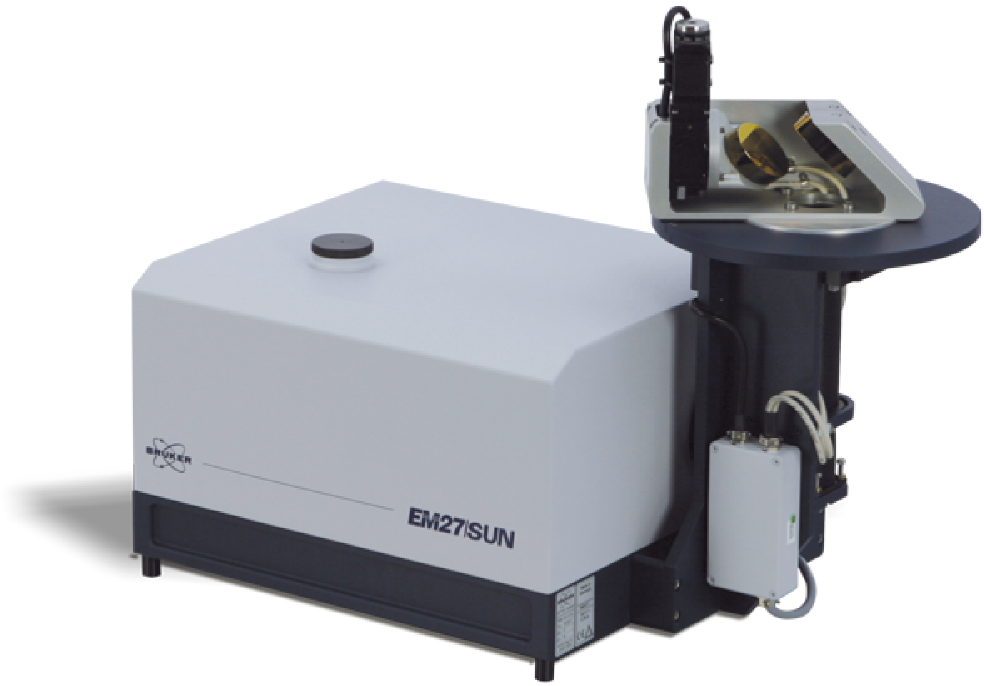


Figure 1.6: EM27/SUN spectrometer[14]

Chapter 2

Context and objectives of the internship

The CNES is the French government space agency. Established in 1961, its role is to shape and implement France's space policy. It conducts research and development in all space-related fields, including satellite technology, launchers, space exploration and Earth observation. The agency collaborates with international partners and organisations such as NASA and ESA. In the context of Earth observation, CNES is developing two major space missions to monitor GHG emissions: MicroCarb and MERLIN.

2.1 MicroCarb

MicroCarb is a microsatellite developed by CNES with the partecipation of UK Space Agency. It will be launched by Vega C in April 2025 from French Guyane and it will observe the atmosphere for five years from a helioshyncrone circular orbit at 650 Km of altitude [15]. The purpose of MicroCarb is to measure CO_2 fluxes on a global scale assuring high performances, with a precision of less than 1 ppm. The main instrument is a passive grating spectrometer in 4 NIR/SWIR bands $(0.76\mu m, 1.27\mu m, 1.61\mu m and 2.05\mu m$. The spectrometer has high resulution ($R = 25000$), high SNR and a pixel footprint of $4.5x9Km^2$ at the Equator. A visible imager completes the payload.

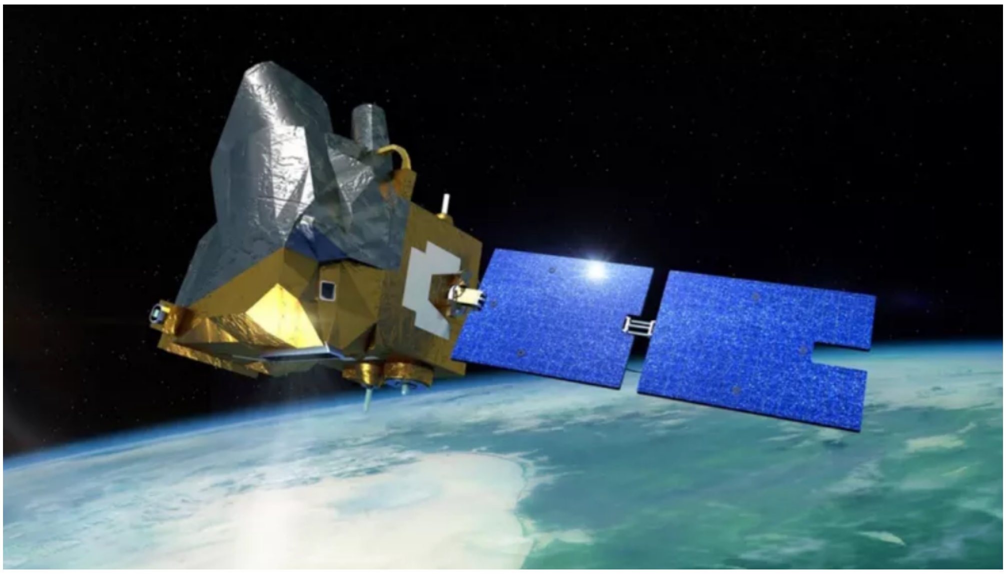


Figure 2.1: Artist's view of MicroCarb [15]

The satellite has a revisit time of 25 days but thanks of its different observation modes every target can be observed once a week[16].

2.2 MERLIN

MERLIN (Methane Remote Sensing Lidar Mission) is an Earth Observation satellite developed by CNES and DLR (german space agency) that will be launched in 2028. The satellite will orbit around the Earth with a polar heliosynchrone orbit at 500 Km and will last for at least three years[17]. The main goal of MERLIN is to study the concentration of CH_4 in the atmosphere with unprecedented precision to bring a significant improvement on the knowledge of methane's emissions and sinks. In order to do so, the satellite embarks the first Integrated Path Differential Absorption LIDAR instrument, developed by DLR. This instrument will provide a precision of 1% on the measurements as well as a targeted accuracy of 0.2% [18]

Figure 2.2: Artist's view of MERLIN[17]

2.3 Levels of data processing

Satellite missions provide enormous amount of data, which requires processing to be exploitable. Data processing involves several levels, each adding a layer of refinement to the original data collected by the instruments [19].

- **Level 0** (L0): unprocessed raw instrument data, at full resolution and with all communication artifacts removed.
- **Level 1** (L1): data processed to sensor unit. Radiometric and geometric correction is applied to raw data.
- **Level 2** (L2): derived geophysical variables. Data is transformet into geophysical quantities using scientific algorithms
- **Level 3** (L3): georeferenced data. Geophysical variables are mapped on uniform space-time grids.
- **Level 4** (L4): this is the most refined data, which is obtained by the assimilation of lower-level data into models.

In table 2.1 the different levels of MicroCarb data are exposed.[20]

To be operational, every satellite needs to undergo the phase of cal/Val (Calibration and validation. The cal/val phase consists in processes that ensure the validity of remotely sensing data. Given the stringent requirements of MicroCarb and MERLIN, the cal/val process, both on ground and in flight, is the key to achieve the needed performances.

Processing level	**Data**
L0	Telemetry of the instrument, AOCS data and orbitography data
L1	Calibrated (spectrally and radiometrically) and geolocalised spectra
L2	Geolocalised concentrations of column-averaged dry-air mole fractions of CO_2 (XCO2)
L3	Global maps of XCO2
L4	Global maps of surface fluxes of CO2

Table 2.1: Data products of MicroCarb

2.4 Calibration

The calibration process is fundamental to obtain a functional instrument and exploitable data. This process is located between L0 and L1. Calibration refers to the process of translating the measurements of the detectors in standard values that can be understood and compared.

2.4.1 Spectral calibration

A detector is composed by a certain number of channels (or pixels) and when observing a spectrum each channel is sensible to a different part of it. Spectral calibration consists to replace channel numbers by wavelenghts (or wavenumbers) along the X axis. This process is usually made by exposing the detector to source with well known emission spectral lines, for example a neon lamp [21].

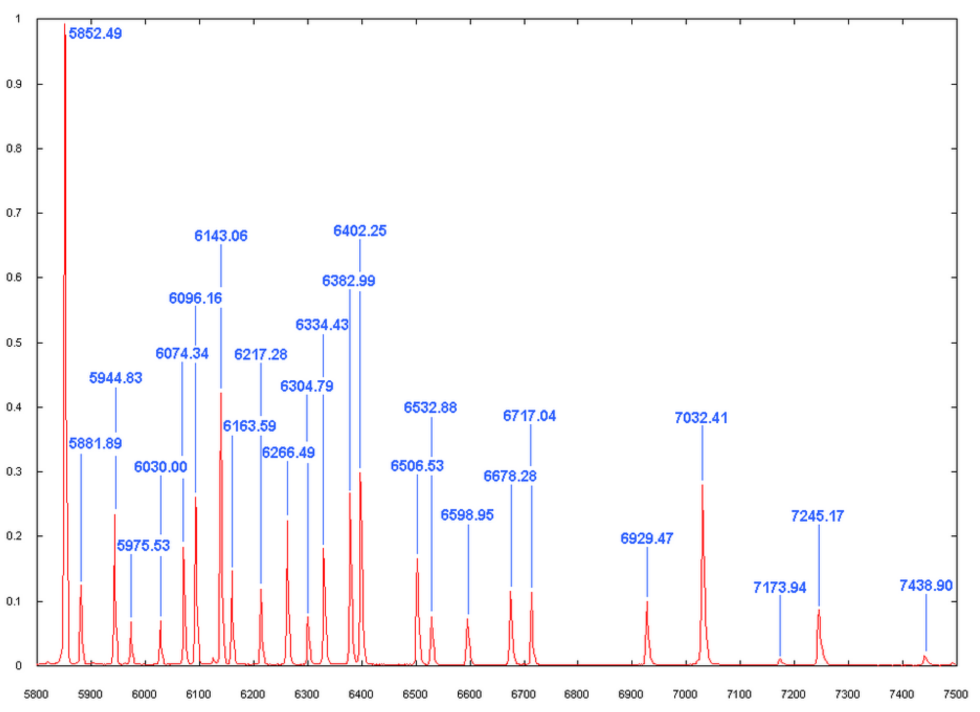


Figure 2.3: Spectrum of a neon lamp. Wavelength[Å] on the X axis[21]

2.4.2 Radiometric calibration

Once the spectral calibration is done, it is necessary to understand what exactly represents the signal received on each channel. If we call X_i the signal coming out from the channel i, and L_i the radiance detected by it, their relation can be expressed as:

$$X_i = f_i(L_i), \forall i = 1, ..., n$$

Where f_i is the function that relies the signal and the radiance and n is the number of channels [22]. A priori this function is different for every channel because of instrumental imperfections. A first step in the calibration process is to "equalise" the response of every channel, which translates into having:

$$X_i = f(L_i), \forall i = 1, ..., n$$

This process heavily depends on the type of instrument. For imagers, for example, if the sensor is a whiskbroom and every pixel is captured by the same detector there will be no need for an equalisation. On the other side, if the sensor is a "push-broom" then it will be necessary to correct the differences between a detector and another, but the imperfections in the resulting image will only appear along one direction.

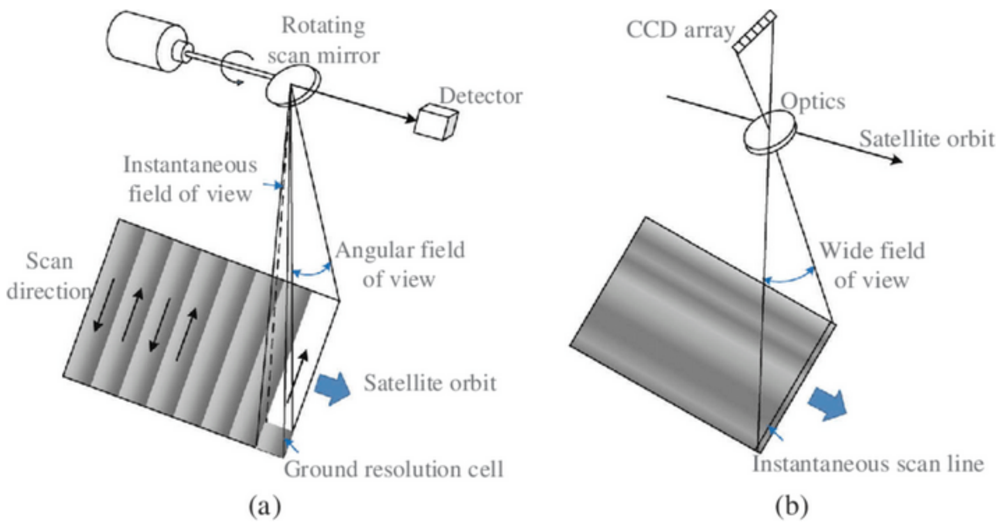


Figure 2.4: (a) Whiskbroom-type imaging instrument and (b) pushbroom-type imaging instrument.[23]

One way of performing the equalisation process is to observe uniform scenes, like the ocean at night or deserts of snow.

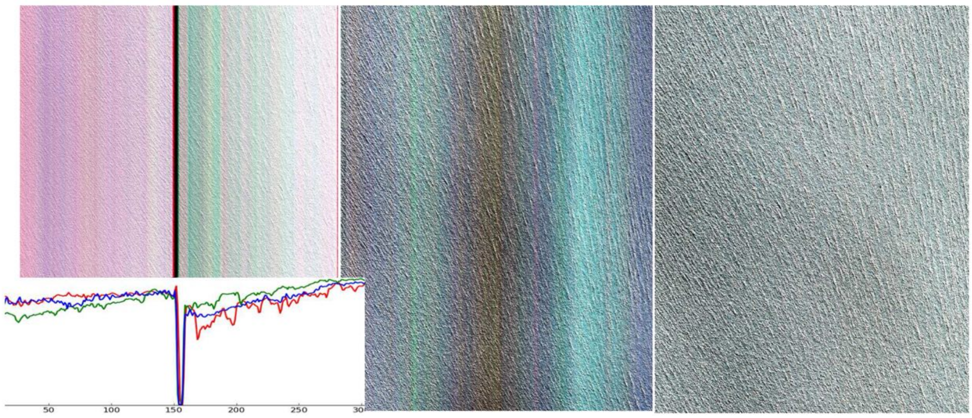


Figure 2.5: Image of Antarctic without equalisation (left), with partial equalisation (center), with complete equalisation (right)[22]

Once the equalisation between the detectors has been made, the next step is to find the relation between the quantity measured and the physical quantity associated. Since to find the function f is a very complicated solution, what is usally done is assuming a linear proportion between X and L, so that the calibration can be obtained by multiplying every channel by a coefficient c_i:

$$L_i = c_i \cdot X_i \forall i = 1, ..., n$$

This process is called absolute radiometric calibration. One of the most common methods to perform it, especially on thermal sensons, is by using a black body. According to Planck's law, the radiance emitted by a black body only depends from the temperature of the body itself:

$$L(\lambda, T) = \frac{2hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}$$

Where h is Planck's constant, c is the speed of light, k is the Boltzmann's constant. Therefore, by precisely knowing the temperature of the blackbody an absolute radiometric calibration can be performed.

Calibration processes are conducted both before and after launch, to assure that the instruments is capable to deliver the expected performances.

For MicroCarb, the equalisation is made by characterising gains inter-pixel observing a flat field. The absolute radiometric calibration is made by observing the Sun. [20]

2.5 Validation

Validation consists in taking processed data (usually L2) coming from the satellite and comparing it with "external truths" to assess its validity and characterize the instrument error. The MicroCarb team will conduct massive operational comparisons between the satellite data and a number of other sources [16]:

- Other satellites, like OCO2 and GOSAT
- CO2 models coming from the Copernicus Atmosphere Monitoring Service (CAMS)
- Ground networks: TCCON, COCCON, AERONET, lidars

2.6 The SA division

I carried mi internship in the "Sondage Atmpspherique" (SA) division of CNES, under the supervision of Christel Guy, performance responsible of MERLIN, and Denis Jouglet, performance responsible of MicroCarb. Within the sub-directorate "Techniques, Performance, Instruments", the expertise of this service covers measurement physics and data processing of optical instruments intended for the observation of the atmospheric column, with scientific or environmental purposes: spectro-imagers, spectrometers, lidars, and other sounders. It also involves mission engineering for projects developed in this field.

The service is responsible for system-level activities focused on product quality: mission technical analyses and translation into system specifications, product simulations, feasibility studies, system modeling, trade-off analyses between on-board and ground systems, establishment and monitoring of system-level performance assessments, specification and validation of processing algorithms, preparation and execution of in-flight system validation, characterization of product performance, and monitoring of systems throughout their operational life.

This service particularly addresses the radiometric and spectral aspects of measurement chains: considering scientific needs, expertise in product quality for scientific purposes, measurement physics, signal processing, calibration of instrumental chains, etc. It is also responsible for absolute calibration and cross-calibration of atmospheric sounders.

Moreover, the service ensures access to the necessary computing tools and resources for all these functions, such as radiative transfer models, inversion codes, calibration databases, etc.

Regarding the "mission engineering" function, it is exercised on behalf of projects within the scope of UV-visible and infrared optics, covering the following activities and responsibilities:

- Monitoring of scientific support actions carried out in laboratories or by users for projects, including "short-loop" iterations on mission specifications (to define the "exact need" for the project), development and monitoring of the downstream part of the CAL/VAL plan, product quality at levels 2-3-4, extension of data use, feedback, and promotion of the project.

- Supporting development projects: participation in interfaces with users and translation of mission needs into system specifications, mastering scientific algorithms (levels 2-3) as needed, evaluating system and mission performance of the project before and after launch (simulations, performance assessments), proposing and monitoring the upstream part of the CAL/VAL plan.

Among other space missions, the engineers in the division are preparing the validation of MicroCarb and MERLIN. In order to do so, a portable fourier transform spectrometer EM27/SUN has been acquired and used for several years already. This spectrometer, EM27/SUN #92, is part of the COCCON network and it is joined by two other spectrometers own by two laboratories in France: the LSCE and the LERMA.

2.7 Objectives of the internship

The objectives of my internships were the following:

- The first main objective was to interface with the EM27/SUN and its processing chain in order to understand and automate it. In particular, the processing chain at the time needed some input files that are available only some days after the measurements. Therefore during the acquisition no information about the results of the measurements were provided. The goal of my work was to be able to obtain in real time the results of the inversions of the spectra acquired and to show them on an interactive dashboard. In order to do so, the whole processing chain needed to be adapted and optimized.

- Correlated to the previous objective was the creation of a GitLab repository where all the necessary to install and run the inversion software, PROFFAST, was uploaded and updated as the project was developed.

- The second main objective was to participate in the measurement campaigns of the EM27/SUN, especially a summer campaign conducted using two other EM27/SUN deployed in different sites. After the campaigns and under the guidance of my tutors, I collected and treated the resulting data and obtained information that would validate the performance of the instruments.

- The third objective was secondary and consisted in helping during measurements conducted in the optical laboratory of CNES and analyze the data to perform spectral and radiometric calibration of the EM27/SUN.

I worked on the first objective during the first three months of the internship, developing a python project that I called "*PROFFAST Dashboard*". The development included various test phases, where the code was tested while having the EM27 carrying out the measurements. The project was uploaded under a GitLab repository and a push was made for every major modification, so that its evolution could be traced. Other files were uploaded in the repository, such as the installer folder containing all the necessary files to install and run PROFFAST and the scripts used to retrieve prior profiles from web databases.

Starting from the month of July I worked on the second main objective, participating in all the campaigns, from setup to maintenance. I then proceeded to perform intercomparisons between the instruments and then analyze the data.

Regarding the calibration of the instrument, I only participated to the measurements at the optic laboratory. The data obtained has not been treated because of lacking of time and it being a secondary objective of the internship. Therefore, the calibration of EM27/SUN will not be discussed in this report.

Figure 2.6: EM27/SUN being deployed at the Pic du Midi (2887 m)

Chapter 3

The EM27/SUN: functioning and treatment of data

This chapter will focus entirely on EM27/SUN's principles of functioning and on how data coming from it is treated.

3.1 Functioning of the instrument

The EM27/SUN is a fourier transform infrared spectrometer. The goal of an absorbtion spectrometer is to measure how much light is absorbed by the detector at each wavelenght. To achieve this, a FTIR incorporates a modified Michelson interferometer that allows to change the Optical path difference (OPD), which is the difference in optical path length between the two arms to the interferometer. There are different ways of modifying the OPD. Rotary movements have proved very successful in doing so and this mechanism is adopted by the EM27/SUN. By varying the OPD an interferogram is obtained (like in figure 3.1). Interferograms are the raw data provided by the spectrometer. To convert this data into an absorbtion spectrum a fourier transform needs to be performed, thus the name of the instrument [24].

Figure 3.2 shows the intern of the EM27/SUN: a set of moving mirrors intercepts the sunlight which enters the instrument. From there the beam passes through a 750 nm long pass filter [25] to block unwanted radiation. The beam is then splitted and reconstructed by the Michelson interferometer. Finally it passes through an aperture of 3mm, which avoids non linear detector response and controls optical aberrations, and it's focused on a GaAs detector. An internal camera is used by a system called CAMTracker to follow the Sun by rotating the mirrors. The EM27/SUN has a spectral range of $4000cm^{-1} - 11100cm^{-1}$ with a resolution of $0.5cm^{-1}$. Its results have been validated by conducting simultaneous measurements with a TCCON station. The comparison shows a difference of $0.12 \pm 0.08\%$ [25] for the averaged dry-air mole fraction of CO_2, XCO_2. It's worth noting that this fraction is calculated using the ratio between the measured CO_2 and O_2 total columns:

$$XCO_2 = \frac{CO_2}{O_2} \cdot 0.2095$$

Where 0.2095 is the dry air mole fraction of O_2. The instrument is equipped with two detectors that cover different spectral bands:

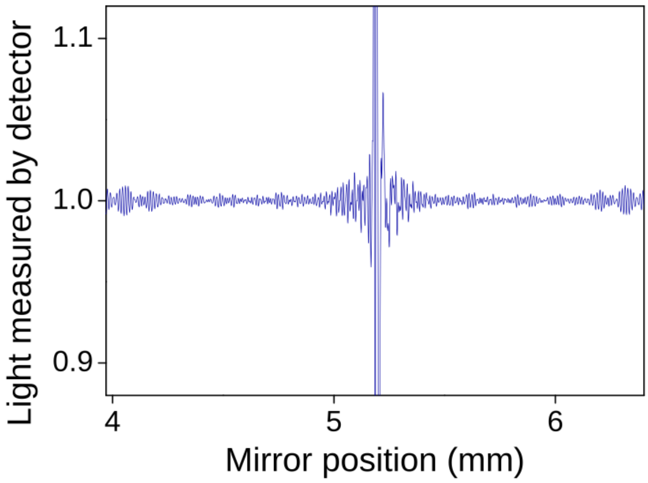


Figure 3.1: Typical interferogram [24]

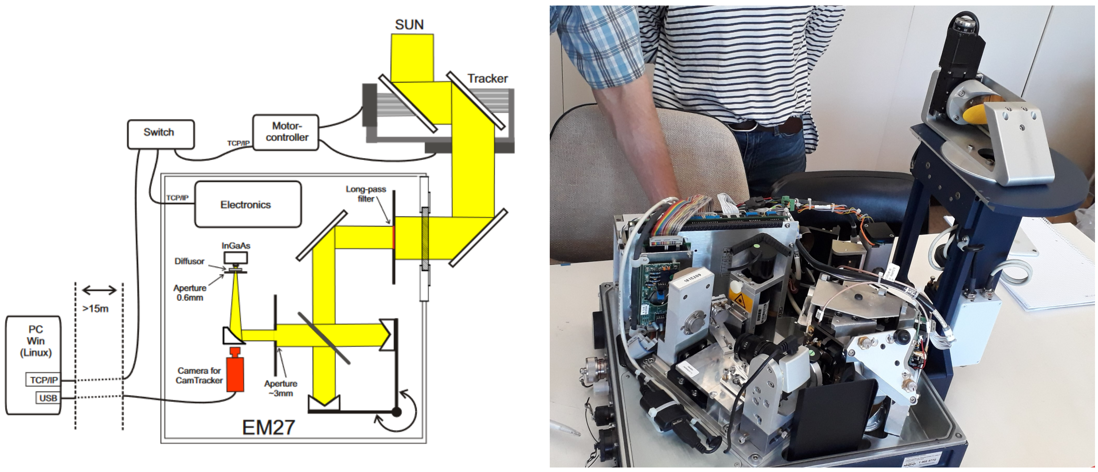


Figure 3.2: Schematic of the EM27/SUN[25] (left) and photo of the inside of the instrument (right)

- Detector 1 covers the region of $5500cm^{-1} - 11100cm^{-1}$ and can detect CH_4 ($5897cm^{-1} - 6145cm^{-1}$), CO_2($6173cm^{-1} - 6390cm^{-1}$), O_2 ($7765cm^{-1} - 8005cm^{-1}$) and H_2O($8353cm^{-1} - 8463cm^{-1}$)

- Detector 2 is sensible from $4000cm^{-1}$ to $5500cm^{-1}$ and can detect CO and CH_4 in the window $7765cm^{-1} - 8005cm^{-1}$ [26]

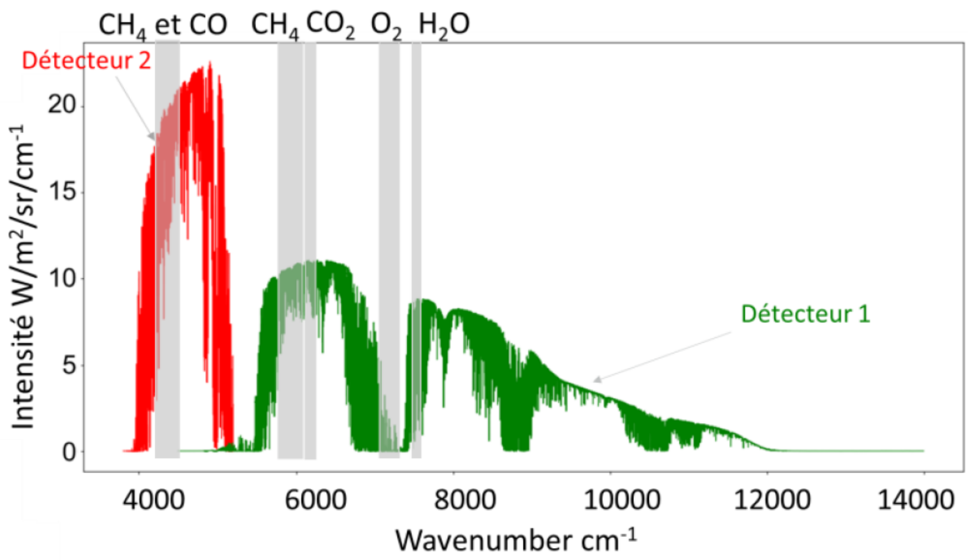


Figure 3.3: Spectrum acquired by the EM27/SUN, with highlighted absorbtion windows[26]

3.2 Data processing

As mentioned earlier, the data coming from the spectrometer is a series of interferograms. The final product, which can be considered as L2 data, are the column averaged dry-air mole fractions of the greenhouse gases above mentioned. Final data is obtained by using two softwares provided by Bruker. The first one, called OPUS, allows to pilot EM27 measurements and automatically converts the interferograms into spectra and saves them on the local disk (with a binary format .BIN). OPUS can also be used to change the configuration of the instrument. This is particularly useful to adapt the acquisitions of the instrument for particular cases (e.g. calibration). Fig. 3.3 shows a typical spectrum obtained by OPUS after an acquisition. It is important to notice that the intensity is expressed in arbitrary units and not in physical units. This is because the spectrum is just an intermediate product and it's not exploited. Nevertheless, radiometric calibration and correction of OPUS spectra has been made in the past at CNES and another calibration is currently ongoing.
To convert the spectra into the concentrations of different GHG, KIT developed an inversion software called PROFFAST. This software is written in FORTRAN but a python wrapped, called PROFFASTpylot, has been developed to make its utilisation easier. An explanation of PROFFAST and its python wrapper is made in the following pages.

3.3 Acquisition of data using EM27/SUN

Being a portable spectrometer, the EM27/SUN is relatively easy to deploy and CNES has used it to perform measurements on various sites of interest, like the "Observatoire du Pic du Midi", the "Centre de recherches atmosphériques" (CRA) and the CNES office in Toulouse. The spectrometer is paired with the following instruments:

- A GPS antenna, that provides the position of the instrument (latitude and longitude) to produce georeferenced data.

- A PTU sensor which measures pressure, temperature and humidity at the surface. This data is essential for the software PROFFAST to perform the inversion of the spectra.
- A raspberry pi module is connected to the antenna and the PTU, allowing the data to be transferred to the PC.
- A PC to pilot the spectrometer and the measurements.
- In the last months, a CO_2 sensor has been deployed with the EM27/SUN in order to quantify the difference between CO_2 concentration at the surface and in the whole atmospheric column.

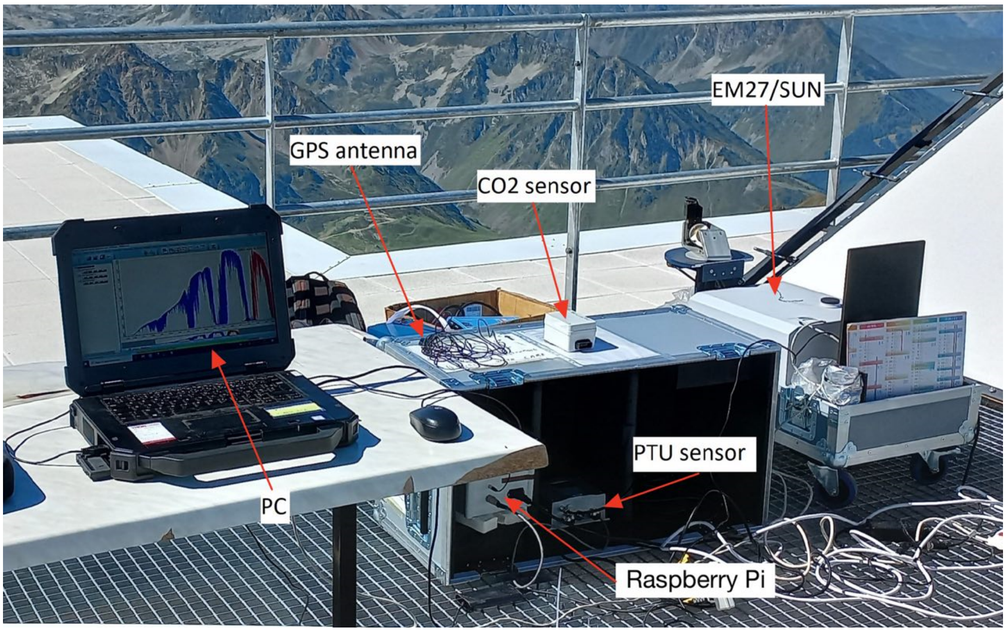


Figure 3.4: Typical setup of the EM27/SUN

Once switched on, the instrument is connected to the PC using an ethernet cable. The PTU sensor and the GPS antenna are connected to the raspberry pi. The raspberry can connect to the PC via WiFi. By using VNC Viewer the acquisition of GPS data can be started. To retrieve data coming from the PTU, the software MobaXterm is used to establish a SFTP. Data are stored in the *.met* format.

The next step is to properly align the mirrors toward the Sun. In order to do so, the instrument is first oriented to the south. Then the software *CamTracker* is opened. CamTracker is a software provided by Bruker that, by using an internal camera (3.2 and GPS data, can continuously orientate the mirrors to correctly face the Sun. Without this correction the instrument would loose track of the Sun within a minute because of Earth rotation.
By using CamTracker the mirrors can be initialized. At this point the operator needs to ensure that the light collected by the mirrors is uniform and correctly enters the instrument, as showed in the image above.

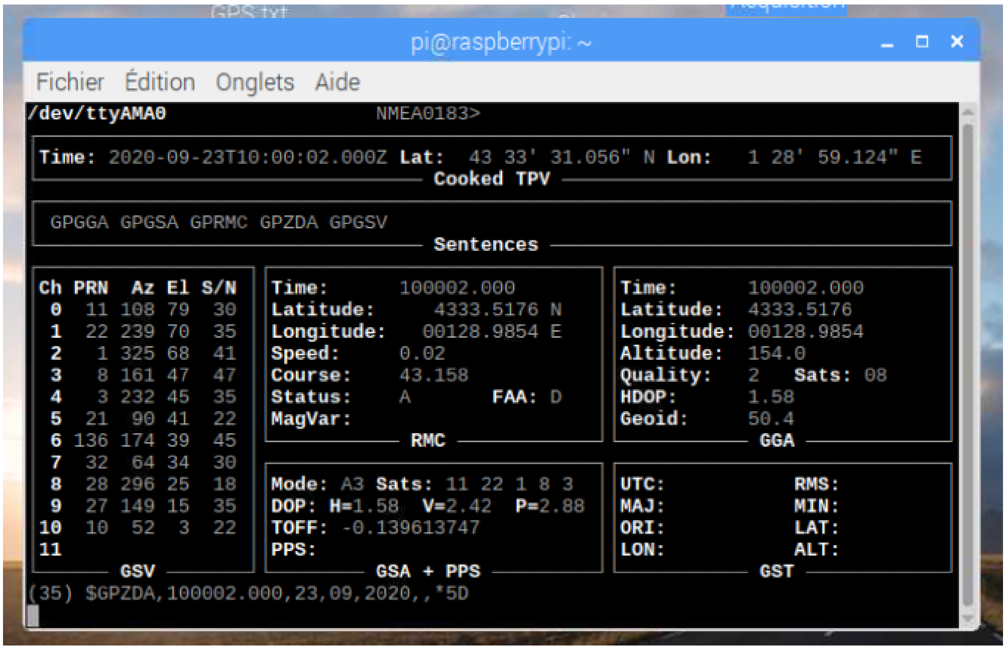


Figure 3.5: GPS data acquisition using VNC Viewer

Figure 3.6: Correct positioning of EM27/SUN's mirrors

At this point the Sun should be in the field of view of the camera. Once manually centered, the software is capable to constantly mantain the centering. If the signal is temporarily lost (e.g.

cloudy sky) CamTracker relies on GPS data and ephemerids to orient the mirrors until the signal is found again.

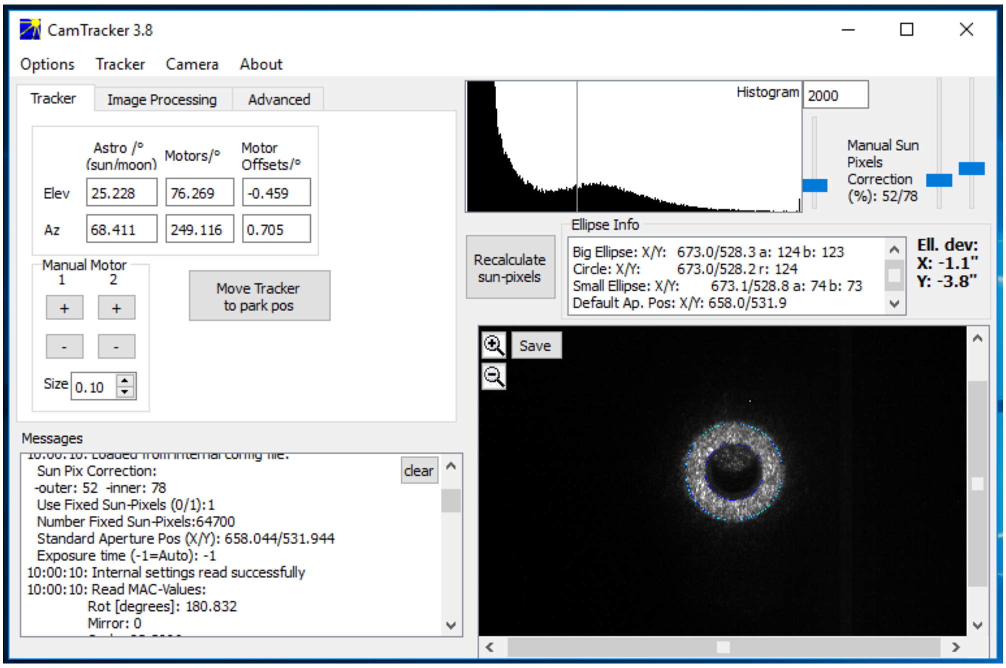


Figure 3.7: view of CamTracker with Sun correctly centered

Once all these steps are done, acquisition of the spectra can be started using the software *OPUS*. OPUS contains various parameters that can be modified to change the behaviour of the instrument. Among these there are the gains of the two detectors, the integration time (time to obtain a spectrum), the scanner velocity. The sensibility of the EM27 to these parameters have been studied in a previous intership in the SA division [26]. For all the measurement campaigns conducted during my internship, a standard configuration has been used. OPUS supports continuous measurements, which is what it's usually done during campaigns. The spectra are saved in binary format in a folder defined in the configuration.

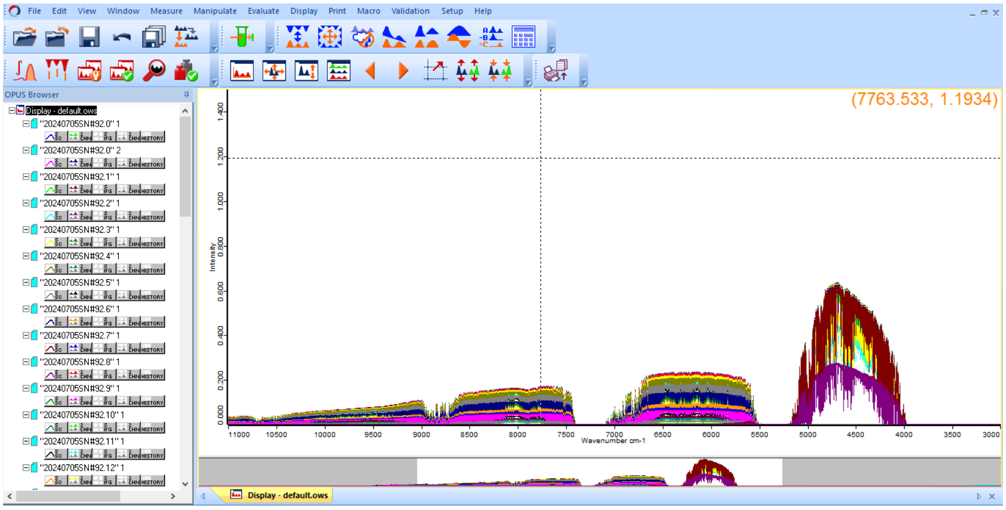


Figure 3.8: view of OPUS during multiple spectra acquisition

Chapter 4

The inversion software PROFFAST

PROFFAST regroups the softwares developed by KIT to treat the spectra coming from the EM27 and obtain the concentration of GHG in the atmospheric column. This is a classic inverse problem and the solution has to be found by using appropriate methods.

4.1 Inverse problems

Inverse problems are a class of problems which goal is to determine the model or the system parameters (input) that produce a given set of observations or data (output). This is in contrast to classic forward problems, where the system is known and it is described by a model so that by using the model input the outcome of the model can be predicted.

4.1.1 Theory

An inverse problem can be mathematically expressed as follows [27]:

$$\mathbf{d} = \mathcal{F}(\mathbf{m}) \tag{4.1}$$

Here:

- $\mathbf{m}$ is the model parameters.
- $\mathbf{d}$ represents the observed data.
- $\mathcal{F}$ is the forward operator that maps the model parameters to the observed data.

Inverse problems are often ill-posed, which means they may violate one or more of Hadamard's conditions for well-posed problems:

- **Existence**: A solution exists.
- **Uniqueness**: The solution is unique.
- **Stability**: The solution depends continuously on the data, meaning small errors in data result in small errors in the solution.

If any of these conditions are not met, the problem is considered ill-posed. For example, small noise in the data $\mathbf{d}$ can lead to large deviations in the estimated parameters $\mathbf{m}$.

To address the ill-posedness, regularization techniques are applied. The idea is to incorporate additional information or constraints to stabilize the solution.

A common approach is Tikhonov regularization, where the inverse problem is reformulated as an optimization problem:

$$\min_{\mathbf{m}} \left\{ \|\mathcal{F}(\mathbf{m}) - \mathbf{d}\|^2 + \alpha \|\mathbf{R}(\mathbf{m})\|^2 \right\} \tag{4.2}$$

Here:

- $\|\mathcal{F}(\mathbf{m}) - \mathbf{d}\|^2$ is the data fidelity term, measuring the difference between observed and predicted data.
- $\|\mathbf{R}(\mathbf{m})\|^2$ is the regularization term, incorporating prior knowledge about the model (e.g., smoothness).
- α is a regularization parameter that balances the trade-off between fitting the data and imposing the regularization.

In many cases, the forward model $\mathcal{F}$ can be approximated as linear, especially for small perturbations around a known state. The problem then reduces to solving a system of linear equations:

$$\mathbf{d} = \mathbf{A}\mathbf{m} \tag{4.3}$$

Where:

- $\mathbf{A}$ is the forward operator (a matrix in the linear case).

The inverse problem then seeks to find $\mathbf{m}$ from the observed $\mathbf{d}$. In the case where $\mathbf{A}$ is not invertible or is ill-conditioned, regularization becomes crucial.

If the forward model is nonlinear, the problem is more challenging. In such cases, iterative methods like the Gauss-Newton method or Levenberg-Marquardt algorithm are often used to find an approximate solution.

4.1.2 Application to the case study

For PROFFAST the goal is to estimate the concentrations of greenhouse gases from the measurements of the atmospheric spectrum.

The forward model is the radiative transfer equation [28], which describes how light propagates through the atmosphere and how it is absorbed by different gases. The intensity of light at a given wavelength λ can be modeled as:

$$I(\lambda) = I_0(\lambda) \exp\left(-\sum_i m_i \sigma_i(\lambda)\right) \tag{4.4}$$

Where:

- $I(\lambda)$ is the observed intensity at wavelength λ.
- $I_0(\lambda)$ is the initial intensity before absorption.
- m_i is the concentration of the i-th gas.

- $\sigma_i(\lambda)$ is the absorption cross-section of the i-th gas at wavelength λ.

Given the observed spectrum $I(\lambda)$, the inverse problem is to estimate the concentrations m_i for each gas. This problem is inherently ill-posed because the absorption spectra of different gases overlap, making it difficult to distinguish between them. Furthermore, measurement noise can lead to large errors in the estimated concentrations.

To regularize the problem, a common approach is to minimize a cost function that combines the fit to the data with a regularization term that enforces smoothness or other physical constraints on the concentration profiles:

$$\min_{\mathbf{m}} \left\{ \sum_{\lambda} \left(I(\lambda) - I_0(\lambda) \exp\left(-\sum_i m_i \sigma_i(\lambda) \right) \right)^2 + \alpha \|\mathbf{Lm}\|^2 \right\} \tag{4.5}$$

Where $\mathbf{L}$ is a regularization operator (e.g., a derivative operator for smoothness).

This inverse problem is solved by using iterative methods, such as Levenberg-Marquardt and Conjugate Gradient.

4.2 Functioning of PROFFAST

PROFFAST is composed by three different executable files written in fortran [29]:

- **preprocess5.F90** is the first executable, which takes as input the interferograms of the EM27/SUN. The script checks the size of the interferograms and other parameters to select only the ones that are suitable for the inversion. It also performs other operations like DC correction, apodization and resampling. The obtained spectra are saved in a binary format.

- The second step of the treatment is done by **pcxs20.f90**. This script generates a daily lookup table of x-sections from line-by-line calculations and column sensitivities for each gas as function of the solar zenith angle.

- Finally, **invers20.f90** performs the inversion of each spectrum and outputs the column-averaged dry-air mole fractions of the gases.

In order to make the utilisation of PROFFAST easier, a python library called PROFFASTpylot was created. This library provides several classes that allow, among other things, to run PROFFAST on a certain set of spectra or to run each part separately. Simon Prady, from Magellium, built on this library by implementing a script that allows to better display PROFFAST's outputs, for example by plotting GHG concentrations over time or GHG vertical profiles.

Fig. 4.1 represents well the workflow of the software. While the interferograms come from the EM27/SUN, the other inputs have different sources:

- Coordinates come from the GPS antenna.

- Pressure data come from the PTU sensor. They are stored in a *.met* file together with temperature, humidity and coordinates.

- map files contain the a priori column profiles of gases concentrations, obtained by a numerical model.

Up until the beginning of my internship map files were obtained from the TCCON servers. The a priori profiles are generated by the TCCON retrieval algorithm, GGG2020, based on NCEP/NCAR analyses coupled with empirical models that were developed from balloon flights, satellites and in situ data [30]. While these map files have the advantage of being used by the TCCON network and therefore assure the coherence of the results, they have the disadvantage of being available with a delay of some day from the date of the observation. One of the objectives of my internship was to modify the processing chain of CNES to allow real-time processing of spectra acquired from the EM27/SUN. A major issue was represented by the impossibility of obtaining TCCON map files before the observation campaigns. In the next chapter, the project I developed during the first months of my internship, *PROFFAST Dashboard*, is presented.

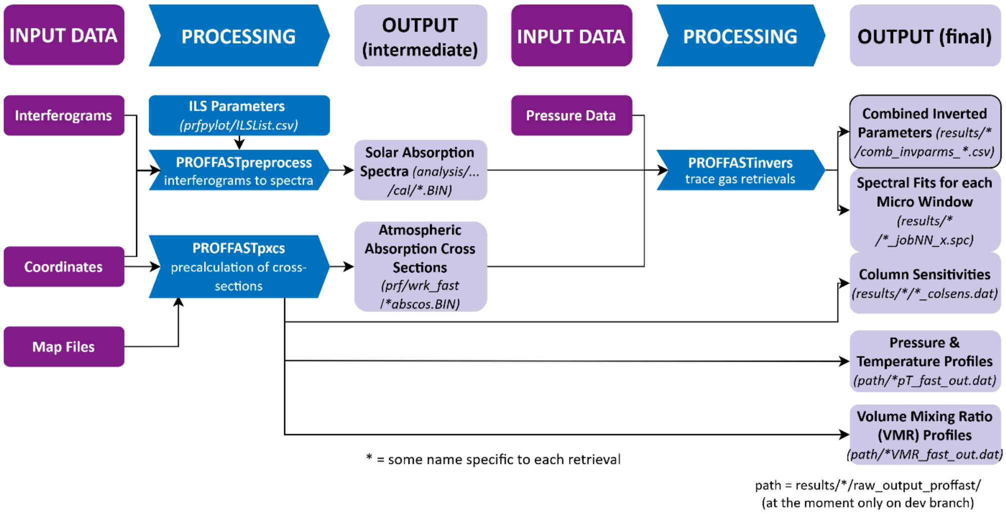


Figure 4.1: Schematic of PROFFAST functioning [31]

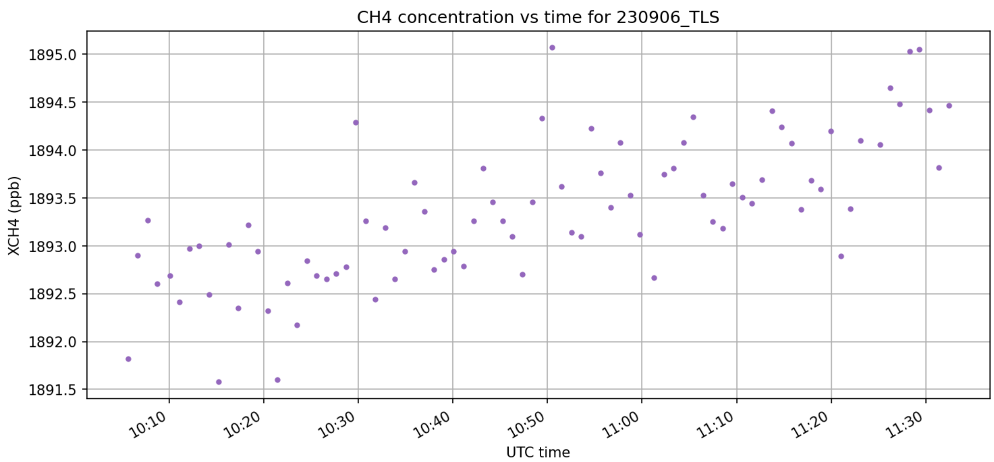


Figure 4.2: Exampe of a plot of CH_4 concentration over time

Chapter 5

PROFFAST Dashboard

Real-time processing and displaying of data during acquisition campaigns has several advantages:

- Instant access to gas concentrations allows to start the interpretation of the results, which is important to direct the observations during the campaign.
- A direct feedback on the correct functioning of the instrument is provided.
- More effective communication of the potentiality of EM27/SUN.

Under the guidance of my tutors Christel Guy and Denis Jouglet, I decided to implement a dashboard coded in Python in the IDE PyCharm. The dashboard needed to meet the following requirements:

- Display data relative to the last acquisition of the instrument.
- Display data coming from map files and the PTU sensor (temperature, pressure and humidity).
- Display coordinates.
- Display plots of temporal evolution of GHG concentrations measured.
- Automatically update the data when the instrument acquires a new spectrum.

5.1 Obtaining map files

As mentioned in the previous chapter, map files based on GGG2020 represented a major issue in the creation of the dashboard due to their unavailability before the day of the measurements. Two solutions could have been pursued:

- the fist one consisted in a climatology approach: by collecting data from the last decades a model could be created to provide averaged trends of GHG concentration profiles. While this approach would not be as precise as TCCON data and it would not be easy to implement, it would have the advantage of obtaining the map files without needing an internet connection.
- the second solution was to use the database of the European Center for Medium-range Weather Forecast. The ECMWF provides accurate forecasting of numerous weather parameters, GHG concentrations included, by implementing the Copernicus Atmosphere

Monitoring Service (CAMS). These forecasts can be downloaded up to five days before the desired date therefore allowing real-time processing of the spectra.

The second option was chosen because of the reliability of the data and because it was easier and faster to implement. In fact, data from ECMWF database (from now on referred as CAMS) can be downloaded by doing a request from its Web API. Data is provided in the form of *.grib* files, therefore it needs to be extracted and converted in a suitable format. This procedure, from the web request to the conversion in map format, has been already implemented by Hippolyte Leuridan from the Laboratoire des Sciences du Climat et de l'Environnement, who accepted to share his work with us. During my work in taking up these scripts, I introduced some modifications especially to adapt the output to the newer version of TCCON prior files.

In fact, up until 2022, TCCON map files were generated by an older version of the inversion software GGG: GGG2014. Version GGG2014 offered the possibility of correcting laser sampling errors detected on several TCCON spectrometers, while the newer version (GGG2020) includes changes to the spectroscopic database, previous profiles, spectral adjustment and post-processing. A major objective of this update was to diagnose and reduce inter-site bias. The product of GGG2014 is a map file per day, with a temperature profile every 1 meter, between 0 - 70 km height. For GGG2020, on the other hand, one file every 3 hours (the first at 00:00) is provided, with a total of 8 files per day. There are 51 levels for each variable between 0 and 70km altitude, with variable resolution in the atmospheric profile. [26] A study on the bias between the results coming from the two models has been conducted by the previous intern in the SA division, N. Montenegro.

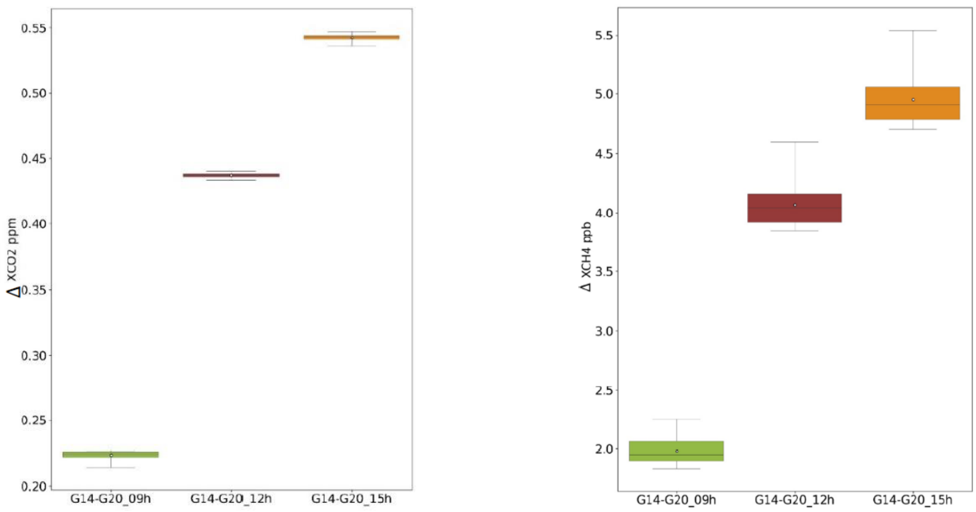
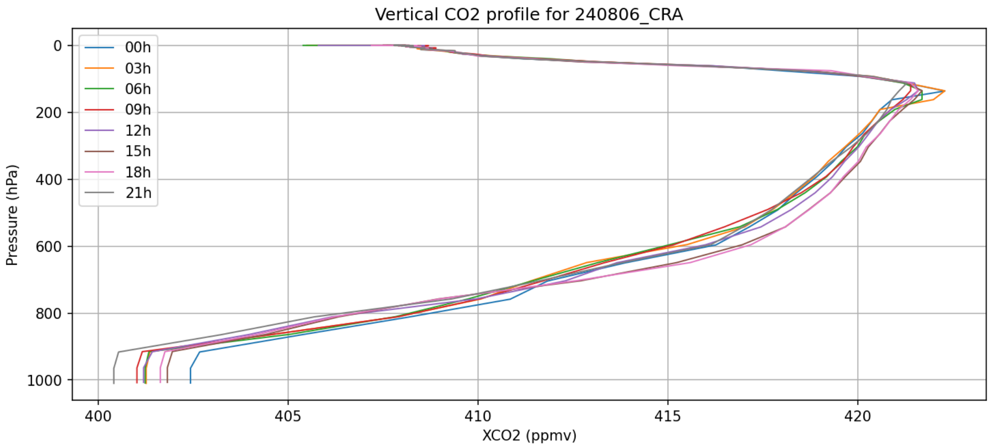


Figure 5.1: Bias CO_2 and CH_4 between GGG2014 and GGG2020

In the figure 5.2 the vertical concentration of CO_2, as modeled by GGG2020 and CAMS, is showed for multiple hours of the 6th August 2024. The plots show very well the difference between the two models: GGG2020 is a climatology model which focuses on long-term averages and global trends rather than the rapid changes captured by meteorological models. It integrates long-term observational data, which explains the stability and smoother curves across different time points. Short-term fluctuations are not captured by this model. On the other hand, CAMS is a meteorological model that captures and predicts short-term changes in atmospheric conditions due to atmospheric mixing, convection, or changing weather patterns. Because of this, CAMS model should be more accurate than TCCON. This will be demonstrated in the following chapters. Fig. 5.3 shows the comparison between the results obtained by using prior profiles coming from



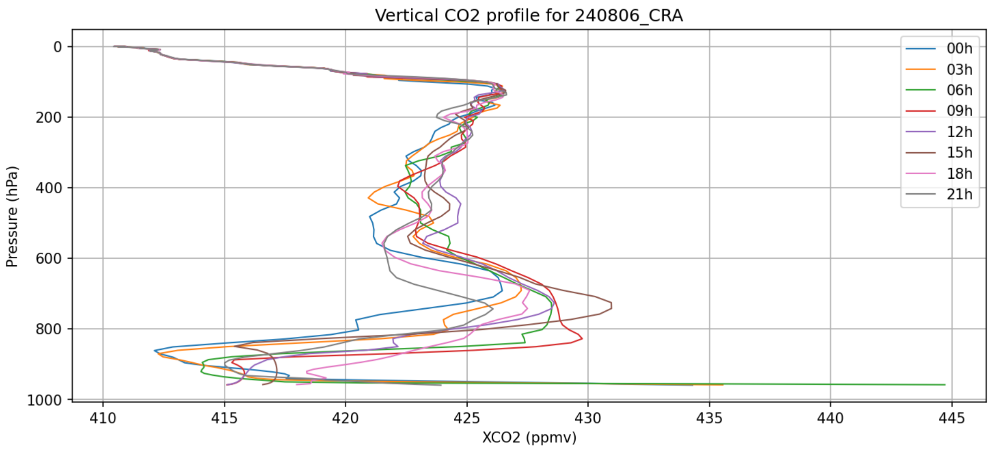


Figure 5.2: CO_2 prior profiles from GGG2020 (top) and CAMS (bottom)

CAMS (in blue) and GGG2020 (in orange). The comparison shows that the concentrations obtained differ of slightly more than 0.12 ppm in the worst case. It is interesting to notice that this difference varies and it doesn't seem to be correlated to the local time. Further analysis would be needed to understand the causes of such variation.

5.2 Processing chain before the dashboard

The processing chain used before the dashboard was conceived to treat all the spectra from one acquisition at the same time. Simon Prady from Magellium shared with me the scripts he created to do that. The processing chain was essentially composed by two main scripts. The first one, *run_proffast_unique.py* , takes as inputs the details of the acquisitions (e.g. date and site) and the directory where the spectra, the met file and the map files are located. The script then calles an instance of the class Pylot from PROFFASTPylot and runs PROFFAST using the functions

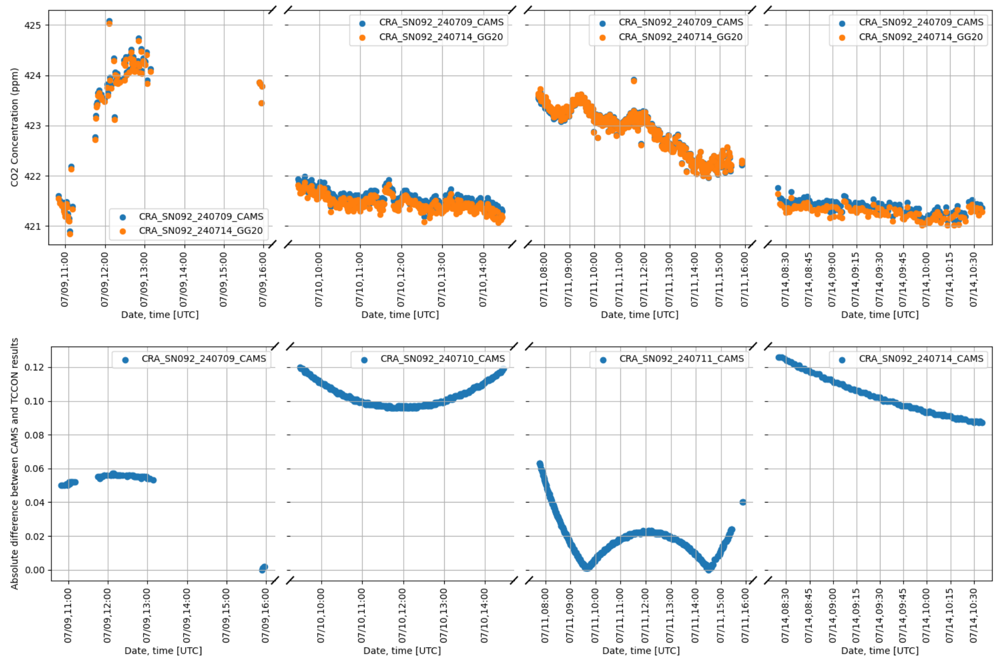


Figure 5.3: Inversions results using the two models (top) and absolute difference (bottom)

of the class. The outputs of PROFFAST are saved by default in the folder where it was installed. The second script, called *traitement_resultats_V*2.4*.py* processes the outputs of PROFFAST, like the name suggests. It takes as inputs the details of the acquisition and the directory where the results of PROFFAST are stored. This script can perform a number of different operations, which can be selected by modifying flag variables in the script:

- Plotting of GHG concentrations and PTU data over time
- Plotting vertical profiles of GHG
- Conversion of spectra from binary to ascii format
- Calibration of spectra
- Plot of uncalibrated and calibrated spectra

All these outputs are saved in the folder where the raw spectra are saved. Calibration of spectra was possible because of the work conducted by the Laboratoire d'Optique Atmosphérique (LOA) who provided the gains and offset to calibrate the signal coming from the two detectors.

5.3 Processing chain of the dashboard

Contrarily to the previous processing chain, the dashboard needed to show the results of the inversion in real time. This means that PROFFAST and the related functions to process PROFFAST's

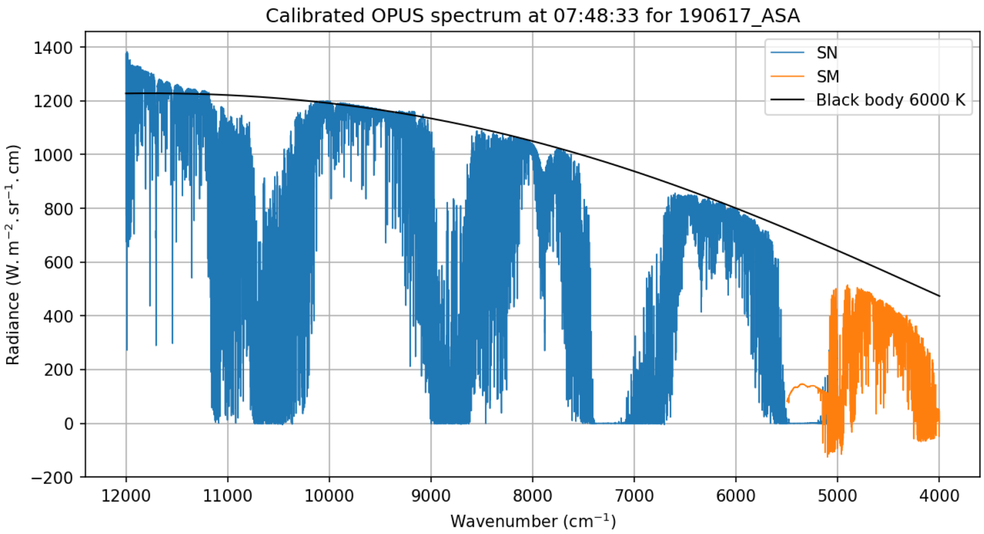


Figure 5.4: Exampe of a calibrated spectrum

data needed to be called everytime a new spectrum was acquired by the EM27. The acquisition time of each scan is one minute. It was therefore essential to optimize the code and the calculations. The most computationally expensive part of PROFFAST is the calculation of the x-sections of GHG, which takes several minutes. Fortunately, before launching the computation PROFFAST checks if the calculation has already been made and if so it skips it. This allows the whole code to be run in less than a minute. All the code was developed by me, but for running PROFFAST and treat its results I built on the code done by S. Prady and I adapted it for the dashboard.

The flow chart in fig. 5.5 illustrates the functioning of the dashboard: once started, all the spectra acquired by the EM27/SUN are processed by PROFFAST. The outputs are then further processed and moved to the folder where the spectra are, to be easily accessible and readable. At this point the dashboard is launched. At the same time, a parallel thread in the code constantly checks if a new spectrum is acquired by the instrument. If this is the case the processing chain starts again and the dashboard is updated. Going more into detail, fig. 5.6 shows an UML diagram that provides an overview on how the code works and the scripts interact:

- *main.py* is the main script. This script takes all the input data needed for the functioning of the dashboard and organises them into a dictionary. *main.py* is divided in two threads. The main thread calls the functions *run_proffast()* and *traitement()*. These two are essentially the function version of the scripts $run\_proffast\_unique.py$ and $traitement\_resultats\_V2.4.py$. The advantage is better modularity in the code structure and the possibility to regroup all input data in the main script. *traitement()* returns a dictionary that contains the paths of the data to be displayed in the dashboard. The dashboard is then launched by the function *main()* contained in the *frontEnd.py* script. The second thread is a daemon thread, which is a thread that runs in the background while the main thread is running and continues running until the main thread ends. In this case the daemon thread is used to detect whether a new spectrum is added to the spectra folder (meaning that a new acquisition has been made). In order to do so, an instance of the class EventHandler is created. This class inherits from FileSystemEventHandler,

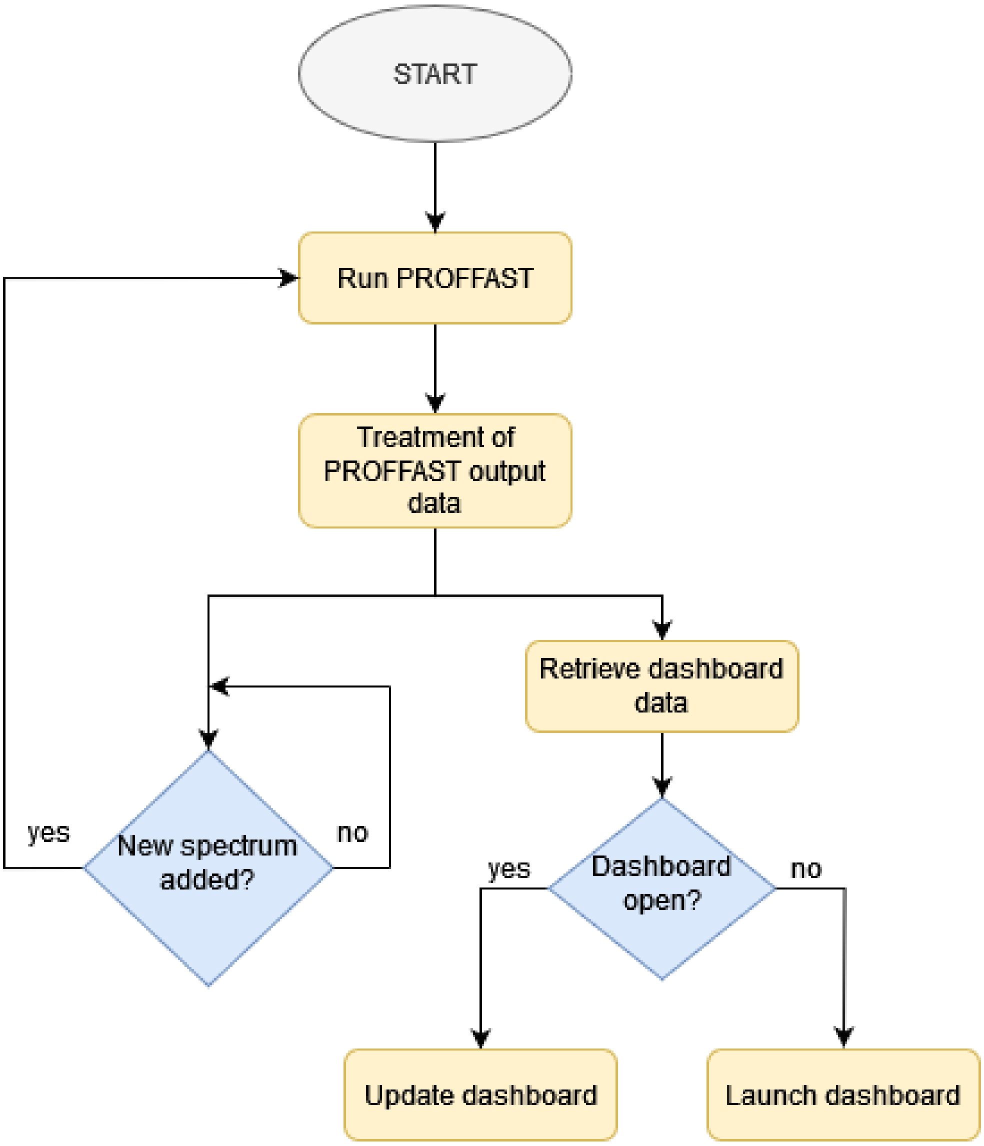


Figure 5.5: Flow chart of the dashboard

a build in class of python library *watchdog*. Every time that a new spectrum is detected the function *on_created* is launched. This function calls *run_proffast()* and *traitement()*, therefore performing again the processing.

- *run_proffast()* takes as input a dictionary containing several information about the observations and the placement of data. It creates an instance of the class *Pylot* from PROFFASTpylot, which contains the methods that launch the different parts of PROFFAST. *run_proffast()* uses also two functions from the script *util.py*: the function *PTURetriaval()* establishes a connection with the raspberry pi that collects the data of the PTU sensor and retrieves the latest met file.

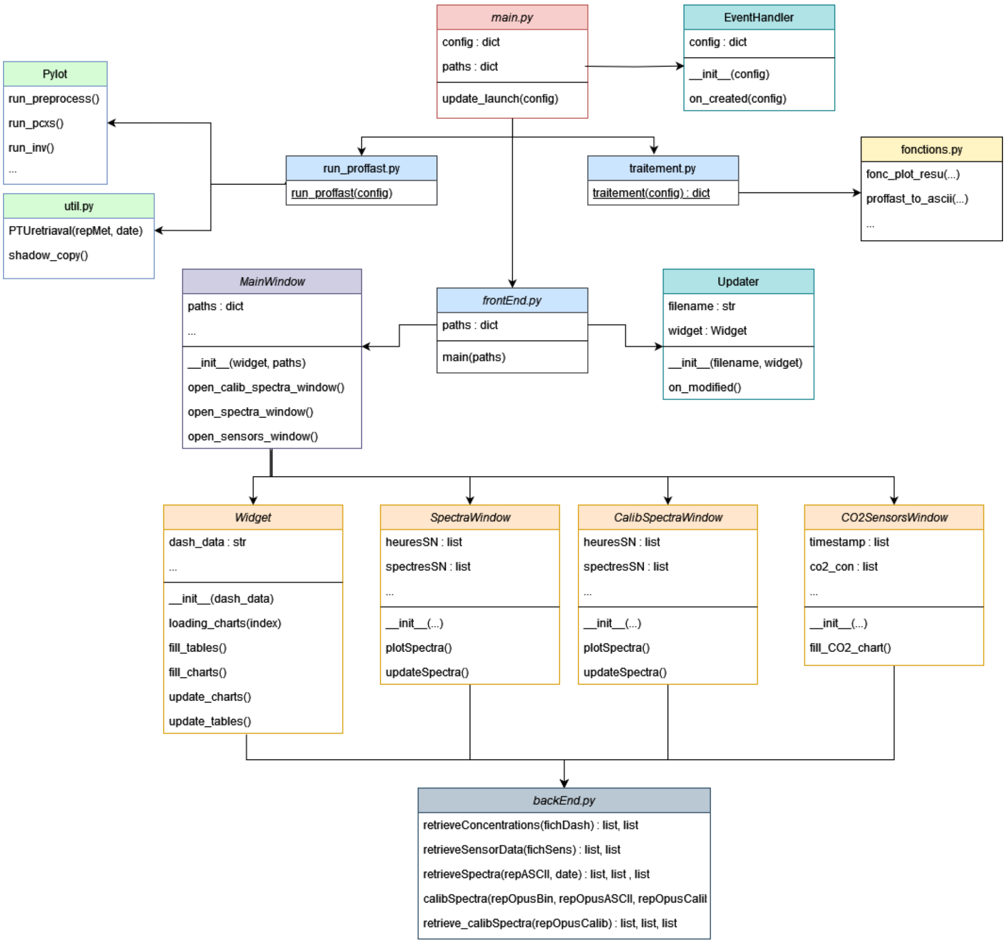


Figure 5.6: UML diagram of the dashboard

One problem encountered during the testing phase was that during the measurements the software OPUS keeps the spectra locked. Therefore it is impossible to access them and therefore to perform the inversion. To bypass this problem the solution of using shadow copies has been adopted: A shadow copy is a backup technology that creates snapshots of a file or volume at a given point in time. This means that even if a file is being modified or used by a program, a shadow copy can capture its state without interfering with the current application. The result is that a "shadow" version of the spectra is accessible by the dashboard. The function *shadow_copy()* performs exactly this task: it creates a shadow copy of the local disk and deletes old shadow copies to avoid filling up memory.

- *traitement()* uses the functions contained in the script *fonctions.py* to perform its operations
- the script *frontEnd.py* contains the code that constitutes the dashboard. This has been coded by using the python library PySide6. This is the python wrapper of the Qt library, conceived to create GUI in C++. PySide6 has been chosen because of its versatility and for its complete documentation. The function *main()* sets an Observer by creating an instance of the class Updater, which inherits from the class FileSystemEventHandler. This class

detects whether the new spectra have been processed, and if so it retrieves the processed data to update the dashboard. The function then creates an instance of the class MainWindow, which is the core of the dashboard.

- MainWindow is a class that inherits from the PySide class QMainWindow. It sets the layout of the dashboard and in turn creates an instance of the class Widget (from QWidget), which displays main information like GHG evolution over time. MainWindow also contains methods that display new windows by creating instances of other classes. In particular: SpectraWindow and CalibSpectraWindow display spectra and calibrated spectra respectively. CO2SensorsWindow displays the CO_2 concentration collected by sensors during the acquisitions.

- finally, the script *backend.py* regroups all the functions that read output files from *traitment()* and organize data in lists so that they can be easily displayed.

Fig. 5.7 shows the main window of the dashboard: on the left side the two tables show data obtained from the inversion of the latest spectra. On the right side this data is plotted in two graphs. The plotted data can be changed by using the comboboxes in the bottom left side. From the "File" button the other windows can be opened. Fig. 5.8 and 5.9 show other two windows of

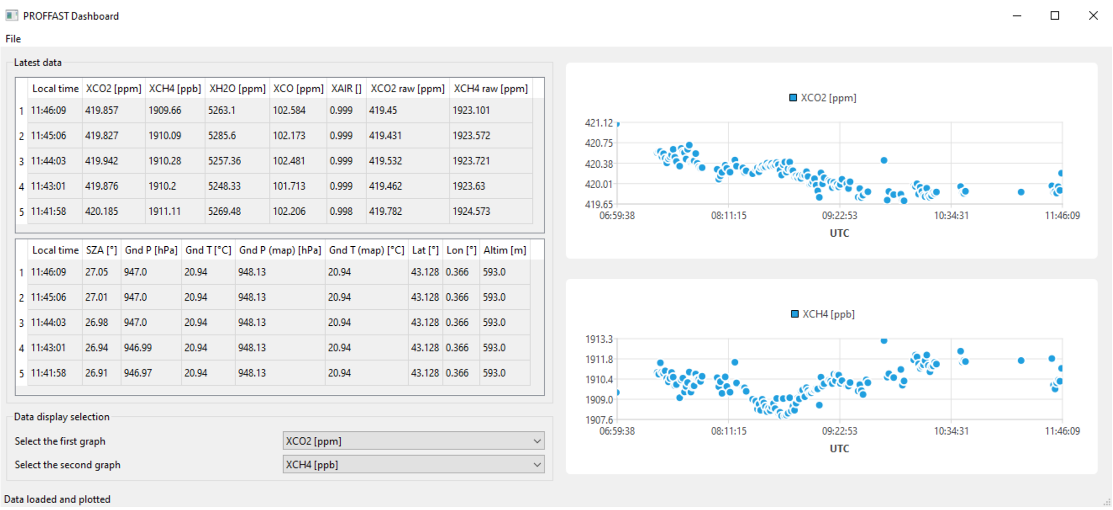


Figure 5.7: View of the main window of the dashboard

the dashboard. It is interesting to notice that the calibrated spectrum showed in 5.9 is quite far from the black body line, in contrast to fig. 5.4. In fact, the EM27/SUN #92, the instrument that acquired these spectra, has been sent to Bruker in order to change the internal camera for the CamTracker. This operation probably resulted in some internal parts being moved. Therefore, the calibration of the instrument done by LOA in 2021 is not valid anymore.

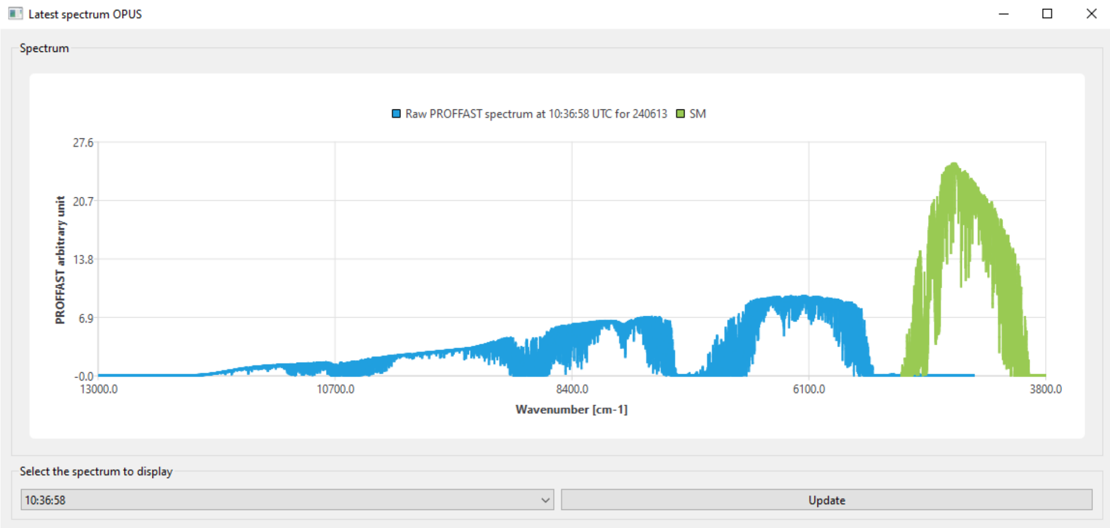


Figure 5.8: View of the spectra window

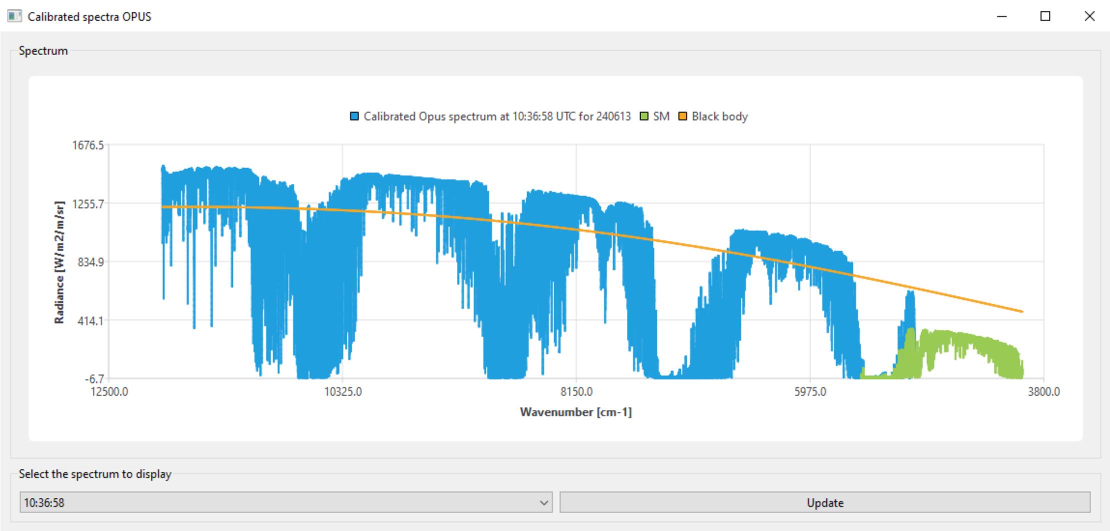


Figure 5.9: View of the calibrated spectra window

Chapter 6

Measurement campaigns

During the summer of 2024, the department of Sondage Atmosphérique decided to start an ambitious campaign of measurement, involving three EM27/SUN deployed in different sites, AirCore balloon flights from CNES base in Aire-Sur-l'Adour and OCO-2 and Sentinel-5P passages.

6.1 Objectives

The campaign had different objectives:

- The most important objective was to prove that, by deploying different instruments at different altitudes, it was possible to increase the vertical resolution of the measurements by obtaining a vertical pseudo-profile of CO_2 concentration.
- Another objective was to consolidate the reliability of EM27/SUNs and to prove that their results were consistent with each other.
- Furthermore, there was interest in obtaining a temporal series of CO_2 concentration in the lower layer of the atmosphere, since it is the most difficult to model.
- Another objective was to compare the pseudo-profiles to the profiles obtained by models in order to highlight any strong difference between the two.

This campaign fits well into MicroCarb's validation plan. Indeed, with the observations that have been made, it is possible to prove that data from the satellite can be validated by multiple deployed EM27/SUNs at the same time. In addition, demonstrating that it is possible to obtain vertical resolution in their measures opens new possibilities for the validation of the vertical profiles estimated by OCO-2 and MicroCarb itself. Finally, being able to measure spatial variability in CO2 concentration provides insight into the limitations of satellite observations. In particular, a MicroCarb pixel has a ground footprint of about 4Km x 9 km. If the CO2 concentration varies greatly within this area, as may be the case in mountainous areas, the pixel could not be representative of the true concentration.

To achieve the objectives stated above, two EM27/SUN have been permanently installed at the Centre de Recherche Atmosphérique (CRA) and at the observatory of Pic du Midi, while a third one has been used to make inter comparisons between the instruments and to perform measurements during AirCore balloons flights. Tables 6.1 and 6.4 provide more information about the instruments and the observation sites.

Instrument	Owner	Site
EM27/SUN #92	CNES	CRA
EM27/SUN #179	LSCE	PDM
EM27/SUN #118	LERMA	None

Table 6.1: EM27/SUN used for the campaigns

Site	Abbr.	Latitude [°]	Longitude [°]	Altitude [m]
CNES, Toulouse	TLS	43.559	1.483	140
CRA, Lannemezan	CRA	43.128	0.366	594
Pic du Midi de Bigorre	PDM	42.936	0.143	2883
CNES, Aire-Sur-l'Adour	ASA	43.706	-0.251	90

Table 6.2: Observation sites

6.2 Casing of EM27/SUN

EM27s typically need to be operated by a person in order to start and end the measurements. This would have represented a huge limit because of the lack of people able to pilot the instrument in the sites of CRA and PDM. A solution to this problem has been developed by the Laboratoire des Sciences du Climat et de l'Environnement, which created a casing that would contain the instrument and automatically manage the measurements. Figs. 6.1 and 6.2 show its main features: firstly, the instrument is placed inside the casing. Inside the casing a PC receives all the inputs coming from the different sensors: PTU sensor, pyranometer and rain detector. The PC is connected to the power grid, but can switch to an internal battery in case of problems. The connection is provided by a 4G router and the PC can be accessed by using the software AnyDesk. Two aeration vents assure that the internal temperature does not reach excessive values.

The acquisitions are started by opening the mirrors' lid only after several conditions are met (sufficient sunshine, no rain and power supply on electric grid). Acquisitions can also be started or stopped manually. After a day of measures, the spectra acquired are preprocessed locally. They are then sent in binary format to the server of LSCE, where they are treated to obtain the concentrations. Results are available after a few days and they're obtained by using both CAMS and TCCON prior files.

6.3 Setup campaign

A setup campaign was organized in the days from the 9th July to the 12th July. The planning was the following: The 12th of July was a day of particular interest since the satellite OCO-2 would pass almost exactly over the site of ASA and not far from the site of PDM. Fig. 6.3 shows it on a map. On that day, all three instruments would have measured GHG concentrations at different altitudes and in conjunction with two other reliable sources: OCO-2 and the AirCore balloon.

Unfortunately these kind of measurements are strongly weather-depended, since if the Sun

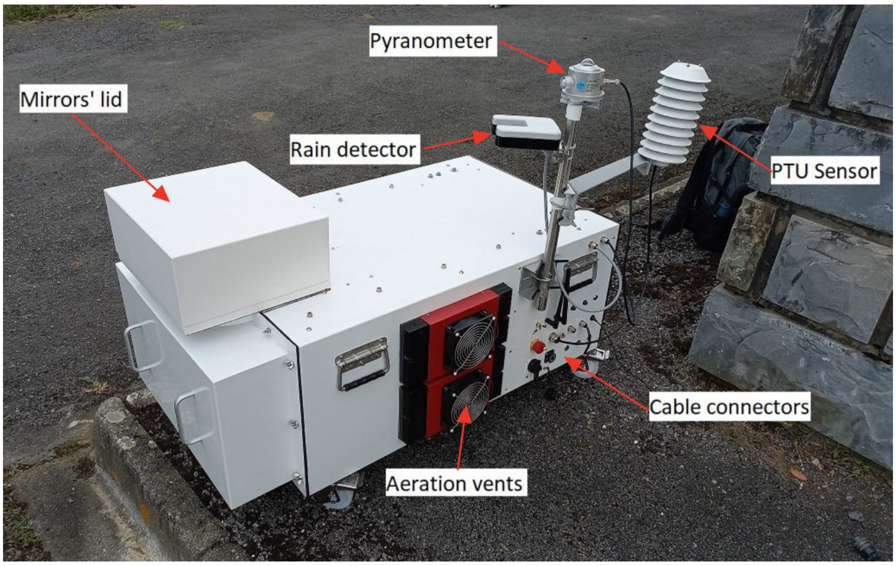


Figure 6.1: EM27/SUN casing deployed in CRA

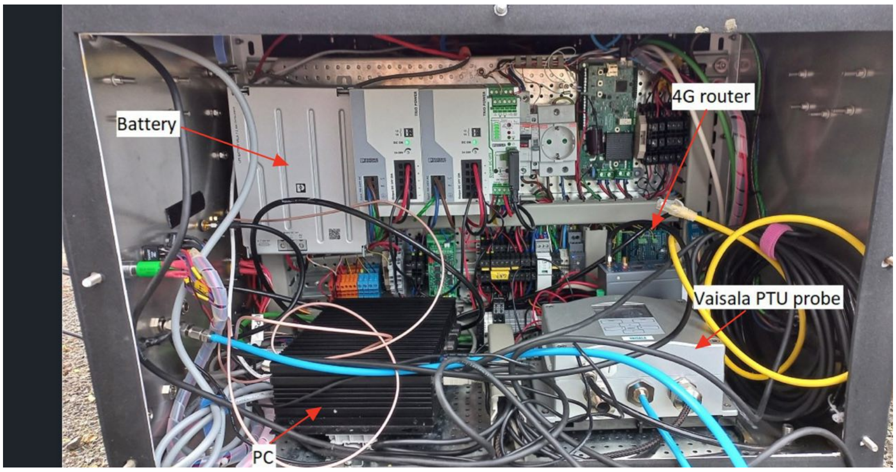


Figure 6.2: Interior of the casing

is covered by clouds no spectrum can be acquired. The 9th July, due to unfavorable weather and problems in the installation and testing of EM27/SUN #92, there were not sufficient spectra acquired by the three instruments to perform an effective intercomparison. On the other side on July 10th everything worked as planned. Finally, the observations planned on July 12nd were not conducted because of bad weather: no AirCore balloon was released and OCO-2 didn't perform measurements over the site of ASA because of cloudy sky.

Day	Objective
9th July	Installation of EM27/SUN #92 in the casing at CRA. Intercomparison between the three EM27/SUN
10th July	Installation of EM27/SUN #179 in the casing at PDM. Intercomparison between #179 and #118
12th July	AirCore Balloon release and deployement of EM27/SUN #118 at ASA, in correspondence with the passage of satellite OCO-2.

Table 6.3: Planning of the setup campaign

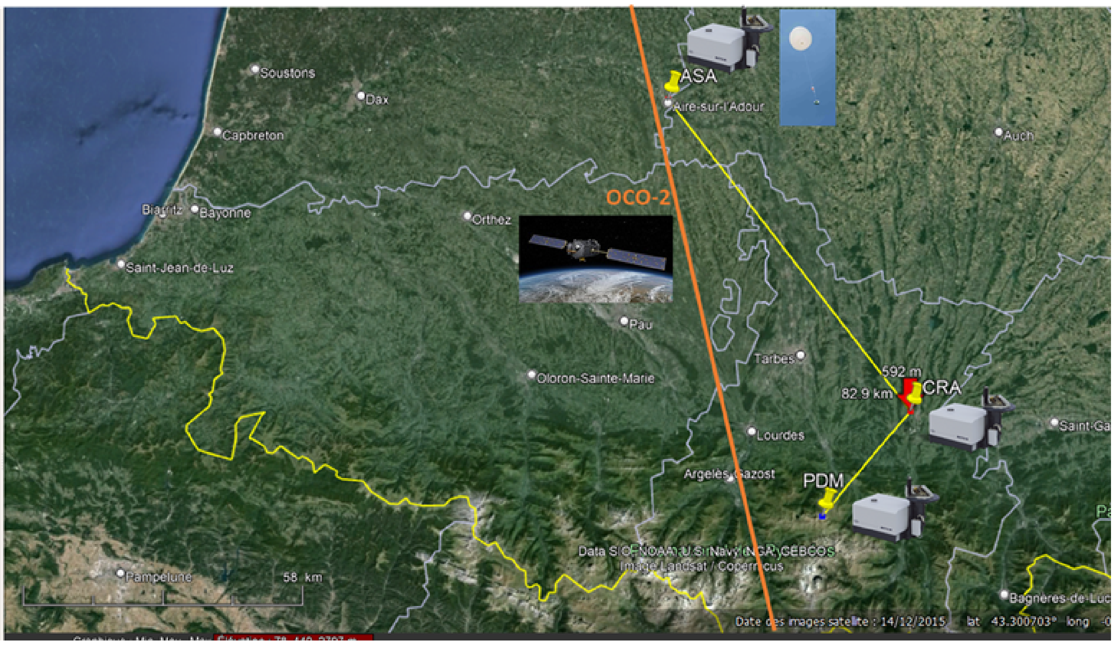


Figure 6.3: Planning of the observations on the 12th July

6.4 Follow-up campaigns

Two other campaigns were conducted during the summer, mainly to perform other intercomparisons and to solve problems related to the casings. In particular, the two instruments had connection problems and couldn't send data over several days. An intervention was needed to solve these issues. The table above shows all the campaigns that were done until the end of August. Sufficient data were collected to perform the intercomparison between the instruments.

Figure 6.4: Deployement of the three EM27/SUN at CRA during the setup campaign

Day	Objective	Outcome
24th July	Intercomparison between #118 and #92 at CRA in the morning. Intercomparison between #179 and #92 at PDM in the afternoon. Verification of the good functioning of the instruments.	Ineffective intercomparison at CRA because of bad weather. Very good intercomparison at PDM. All instruments were functioning properly
25th July	Deployement of #118 on the site of ASA in conjunction with AirCore balloon release	AirCore released and successfully retrieved. The sky was cloudy but some measurements were possible
6th August	Solve #92 and #179 connection issues and perform an intercomparison between #118 and #92	Connection problems solved and intercomparison made with a clear sky

Table 6.4: Planning of follow-up campaigns

Chapter 7

Data Processing and Analysis

7.1 Preliminary Work

Data coming from EM27/SUNs #92 and #179 are treated by LSCE and uploaded on their server, whereas data from #118 are treated by me (using CAMS and TCCON files) and Simon Prady (using TCCON files). In order to ensure the consistency of the data it was necessary to check that the treatments of CNES and LSCE lead to the same results if given the same inputs. This was not the case for the following reasons:

- LSCE performs the inversions using PROFFAST V2.3, which is not its latest version. Furthermore, when using CAMS, they use only one file (at 12h of the considered day) instead of eight (one every three hours of the considered day)
- CNES performs the inversions using PROFFAST V2.4, which is its latest version. Furthermore for both ECMW and TCCON, eight files are used.

Because of this, starting from the same raw data the two processing chains would provide results that differed up to 0.30 ppm one to another. After further processing of the data, I ascertained that:

- The raw atmospheric spectra that I treated with PROFFAST V2.3 and one CAMS map file give the same XCO_2 than LSCE treatments based on CAMS models.
- The raw atmospheric spectra that I treated with PROFFAST V2.3 and eight TCCON map files give the same XCO_2 than LSCE treatments based on TCCON models.

The comparisons for 9th and 10th July (using CAMS prior data) are showed in fig.7.1. As it can be seen they're identical except for some points that are considered as anomalies. For the first part of my analysis (inter-comparisons of EM27/SUNs on the same site) I carried on comparisons with both treatments (one file CAMS and eight files TCCON). For the following analysis though, I used the results coming from LSCE treatments based on TCCON models.

7.2 Analysis of the inter-comparisons

Concentrations coming from the different spectrometers needed to be pre-treated before being compared between each other. There are essentially two pre-treatements that need to be done:

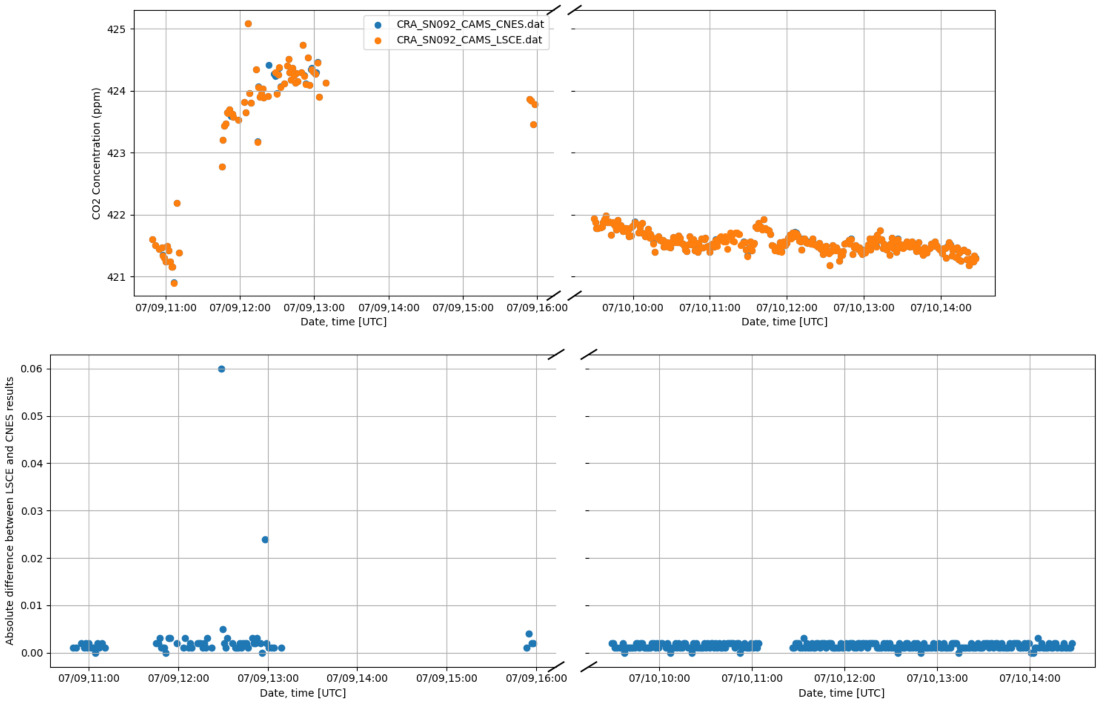


Figure 7.1: Plot of CNES and LSCE results (top) and absolute difference between LSCE and CNES "onemap" (bottom)

- **Smoothing of the data:** measurements are affected by noise. In order to reduce it a moving average has been applied on them, with a width of thirty minutes.
- **Resampling of the data:** the EM27 acquires a spectrum approximately every minute, but the time of the measurements does not match between the instruments. Therefore every set of data was aligned and downsampled to have a point every five minutes.

7.2.1 Inter-comparisons in the same site

Once validated the treatments, the following step was to plot every day where there were two or more EM27/SUN deployed in the same site. The goal is to derive a correction factor to be applied to SN092 and SN179 to eliminate the bias between the two instruments. Unfortunately, the only day where all three instruments were deployed (9 July 2024) not enough spectra were measured by #179 to obtain a proper comparison. Therefore, the correction factor was to be derived by analyzing the differences between SN092-SN118 and SN179-SN118. Fortunately, the inter-comparisons showed a bias in the measurements of around 0.1ppm for #179 and #118 and less than 0.1ppm for #92 and #118 . Since these results are above the required precision, no correction factor was applied to the following comparisons. Figures 7.2 and 7.3 show the results. One anomaly was found for the 9th July in CRA: after 11.30 AM the mirrors of the EM27/SUN #92 stopped pointing correctly the Sun. This could be ascertained because the software that pilots the mirrors has an internal camera that takes pictures of the pointing every minute. Because of the misalignment the results of the inversions of the spectra were not valid

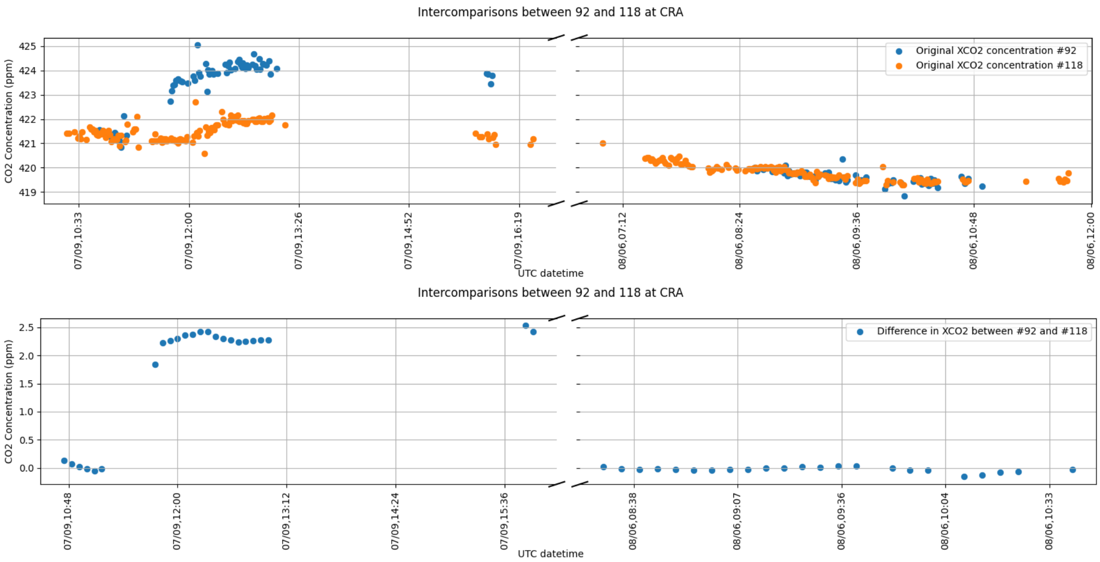


Figure 7.2: Inter-comparison at CRA between #118 and #92 (top) and difference between the measurements (bottom)

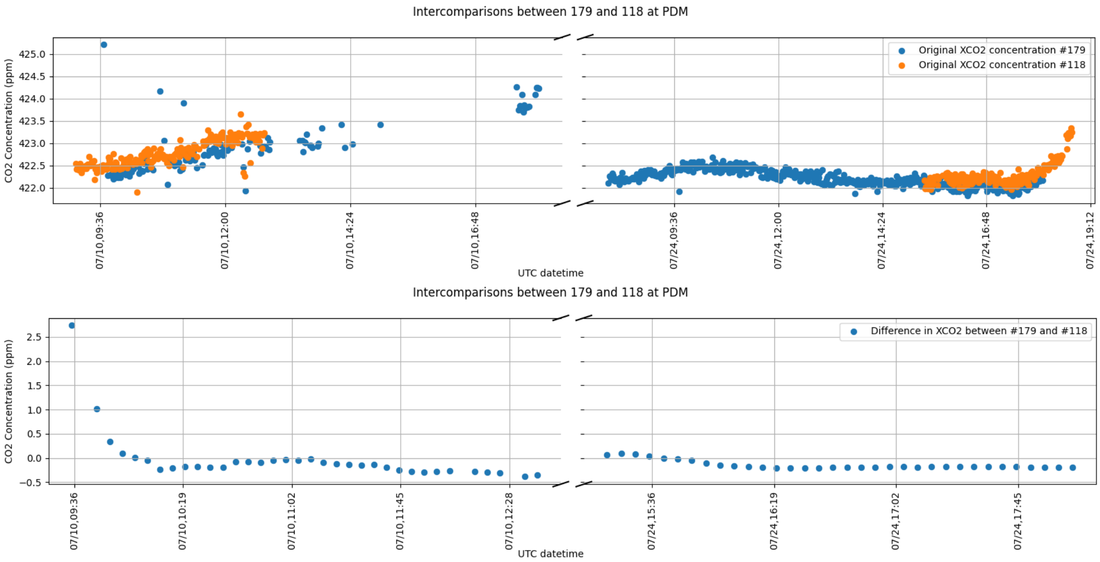


Figure 7.3: Inter-comparison at PDM between #118 and #179 (top) and difference between the measurements (bottom)

and that explains the difference between the two instruments showed in 7.2.

7.2.2 Inter-comparison between the sites of CRA and PDM

Thanks to the automatic casing of LSCE, the instruments deployed at CRA and PDM could acquire spectra of the atmosphere for two months. The concentrations acquired in these days were plotted in order to observe the XCO_2 differences between the two sites with high temporal resolution. In fig.7.4 it can be seen that a lot of days are not presents and some of the days that

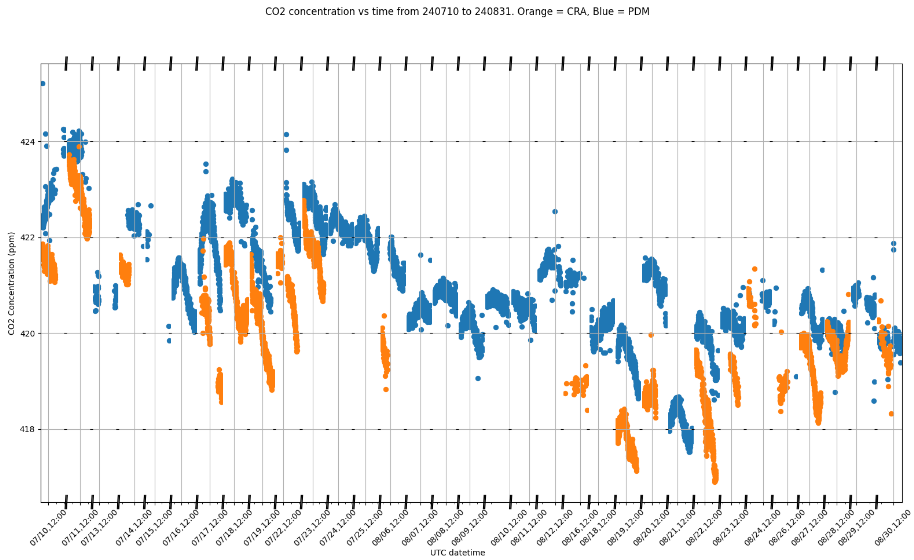


Figure 7.4: CO_2 concentrations at PDM (blue) and CRA (orange) in the two months of inter-comparisons

are plotted do not contain measurements from the two sites. Indeed, two factors limited the number of observations: bad weather and connection problems with the casing. Because of that, there were only 18 days where both instruments worked at the same time.

What can also be noted is that CO_2 concentration is always lower for CRA than for PDM. This result fits well with the theory, since the air closer to the surface is expected to have less CO_2 during the summer because of the photosynthesis of the vegetation. Furthermore, the mean concentration on the graph constantly decreases over the two months, in accordance with the yearly trends. Fig. 7.5 shows the plots of the difference in XCO_2 between the two sites. This difference is not constant over time, with a minimum of $0.27ppm$, a maximum of $2.67ppm$ and an average of $1.41ppm$ over all the days. Overall it can be seen that the difference tends to increase during the day. A possible reason for this is, again, photosynthesis: the site of CRA, situated in the countryside, is surrounded by vegetation that as the day goes on absorbs more and more CO_2. The site of PDM is situated on a mountain, therefore with very few vegetation. Nevertheless this lowering in XCO_2 concentration is still observed, albeit in smaller amounts. In this case, however, it is due to convection currents carrying air from the plains to the mountains during the day.

7.3 Pseudo-Profile Construction

Thanks to the difference in CO_2 concentration obtained between the two sites, it was possible to construct a pseudo vertical profile of XCO_2 :

- **High layer:** EM27/SUN #179 at PDM associated with 2800m-80km (end of atmosphere).
- **Low layer:** EM27/SUN #179 at CRA associated with 598m - 2800m.

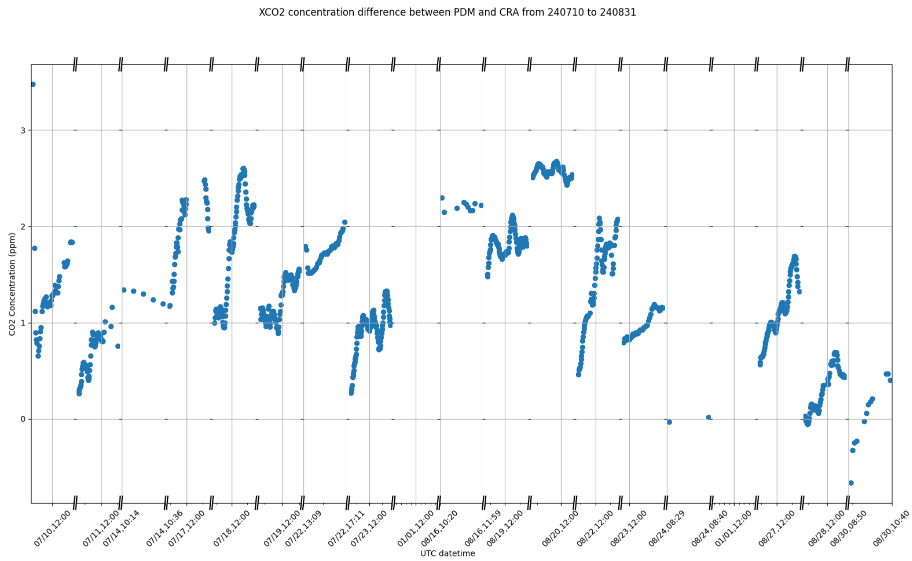


Figure 7.5: Difference between CO_2 concentrations between PDM and CRA in the two months of inter-comparisons

The primary objective here is to demonstrate that differential measurements can provide access to the vertical gradient of the gases, which is a key aspect of this study. To obtain the XCO_2 in the low layer the following formula was used:

$$XCO2_{CRA} = \frac{P_{PDM}}{P_{CRA}} \cdot XCO2_{PDM} + \frac{P_{CRA} - P_{PDM}}{P_{CRA}} \cdot XCO2_{CRA-PDM}$$

Where:

- $XCO2_{CRA}$ is the dry-air mole fraction averaged on the atmospheric column from CRA's altitude to free space.
- $XCO2_{PDM}$ is the dry-air mole fraction averaged on the atmospheric column from PDM's altitude to free space.
- $XCO2_{CRA-PDM}$ is the dry-air mole fraction averaged on the atmospheric column from CRA's altitude to PDM's altitude.
- P_{CRA} is the pressure measured at CRA
- P_{PDM} is the pressure measured at PDM

From this, the CO_2 concentration between CRA and PDM can be obtained. The same formula could be used to retrieve the concentration between ASA and CRA by substituting PDM with CRA and CRA with ASA. This would allow to obtain an even lower layer, associated with 78m – 598m. Unfortunately, during the two months there was only one day of measurements at ASA. In the same day the site of CRA didn't acquire any spectrum due to bad weather so no pseudo-profile could be traced. Once obtained the integrated values of XCO_2, a vertical pseudo-profile of CO_2 concentration with respect to atmospheric pressure can be created.

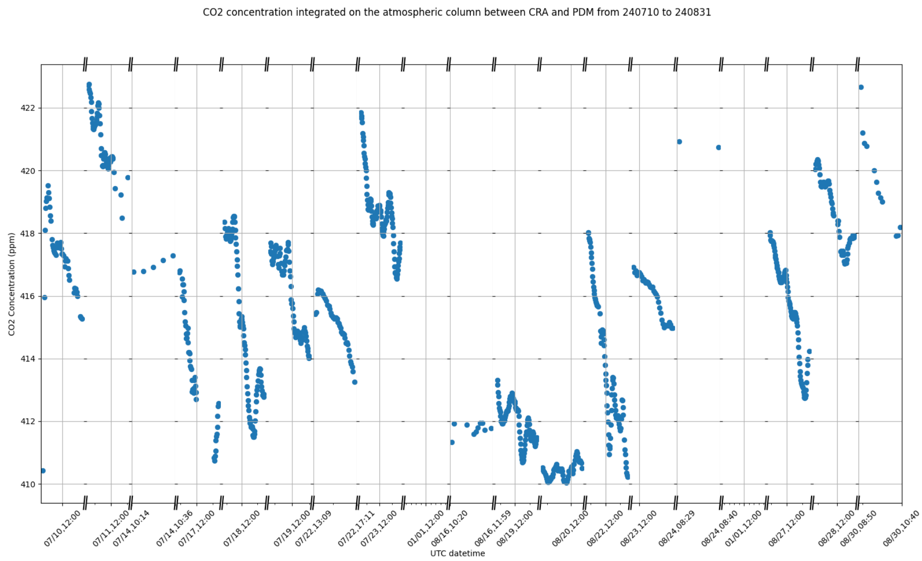


Figure 7.6: XCO_2 concentration in the atmospheric column between CRA and PDM

Figure 7.6 further confirms the hypothesis stated in the previous section: the CO2 concentration between CRA and PDM decreases throughout the day due to photosynthesis and it can explain the variations shown in fig. 7.5. It can be observed that, in general, the pseudo-profiles

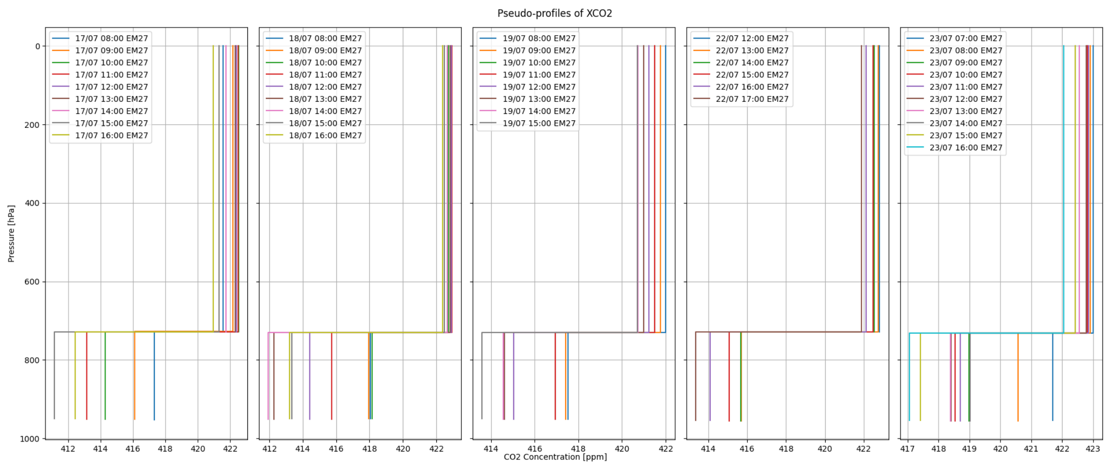


Figure 7.7: Pseudo-profiles of CO_2 obtained from measurements at CRA and PDM

are in agreement with the theory since the concentration decreases throughout the day and that this decrease is far more pronounced in the low layer than in the high layer.

7.4 Comparison with CAMS Profiles

The map files obtained from the ECMWF database (CAMS) contain prior vertical profiles of major GHG, like CO_2. These profiles are traced every three hours of the day and have a high vertical resolution. Once obtained the pseudo-profiles, the objective is to compare them with the actual CAMS profiles to verify whether the model accurately reproduces the temporal variations. In order to do so, the high vertical resolution of the profiles needed to be reduced to two points. In other words we needed to obtain the column-averaged dry-air mole fraction of CO_2 between CRA and PDM and from PDM to space. The theoretical formula to obtain it is the following:

$$XCO_2 = \frac{\int_{H_1}^{H_2} CO_2(z) \cdot p(z)\, dz}{\int_{H_1}^{H_2} p(z)\, dz}$$

Where:

- $CO_2(z)$ is the mole fraction of the gas at altitude z (as a function of altitude).
- $p(z)$ is the atmospheric pressure at altitude z.
- z is the altitude..
- dz represents an infinitesimal change in altitude.

This is well approximated by summing the pressures across all altitudes:

$$XCO_2 = \frac{\sum (CO_2(z) \cdot \Delta p(z))}{\sum \Delta p(z)}$$

With this formula map files can be used to create a "prior pseudo-profile", that can be compared with the one obtained from the measurements. In order to display the information as clear as possible, hourly plots were created where pseudo-profiles created from EM27/SUN measurements are compared with pseudo-profiles created from CAMS profiles. In the same plot I added the XCO_2 concentration integrated on the whole atmospheric column (from CRA to space) and ground measurements coming from PICARRO CO_2 sensors (two at CRA at 30m and 60m from the ground and one at PDM at 20m from the ground). Two plots for 10th and 11th July at 12h are showed in fig. 7.8. It has to be noted that the CAMS profiles used for this analysis refer to the site of CRA. Table 7.1 summarizes the results of this comparisons for every day where measurements in both sites were available. Furthermore, only data referring to an hour were the corresponding map profile was available is presented in the table. For reasons of space, the letter "L" in the table indicates values that refer to the lower layer (the atmospheric column between CRA and PDM). The letter "U" refers to the upper layer (the atmospheric column between PDM and the space). The unit of measurement is ppm (parts per millions)

Datetime	XCO2 L	XCO2 U	XCO2 L CAMS	XCO2 U CAMS	Δ XCO2 L	Δ XCO2 U
10/07 - 9h	417.318	423.087	421.305	423.169	-3.987	-0.082
10/07 - 12h	416.988	422.797	418.390	423.957	-1.402	-1.159
11/07 - 9h	421.773	423.908	423.763	425.708	-1.989	-1.800
11/07 - 12h	419.797	423.890	420.556	425.220	-0.759	-1.330
17/07 - 9h	416.094	422.161	421.187	422.595	-5.093	-0.434
17/07 - 12h	–	422.383	–	423.117	–	-0.734
17/07 - 15h	411.167	421.305	413.018	422.834	-1.852	-1.529
18/07 - 9h	417.960	422.734	423.897	424.165	-5.937	-1.431
18/07 - 12h	414.416	422.888	420.431	423.870	-6.014	-0.982
18/07 - 15h	413.349	422.478	422.835	424.634	-9.486	-2.156
19/07 - 9h	417.412	421.765	419.824	423.536	-2.411	-1.771

Datetime	XCO2 L	XCO2 U	XCO2 L map	XCO2 U map	Δ XCO2 L	Δ XCO2 U
19/07 - 12h	415.046	421.234	416.666	422.993	-1.620	-1.758
19/07 - 15h	413.582	420.697	415.069	422.642	-1.486	-1.945
22/07 - 12h	–	422.812	–	425.351	–	-2.539
22/07 - 15h	415.094	422.484	416.765	424.838	-1.671	-2.354
23/07 - 9h	418.981	422.824	417.242	425.186	1.738	-2.362
23/07 - 12h	418.407	422.789	416.258	425.136	2.148	-2.347
23/07 - 15h	417.422	422.423	415.811	424.861	1.611	-2.438
16/08 - 12h	411.491	421.201	415.566	422.708	-4.074	-1.507
20/08 - 9h	410.182	421.300	409.323	422.830	0.859	-1.531
20/08 - 12h	410.700	421.400	409.718	423.100	0.982	-1.700
22/08 - 9h	416.632	420.110	421.739	421.553	-5.107	-1.443
22/08 - 12h	412.114	419.968	418.533	421.518	-6.420	-1.550
22/08 - 15h	411.809	419.248	417.615	421.819	-5.807	-2.570
23/08 - 12h	416.493	420.252	422.612	421.969	-6.120	-1.717
24/08 - 9h	–	420.799	–	–	–	-3.805
26/08 - 9h	414.723	420.426	415.343	419.931	-0.619	0.494
27/08 - 9h	417.552	420.564	418.888	421.160	-1.336	-0.596
27/08 - 12h	415.812	420.555	419.160	421.534	-3.347	-0.979
27/08 - 15h	413.047	419.964	415.436	422.220	-2.388	-2.256
28/08 - 9h	419.752	420.099	420.605	421.187	-0.853	-1.088
28/08 - 12h	417.820	419.920	420.979	421.409	-3.160	-1.489
30/08 - 9h	419.830	419.858	423.556	421.861	-3.726	-2.004

Table 7.1: Comparison between measured and CAMS predicted XCO2

From the two tables, it can be inferred that the predictions are better for the uppermost layer, with an average distance from observations of 1.633 ppm. For the lowest layer, there are deviations up to almost 10 ppm, with a mean of 3.133 ppm. Furthermore, the same analysis was

	Min [ppm]	Max [ppm]	Average [ppm]
$\|\Delta XCO2\ L\|$	0.619	9.486	3.133
$\|\Delta XCO2\ U\|$	0.082	3.805	1.633

Table 7.2: Summary statistics

conducted by using CAMS profiles of PDM and the differences obtained were larger. A table similar to 7.1 can be found in the appendix (A.1). The summary tab provided here shows that there is not significant change from the two sites.

	Min [ppm]	Max [ppm]	Average [ppm]
$\|\Delta XCO2\ L\|$	0.127	8.741	3.524
$\|\Delta XCO2\ U\|$	0.028	2.623	1.300

Table 7.3: Summary statistics using CAMS models at PDM

All these results could be explained in several ways:

- The XCO_2 relative to the atmospheric column between CRA and PDM was calculated with concentrations obtained at two different sites (CRA and PDM precisely). The result therefore suffers from some error.
- Generally speaking, models are not very capable of accurately predicting concentrations in the lower atmosphere. The fact that the differences are more pronounced in the lower layer is probably due to this.
- Finally, models have lower accuracy in mountainous areas. Forecasting in these regions is more challenging due to the high variability of weather conditions.

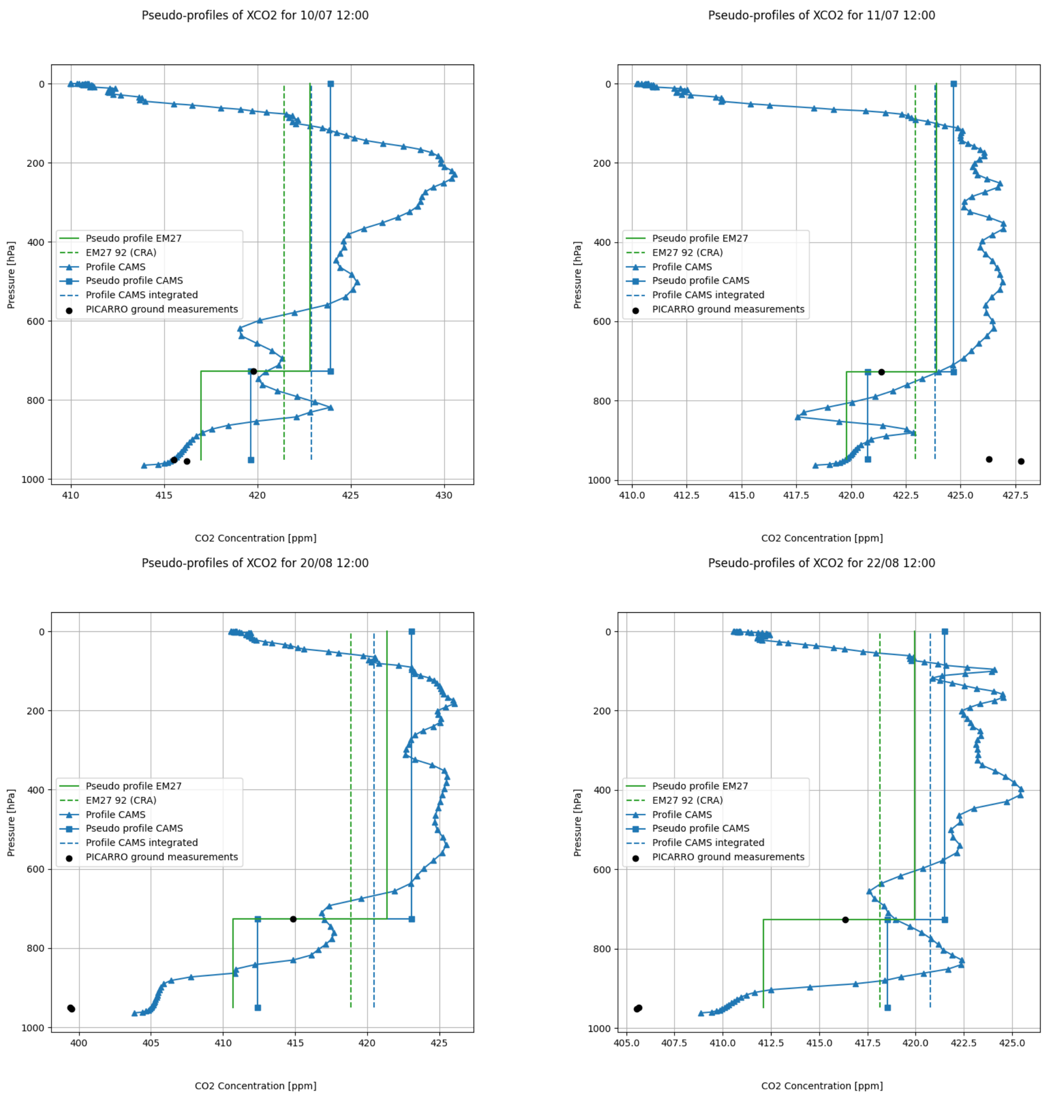


Figure 7.8

Furthermore, it is interesting to notice that PICARRO ground measurements are almost always very different from model predictions, suggesting that models are not very effective for local measurements. In B.3, the variation of PICARRO ground measurements over time is showed.

7.5 Comparison with GGG2020 profiles

The same analysis were conducted by using prior profiles obtained from the GGG2020 model. Some of the results are showed in fig. 7.9. A complete table showing the results for every day of comparison can be found in A.2, whereas a summary table can be found in 7.4. What the plots and the tables show is that, in general, the meteorological approach of CAMS models works

	Min [ppm]	Max [ppm]	Average [ppm]
|ΔXCO2 L GGG2020|	0.031	11.144	4.781
|ΔXCO2 U GGG2020|	0.792	3.906	2.548

Table 7.4: Summary statistics using GGG2020 models at CRA

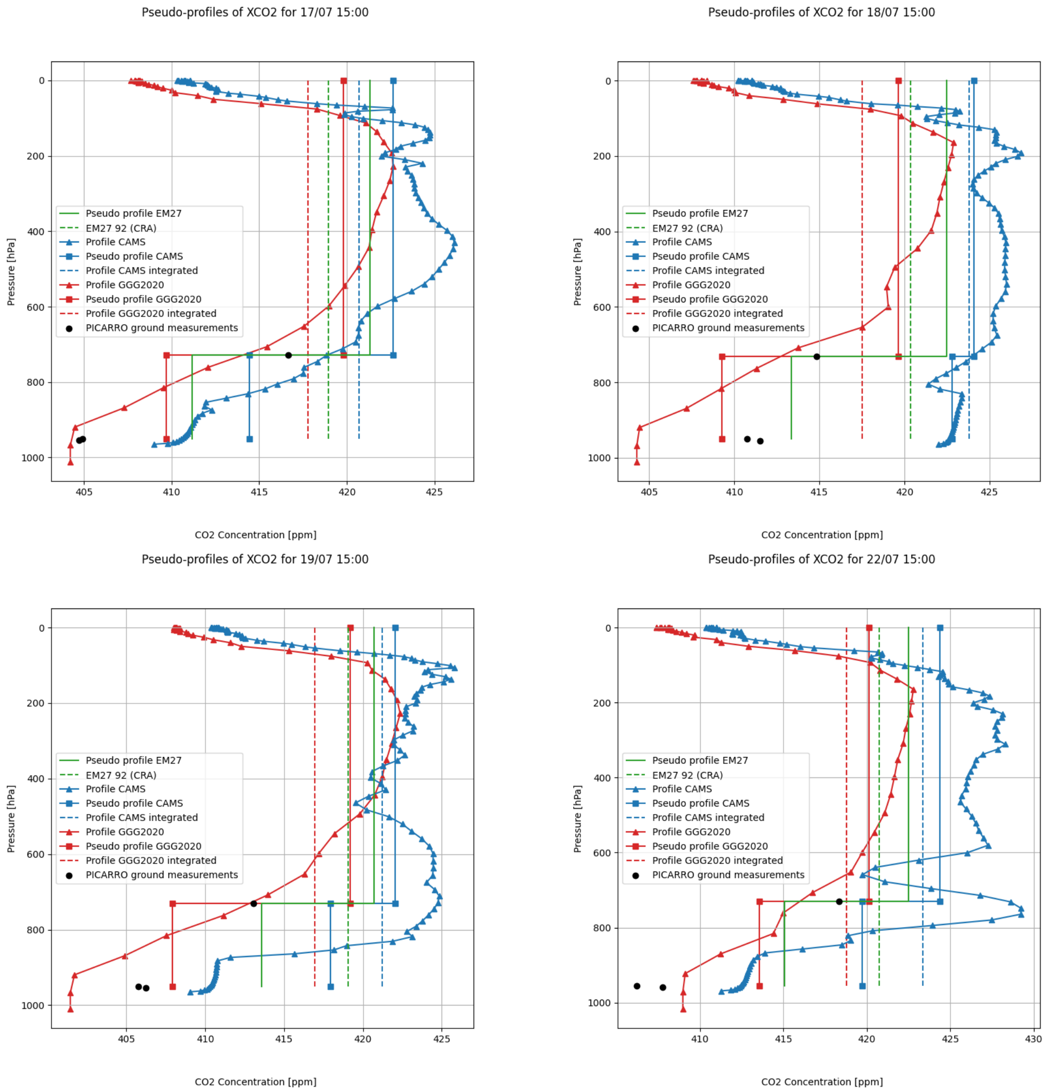


Figure 7.9

better than the climatological approach of GGG2020 models. This would mean that by using CAMS models as prior map files for the inversion, the resulting concentrations would be more accurate. On the other hand, when dealing with large networks of instruments whose products largely depend on the inputs of the processing softwares, it is absolutely essential to ensure that the raw data follow the same processing chain. The TCCON network uses GGG2020 map files for its inversion and this is the main reason why COCCON network and CNES still continue to use this model instead of CAMS. Nevertheless, it is important to be aware of the limitations of this approach when conducting studies and preparing operations like the Cal/Val phase of MicroCarb and MERLIN which require high precision and accuracy.

7.6 Comparison of CH_4 pseudo-profiles with models

The other major GHG measured by the EM27/SUN is the methane. Following the same methods used for CO_2, pseudo-profiles created from EM27 measurements were compared with CAMS and GGG2020 models. Figures 7.10 and 7.11 show that methane has a very different behavior than

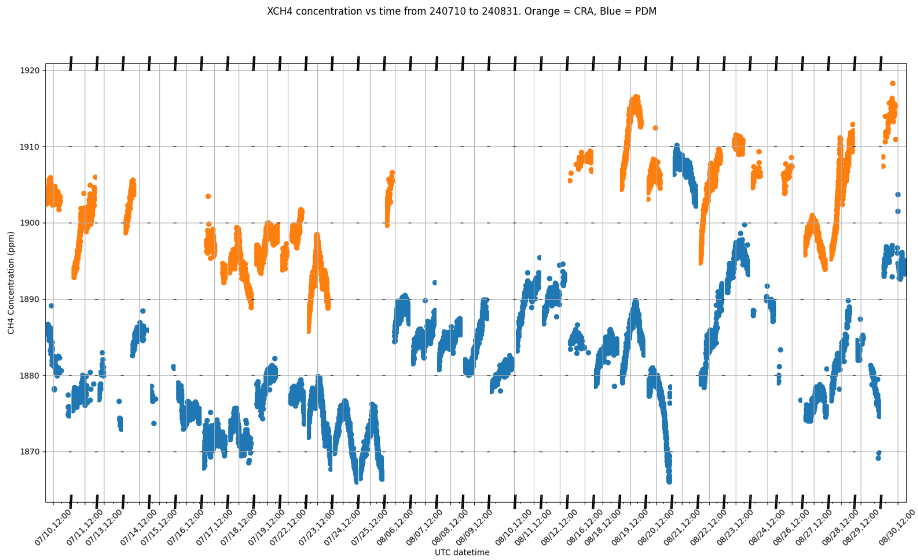


Figure 7.10: CH_4 concentrations at PDM (blue) and CRA (orange) in the two months of inter-comparisons

CO_2. Vegetation doesn't affect CH_4 concentration at the ground, therefore the CH_4 measured at PDM is always lower than CRA. The lower layer of the atmosphere has a higher concentration of CH_4. From the graphs we can see that there is a certain evolution over the course of the day. The detailed comparison of measurements and models can be found in A.3. A summary table is provided here; The difference between CAMS and GGG2020 models is much more pronounced in

	Min [ppb]	**Max [ppb]**	**Average [ppb]**
|ΔXCH4 L CAMS|	2.197	50.205	23.817
|ΔXCH4 U CAMS|	6.066	36.247	20.556
|ΔXCH4 L GGG2020|	28.824	87.207	63.192
|ΔXCH4 U GGG2020|	48.148	66.478	57.764

Table 7.5: Summary statistics of Δ XCH4 using GGG2020 and CAMS prior profiles

this case, with the latter demonstrating far worse results than the former. In general
Some of the results are showed in the plots of fig. 7.13.

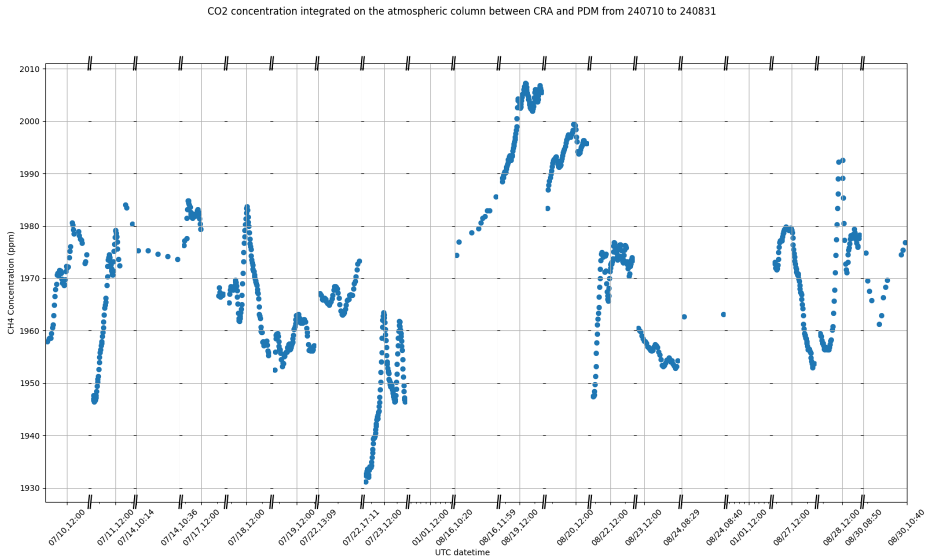


Figure 7.11: XCO_4 concentration in the atmospheric column between CRA and PDM

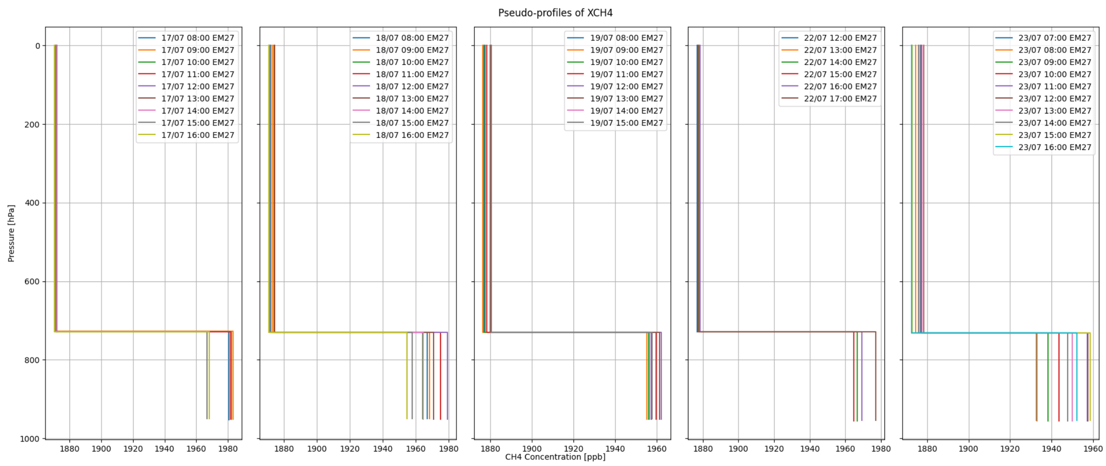


Figure 7.12: Pseudo-profiles of CH_4 obtained from measurements at CRA and PDM

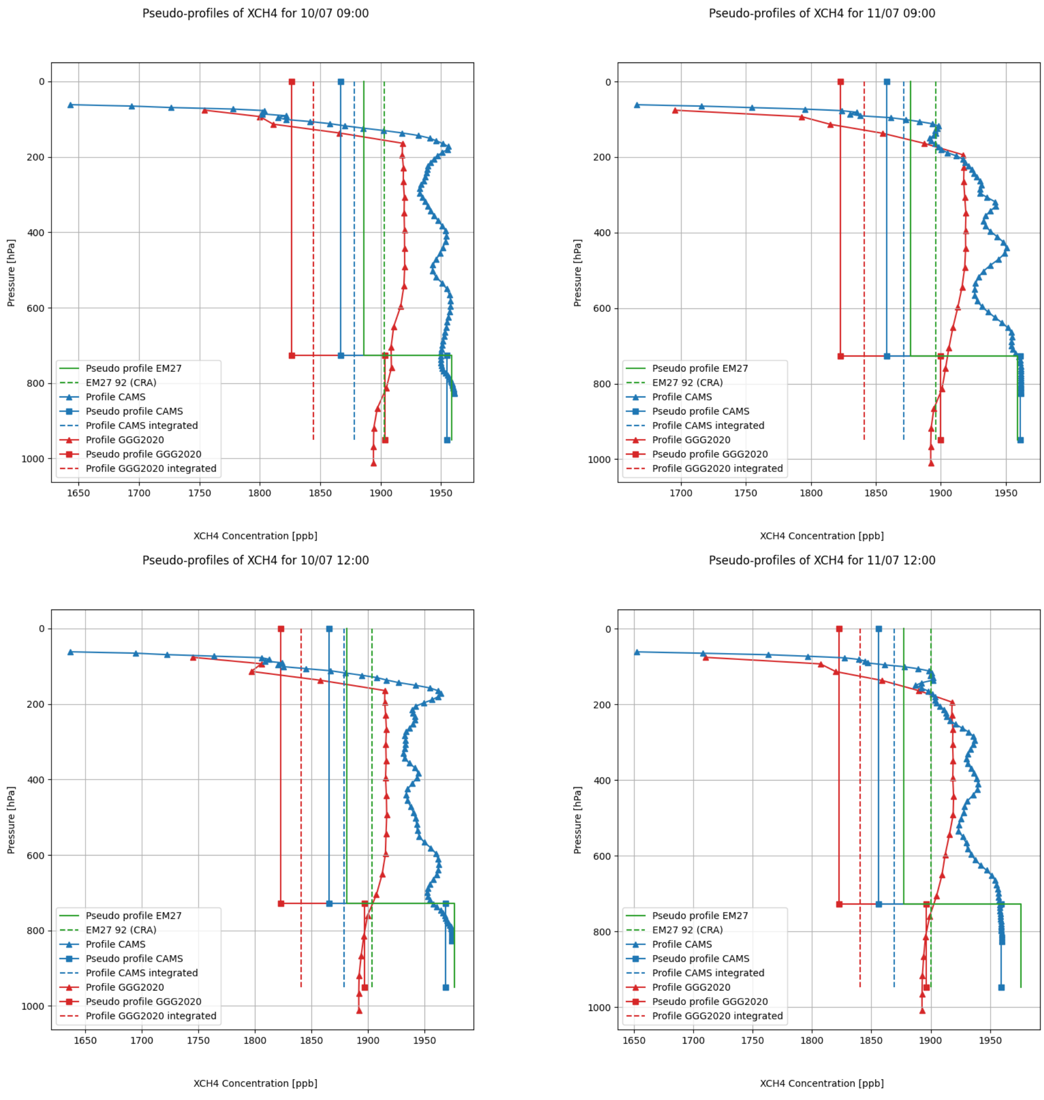

Pseudo-profiles of XCH4 for 10/07 09:00
Pseudo-profiles of XCH4 for 11/07 09:00
Pseudo-profiles of XCH4 for 10/07 12:00
Pseudo-profiles of XCH4 for 11/07 12:00
Pressure [hPa]
XCH4 Concentration [ppb]
Pseudo profile EM27
EM27 92 (CRA)
Profile CAMS
Pseudo profile CAMS
Profile CAMS integrated
Profile GGG2020
Pseudo profile GGG2020
Profile GGG2020 integrated

Figure 7.13

Chapter 8

Conclusions and perspectives

Before drawing the conclusions of my work, I'd like to express my gratitude for the welcoming and support provided by the SA division, in particular from my tutors Christel Guy and Denis Jouglet. Their help was essential during these months and without their guidance I would not have been able to get this far. I'd also like to express a special thanks to Simon Prady, who helped me out a lot in the first phases of the internship.

My internship focused on two main objectives: the first was to create the PROFFAST dashboard so that GHG concentrations could be obtained in real time during measurements with the EM27/SUN. The second was to actively participate in the summer observation campaign and take care of data processing.

8.1 PROFFAST dashboard

The dashboard is a powerful tool to quickly gain insight on the ongoing observation. Its development stopped in July when the summer campaign started because the EM27/SUN possessed by CNES (#92) was deployed on the site of CRA for the following two months. In fact, the dashboard was installed on the computer associated to the instrument but in CRA another computer, managed remotely by the LSCE. When the campaign started, the dashboard possessed the following features:

- Real time plotting of GHG concentrations.
- Real time results of the inversions on a table.
- Plot of the acquired spectra, calibrated and not.
- Plot of data coming from the CO2 sensor.

The following features could be added to make the data presented more complete:

- Plot of vertical profiles of GHG (from map files).
- Plot of data coming from the PTU sensor (pressure, temperature and humidity).

Furthermore, all the necessary scripts and files to run the dashboard have been uploaded to a GitLab repository. Possible future developments for PROFFAST dashboard involve its utilization in a broader context:

- Installing the dashboard on the computers of other laboratories using EM27/SUN for GHG measurements.
- Integrate the dashboard in the set of software that LSCE uses for the casings of EM27/SUN.
- Create a different version of the dashboard to be used for post data treatment, for example to compare the results coming from two different sets of map files (TCCON vs CAMS). I started developing this version of the dashboard under the name of "PRFdash home" but didn't continue the development to concentrate on the summer campaign.

The problem with this type of measurements is that they are heavily dependent on the weather. Therefore, automatic and continuous measurements are preferred to measurements made on a one-time basis (e.g., during satellite passes) and operated by humans. In this case, a dashboard like the one I developed would be very useful for the monitoring of the instruments and for the interpretation of the results.

8.2 Summer campaign

The summer campaign was conducted during the months of July and August with three EM27/SUN (#92, #118, #179) and provided continuous and automatic measurements on the sites of CRA and PDM. The treatment of these data showed the following results:

- The three instruments had an inter-bias between them of ≈ 0.1 ppm.
- CO_2 concentrations are consistent with annual observed trends.
- CO_2 concentration is lower in the lowest layer of the atmosphere because of the photosynthesis of vegetation.
- Confrontation of model-based vertical pseudo-profiles with actual measurements showed a lack of precision of these models to predict CO_2 concentration in the lower layer of the atmosphere.
- In general, the CAMS model proved to be more accurate than the GGG2020 model.

To push further the analysis of this data, other tools and sources of CO_2 measurements can be used. For the two months of campaign the following data was available but hasn't been used for lack of time, therefore their analysis was not included in the report.

- Wind maps that show the direction of the wind at CRA and PDM [B.1]. This data would be useful to analyze the phenomenon of air convection between the plain and the mountain, which could explain the variation in the difference between CO_2 concentrations between CRA and PDM.
- Cloud index graphs for CRA and PDM [B.2] would provide insights on cloudy days. The impact of clouds in the results of the inter-comparisons could be studied.
- In-situ CO_2 measurements provided by PICARRO analyzers at CRA and PDM [B.3] provide further data to be compared with EM27/SUN's measurements
- AirCore measurements of CO_2 coming from atmospheric balloons are very precise and accurate. During one of these measurements in the site of ASA an EM27/SUN was deployed. The comparison of the two sets of data would provide very interesting insights.

- Finally, satellite observations from OCO-2 B.4 were available for a number of days. Unfortunately, during sunny days, the satellite didn't pass exactly over our sites of interest. Nevertheless, its data is very precious to confirm the validity and precision of our measurements.

All these efforts were made to ensure the highest possible accuracy and precision during the Cal/Val phase of the MicroCarb and MERLIN missions, which is crucial for the satellites to perform as expected once in orbit. These two missions are at the forefront of Earth observation, pushing the limits of what can be achieved. Their impact on global efforts to combat climate change will be significant, and I am proud to have contributed, even for a small part, to this important CNES initiative.

Appendix A

Additional results of the summer campaign

A.1 Comparison between EM27/SUN measurements and CAMS profiles from PDM

Datetime	XCO2 L	XCO2 U	XCO2 L CAMS	XCO2 U CAMS	Δ XCO2 L	Δ XCO2 U
10/07 - 9h	417.318	423.087	421.731	420.770	-3.452	0.028
10/07 - 12h	416.988	422.797	421.435	412.508	4.479	-1.061
11/07 - 9h	421.773	423.908	423.412	425.559	-3.786	-1.378
11/07 - 12h	419.797	423.890	422.939	423.643	-3.846	-1.053
17/07 - 9h	416.094	422.161	420.733	416.495	-0.401	-0.495
17/07 - 15h	411.167	421.305	418.938	408.177	2.990	-0.932
18/07 - 9h	417.960	422.734	421.629	421.670	-3.710	-0.966
18/07 - 12h	414.416	422.888	420.928	422.678	-8.261	-0.953
18/07 - 15h	413.349	422.478	420.366	416.524	-3.175	-1.869
19/07 - 9h	417.412	421.765	420.753	420.532	-3.120	-1.152
19/07 - 12h	415.046	421.234	419.797	419.291	-4.245	-1.115
19/07 - 15h	413.582	420.697	419.047	419.056	-5.474	-1.380
22/07 - 15h	415.094	422.484	420.738	417.048	-1.953	-1.988
23/07 - 9h	418.981	422.824	421.923	418.126	0.854	-1.837
23/07 - 12h	418.407	422.789	421.764	409.666	8.741	-1.905
23/07 - 15h	417.422	422.423	421.256	410.164	7.258	-1.814
16/08 - 12h	411.491	421.201	418.905	413.422	-1.930	-1.261
19/08 - 9h	410.182	421.300	418.022	413.132	-1.108	-0.444
19/08 - 12h	410.700	421.400	417.949	407.837	4.424	-0.699
19/08 - 15h	413.349	422.478	417.462	406.767	4.728	-1.549
20/08 - 9h	410.182	421.300	418.677	413.602	-3.420	-1.431
20/08 - 12h	410.700	421.400	418.880	406.745	3.955	-1.297
22/08 - 9h	416.632	420.110	419.900	419.297	-3.268	-2.049
22/08 - 12h	412.114	419.968	418.134	414.190	-2.077	-1.546
22/08 - 15h	411.809	419.248	417.511	411.936	-0.127	-2.623
23/08 - 12h	416.493	420.252	419.373	411.932	4.561	-1.720
26/08 - 9h	414.723	420.426	419.068	417.851	-3.128	0.443

Table A.1: Comparison between measured and CAMS predicted XCO2 (PDM data)

A.2 Comparison between EM27/SUN measurements and GGG2020 profiles from CRA

	Min	Max	Average
|ΔXCO2 L|	0.127	8.741	3.524
|ΔXCO2 U|	0.028	2.623	1.300

Table A.2: Summary statistics

Datetime	XCO2 L	XCO2 U	XCO2 L GG20	XCO2 U GG20	Δ XCO2 L GG20	Δ XCO2 U GG20
10/07 - 9h	417.318	423.087	411.232	420.202	6.086	2.885
10/07 - 12h	416.988	422.797	411.464	420.375	5.524	2.422
11/07 - 9h	421.773	423.908	410.630	420.027	11.144	3.881
11/07 - 12h	419.797	423.890	409.866	420.022	9.931	3.868
17/07 - 9h	416.094	422.161	411.262	419.862	4.832	2.299
17/07 - 15h	411.167	421.305	408.177	419.792	1.483	1.513
18/07 - 9h	417.960	422.734	409.952	419.704	8.008	3.030
18/07 - 12h	414.416	422.888	409.682	419.683	4.735	3.205
18/07 - 15h	413.349	422.478	409.291	419.627	4.058	2.851
19/07 - 9h	417.412	421.765	408.966	419.269	8.446	2.496
19/07 - 12h	415.046	421.234	408.290	419.127	6.756	2.107
19/07 - 15h	413.582	420.697	409.086	419.203	5.662	1.495
22/07 - 15h	415.094	422.484	413.579	420.138	1.516	2.346
23/07 - 9h	418.981	422.824	410.709	419.006	8.271	3.817
23/07 - 12h	418.407	422.789	410.473	418.991	7.934	3.797
23/07 - 15h	417.422	422.423	410.661	419.398	6.761	3.025
16/08 - 12h	411.491	421.201	407.845	417.295	3.646	3.906
19/08 - 9h	410.182	421.300	412.056	417.999	-0.031	1.893
19/08 - 12h	410.700	421.400	412.896	418.306	-0.635	1.401
19/08 - 15h	413.349	422.478	413.300	418.508	-1.805	0.792
20/08 - 9h	410.182	421.300	412.196	418.472	-3.661	2.828
20/08 - 12h	410.700	421.400	412.863	418.225	-2.163	3.174
22/08 - 9h	416.632	420.110	411.519	417.873	5.114	2.237
22/08 - 12h	412.114	419.968	410.032	417.874	2.081	2.094
22/08 - 15h	411.809	419.248	409.306	417.918	2.503	1.331
23/08 - 12h	416.493	420.252	420.821	417.298	6.872	2.954
26/08 - 9h	414.723	420.426	419.693	417.040	4.775	3.386

Table A.3: Comparison between measured and GGG2020 predicted XCO2

A.3 Comparison between EM27/SUN measurements of CH_4 and prior models

For reasons of space, the unit of measurement has been omitted in the tables. For CH_4 is ppb (parts per billion).

A.3.1 CH_4 measurements vs CAMS profiles at CRA

Datetime	XCH4 L	XCH4 U	XCH4 L CAMS	XCH4 U CAMS	Δ XCH4 L	Δ XCH4 U
10/07 - 9h	1903.197	1866.932	1954.595	1866.932	3.991	19.236
10/07 - 12h	1903.520	1865.574	1968.947	1865.574	7.515	15.605
11/07 - 9h	1896.029	1858.192	1961.284	1858.192	-2.197	18.738
11/07 - 12h	1900.157	1856.105	1959.119	1856.105	16.816	21.125
17/07 - 9h	1898.260	1858.717	1957.876	1858.717	25.277	13.416
17/07 - 15h	1893.500	1857.388	1959.850	1857.388	6.816	13.829
18/07 - 9h	1895.196	1853.481	1944.803	1853.481	23.686	19.634
18/07 - 12h	1896.360	1855.199	1944.265	1855.199	34.897	16.235
19/07 - 9h	1894.728	1870.315	1947.793	1870.315	7.469	6.066
19/07 - 12h	1899.058	1871.220	1947.210	1871.220	15.080	8.709
22/07 - 15h	1897.881	1857.255	1934.068	1857.255	30.604	19.959
23/07 - 9h	1891.146	1850.190	1925.667	1850.190	12.519	26.564
23/07 - 12h	1896.906	1851.821	1932.189	1851.821	25.418	26.564
16/08 - 12h	1908.487	1859.475	1943.625	1859.475	45.634	24.002
19/08 - 9h	1907.428	1860.020	1957.369	1860.020	34.097	21.200

19/08 - 12h	1914.810	1862.994	1954.302	1862.994	50.205	24.085
22/08 - 9h	1898.893	1849.593	1939.308	1849.593	21.684	30.376
22/08 - 12h	1905.585	1854.925	1940.931	1854.925	33.433	29.710
22/08 - 15h	1908.130	1856.199	1944.020	1856.199	30.076	31.832
23/08 - 12h	1910.511	1860.059	1950.680	1860.059	6.396	36.247
26/08 - 9h	1906.330	1862.340	1946.405	1862.340	41.828	18.398

Table A.4: Comparison between measured and CAMS predicted XCH4

A.3.2 CH_4 measurements vs GGG2020 profiles at CRA

Datetime	**XCH4 L**	**XCH4 U**	**XCH4 L GG20**	**XCH4 U GG20**	**Δ XCH4 L**	**Δ XCH4 U**
10/07 - 9h	1903.197	1866.932	1903.679	1826.123	54.906	60.045
10/07 - 12h	1903.520	1865.574	1896.844	1823.160	79.617	58.019
11/07 - 9h	1896.029	1858.192	1899.771	1822.530	59.316	54.401
11/07 - 12h	1900.157	1856.105	1896.346	1823.006	79.589	54.225
17/07 - 9h	1898.260	1858.717	1906.524	1821.903	76.629	50.230
17/07 - 15h	1893.500	1857.388	1899.181	1820.007	67.484	51.210
18/07 - 9h	1895.196	1853.481	1902.437	1824.443	66.052	48.672
18/07 - 12h	1896.360	1855.199	1900.369	1823.286	78.793	48.148
19/07 - 9h	1894.728	1870.315	1905.829	1818.149	49.433	58.232
19/07 - 12h	1899.058	1871.220	1903.082	1819.397	59.208	60.531
22/07 - 15h	1897.881	1857.255	1905.014	1822.600	59.658	54.615
23/07 - 9h	1891.146	1850.190	1909.362	1823.892	28.824	52.861
23/07 - 12h	1896.906	1851.821	1908.049	1821.952	49.558	56.433
16/08 - 12h	1908.487	1859.475	1906.768	1827.891	82.491	55.586
19/08 - 9h	1907.428	1860.020	1917.830	1824.120	73.637	57.100
19/08 - 12h	1914.810	1862.994	1917.300	1822.149	87.207	64.929
22/08 - 9h	1898.893	1849.593	1916.733	1815.662	44.260	64.307
22/08 - 12h	1905.585	1854.925	1911.375	1818.975	62.988	65.661
22/08 - 15h	1908.130	1856.199	1906.724	1821.800	67.372	66.231
23/08 - 12h	1910.511	1860.059	1909.999	1829.828	47.078	66.478
26/08 - 9h	1906.330	1862.340	1918.704	1826.857	69.528	53.881

Table A.5: Comparison between measured and GGG2020 predicted XCH4

Appendix B

Additional data for the summer campaign

B.1 Wind charts

The following wind charts are available for the site of CRA. These measurements are made by a UHF radar and can be used to interpret the variations in XCO_2 at PDM, since winds that rise from the plain to the mountains could uniform CO_2 concentrations.

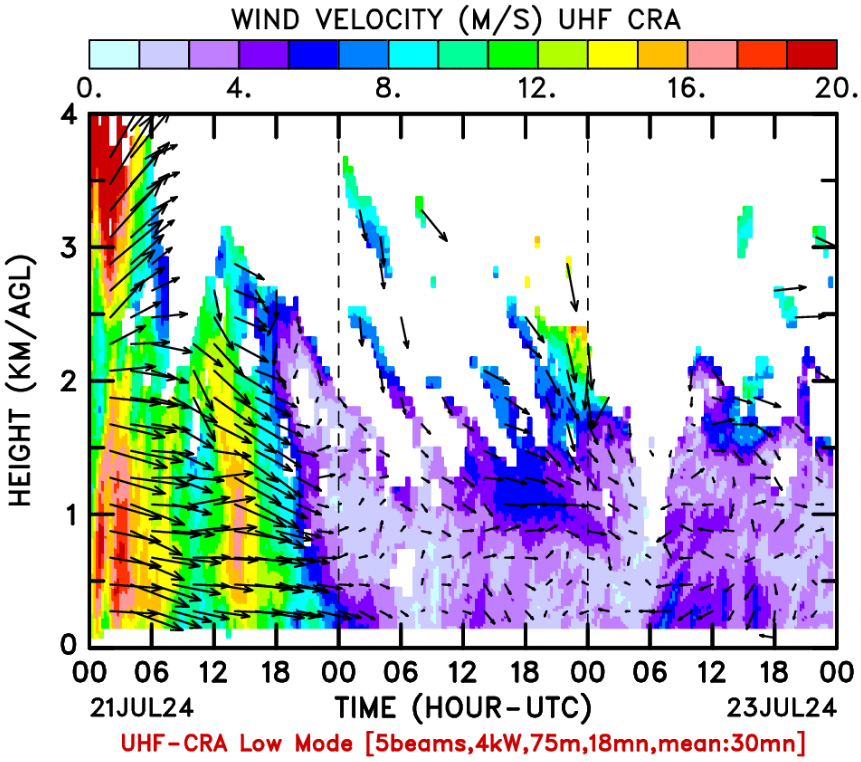


Figure B.1: Horizontal wind velocity

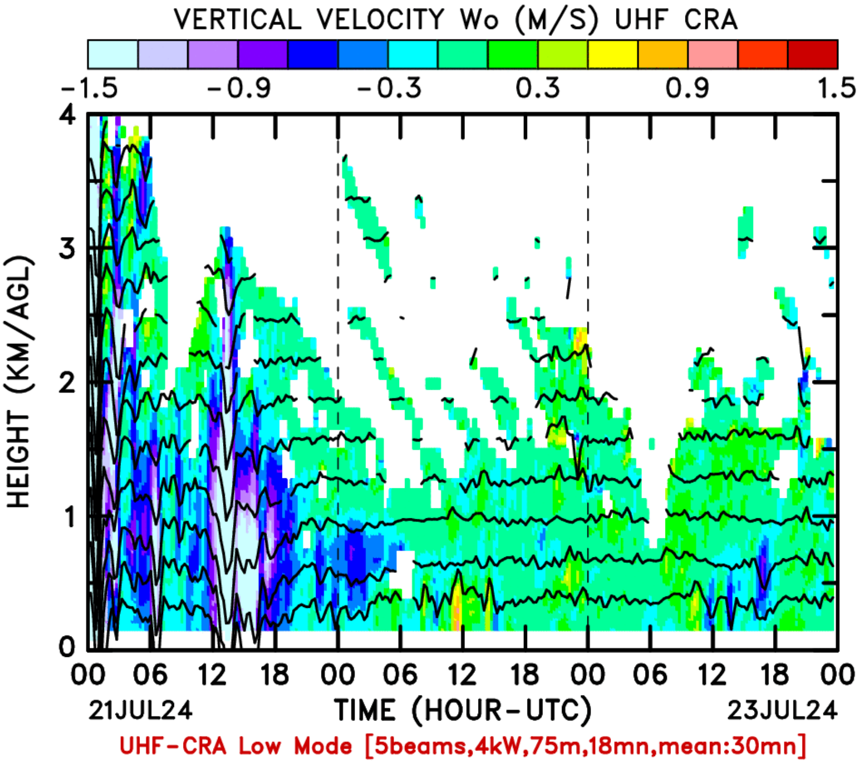


Figure B.2: Vertical wind velocity

B.2 Cloud indexes and webcams

The cloud indexes and the images of the sky can be useful to assess whether the measurements were affected by clouds and observe if there is any significant variation in the results.

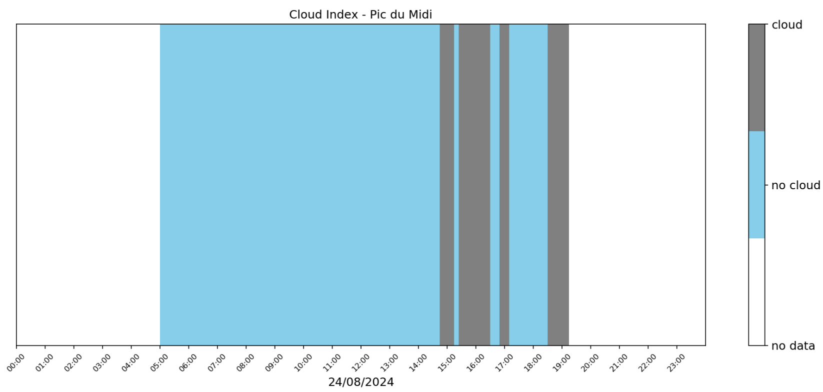


Figure B.3: Cloud index for a partially cloudy day

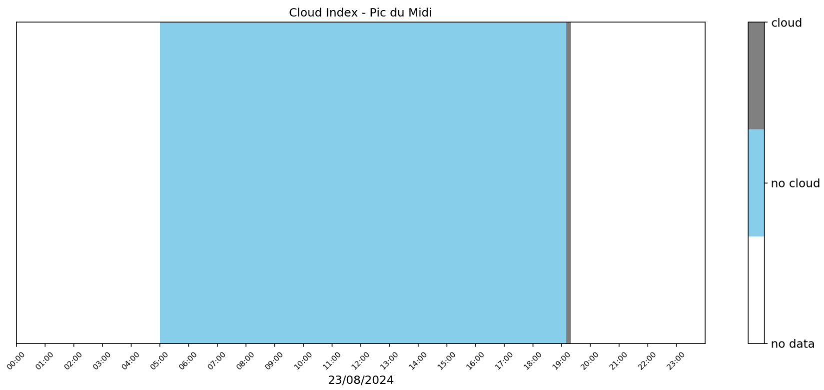
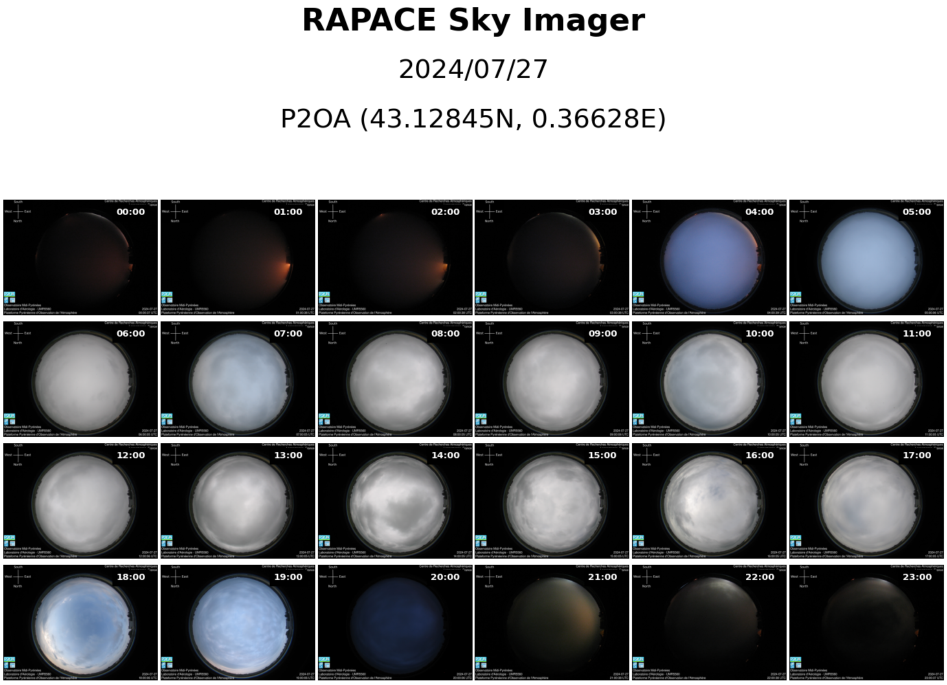


Figure B.4: Cloud index for a sunny day



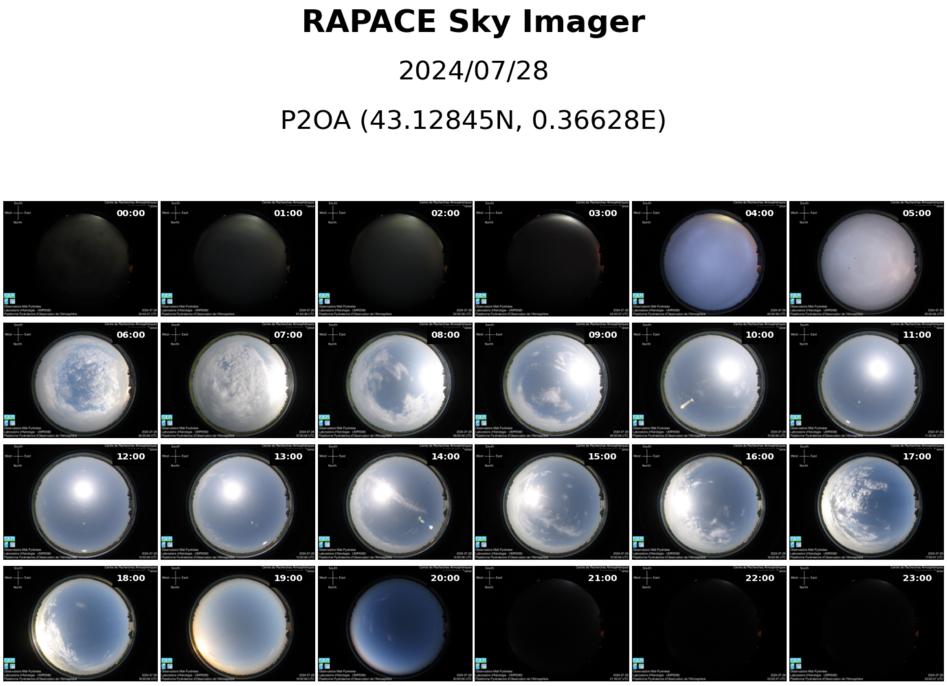


Figure B.5: Sky imager at CRA for a cloudy and sunny day

B.3 PICARRO measurements

In-situ measurements coming from PICARRO sensors can be useful to assess the validity of map prior vertical profiles of CO_2 concentration. Generally speaking, the CO_2 concentration on the ground lowers during the day to rise again during the night. In fig. B.6 it can be seen that during the night there is a slight difference in the CO_2 measured at 30m and 60m. This is again due to the vegetation: at night, vegetation releases CO2, resulting in a higher concentration at 30 meters compared to 60 meters. On the otherhand, during the day, solar-driven convection causes the CO2 concentration to become more uniform.

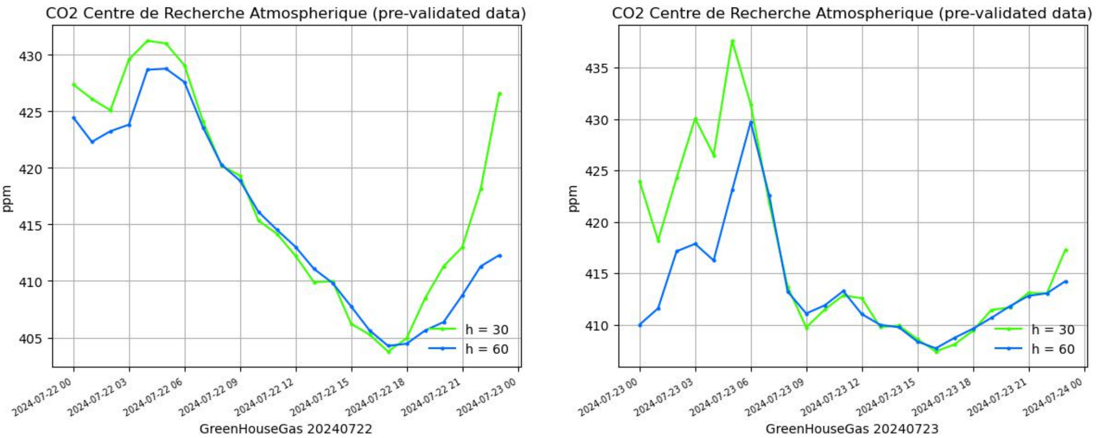


Figure B.6: PICARRO data for CRA

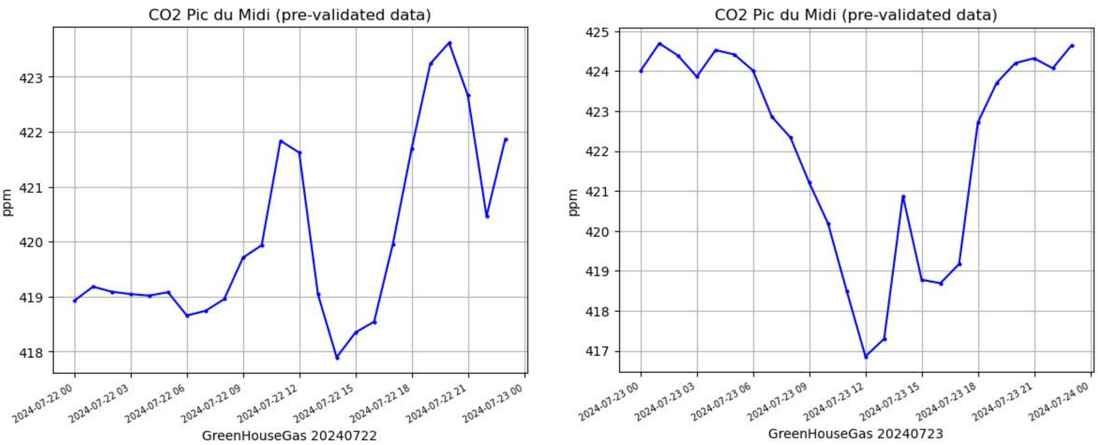


Figure B.7: PICARRO data for PDM

B.4 Satellite observations

We dispose of different satellite observations: OCO-2 provides CO_2 measurements, while S5P, with its instrument TROPOMI, provides CH_4 measurements. To have a comprehensive view of the clouds over the region of interest, Eumetsat views can be exploited.

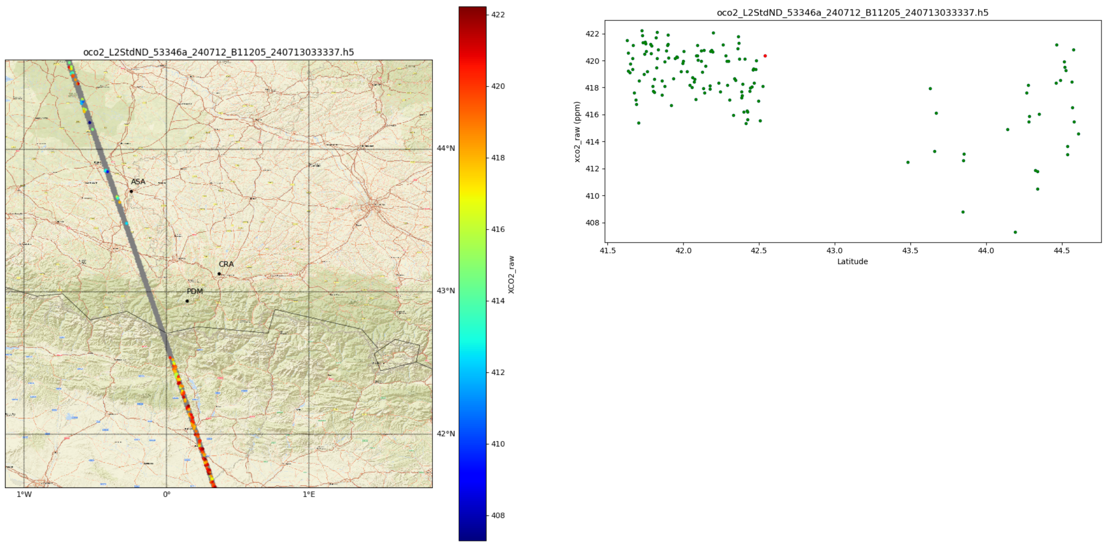


Figure B.8: OCO-2 L2 map for a passage over the sites of CRA, PDM and ASA (left) and relative CO_2 measurements

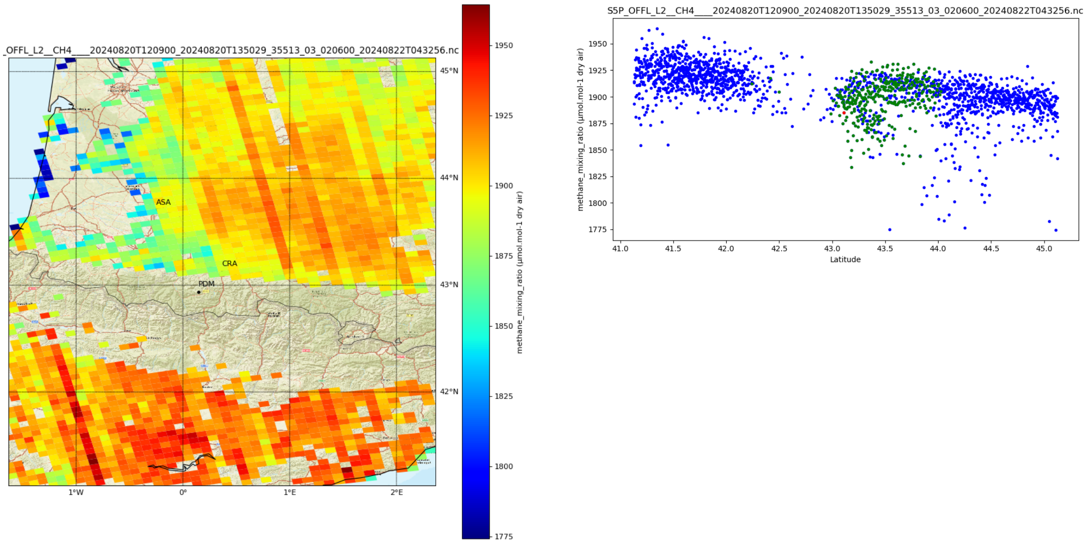


Figure B.9: TROPOMI (Sentinel 5P) L2 map for a passage over the sites of CRA, PDM and ASA (left) and relative CO_2 measurements

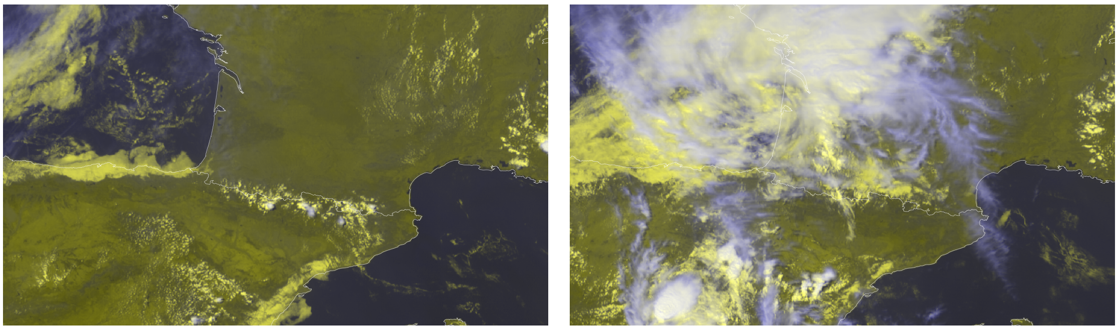


Figure B.10: Eumetsat view of the sites of CRA, PDM and ASA

Bibliography

[1] Timothy J. Wallington et al. *ENVIRONMENTAL AND ECOLOGICAL CHEMISTRY - Vol. 1 - Greenhouse gases and global warming.* 2009 (cit. on p. 13).

[2] uoregon.edu. «The Earth's radiation balance». In: () (cit. on p. 14).

[3] NOAA. *Global Monitoring Laboratory.* URL: `https://gml.noaa.gov/ccgg/` (cit. on p. 14).

[4] APEC. «An overview of GHG monitoring: objectives and technologies». In: (Apr. 2024) (cit. on p. 15).

[5] ESA. «The Permian Basin: wind direction». In: () (cit. on p. 15).

[6] URL: `https://www.eoportal.org/satellite-missions/oco-2#eop-quick-facts-section` (cit. on p. 15).

[7] URL: `https://sentiwiki.copernicus.eu/web/s5p-mission` (cit. on p. 15).

[8] URL: `https://gml.noaa.gov/ccgg/aircore/` (cit. on p. 16).

[9] Anna Karion et al. «AirCore: An Innovative Atmospheric Sampling System». In: (Nov. 2010) (cit. on p. 16).

[10] Wunch et al. «The Total Carbon Column Observing Network». In: (May 2011) (cit. on p. 16).

[11] URL: `https://en.wikipedia.org/wiki/Total_Carbon_Column_Observing_Network` (cit. on p. 17).

[12] Frey et al. «Calibration and instrumental line shape characterization of a set of portable FTIR spectrometers for detecting greenhouse gas emissions». In: (July 2015) (cit. on p. 17).

[13] URL: `https://www.imk-asf.kit.edu/english/COCCON.php` (cit. on p. 17).

[14] URL: `https://www.bruker.com/fr/products-and-solutions/infrared-and-raman/remote-sensing/em27-sun-solar-absorption-spectrometer.html` (cit. on p. 18).

[15] URL: `https://cnes.fr/projets/microcarb` (cit. on pp. 19, 20).

[16] Denis Jouglet. «MicroCarb status, validation needs, EM27 works». In: *TCCON / COCCON annual meeting* (July 2024) (cit. on pp. 20, 25).

[17] URL: `https://cnes.fr/projets/merlin` (cit. on pp. 20, 21).

[18] Deniel et al. «MERLIN and MICROCARB : Preparation of 2 space missions for CO2 and CH4». In: (Apr. 2017) (cit. on p. 20).

[19] URL: `https://help.marine.copernicus.eu/en/articles/5046705-which-levels-are-used-for-data-processing` (cit. on p. 21).

[20] Jouglet Denis. «Introduction au sondage atmosphérique». In: (Nov. 2023) (cit. on pp. 21, 25).

[21] URL: http://www.astrosurf.com/buil/us/spe2/hresol4.htm (cit. on pp. 22, 23).

[22] CNES Fanny Le Mer. «Télédétection et traitement des images: radiométrie des images». In: (Jan. 2024) (cit. on pp. 23, 24).

[23] URL: https://www.researchgate.net/figure/a-Whiskbroom-type-imaging-instrument-and-b-pushbroom-type-imaging-instrument_fig2_318669749 (cit. on p. 24).

[24] URL: https://en.wikipedia.org/wiki/Fourier-transform_infrared_spectroscopy (cit. on pp. 28, 29).

[25] Gisi et al. «XCO2-measurements with a tabletop FTS using solar absorption spectroscopy». In: (Aug. 2012) (cit. on pp. 28, 29).

[26] Nicole Montenegro Varola. «Acquisition et analyse de données pour la validation des missions satellitaires de suivi des gaz à effet de serre». In: (July 2021) (cit. on pp. 29, 30, 33, 40).

[27] Selime Gurol (CERFACS) - ISAE-Supaero. «Problèmes inverses». In: () (cit. on p. 35).

[28] A. LIGHTMAN G. RYBICKI. «Fundamentals of radiative transfer». In: () (cit. on p. 36).

[29] Hase F. «PROFFAST Ver 2.0». In: () (cit. on p. 37).

[30] Wunch et al. «The Total Carbon Column Observing Network». In: (2011) (cit. on p. 38).

[31] Bruker KIT ESA. «PROFFAST User Manual». In: () (cit. on p. 38).