**Università degli Studi di Basilicata**

DOTTORATO DI RICERCA IN
Ingegneria per l'Innovazione e lo Sviluppo Sostenibile

# An End-to-end simulator for the performance assessment of space Raman Lidar systems

FIS-01
FIS-06

Dottoranda:
**Noemi Franco**

Relatore:
**Prof. Paolo Di Girolamo**

Correlatore:
**Dott. Donato Summa**

Coordinatrice del Dottorato:
**Prof.ssa Aurelia Sole**

**XXXVI Ciclo**

# Contents

# Introduction

One of the most pressing and widely discussed topics today is climate change. This topic transcends geographical boundaries, impacting ecosystems, economies and societies around the world and requiring increasingly interdisciplinary research, involving different fields with a clear focus on atmospheric science, oceanography, geophysics, ecology. In the face of escalating environmental challenges, a deep understanding of the planet's climatic processes and the underlying mechanisms driving these transformations has never been more necessary. A comprehensive monitoring of the atmosphere is essential to increase our comprehension of the Earth system. A systematic and accurate collection of atmospheric data with a global coverage allows to study processes and to improve models and tools used in forecasting, allowing to anticipate and promptly respond to extreme events with more precision and timeliness.

Over the past decade, the Lidar technique has become more and more competitive thanks to the advancement in many technological fields, including the development of high-power stable laser, with short pulse duration and narrow linewidth; the increased sensitivity and speed of photodetectors; the advancement in optical filters' manifacturing; as well as the miniaturization of components and the reduction of production costs.

Unlike traditional passive methods, Lidars directly provide vertically resolved profiles, bypassing the complexity of inversion algorithms. The high temporal and spatial resolution allows to retrieve vertical gradients in water vapor mixing ratio and temperature profile and is crucial to determine the vertical distribution of aerosols, clouds and other atmospheric constituents. Moreover, the acquisition of measurements requires short time interval, making near real-time monitoring possible.

The need to fill gaps that are still present in our observational capabilities - especially in the low troposphere - together with new possibilities brought by improved instrument and technologies, converge to the idea of exploiting the space-borne Raman Lidar technique. The availability of high vertically resolved thermodynamic and/or optical profiles would bring an entirely new globally-sampled information,

also in under-observed parts of the planet and atmosphere. The contribution of these new data would have strong benefits in several research areas, such as radiative transfer models, land-atmosphere feedbacks, mesoscale circulations, data assimilation etc.
In this context, it's important to have a tool capable to assess the performance of a space-borne Raman lidar, to define its experimental parameters, as well as to estimate the impact of retrieved products, in order to have a useful support for the development phase and feasibility studies. For this purpose, an end-to-end simulator has been developed.
The simulator allows to simulate the lidar signals, taking into account all the experimental and environmental conditions and then to analyze them in order to retrieve thermodynamic and optical parameters, together with their statistical and systematic uncertainties. The purpose of this thesis is to describe the procedures on which the simulator is based, providing a detailed description of the strategies used and the theoretical background.

Specifically, the first chapter describes the structure of lidar systems, with a brief overview on the main components. A section is also dedicated to two ground based Raman lidar - Concerning and Marco - designed and developed during the second year of the PhD course. The development of these systems was useful to better understand all the mechanisms that had to be simulated and also to have a reference to test the theoretical model included in the simulator. Moreover, many of the technological solutions used for the systems are a starting point for the development of the space-borne prototype.
Concerning and Marco were also used in the field campaign WaLiNeAs, during which they operated in a network of lidars located in the south of France starting from October 2022. The aim was to provide near real-time water vapor profile to be assimilated in a NWP model, in order to improve the capability to predict extreme precipitation events. Concerning operated continuously until January 2023, providing very accurate measurements during both day and night and in all weather conditions. Marco is based on a micro-pulse lidar, so it is less performing than Concerning, but with more than a year of continuous operations has demonstrated the capability to provide useful products with an ultra compact and economic design.

The second chapter is focused on the end-to-end simulator. The model is divided into two separate modules: the forward module estimates the lidar signals in terms of detected photons, by simulating the propagation of the laser beam in the atmosphere, its interaction with the atmospheric constituents (molecules and particles) and the behavior of all the devices included in the receiving system. The

structure of the simulator allows to consider both space-borne and ground-based lidar systems and to specify all the experimental parameters, so that sensitivity studies can be easily conducted.
Specifically, it is possible to simulate the elastic signal, i.e., the amount of radiation elastically backscattered by atmospheric constituents at the same wavelength of the incident radiation, the Raman $N_2$ and $H_2O$ roto-vibrational signals and pure-rotational $N_2 - O_2$ signals. Simulated signals include a background term that depends on the solar radiation received together with the backscattered signal. An improved theoretical model is proposed to better evaluate this contribution with sun zenith angles near 90°, values with which traditional models, based on parallel-plane approximation, fail. The improved algorithm is particularly suited to evaluate the performance of space-borne lidar on dawn-dusk orbits. To simulate the statistical fluctuation of the number of photons typical of the shot-noise, the signals are perturbed using the Poisson statistics.
Finally, a section is dedicated to simulation in presence of clouds. Clouds contribute attenuating the signal, but also increasing the background due to cloud reflectivity. Moreover, in case of discontinuous clouds, the simulation consider the capability of the lidar to exploit shots between clouds, using weights linked to cloud distribution and the combination of possible paths.
The second part of the chapter is dedicated to the retrieval module, used to analyze the lidar signals in order to retrieve vertically resolved profiles of water vapor mixing ratio, temperature, backscatter and extinction coefficient, using consolidated techniques, a description of which is provided. The end-to-end structure allows to estimate both the statistical and systematic uncertainties, by comparing the retrieved profiles with the data used as input.
The third chapter shows the results used in two ongoing space lidar projects - Atlas and Caligola - obtained through the simulator. Atlas is a space-borne Raman lidar able to measure atmospheric temperature and water vapor mixing ratio with high temporal and spatial resolution. The mission concept was proposed in the frame of "Earth Explorer-11 Mission Ideas" and was also submitted in an improved version to the call "Earth Explorer-12". The simulator was used in this context to verify the performance of the lidar in different conditions, e.g., changing some experimental parameters, in different illumination conditions, using different data as input. The simulation were performed considering both atmospheric models and data extracted from the NASA's GEOS-5 dataset in order to evaluate the performance along multiple orbits around the Earth.
On the other hand, Caligola is a three-wavelength Raman lidar with mission founded by the Italian Space Agency (ASI), with objectives that includes both atmospheric and oceanic observations. The project involves the collaboration of several organization: the scientific studies are lead by the University of Basilicata

together with ISMAR-CNR, technological feasibility studies are on-going at the Leonardo S.p.A and lately NASA has showed a strong interest in the mission and has initiated a pre-formulation study at Langley with the possibility to be maintained within the Earth Systematic Mission Program Office at NASA Goddard Space Flight Center. Within this project, a strong cooperation with Leonardo required several simulations to define the experimental setup of Caligola, including the selection of channels, sensitivity studies for the choice of the field-of-view, considerations about the expected performance depending on the local passage time etc.

The end-to-end simulator is therefore a powerful instrument that can be used in different contexts and for different purposes, from conceptual design to verification and validation. It has certainly been a useful tool for missions and projects that, if achieved, would make a contribution to research and knowledge.

# Chapter 1

# The Lidar system

## 1.1 Configuration of a lidar system

A lidar (LIght Detection And Ranging) is an instrument used for remote monitoring, which exploit the interaction of radiation with the atmospheric components to derive thermodynamic, optical and compositional properties of the atmosphere [1].
In its simplest configuration, a Lidar consists of a laser source and a receiver (Figure 1). As the laser pulses propagate through the atmosphere, part of their energy is backscattered, elastically or inelastically, to the instrument by particles and molecules. The backscattered radiation is collected by a telescope, detected by sensors and sampled in a vertically resolved signal. Since each different species produces a specific frequency shift, the Raman lidar technique exploits the spectrally resolved pure-rotational and roto-vibrational Raman scattering phenomena, to measure a wide range of compositional properties. The receiver, in fact, includes a number of spectrally separated channels, each one tuned on a specific wavelength maximizing the backscattering from the examined atmospheric species.
A brief description of the lidar system's components of a Lidar is here provided.

### 1.1.1 Laser source

The laser sources used in lidar application are usually selected in order to have high power. Solid-state laser such as Nd:YAG, have single pulse energy of 0.1-1 J with pulse duration of 10-30 ns, with an emitted fundamental wavelength of 1064 nm, which can be doubled (532 nm) or tripled (354.7 nm). The amount of energy

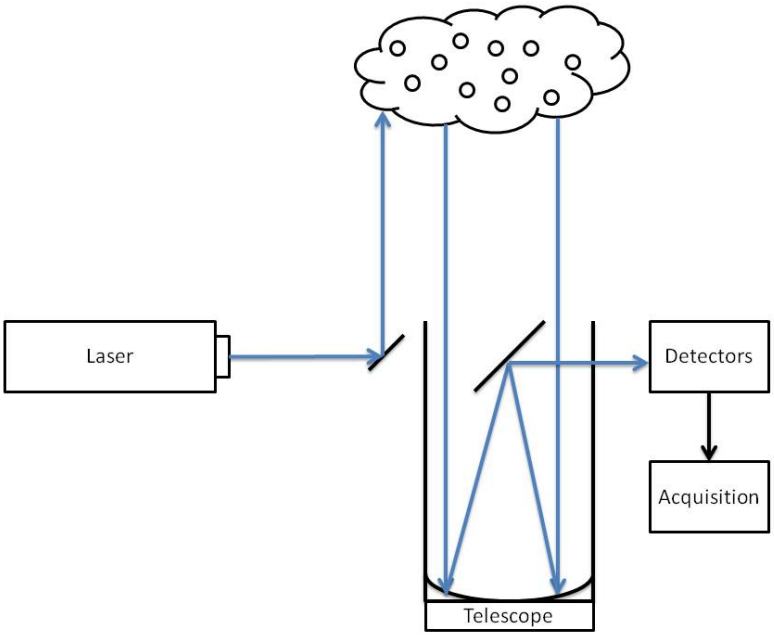



Figure 1: Schematic configuration of a Lidar system

elastically or anelastically scattered by atmospheric molecules is inversely proportional to the fourth power of the wavelength, therefore the scattering is stronger at visible and ultraviolet wavelength than in the infrared [2]. For measurement of atmospheric temperature, it is also required a good stability in frequency. In fact, the retrieval of atmospheric temperature is based on the capability to select specific rotational lines, this capability being dependent on the spectral stability of on the stimulating wavelength. A drift in the emitted frequency causes a shift of the rotational lines and a variation in intensity of the collected signals which produces a systematic error. High frequency stability are obtained with laser equipped with "injection-seeding", a device allowing to obtain a high monochromatic laser emission, based on on the selection of a single longitudinal mode. This choice allows to obtain a laser line of width (FWHM) smaller than 0.05 pm, together with a frequency stability of the order of 1 pm.
A laser beam is characterized by an intrinsic divergence due to diffraction which is proportional to the wavelength and inversely to the beam diameter. Keeping the beam divergence small is important in order to consider small field-of-view and reduce the collected environmental background (primarily associated with solar and lunar illumination). For this purpose, the laser beam can be expanded through a beam-expander.
Beam expanders are devices made of two lenses with positive or negative focal

lengths (Keplerian or Galileian configuration respectively), able to increase the diameter of a collimated input beam to a larger collimated output beam. They are therefore characterized by the magnifying power, that is the ratio between input and output divergence or, analogously, between output and input beam diameter. The expanded laser beam is than deflected towards the zenith through a mirror with high reflectivity and good damage threshold, to withstand the laser power density.

### 1.1.2 Photon collection and detection

The telescope has the role to collect the lidar signals, therefore high imaging capabilities are not strictly required and an optical quality of the telescope primary mirror corresponding to a surface wave-front error of the order of $\lambda$ RMS is sufficient. Receiver telescopes have usually diameters of 0.5-1 m and f-number, i.e., the ratio between focal length and aperture diameter, between f/3 and f/10. Different telescope configurations can be used: among others, Cassegrainian, Newtonian or Dobsonian type. It is possible to consider a coaxial setup, with the laser beam on the telescope optical axis, or biaxial setup, with laser beam off the optical axis and possibly tilted against it. The field-of-view is determined by the field-stop in the focal point of the telescope and is usually slightly larger than the beam divergence, in order to reduce the collected background but without losing the stability of the laser beam within the FOV [3].
The collected signal is then divided through the various channels using dichroic beamsplitters, optical devices that reflect specific wavelengths and transmit a different wavelength range. Beam splitters can be polarizing or non-polarizing, depending if they split the light into reflected S-polarized and transmitted P-polarized beams or if they maintain the original polarization state.
Spectral selection can be obtained using interferential filters, which transmit the wavelength within a specific bandwith, and suppress the wavelength outside the filter pass-band spectral region.
The photon detection is usually based on photodiodes or photomultipliers, devices that convert incident photons into electrical signals. High quantum efficiency and low noise are required.

### 1.1.3 Data Acquisition

The signal are then recorded as a function of time: the data acquisition starts when the laser pulse enters the atmosphere, using a trigger signal that senses a fraction of the outgoing pulse, and ends after a time interval corresponding to the range

resolution. For example a 50 ns acquisition corresponds to a vertical resolution of 7.5 m. The electric signal can be sampled using both photon-counting or analog detection. In photon-counting, individual photons are counted, so this technique is optimized for low level light intensities, i.e., to detect the signal at high altitude, but at higher levels this approach results in non linear signal responses. To increase the dynamic range, it is possible to use the analog detection, in which the signal is amplified and digitized by a 16-Bit 40/80 MHz analog-to-digital converter [4]. The acquired signals are therefore stored on a computer and analyzed applying consolidated techniques. The algorithm used to retrieve physical properties from lidar signals are explained in section 2.6.

## 1.2 Experimental development

The research activities included, in addition to a main part of theoretical research, also a part dedicated to experimental development, which led to the design and realization of two new and innovative ground-based Raman lidars.
The lidar CONCERNING (COmpact RamaN lidar for atmospheric CO2 and thERmodyNamic profilING) was developed thanks to italian funding FISR (Fondo Integrativo Speciale per la Ricerca) and is a compact Raman lidar capable, at present, to realize measurements of water vapor mixing ratio, temperature, backscatter and extinction profiles with high temporal and spatial resolution. Experimental measurements of $CO_2$ are currently ongoing and will be subject of study for the following months.
The second lidar is MARCO (Micropulse Atmospheric Optical Radar for Climate Observations, in honor of Prof. Marco Cacciani), the first micro-pulse lidar able to provide water vapor mixing ratio, backscatter and extinction coefficients profiles and, in the next future, also temperature.
Both lidars have very compact configuration and limited dimensions that make them easily transportable. Moreover, they are completely autonomous, operable and monitored from remote and capable to provide real time measurements of a variety of geophysical parameters. A brief description of the two systems, together with few examples of their measurements is provided in the following. The specifications of the two systems are listed in Table 1.

### 1.2.1 Concerning

Concerning is a Raman lidar that can operate at three wavelengths (354.7, 532.15 and 1064.1 nm) and potentially equipped with twelve channels, eight of these currently fully operational: the elastic signal at 354.7 nm, together with parallel and

cross polarization, the roto-vibrational $N_2$ and $CO_2$ Raman signals and the rotational Raman $N_2 - O_2$ signals at low and high quantum number J. Other four channels are dedicated to the total and parallel elastic signal at 532 nm, the total signal at 1064 nm and the roto-vibrational $N_2$ signal at 607.4 nm and are used when the system is operated at the three wavelengths simultaneously.
The laser source, a customized version of Merion MW 7-100 produced by Lumibird, is a diode-pumped Nd:YAG with high frequency stability ($<$ 5 pm) and narrow laser line ($< 0.005$ cm$^{-1}$). The latter property is obtained through injection seeding through a diode laser. The single single pulse energy is 110 mJ at 355 nm, that operated at a repetition rate of 100 Hz gives an average power of 11 W.
The laser beam is expanded by a reflective Cassegrain beam expander optimized for the three wavelengths and produced by Trioptics, that reduces the beam divergence from 0.39 mrad to 0.12 mrad. The laser is deflected vertically by a motorized mirror that can be controlled by a computer, allowing for boresight alignments, i.e., the co-alignment of the laser bems and the telescope field of view. The signal is then collected by a Dobsonian telescope with a 500 mm diameter aperture and focal length f/3.6.
The optical layout includes several beam splitters chosen to efficiently divide the collected signal (Figure 2). For the spectral selection, accurate sensitivity studies led to the selection of specific interferential filters with narrow bandwidths (0.1-0.3 nm) acquired by Alluxa.
The signal detection is made by photomultipliers with high quantum efficiency ($>$30 % in the UV) and high gain ($\approx 10^6$) produced by Hamamatsu. The signal is then acquired in both photon-counting (800 MHz) and analog detection (16 bit, 40MHz) using a Transient Recorder produced by Licel.
The entire system is located inside a air-conditioned cabinet of 2600 x 1340 x 2265 mm, equipped with a quartz window, which allows to operate in any weather condition. The system is therefore easily transportable and suitable for field campaigns.

### 1.2.2 Marco

Marco [5] is an ultra-compact micropulse lidar, specifically designed to have high mobility without losing performance. In fact the entire system, placed inside a cabinet with similar characteristics of the one hosting the system Concerning, is only 135 x 134 x 173 cm. The fiber optic micro-pulse laser operated at 354.7 nm has a single pulse energy of 250$\mu$J and a repetition rate of 20 kHz, corresponding to $\approx$ 5 W in the UV. To compensate the background contribution due to the high repetition rate, a Ritchey-Chretien with a 406 mm aperture and focal length f/8 was used, in order to allow an acceptable reduction of the field-of-view (0.31 mrad,

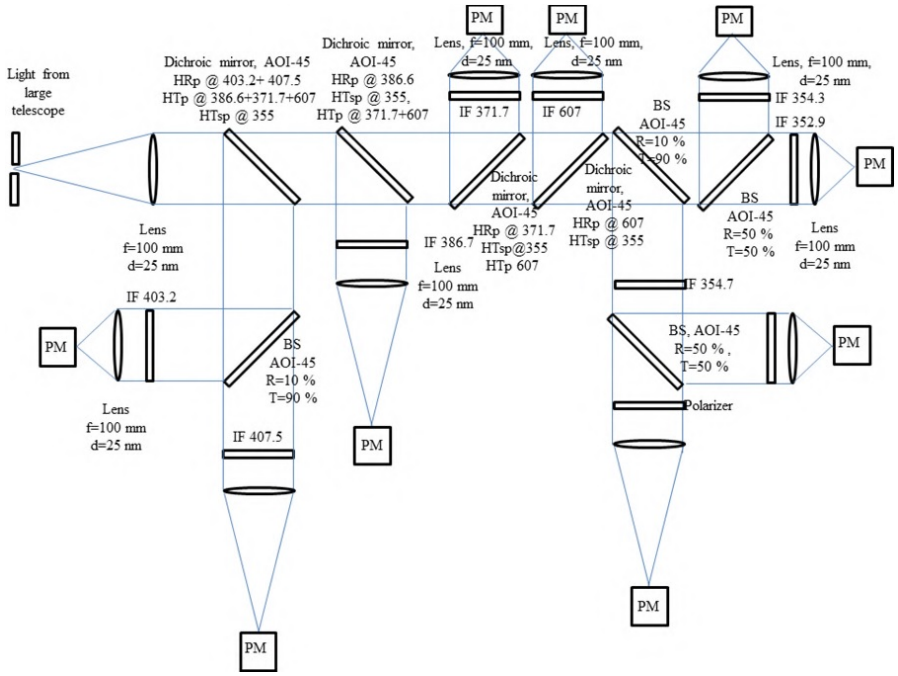



Figure 2: Optical layout of Concerning

with a beam divergence of 0.30 mrad).
At present, five channels are operative: the elastic backscatter signal at 354.7 nm, together with the parallel and cross polarizations and the two roto-vibrational signal of $N_2$ and $H_2O$ at 386.7 and 407.5 nm respectively. Another four channels dedicated to rotational high and low quantum number signals, the $CO_2$ roto-vibrational signal (371.7 nm) and a fluorescence backscatter signal (450-460 nm) will be added in the next future.
As in the system Concerning, the receiving system is based on interferential filters and photomultipliers for spectral selection and detection and the data acquisition is carried out in both analog and digital sampling using a Transient Recorder.

### 1.2.3 The field-campaign WaLiNeAs

The systems Concerning and Marco have taken part of the field campaign WaLiNeAs (Water vapor Lidar Network Assimilation) for the improving of heavy precipitation forecasting in southern France. The campaign, started in October 2022, with the participation of French, German, Spanish and Italian researchers to develop a lidar network capable to provide high-resolution measurements of water vapor. These profiles have to be assimilated in near real-time, with updates every

| Specification | Concerning | Marco |
|---|---|---|
| Laser source | Diode-pumped Nd:YAG<br>Merion MW7-100 | Fiber optic Nd:YAG<br>Sintec AO-V-355-Water |
| Wavelength | 355-532-1064 nm | 355 nm |
| Repetition rate | 100 Hz | 10-40 kHz |
| Single Pulse Energy | 110 mJ @355<br>160 mJ @532<br>300 mJ @1064 | 190 $\mu$J @10kHz<br>250 $\mu$J @20kHz<br>200 $mu$J @30kHz<br>140 $\mu$J @40 kHz |
| Beam diameter | 6.5 mm | 8 mm |
| Beam divergence | 0.39 mrad<br>0.12 mrad (after BEX) | 0.30 mrad |
| Telescope | Dobsonian | Ritchey-Chretien |
| Aperture | 500 mm | 406 mm |
| Focal length | f/3.6 | f/8 |
| FOV (FWHM) | 0.55 mrad | 0.31 mrad |
| Operative channels | Elastic tot 354.7<br>Elastic $\parallel$ 354.7<br>Elastic $\perp$ 354.7<br>Roto-vib Raman $H_2O$ 407.5<br>Roto-vib Raman $N_2$ 386.63<br>Roto-vib Raman $CO_2$ 371.7<br>Rot $N_2/O_2$ LoJ 354.3 nm<br>Rot $N_2/O_2$ HiJ 352.9 nm | Elastic tot 354.7<br>Elastic $\parallel$ 354.7<br>Elastic $\perp$ 354.7<br>Roto-vib Raman $H_2O$ 407.5<br>Roto-vib Raman $N_2$ 386.63 |
| Other channels | Elastic tot 532.05 nm<br>Elastic $\parallel$ 532.05 nm<br>Elastic total 1064.1 nm<br>Roto-vib $N_2$ 607.4 nm | Rot $N_2 - O_2$ LoJ 354.3 nm<br>Rot $N_2 - O_2$ HiJ 352.9 nm<br>Rot-Vib $CO_2$ 371.7 nm (?)<br>Fluorescence (?) |
| Detector | Photomultiplier tubes | Photomultiplier tubes |
| Acquisition | 16Bit 40MHz analog<br>800 MHz digital | 16Bit 40Mhz analog<br>800MHz digital |
| Dimensions | 260x134x226 cm | 135x134x173 cm |

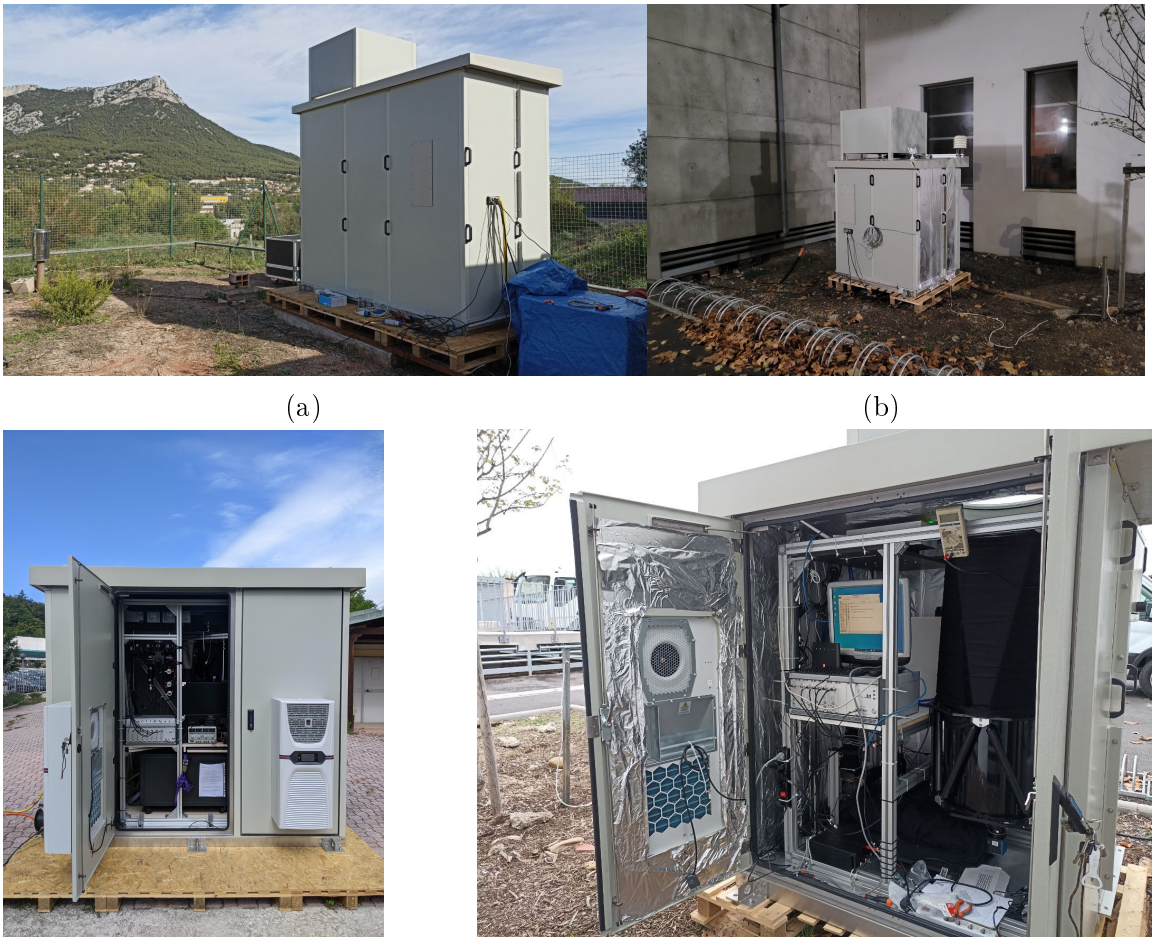
Table 1: Concerning and Marco system specifications


Figure 3: Sites of Concerning and Marco in the southern France

15-30 minutes, in the AROME-France model (French Application of Research to Operations at Mesoscale). The aim of the campaign is to fill an observational gap in the low troposphere and to enhance the capability to predict extreme heavy precipitation events which are frequent along the Mediterranean french coast. The implementation of an integrated prediction tool, coupling network measurements of WV profiles and a NWP model would help to estimate the amount, timing and locations of rainfall associated with these events up to 48 h in advance [6], which would be extremely beneficial in order to reduce the social and economical impact, as well as the threat to human lives.

The system Concerning was deployed in Toulon-La Garde (43.136°N 6.012°E, 65 m elevation) together with a Wind Lidar and started operation on 29 September 2022 - 15:35 UTC, performing 118 days of continuous measurements till 25 January 2023 - 15:00 UTC, when a laser failure forced us to stop the operations. The second system Marco was deployed in Port-Saint-Louis-du-Rhone (43.393°N 4.813°E, 5 m elevation) and started collecting measurements on 19 October 2022 - 21.02 UTC and presently it is still operational, with more than a year of continuous measurements (Figure 3-4).

Both systems are are completely automatic and designed to be operated remotely and in any weather condition. Every laser operation (i.e., alignment, switching on/off, power check) is managed through dedicated software applications (Figure 5). All the system parameters (internal temperature and humidity, laser status ecc...) are constantly monitored and any system failure is reported within 30 minutes with an email advice. Moreover, the most recent acquisitions are automatically uploaded on a dedicated website (web.unibas.it/lablidar) which



(a) (b)

(c) (d)

Figure 4: Concerning (a-c) and Marco (b-d) during the field campaign

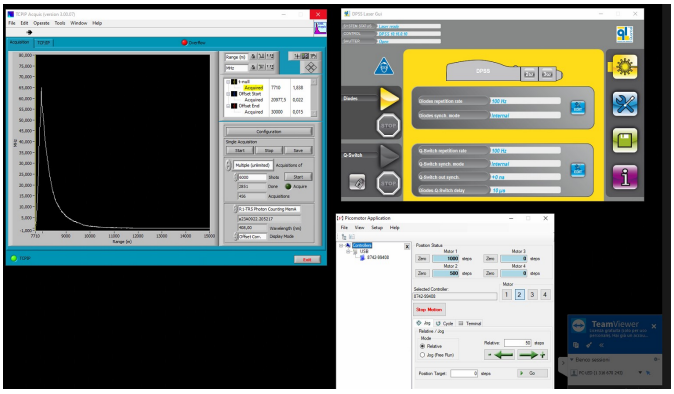


Figure 5: Remote control of Concerning with the current lidar signal, laser status and picomotor controller

shows the last 12 hours of water vapor measurements, updated every 3 hours, together with the last 7 days of range corrected elastic signal and water vapor measurements, updated every day. In this way, it is possible to intervene immediately if adjustments are required. The website is also a repository for all the past measurements. For the WaLiNeAs campaign, Concerning was operated with a raw vertical resolution of 7.5 m and temporal resolution of 10 s up to 70 km, while Marco has a vertical resolution of 15 m and temporal resolution of 3 s up to 5 km. The data are uploaded in a Network Attached Storage (NAS) and analyzed every 15 minutes for Concerning and every 30 minutes for Marco in order to provide near real-time measurements of water vapor mixing ratio. The data, averaged on 15 minutes and 30 minutes for Concerning and Marco, respectively, and with a vertical resolution of 100 m. Data are then uploaded on the Meteo-France server for assimilation.
Figure 6 shows four consecutive months of water vapor measurements obtained with Concerning. The system has shown excellent performance in both night and day time and recorded several precipitation events, including three weather watch-alert in the Gard department on 20-21 October, 31 October, 8-9 November and 14 November. Especially the second half of October shows high concentrations of water vapor, while moving towards the winter months the quantity naturally decreases.
Figure 7 shows 11 months measurements performed with Marco. The figure clearly reveals the transition between winter and summer. The system shows good performance in the night time, even if the system performance suffers a natural decline over the months due to laser power loss as well as dirt accumulated by precipitations. In the day time it is necessary to average more data to extract acceptable results up to 0.5-1 km, due to the high background.
The data collected from the two systems will be part of deeper studies in the next research activities. In addition to climatological studies, latent flow measurements are ongoing, using simultaneous data of water vapor mixing ratio from Concerning and vertical wind components from the co-located wind lidar [7]. Moreover, dedicated studies will be conducted about several dark-bands phenomena [8] that were observed with Marco, sometimes with an extension of even a couple of days (Figure 8).

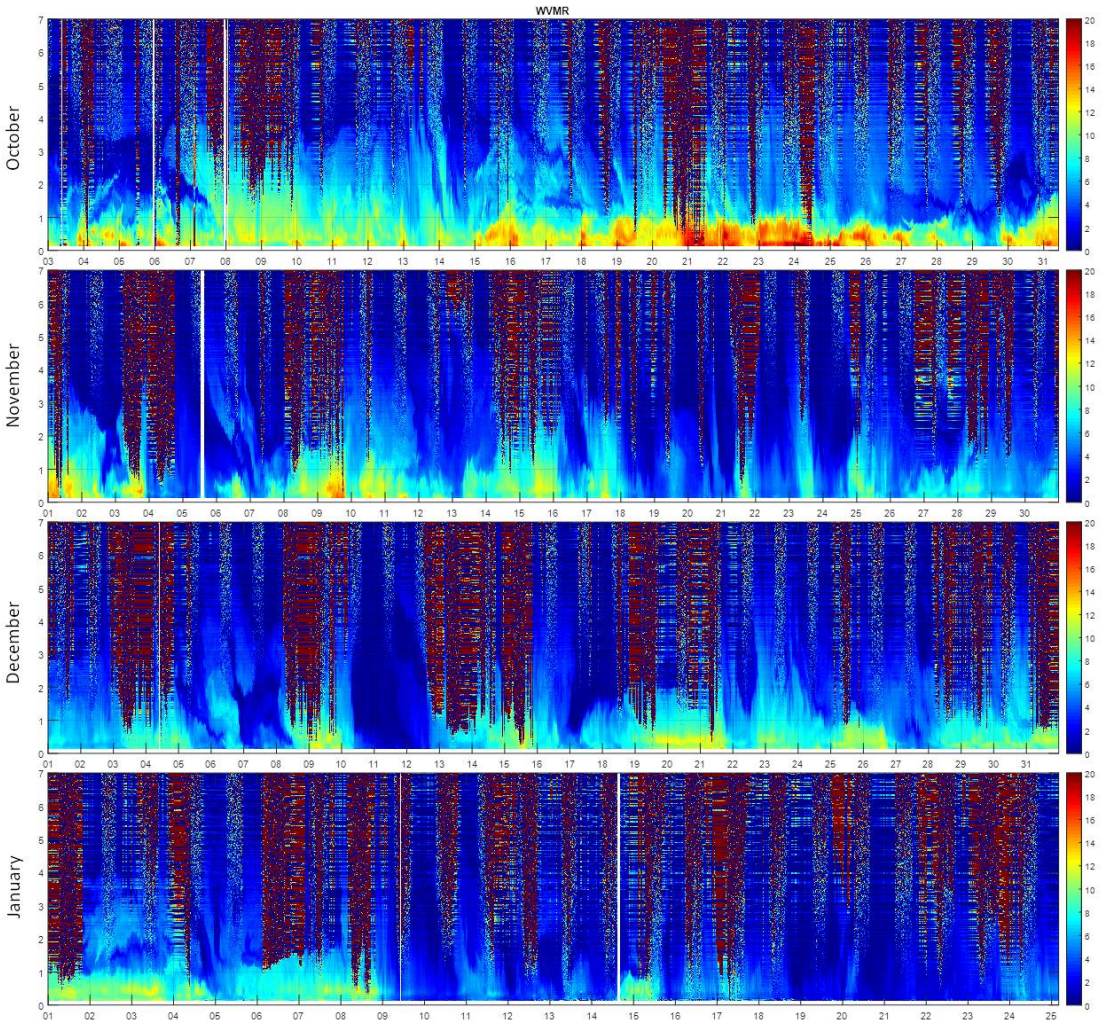



Figure 6: WVMR measurements with Concerning ($\Delta z = 7.5\,m$, $\Delta t = 15\,min$)

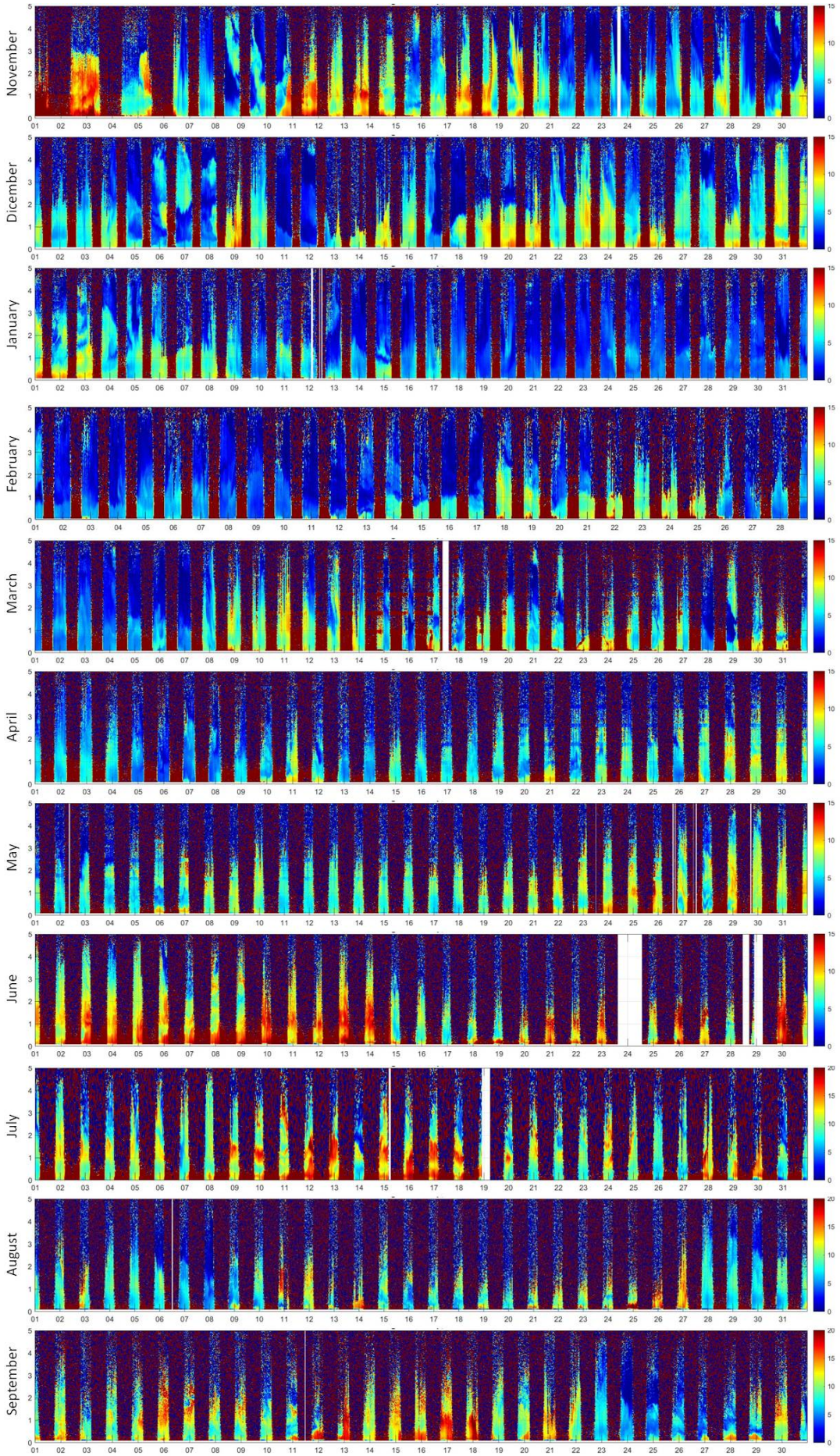
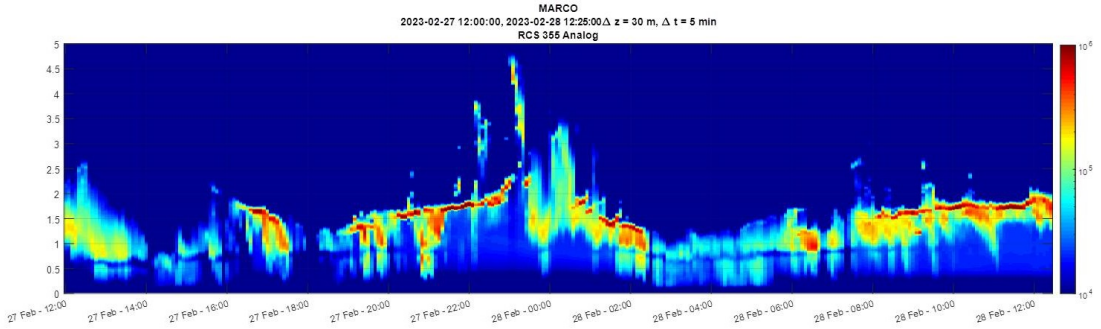
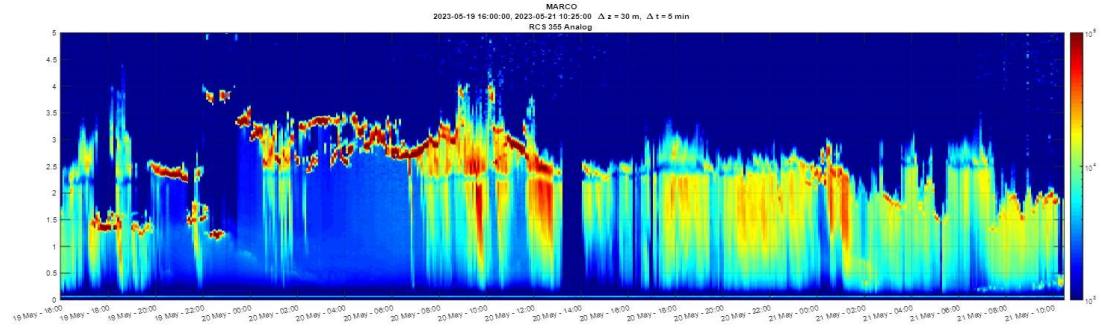



Figure 7: WVMR measurements with Marco ($\Delta z = 30\,m$, $\Delta t = 30\,min$)




(a)




(b)

Figure 8: Examples of dark band observed with Marco on (a) 27-28 February 2023 and (b) 19-21 May 2023

# Chapter 2

# The End-to-End Simulator

End-to-end performance simulators are frequently used to size remote sensors. These numerical tools are particularly important when designing a complex remote sensor, as a Raman lidar, especially when this is aimed to be deployed on a space platform. In fact the definition of the technical specifications of the different subsystems is in this case often relying on limited experimental verification and mostly depends on theoretical models and simulations. A comprehensive understanding of the characteristics of Raman lidar and an estimation of its performances plays therefore an important role in designing the system, from laser transmitters to receivers.

## 2.1 The role of simulations

During the design and development phase of a space Raman lidar, it is important to estimate the performance and the potential impact of the system, with a detailed assessment of the major statistical and systematic error sources affecting the measurements. The choice of a possible experimental setup depends on the scientific objectives that must be reached, but it is also strongly influenced by the technological solutions available. The need of minimizing the size and weight, especially for space-oriented mission, the availability of components with an acceptable TRL (Technology Readiness Level) as well as some physical or practical limitations can have a strong impact on the real feasibility of a new instrument. In most cases, a trade-offs are essential to find the best compromise between different observational requirements and scientific needs.

For example, an estimate of the signal intensity is useful for the optimization of the receiver system in terms of optical layout, available channels, damage thresholds, minimization of the background. The quality of the observational products can be evaluated based on the estimate of the statistical and systematic errors affecting the measurements. Performance can be estimated considering different technical specifications for the sub-systems included in the experimental setup, in order to verify if the scientific objectives can be reached. Moreover, the performance can be influenced by external factors, i.e., different environmental illumination conditions result in different noise amounts, whose variability must be taken into account.
For this purpose, an End-to-End simulator has been developed in order to estimate the performance of space-borne, air-borne or ground-based Raman lidars. The "end-to-end" structure, i.e., the simulation of the full chain of mechanisms from the production of the signals to the output products, is particularly effective in these studies because it allows to estimate not only the statistical uncertainties, but also the systematic error sources. Furthermore, the scalability of the experimental parameters is a functional way to verify the performance according to the characteristics of the different sub-systems.
The simulator consists of two distinct modules: the forward module, which simulates the propagation of the laser beam in the atmosphere, considering all possible interaction mechanisms with atmospheric constituents, the behavior of all devices included in the experimental setup and the background radiation. The practical graphical interface allows to specify all the experimental parameters to be considered in the simulation, e.g., the laser source specifications (emitting wavelengths, single pulse energy, repetition rate, beam divergence, etc.), the receiver specifications (telescope characteristics, receiving field-of-view, detector efficiency and gain, interference filters' transmission and blocking characteristics, etc.), temporal and spatial resolutions and for space-borne sensors also orbital parameters (satellite height and speed, equatorial crossing time, viewing geometry).
The retrieval module performs the analysis of the simulated lidar signals to determine the vertical profiles of products such as temperature, water vapour mixing ratio, backscatter and extinction coefficients. Therefore, it is based on the application of the algorithms which are typically used also in the case of real measurements to determine the atmospheric quantities.
A detailed explanation of the equations and algorithms used in the simulator is here provided.

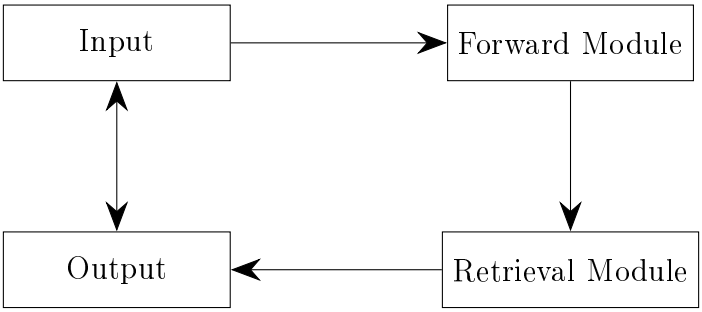



Figure 9: Block structure of the End-to-End Simulator

## 2.2 Forward Module

The forward module of the End-to-End simulator provides synthetic lidar signals calculated using the physical and optical parameters provided by the user and considering the specified experimental setup. The algorithm simulates the propagation of the beam from the laser source through the atmosphere and back to the receiver, considering the atmospheric attenuation according to the Beer-Lambert Law. In particular, the simulator can estimate the elastic backscattered signal at multiple wavelength, the roto-vibrational $H_2O$ and $N_2$ Raman signals and the pure rotational Raman signals of $O_2$ and $N_2$. The simulation takes into account an additional term for the solar background, estimated as a function of the solar zenith angle. Furthermore, the simulated signals are perturbed using the Poisson statistics in order to generate the typical shot-noise affecting Raman signals.

### 2.2.1 Input parameters

As input parameters for the initialization of the simulator, some thermodynamic and optical parameters are used; in particular pressure, temperature (T) and water vapour mixing ratio (WVMR) profiles, as well as aerosol and cloud optical parameters' profiles. The air number density can be either an input parameter or can be calculated through the ideal gas law, with the height that can be obtained from the hydrostatic equation. The simulator can also consider relative humidity as input humidity parameter, instead of water vapour mixing ratio. In this case the WVMR is calculated as:

$$\chi_{H_2O}(z) = \frac{e(z)}{p(z) - e(z)}\varepsilon \tag{2.1}$$

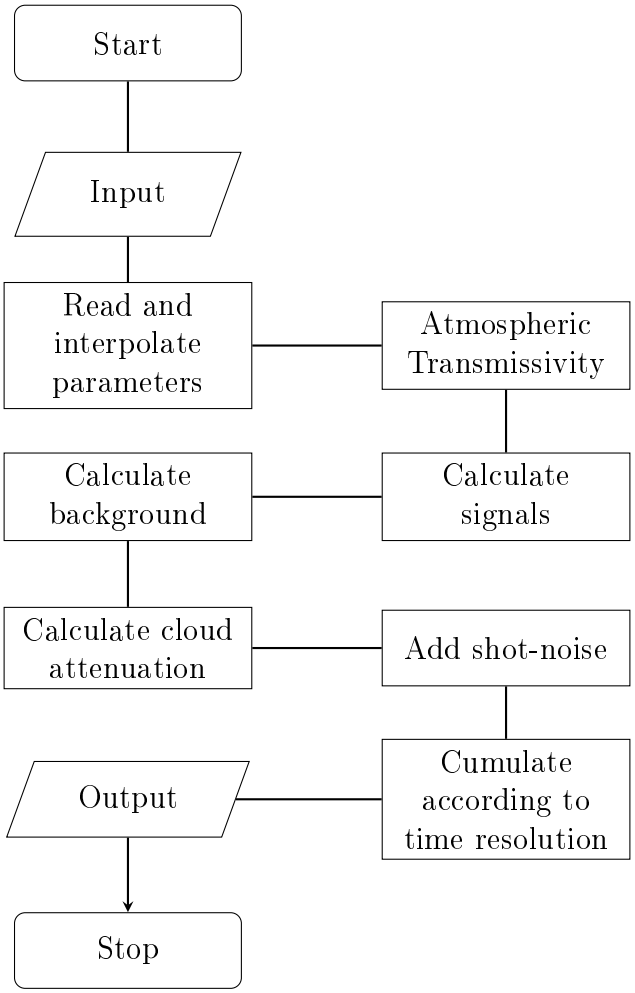



Figure 10: Flow-diagram of the forward module

where $\varepsilon = 0.622$ is the ratio of the water molar mass (18.02 kg/mol) divided by the dry air molar mass (28.97 kg/mol). The vapour pressure is obtained from the saturation vapour pressure as $e(z) = e_{sat}(z) \cdot RH(z)$. In turn, $e_{sat}$ is calculated through the Magnus equation [9]:

$$e_{sat} = \left[1.0016 + 3.15 \times 10^{-6} p(hPa) - \frac{0.074}{p(hPa)}\right] \cdot 6.112 \exp\left(\frac{17.62 \cdot t(°C)}{243.12 + t(°C)}\right) \quad (2.2)$$

Additional input parameters are the aerosol backscatter and/or extinction coefficient and some cloud optical properties, e.g., optical thickness and cloud fraction. Below 100 km, the molecular backscattering coefficient can be obtained from the Rayleigh approximation [2]:

$$\beta_{mol,\lambda}(z) = n(z)\frac{d\sigma(\pi)}{d\Omega} \approx n(z)\left(\frac{550}{\lambda}\right)^4 \cdot 5.45 \times 10^{-28} \quad (2.3)$$

where $n(z)$ is the air number density and is the considered wavelength.
The molecular extinction coefficient is obtained from the backscatter coefficient as

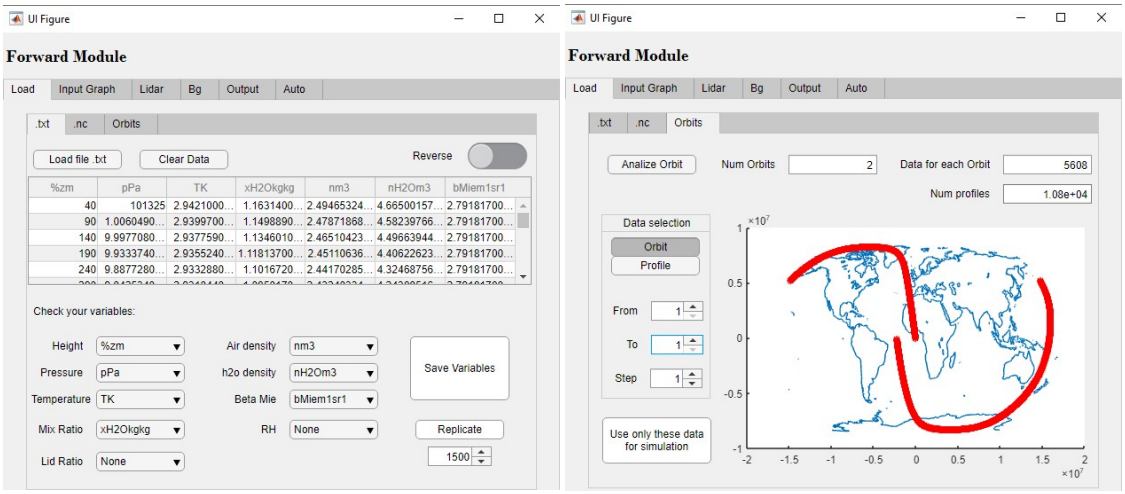
$$\alpha_{mol,\lambda}(z) = \frac{8\pi}{3}\beta_{mol,\lambda}(z) \quad (2.4)$$

while the particle extinction, if not provided, is calculated using the lidar ratio:

$$\alpha_{par,\lambda}(z) = \beta_{par,\lambda}(z) \cdot LR \quad (2.5)$$

Unless differently specified, the lidar ratio value considered in the simulator is $LR = 40$, that is a typical value for the boundary layer and free troposphere.

The input mask (Figure 11a) offers a practical way to import data in different file formats (.txt or .nc) with an automatic recognition of the parameters. The simulator can be used for both single-profile simulation, e.g., based on atmospheric models or radiosounding data, or applied to a temporal sequence of several input profiles to generate a simulated measurement sequence and, ultimately, orbital simulations.
The single-profile simulation is useful to evaluate the average climatological performance of the instrument. Simulations based on radiosounding data have been used, for example, to compare real lidar signals with simulated ones in order to evaluate the level of coupling between the telescope and the receiving channels.




(a) (b)

Figure 11: (a) Input mask showing the imported parameters and (b) orbit visualization

The time sequence simulation allows to evaluate the performance stability and is very useful in space-borne lidars simulations, in which atmospheric conditions and solar background are largely varying along the orbit.
When latitude and longitude are provided together with the input data, it's possible to visualize the spot or the orbit track on a map and, eventually, select a smaller portion of data (Figure 11b).

### 2.2.2 Sub-orbital track calculation

For space-borne simulations is sometimes useful to calculate the suborbital track of the satellite. In fact, thermodynamic parameters can be extracted from global dataset once the orbit is known in terms of latitude-longitude as a function of time. Moreover, it is useful to estimate the variability of solar background as a function of the solar zenith angle (SZA), the latter depending on the kind of orbit and time of the ascending node.
The simulator provides a tool to calculate the suborbital track of sun-synchronous orbits, i.e., nearly polar orbit with passage on any given point of the Earth's surface at the same local time.

The orbital period or Keplerian period of a spacecraft is calculated as:

$$T = 2\pi\sqrt{\frac{a^3}{\mu}} \tag{2.6}$$

where $a$ is the semi-major axis of the orbit and $\mu$ is the standard gravitational parameter ($\mu = 398600.440\,km^3/s^2$ for Earth). In a sun-synchronous orbit, the angular speed of the Earth must be equal to the nodal precession rate $\dot{\Omega}$, the latter calculated as [10]:

$$\dot{\Omega} = -\frac{3}{2}J_2\sqrt{\frac{\mu}{R^3}}\left(\frac{R}{a}\right)^{7/2}\cos(i) = -K_0 \cdot \eta^{-7/2}\cos(i) \tag{2.7}$$

where $R$ is the Earth's radius, $J_2 = 1.08263 \times 10^{-3}$ is the coefficient for the second zonal term and $i$ the orbit inclination. The orbit inclination therefore can be calculated as:

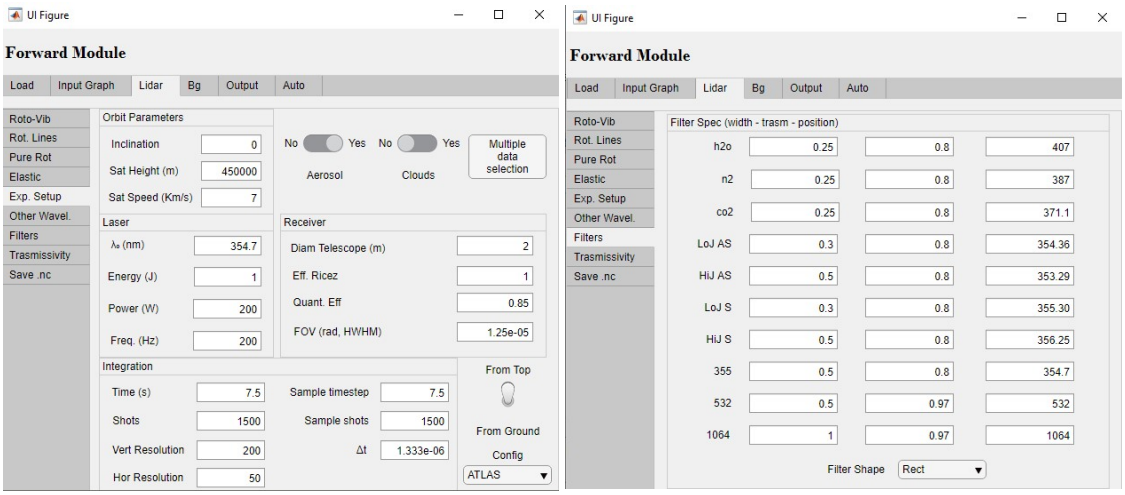
$$\cos(i) = -\frac{\dot{\Omega}_{rev}}{K_0}\eta^{7/2} \tag{2.8}$$

where the revolution Earth speed is $\dot{\Omega}_{rev} \approx 1.99099299 \times 10^{-7} rad/s$. Once the orbital elements are known (the semi-major axis $a$, the eccentricity $e$ and the eccentric anomaly $E$) the satellite's coordinates in the orbital plane are:

$$\begin{cases} X = a(\cos E - e) \\ Y = a\sqrt{1-e^2}\sin E \\ Z = 0 \end{cases} \tag{2.9}$$

To refer the coordinates to the equatorial planes, it is possible to use the Euler angles:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P(\Omega, i, \omega)\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{2.10}$$




(a) (b)

Figure 12: (a) Main experimental setup mask and (b) filters mask

where the transformation matrix $P$ depends on the longitude of ascending node $\Omega$, the argument of periapsis $\omega$ and the inclination $i$ and is:

$$P = \begin{pmatrix} \cos\Omega\cos\omega - \sin\Omega\sin\omega\cos i & -\cos\Omega\sin\omega - \sin\Omega\cos\omega\cos i & \sin\Omega\sin i \\ \sin\Omega\cos\omega + \cos\Omega\sin\omega\cos i & -\sin\Omega\sin\omega + \cos\Omega\cos\omega\cos i & -\cos\Omega\sin i \\ \sin\omega\sin i & \cos\omega\sin i & \cos i \end{pmatrix} \tag{2.11}$$

The equatorial coordinates are then used to calculate latitude and longitude $(\psi, \lambda)$ as:

$$\psi = \arcsin\left(\frac{z}{a}\right) \tag{2.12}$$

$$\lambda = 2\arctan\left(\frac{y}{\sqrt{x^2+y^2}+x}\right) - \dot{\Omega}_{rot}t \tag{2.13}$$

where the term $\dot{\Omega}_{rot}t$ takes in account the rotation of Earth, with $\dot{\Omega}_{rot} = 7.292 \times 10^{-5}\, rad/s$ being the Earth angular speed.

### 2.2.3 Experimental setup

All the experimental parameters used in the simulation can be set and changed, so that it is very straightforward to do sensitivity studies and trade-off (Figure 12).

Specifically, the experimental setup mask allows to set the laser parameters (wavelength, single-pulse energy, power and repetition rate) and the receiver's main characteristics (telescope diameter, field-of-view (FOV), detector efficiency, photomultipliers' quantum efficiency).
A simple switch button allows to choose between a ground-based or space-borne lidar simulation. In the second case, the satellite's height and speed must be specified. By default, the simulator considers a nadir viewing geometry, but it is possible to specify a different pointing inclination.
The simulation resolution is indicated in terms of vertical and temporal (or horizontal) resolution. For time-sequence simulations, the sample timestep indicates the time interval between two input profiles, so that the input parameters are kept constant in that timeframe. The number of sample shots, which depends on the acquisition time, indicates how many shots must be cumulated in a single simulated signal. For example, if we consider a $10\,s$ sample timestep with an acquisition time of $1\,s$ the results will be a time-sequence of ten profiles, based on the same input parameters. Vice versa, a sample timestep of $1s$ with an acquisition time of $10\,s$ produces a single output, obtained by cumulating ten profiles based on different input parameters.
It is possible to set an integration time different from the acquisition time: in this case, the simulator automatically average multiple simulated signals in a single output, providing also the respective standard deviation. This tool is useful to estimate the impact of statistical fluctuations and the representation error. For example, if we have a timestep of $1\,s$, an acquisition time of $0.1\,s$ and an integration time of $1\,s$, the output will be a single profile, obtained by averaging ten different profiles based on the same input, but that will show a different count of photons due to the random shot noise applied to each profile. The standard deviation calculated together with the output profile determines the statistical uncertainty of the signals and is used in the retrieval module to estimate the statistical uncertainty of the products, through error propagation.
A dedicated mask is used to specify the characteristics of the interferential filters in terms of central wavelength, filter width (FWHM) and peak transmission. If available, it's possible to import the real transmission curve of a specific filter to have a realistic estimation of the transmitted signal. Otherwise the transmission curve can be estimated considering a rectangular or a Gaussian shape (Figure 13):

$$\tau_{rect}(\lambda) = \begin{cases} \tau_{max}, & \text{if } |\lambda - \lambda_0| \leq \text{FWHM} \\ 0, & \text{if } |\lambda - \lambda_0| > \text{FWHM} \end{cases} \tag{2.14}$$

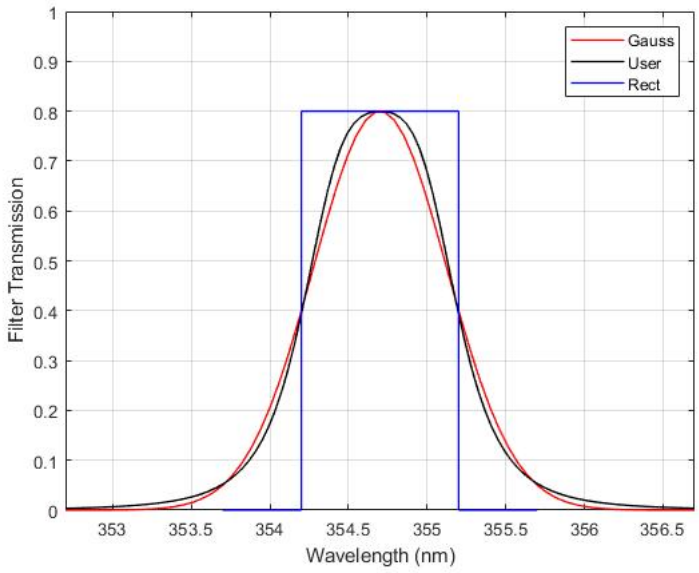



Figure 13: Transmission curve for different filter shapes: a Gaussian profile, a user defined profile and a rectangular profile

$$\tau_{gauss}(\lambda) = \tau_{max} \cdot \exp\left[ -\frac{(\lambda - \lambda_0)^2}{2} \cdot \left( \frac{2\sqrt{2\ln 2}}{\mathrm{FWHM}} \right)^2 \right] \tag{2.15}$$

### 2.2.4 Elastic and Roto-vibrational signals

Input parameters are used to estimate through the lidar equation the number of photons reaching the different receiving channels. Specifically, the elastic backscattered signal at the laser wavelength $\lambda_0$, expressed as number of photons per shot, is calculated as:

$$P_{\lambda_0}(z) = P_0 \cdot \eta\tau_{\lambda_0} \cdot \frac{c\Delta t}{2} \cdot \frac{A_{tel}}{z^2} \cdot [\beta_{mol,\lambda_0}(z) + \beta_{par,\lambda_0}(z)] \cdot T^2_{\lambda_0}(z) \tag{2.16}$$

while water vapour and nitrogen roto-vibrational Raman signals are estimated as:

$$P_{\lambda,H_2O/N_2}(z) = P_0 \cdot \eta\tau_{\lambda} \cdot \frac{c\Delta t}{2} \cdot \frac{A_{tel}}{z^2} \cdot n_{H_2O/N_2}(z) \frac{d\sigma_{H_2O/N_2}}{d\Omega} \cdot T_{\lambda_0}(z) T_{\lambda}(z) \tag{2.17}$$

where $\lambda$ is the receiver wavelength. For a laser wavelength $\lambda_0 = 354.7\,nm$, such as the one emitted by a tripled-frequency Nd:YAG laser source, the signals are collected at $\lambda_{H_2O} = 407.5\,nm$ and $\lambda_{N_2} = 386.7\,nm$.
In the equations, $P_0$ is the number of emitted photons, calculated as the single pulse energy $E_0$ divided by the photon energy $\not{h}c/\lambda_0$, $\eta$ is the overall transmission-receiving efficiency (excluding the IF transmission), $\tau_\lambda$ is the filter transmission and $A_{tel}$ is the area of the telescope.
The signal depends on the distance $z$ between the scattering volume and receiver. For a ground based lidar, $z$ corresponds to the altitude $h$, while for a space-borne system is calculated from the satellite height as $z = h_{sat} - h$. If the pointer is tilted by an angle $\alpha$, then $z = h/cos\alpha$.
The vertical resolution $c\Delta t/2 = \Delta z$ is the height of the scattering volume, calculated considering the round trip made in the time interval $\Delta t$, corresponding to the signal sampling.
The amount of backscattered radiation depends on the backscatter coefficient of molecules $\beta_{mol}(z)$ and particles $\beta_{par}(z)$. For the elastic backscatter lidar equation the molecular backscatter coefficient is calculated through the Rayleigh approximation (2.3), while the particle backscatter mainly depends on the presence of aerosol, therefore it can be imported as input parameter or estimated using a median profile, such as the ESA Aerosol refence Model of the Atmosphere (ARMA) [11].
In the roto-vibrational equation, the backscatter coefficient is calculated from the number density of considered the chemical species, multiplied by the corresponding roto-vibrational Raman differential cross section. In the simulator, the values $\sigma_{H_2O} = 7.0 \times 10^{-34}\,m^2 sr^{-1}$ and $\sigma_{N_2} = 2.5 \times 10^{-34}\,m^2 sr^{-1}$ are used [12].
The attenuation of the laser beam and the backscattered signal depends on the atmospheric transmissivity $T_\lambda(z)$, calculated from the extinction coefficient $\alpha_\lambda(z) = \alpha_{mol,\lambda}(z) + \alpha_{par,\lambda}(z)$ as:

$$T_\lambda(z) = \exp\left[-\int_0^z \alpha(t)dt\right] \text{ or } T_\lambda(z) = \exp\left[-\int_z^{z_{TOA}} \alpha(t)dt\right] \tag{2.18}$$

for ground based or space-borne simulations respectively, with TOA being the top of the atmosphere.
In the simulator, the integral is approximated as a discrete sum as:

$$T_\lambda(z) = \exp\left[-\sum_{i=1}^{N} \alpha(z_i)\Delta z\right] \tag{2.19}$$

where $\Delta z$ is the vertical resolution of the input data.
The dependence of the particle extinction coefficient on the wavelength is considered using as Angstrom coefficient $y = 1$:

$$\alpha_{par,\lambda_1}(z) = \alpha_{par,\lambda_0}(z) \cdot \left(\frac{\lambda_0}{\lambda_1}\right) \tag{2.20}$$

It is also possible to estimate the polarized elastic signals $P_{\lambda_0,\parallel}$ and $P_{\lambda_0,\perp}$, replacing the total backscatter and extinction coefficients with the corresponding polarized terms in (2.16). If co- and cross backscatter coefficient are not available, an estimation can be done using the depolarization ratio $\delta$ as:

$$\beta_{\lambda,\perp}(z) = \frac{\delta\beta_{\lambda,tot}(z)}{1+\delta} \text{ and } \beta_{\lambda,\parallel}(z) = \beta_{\lambda,tot}(z) - \beta_{\lambda,\perp}(z) \tag{2.21}$$

Typical values are $\delta_{mol} = 0.005$ and $\delta_{par} = 0.30$ [13].

### 2.2.5 Pure-rotational signals

Pure rotational Raman signals of $N_2$ and $O_2$ at high and low quantum number J (HiJ and LoJ) can be estimated as:

$$P_{HiJ/LoJ}(z) = P_0 \cdot \eta \cdot \frac{c\Delta t}{2} \cdot \frac{A_{tel}}{z^2} \cdot \sum_{i=N_2,O_2} \sum_{J_i=1}^{N} \left[\tau_{J_i} F_{J_i,T} \cdot n_i(z) \frac{d\sigma_{J_i \to J_i'}}{d\Omega}\right] T_{\lambda_0}(z) T_{HiJ/LoJ}(z) \tag{2.22}$$

The expression in brackets takes into account the sum of the intensities (m-1 sr-1) of each rotational line $I_{J_i \to J_i'}$ that arise from a thermally populated rotational states around the central exciting line, convolved with the transmission curve $\tau_{J_i}$ of the interferential filter [14].
For simple linear molecules such as $N_2$ and $O_2$ that have no electronic angular momentum coupled to the scattering, the allowed transitions in the pure rotational Raman scattering are $J_i \to J_i + 2$ (Anti-Stokes) and $J_i = J_i - 2$ (Stokes). This means that the Anti-Stokes Raman lines are scattered photons with an increased energy due to the interaction with a molecule that gives up part of its energy when returning from the virtual state to the ground state. Vice versa, for Stokes line the photons interacts with a ground state molecules that ends up in a rotational state, absorbing part of the photon energy (Figure 14).

The differential Raman cross section from the state $J$ to $J'$ can be expressed

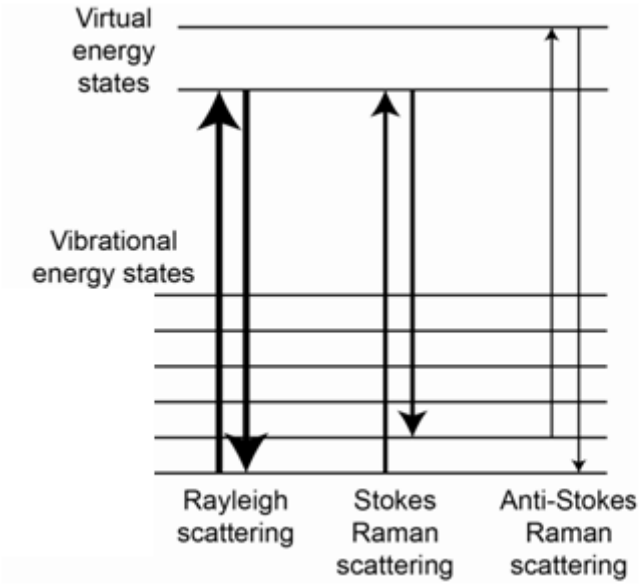



Figure 14: Energy level diagram of the elastic Rayleigh scattering and inelastic Stokes and Anti-Stokes scattering

as:

$$\frac{d\sigma_{J\to J'}}{d\Omega} = \frac{64\pi^4}{45} b_{J\to J'} (\nu_0 + \Delta\nu_{J\to J'})^4 \gamma^2 \tag{2.23}$$

where $\gamma^2$ is the square of the anisotropy of the molecular polarizability tensor, whose value is $0.509 \times 10^{-48} cm^6$ for $N_2$ and $1.27 \times 10^{-48} cm^6$ for $O_2$.
The term $b_{J\to J'}$ is the Placzek-Teller coefficient and for Stokes and Anti-Stokes lines it can be calculated as:

$$b_{S,J\to J+2} = \frac{3(J+1)(J+2)}{2(2J+1)(2J+3)} \tag{2.24}$$

and

$$b_{AS,J\to J-2} = \frac{3J(J-1)}{2(2J+1)(2J-1)} \tag{2.25}$$

The shifts of the rotational lines are constant on the frequency scale and are [15]:

$$\Delta\nu_{S,j\to J+2} - 2B_0(2J+3) + D_0[3(2J+3) + (2J+3)^3] \tag{2.26}$$

and

$$\Delta\nu_{AS,j\to J-2} + 2B_0(2J-1) - D_0[3(2J-1) + (2J-1)^3] \tag{2.27}$$

where $B_0$ is the ground state rotational distortion ($B_0 = 1.989500\,cm^{-1}$ for $N_2$ and $1.437682\,cm^{-1}$ for $O_2$), while $D_0$ is the ground state centrifugal distortion ($D_0 = 5.48 \times 10^{-6}\,cm^{-1}$ for $N_2$ and $4.85 \times 10^{-6}\,cm^{-1}$ for $O_2$).

The term $F_J(T)$ indicates the fraction of molecules in a state $J$ for a gas in thermal equilibrium at a certain temperature $T$ and depends on the rotational energy $E_J$:

$$F_J(T) = \frac{1}{Q(T)} g_J(2J+1) \exp\left(-\frac{E_{rot,J}}{k_b T}\right) \tag{2.28}$$

with

$$E_{rot,J} = [B_0 J(J+1) - D_0 J^2(J+1)^2]hc \tag{2.29}$$

In equation (2.28), $Q$ is the rotational partition function determined normalizing the sum of $F$ on all the states $J$ and can be approximated as:

$$Q(T) = \frac{(2I+1)^2 k_b T}{2hcB_0} \tag{2.30}$$

where $I_{N_2} = 1$ and $I_{O_2} = 0$ is the nuclear spin quantum number, while $g_J$ is a statistical weight factor whose values for $N_2$ are $g_J = 6$ for $J$ even and $g_J = 3$ for $J$ odd; while for $O_2$ are $g_J = 0$ for $J$ even and $g_J = 1$ for $J$ odd.

Exploiting their dependence from temperature, the atmospheric temperature profile can be determined from HiJ to LoJ signals ratio $R(T)$ using the relation:

$$R[T(z)] = \frac{P_{HiJ}(z)}{P_{LoJ}(z)} = \exp\left(a + \frac{b}{T(z)} + \frac{c}{T^2(z)}\right) \tag{2.31}$$

where $a$, $b$ and $c$ are calibration constants, that are usually obtained from a comparison of retrieved data with co-located profiles from other instruments, e.g., radiosoundings.

A theoretical estimation of the calibration constants is provided by the simulator, fitting the function $R = f(T)$. Based on the application, it is possible to choose between a 2-parameter ($c = 0$) or 3-parameter fitting function. The choice of a 3-parameter fitting function results in a more accurate calibration over a larger temperature interval than the one obtainable with the traditional 2-parameter fitting function. Thus, the 3-parameter fitting function is more suited than the 2-parameter one when dealing with the large atmospheric temperature variability

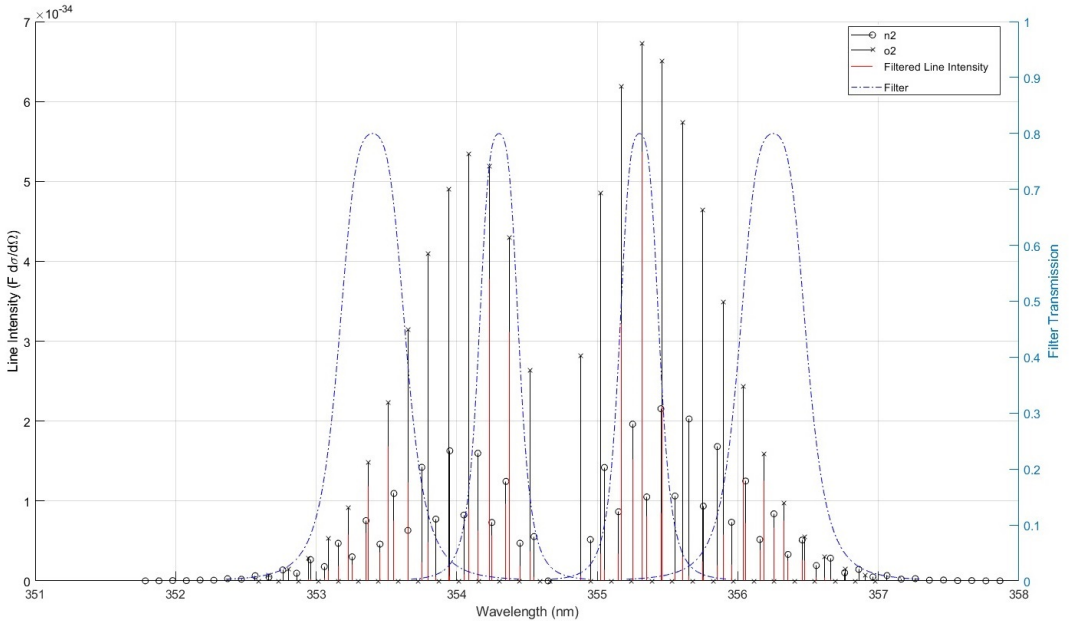



Figure 15: Simulated Raman Anti-Stokes and Stokes branches lines of $N_2$ and $O_2$ at 289 K, together with the HiJ and LoJ filters. The transmitted intensity for each line is indicated in red.

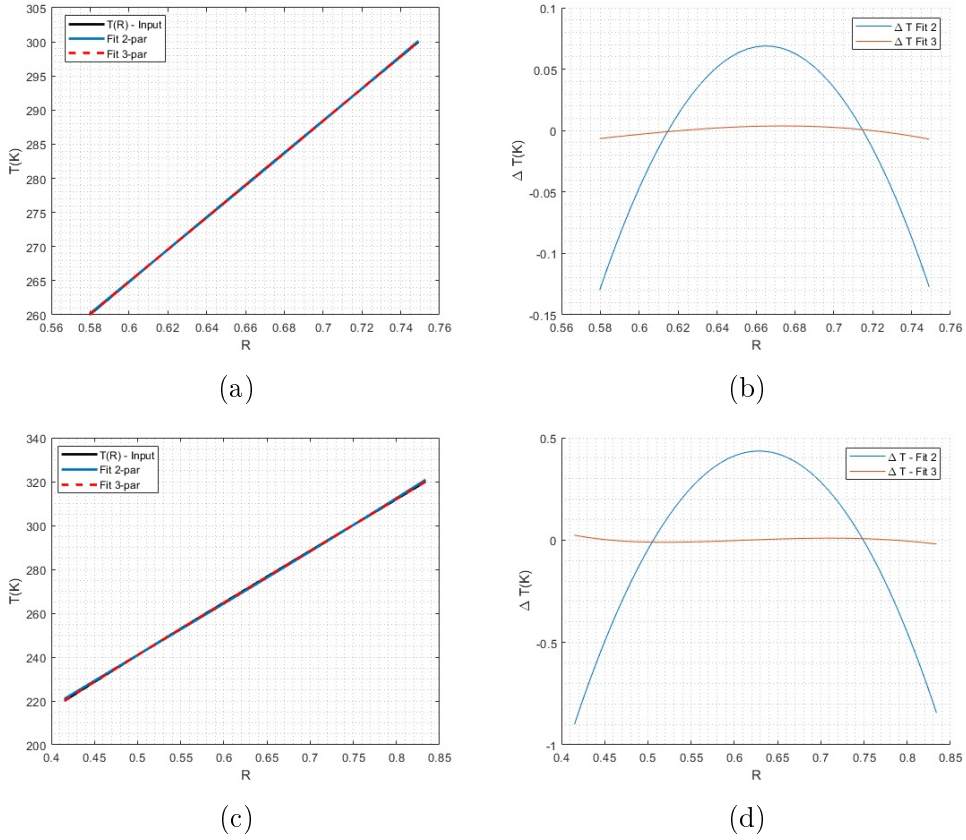
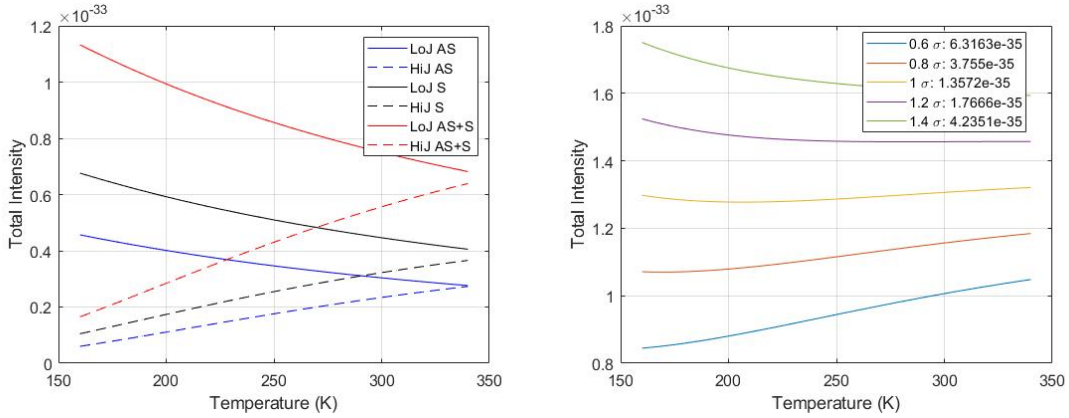



Figure 16: Fitting function in the range 260-300 K (a) and 220-320 K (c) and respective systematic errors (b-d)




(a) (b)

Figure 17: (a) Variation of the total intensity with respect to temperature of the LoJ and HiJ (Stokes and Anti-Stokes) signals and (b) variation with temperature of the weighted sum.

encountered when profiling the Earth atmosphere from space along entire orbits, while the 2-parameter fitting function provides sufficiently high accuracy when dealing with the temperature variability typically sounded by a ground-based Raman lidar system.
As shown in Figure 16, the 2-parameter fit returns enough accurate temperature measurements over the temperature range 260-300 K, with a systematic error not exceeding 0.15 K. The systematic error becomes as large as 1 K if a larger temperature variability interval (220-320 K) is considered. However, the systematic error keeps smaller than 0.02 K over this wider temperature variability interval if a 3-parameter fitting function is considered.

A weighted sum of the HiJ and LoJ signals can be also used as reference signal instead of the roto-vibrational $N_2$ signal, with the weights chosen in order to nullify or reduce to a minimum the dependence on temperature:

$$P_{ref}(z) = P_{HiJ}(z) + wP_{LoJ}(z) \tag{2.32}$$

The simulator automatically provides an estimation of the weight $w$ to be used, evaluating the variation of the reference signal with temperature (Figure 17).

## 2.3 Background Estimation

In the simulated signals, it is possible to include also a daytime background term due to solar radiation. This is calculated as a function of the solar zenith angle and solar irradiance at a certain wavelength $\lambda$, considering three distinct contributions, i.e., the contribution associated with the radiation reflected by the Earth surface and, when present, by clouds and a contribution associated with diffused atmospheric molecules [16]:

$$bg(\lambda, \theta) = bg_{surf} + bg_{clouds} + bg_{atm} \tag{2.33}$$

The total radiant flux (W) obtained from (2.33) is converted in number of photons and added to the simulated signals, considering the specifications of the receiver (total efficiency, filter transmission and width):

$$bg_{phot}(\lambda, \theta) = bg(\lambda, \theta)\Delta t \frac{\lambda}{hc} \eta \tau_{\lambda} \mathrm{FWHM} \tag{2.34}$$

where $\Delta t$ is the signal sampling duration.

### 2.3.1 Sun zenith angle calculation

The background depends mainly on the sun zenith angle $\theta$, that can be calculated from latitude, longitude $(\lambda, \varphi)$ and time (seconds from midnight) as:

$$\cos\theta = sin(\lambda)\sin(\delta) + \cos(\lambda)\cos(\delta)\cos(\mathrm{HA}) \tag{2.35}$$

where $\delta$ is the solar declination, that can be approximated as:

$$\begin{aligned} \delta(\gamma) = 0.006918 &- 0.399912\cos(\gamma) + 0.070257\sin(\gamma) \\ &- 0.006758\cos(2\gamma) + 0.000907\sin(2\gamma) \\ &- 0.002697\cos(3\gamma) + 0.00148\sin(3\gamma) \end{aligned} \tag{2.36}$$

depending on the fractional year $\gamma$:

$$\gamma = \frac{2\pi}{365} \cdot (day - 1) \tag{2.37}$$

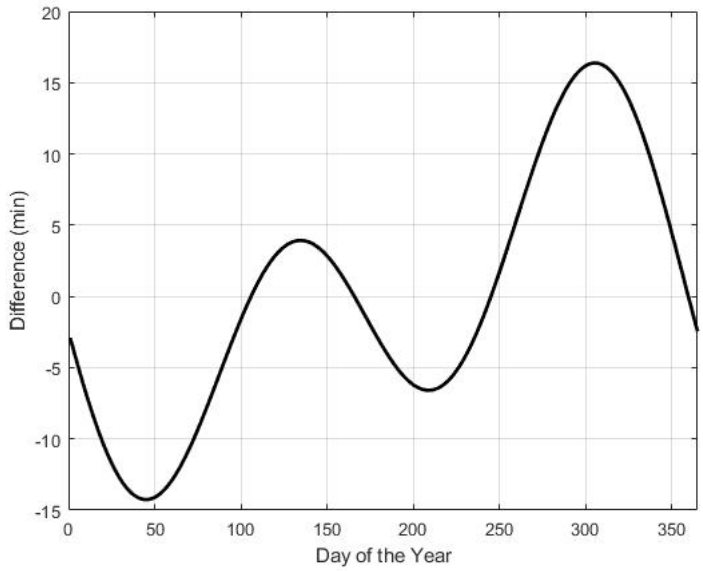



Figure 18: Equation of time function or analemma

while HA is the hour angle:

$$\mathrm{HA} = \left(\frac{tst}{4} - 180\right) \tag{2.38}$$

with $tst$ being the true solar time that is expressed in minutes as:

$$tst(t) = \frac{t}{60} + 4\phi + EoT \tag{2.39}$$

The true solar time should be corrected using the analemma described by equation of time (Figure 18), that considers the difference between apparent and mean solar time during the year:

$$\begin{aligned} EoT(\gamma) = 229.18 \cdot [0.000075 + 0.001868\cos(\gamma) - 0.032077\sin(\gamma) + \\ - 0.014615\cos(\gamma) - 0.040849\sin(2\gamma)] \end{aligned} \tag{2.40}$$

### 2.3.2 Surface contribution

For space-borne simulations, the surface contribution to the background is estimated using a radiative transfer model as:

$$bg_{surf}(\lambda, \theta) = S_0(\lambda) \cos\theta A_{tel} FOV^2 R_{surf} T_{sun}(0) T_{rec}(0) \tag{2.41}$$

The equation calculates the amount of radiant flux in terms of solar irradiance $S_0(\lambda)$ received by the surface of the telescope $A_{tel}$. For simplicity, an ideal diffusely reflecting surface is considered, so that the observed intensity follows the Lambert's cosine law $S = S_0 \cos\theta$, with $\pi FOV^2$ being the solid angle of the detector, where FOV is the half-angle field of view, and $R_{surf}/\pi$ the Lambertian BRDF (Bidirectional Reflectance Distribution Function).
The albedo $R_{surf}$ indicates the fraction of lights reflected by the surface and usually varies from 0.1 (sea surface) to 0.95 (fresh snow), with an average of about 0.3 on Earth.
The radiation is attenuated along the path from the sun to the ground, inclined by an angle $\theta$ corresponding to the solar zenith angle, and then from the ground up to the receiver, that can be nadir-viewing or can form an angle $\theta'$ with the vertical. The atmospheric transmissivity is therefore calculated from $z = 0$ to $z = z_{TOA}$, taking into account the inclination of the path.
The surface contribution is considered null for solar zenith angles bigger than 90° since the solar radiation cannot directly affects the surface.

### 2.3.3 Cloud contribution

The background term due to the presence of clouds is considered for space-borne simulations and, similarly to the surface contribution, is estimated as:

(2.42)

where $N$ is number of clouds. In this case, the equation describes the path traveled by the radiation from the sun to the clouds and then reflected to the receiver. The cloud reflectivity $R_{cloud}$ (Figure 19) is a function of the solar zenith angle $\theta$ and the cloud optical thickness $\tau$. This term can be estimated through the expression is [17, 18]:

$$R_{cloud}(\theta_{obs}, \theta, \varphi, \tau) = R_{\infty}(\theta_{obs}, \theta, \varphi) - t(\tau) K_0(\theta_{obs}) K_0(\theta) \tag{2.43}$$


Figure 19: Cloud reflectivity estimated as a function of the sun zenith angle and cloud optical thickness

where $t(\tau)$ represents the global cloud transmittance, $K_0$ is the escape function, $R_\infty$ is the reflectivity of an infinite cloud, $\theta_{obs}$ and $\theta$ are the observation and the solar zenith angle, respectively, while $\varphi$ is the azimuthal angle identified by the directions of the sun and the observer. The global cloud transmittance is calculated from the optical thickness as:

$$t(\tau) = \frac{1}{0.75\tau(1-g)+1.07} \tag{2.44}$$

in which the asymmetry parameter $g$ is $\approx 0.8843$. The escape function is given by [19]:

$$K_0(\theta) = \frac{3}{7}(1+2\cos\theta) \tag{2.45}$$

$R_\infty$ can be approximated as:

$$R_\infty(\theta_{obs},\theta,\varphi) = \frac{b_1 + b_2\cos\theta_{obs}\cos\theta + p(\Theta)}{4(\cos\theta_{obs}+\cos\theta)} \tag{2.46}$$

| i | $b_i$ | $\beta_i$ | $\Theta_i$ |
|---|---|---|---|
| 1 | 1744.0 | 1200.0 | 0.0 |
| 2 | 0.17 | 75.0 | 2.5 |
| 3 | 0.30 | 4826.0 | $\pi$ |
| 4 | 0.20 | 50.0 | $\pi$ |
| 5 | 0.15 | 1.0 | $\pi$ |

Table 2: Constants used in equation (2.47)

For a nadir-observer, $b_1 = 1.48$ and $b_2 = 7.76$. The phase function $p(\Theta)$, with $\cos\Theta = (-\cos\theta_{obs}\cos\theta + \sin\theta_{obs}\sin\theta\cos\varphi)$, is estimated as:

$$p(\Theta) = 17.7\exp(-3.9\,\Theta) + \sum_{i=1}^{5} b_i \exp[-\beta_i(\Theta - \Theta_i)^2] \tag{2.47}$$

with the constants listed in Table 2.

Similarly to surface background, the cloud contribution is neglected for sun zenith angles larger than 90°.
In addition to the background, the presence of clouds is considered also in the simulation of the signals. A detailed explanation is provided in Section 2.5.

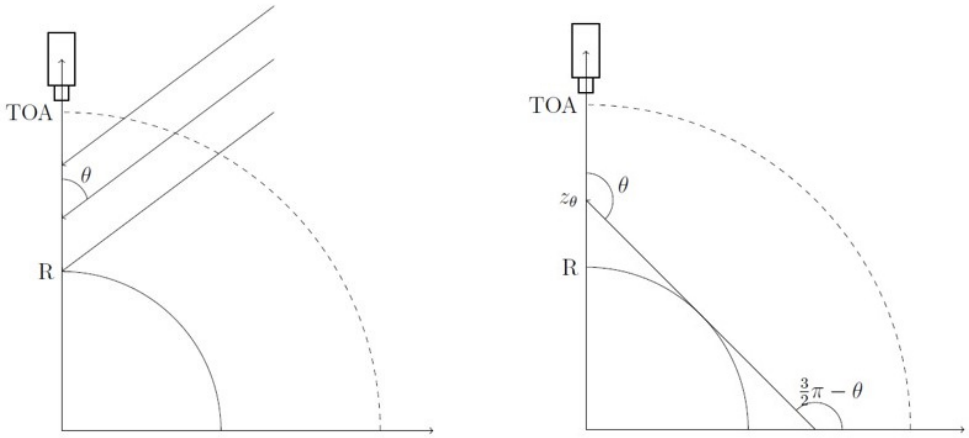
### 2.3.4 Atmospheric contribution

The atmospheric contribution is considered both in space-borne and ground-based simulations and it is obtained by integrating the contribution of the solar radiation scattered at an angle $\theta$, corresponding to the sun zenith angle, by the atmospheric constituents along the vertical path, from a starting point $z_\theta$, to the receiver:

$$bg_{atm}(\lambda,\theta) = S_0(\lambda)A_{tel}\pi FOV^2 \int_{z_\theta}^{z_{TOA}} \left[\beta_{mol}(\lambda,z)\frac{P_{Ray}(\theta)}{P_{Ray}(0)} + \right. \\ \left. + \beta_{par}(\lambda,z)\frac{P_{HG}(\theta))}{P_{HG}(0)}\right] T_{sun}(z)T_{rec}(z)dz \tag{2.48}$$

The scattering coefficients at the angle $\theta$ are obtained from the backscatter coefficients $\beta_{mol}$ and $\beta_{par}$ ($\theta = \pi$) using respectively the Rayleigh phase function [20]:

$$P_{Ray}(\theta) = \frac{3}{4}(1 + \cos^2\theta) \tag{2.49}$$




(a) (b)

Figure 20: Geometric representation of the scattered solar radiation. If the solar zenith angle is $\theta \leq 90°$, the radiation includes the whole vertical interval from surface to the top of the atmosphere and the integral is between $[0, TOA]$. If $\theta > 90°$, only part of the vertical contributes, and the integral is between $[z_\theta, TOA]$.

and the Henyey-Greenstein approximation [21] for the Mie phase function:

$$P_{HG}(\theta) = \frac{1}{2} \frac{1 - g^2}{(1 - 2g\cos\theta + g^2)^{3/2}} \tag{2.50}$$

where the asymmetry parameter is set $g = 0.987$.
The extreme of integration $z_\theta$ depends on the solar zenith angle. If $\theta < 90°$, the integral considers the contribution on the whole vertical path so that $z_\theta = 0$. For $\theta > 90°$ (sunset-sunrise and early morning or night-time), solar radiation can still reach the upper part of the atmosphere and be scattered from molecules above $z_\theta > 0$.

If we assume the Earth to be a perfect sphere with radius $R$ and that only the direct radiation can be scattered by the atmospheric constituents, $z_\theta$ represents the altitude intercepted by the tangent line forming an angle $\theta$ with the vertical to the satellite:

$$z_\theta = R\left[\sqrt{\tan^2\left(\frac{3}{2}\pi - \theta\right) + 1} - 1\right] \tag{2.51}$$

The atmospheric contribution is therefore considered null only when $z_\theta \geq z_{TOA}$ and not when $\theta \geq 90°$. For $R \approx 6371\,km$ and considering $z_{TOA} \approx 60\,km$ this

happens only when $\theta > 98°$.
The choice to integrate only up to $60\,km$ and no further, derives from the consideration that the contribution over this altitude is negligible, e.g., integrating up to 100 km the difference in the background estimation is less than $0.7\,\%$ with respect to the integration up to $60\,km$, but is much more time-consuming.

### 2.3.5 Improved background algorithm

For all three background contributions, the calculation of atmospheric transmissivity from the sun to the scattering volume $T_{sun}(z)$ and then to the receiver $T_{rec}(z)$ is required. For a vertical optical path, the atmospheric transmissivity profile is calculated with the equation (2.19), that for small inclination $\theta$ becomes:

$$T(z,\theta) = \exp\left[-\sum_{i=1}^{N}\alpha(z_i)\frac{\Delta z}{\cos\theta}\right] \tag{2.52}$$

where $\Delta r = \Delta z/\cos\theta$ is the sideways displacement and considering $z_1 = z$.
This approximation is valid for small angles, so it can be always used for the estimation of $T_{rec}(z)$, since the receiver geometry is usually nadir/zenith-viewing or, at the most, slightly inclined.
For high zenith angles, e.g., $\theta \geq 60°$, this approximation is too rough, since the traveled path $\Delta r$ is approximated by excess, so that the atmospheric transmissivity is underestimated together with the solar background atmospheric contribution. Furthermore, a dawn-dusk orbit, which is particularly suitable for space-borne lidar applications because of the reduced solar background noise, is characterized by sun zenith angles that oscillate around 90°, while it is well clear that the expression (2.52) returns valid results only for $\theta < 90°$. Usually, the atmospheric background beyond this value is simply considered null. In this regard, it is to be specified that the parallel-planes approximation is acceptable up to 85°, with difference between real and approximated solar background radiation $< 10\%$, while below 60° the difference is negligible.
To properly estimate atmospheric transmissivity for any value of $\theta$ is therefore necessary to integrate the contribution along the slant path travelled by the solar radiation through the atmosphere, renouncing to the parallel-planes assumption.
For simplicity, we can assume that the atmosphere is horizontally homogeneous, since the horizontal displacement is never too large. However, the variability with altitude of the distance $\Delta r_i$ traveled in each layer $\Delta z = [z_i, z_{i+1})$ is not negligible and has to be accounted for (Figure 21). We obtain from trigonometric calculations

that:

$$\Delta r_i = \frac{|\sin(\gamma_{i+1} - \gamma_i)|}{\sin \gamma_i} L_i \tag{2.53}$$

where $L_i$ is the distance of the intercepted atmospheric bin from the center of the Earth, this latter being assumed to be a sphere of radius $R$, calculated iteratively as:

$$\begin{cases} L_0 = R + z \\ L_i = L_{i-1} - \Delta z, \text{ if } \gamma_i < 90° \\ L_i = L_{i-1} + \Delta z, \text{ if } \gamma_i \geq 90° \end{cases} \tag{2.54}$$

and $\gamma_i$ is the angle formed by the slant path with the segment $L_{i-1}$, calculated as:

$$\begin{cases} \gamma_1 = \pi - \theta \\ \gamma_i = \pi - \arcsin \left[ \dfrac{L_{i-2}}{L_{i-1}} \sin \gamma_{i-1} - 1 \right] \end{cases} \tag{2.55}$$

Thus, the transmissivity profile can be calculated as:

$$T(z) = \exp \left[ - \sum_{i=1}^{N} \alpha(z_i) \Delta r_i \right] \tag{2.56}$$

where $z_i = (\max\{L_{i-1}, L_i\} - R)$ is the altitude corresponding to the displacement $\Delta r_i$.

### 2.3.6 Degree of polarization

Solar radiation has a polarization pattern due to the Rayleigh scattering by air molecules. When estimating the polarized signal $P_{\parallel}$ and $P_{\perp}$, the degree of polarization of skylight must be taken in account. In the Rayleigh sky model, it can be estimated as [22]:

$$\delta = \delta_{max} \frac{\sin^2 \theta}{1 + \cos^2 \theta} \tag{2.57}$$

where $\theta$ is the scattering angle and the direction of polarization is parallel to the horizon.
Equation (2.57) is valid with $\delta_{max} = 1$ only for a spherical symmetrical dipolar scatterer. In the atmosphere, the presence of non-symmetric molecules, particles and multiple-scattering phenomena decreases the maximum polarization degree. A typical value which can apply for a clear-sky simulation is $\delta_{max} = 0.80$.
It should be noticed that the maximum value of polarization happens for $\theta = 90°$.

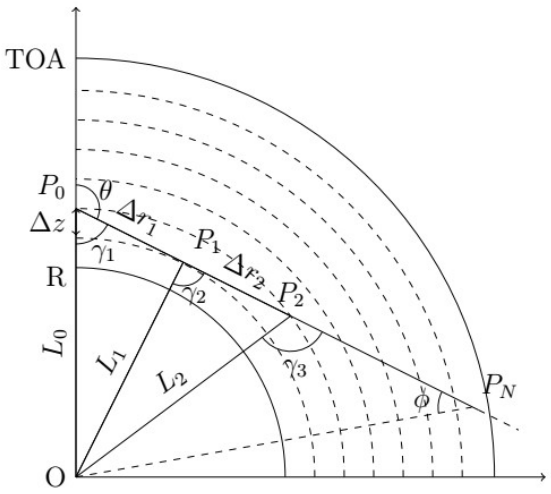



Figure 21: Schematic representation of the slant path traveled by the sun radiation: for each atmospheric bin, the real traveled path $\Delta r_i$ is iteratively calculated once the segments $L_{i-1}$,$L_i$ and the angle $\gamma_i$ are known.

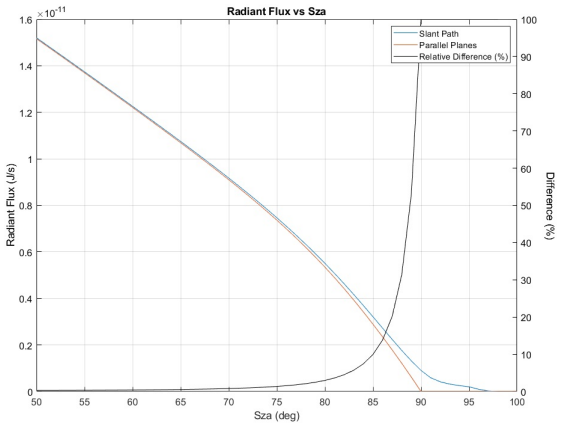



Figure 22: Radiant flux collected by the telescope as a function of the solar zenith angle. The values are calculated considering a telescope diameter $d = 2\,m$, $FOV = 12.5\mu rad$ and a solar irradiance at $355\,nm$ of $S_0 = 3.91 \times 10^8\,W/m^2$ . The figure shows the results obtained using both the improved algorithm (blue line) and the parallel-planes approximation (red line), with their relative difference (black line). Below 60° the difference is negligible. The parallel-planes approximation is acceptable up to 85° (difference $< 10\%$).

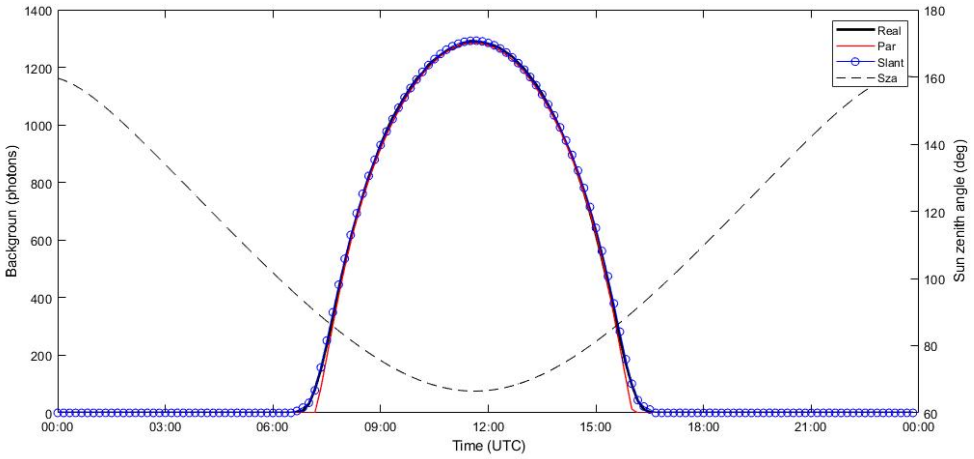
Since for a dawn-dusk orbit, the sun zenith angle has values around 90 degrees, the amount of solar background can be reduced using a polarizing filter.

### 2.3.7 Model validation

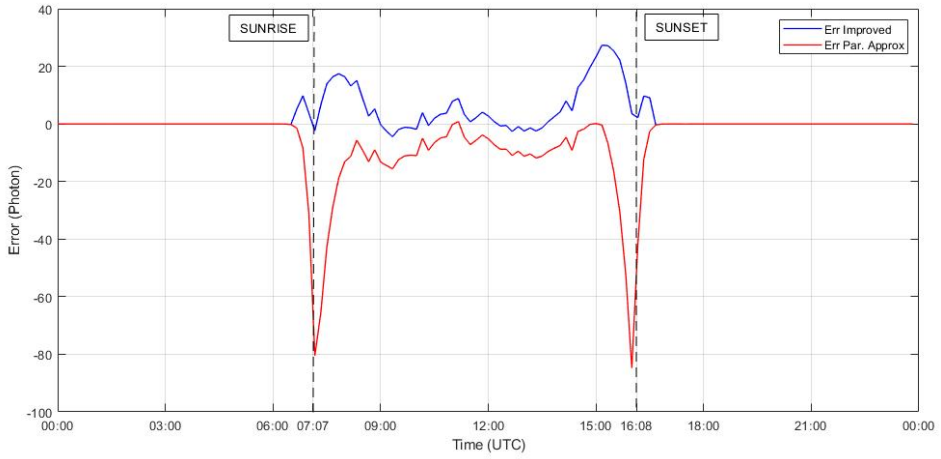
The algorithm for background estimation has been validated through the comparison with real data obtained from the ground-based Raman Lidar "Concerning" during the field campaign WaLiNeAs.
To test the algorithm, the day 27 December 2022 was selected, since it was a cloud-free day with 24 hours of continuous measurements. The solar background was extracted from signals acquired in photon-counting mode using a subtraction scheme. For the simulation, all the experimental parameters of Concerning were considered. A detailed description of the experimental setup is given in Section 1.2.1.
The background was calculated using both the parallel planes approximation and the improved algorithm. Figure 23 shows the solar background obtained from the roto-vibrational nitrogen signal during the day, together with the values obtained from the simulator. A normalization term $k = 0.7$ was used to adjust simulated values to real data. The error is calculated as the difference between simulated and real background. Results show that the the improved algorithm slightly overestimate the background, while the parallel-planes approximation produces a conservative estimate, with values acceptable until $\theta < 85°$. It is clear that during sunrise and sunset, only the improved-algorithm should be used.

## 2.4 Shot-noise

Photon numbers estimated through equations (2.17), (2.22) and (2.16) refer to single shot echoes. Simulated signals are then obtained by integrating single-shot echoes over a predefined acquisition time, i.e. the atmosphere is considered frozen, or homogeneous, over a horizontal area corresponding to the integration time. Input parameters are considered constant during the integration time, this latter typically defining the horizontal resolution. Consequently, the integrated signal is simply a multiple of the single-shot contribution. Therefore signals should be perturbed to account for the statistical fluctuations in the photon number reaching the detectors during the considered integration time to have a more realistic simulation.
The fluctuation of photon number that is typically observed in photon detection is defined as shot-noise or Poisson noise [23]. In fact, the counting of single photons describes a statistical process in which each event is independent from the others




(a)




(b)

Figure 23: (a) Real and simulated background signal associated with solar illumination together with sun zenith angle and (b) difference between real and simulated values

and therefore it is to be considered as stochastic Poisson process.
The Poisson distribution with mean and variance $\mu$ is given by:

$$Poiss_{\mu}(n) = \frac{\mu^n}{n!}e^{-\mu} \tag{2.58}$$

and represents the probability to have a number of events $n$, with $\mu$ being the average number of events in a time interval. If $N_s$ is the number of single-shot echoes, the perturbed total signal $P_{tot}^{noise}$ can be obtained by extracting $N_s$ random values from a Poisson distribution with the mean being equal to the single-shot signal $P_{ss}(z)$, which is thus considered as an average signal. Hence, the integrated signal will be the sum of all perturbed contributions:

$$P_{tot}^{noise}(z) = \sum_{i=1}^{N_s} [Poiss_{P_{SS}(z)}(n)]_i \tag{2.59}$$

where $[Poiss_{P_{ss}(z)}(n)]_i$ is the i-th value randomly extracted from a Poisson distribution with mean $P_{ss}(z)$.
It should be noted that the Poisson distribution is a discrete distribution defined on natural numbers, therefore all the extracted values, as well as the total intensity of the signal, are positive integers. This behavior is useful to simulate the quantization of the signal.
Considering that the sum of independent Poisson random variables is also a Poisson variable with mean equal to the sum of the single parameters, the perturbed integrated signal can be obtained in a faster and less memory-consuming way, extracting a single random number from a Poisson distribution with mean $\mu = N_s \cdot P_{ss}$:

$$P_{tot}^{noise}(z) = Poiss_{[N_s \cdot P_{SS}(z)]}(z) \tag{2.60}$$

The simulator applies equation (2.60) when it calculates the signal acquired during a specified acquisition time. The signal uncertainty is estimated through Poisson statistics as $\Delta P = \sqrt{\mu}$.
When the total signal is calculated over the time integration, equation (2.59) is used. In this case the total signal uncertainty is calculated using the standard deviation of the $n$ summed profiles as $\Delta P = \sigma \cdot \sqrt{n}$.
It is important to notice that the latter method offers a more accurate estimation of the error, but it's much more time-consuming, therefore its use is recommended for sensitivity studies or simulations involving only a few data.

## 2.5 Simulation with clouds

Is it possible to consider simulations in clear sky conditions or including the presence of clouds. Raman Lidar are able to penetrate through cirrus clouds with an optical depth up to 0.3 [24]. Furthermore, a strong point of Lidars is the capa-

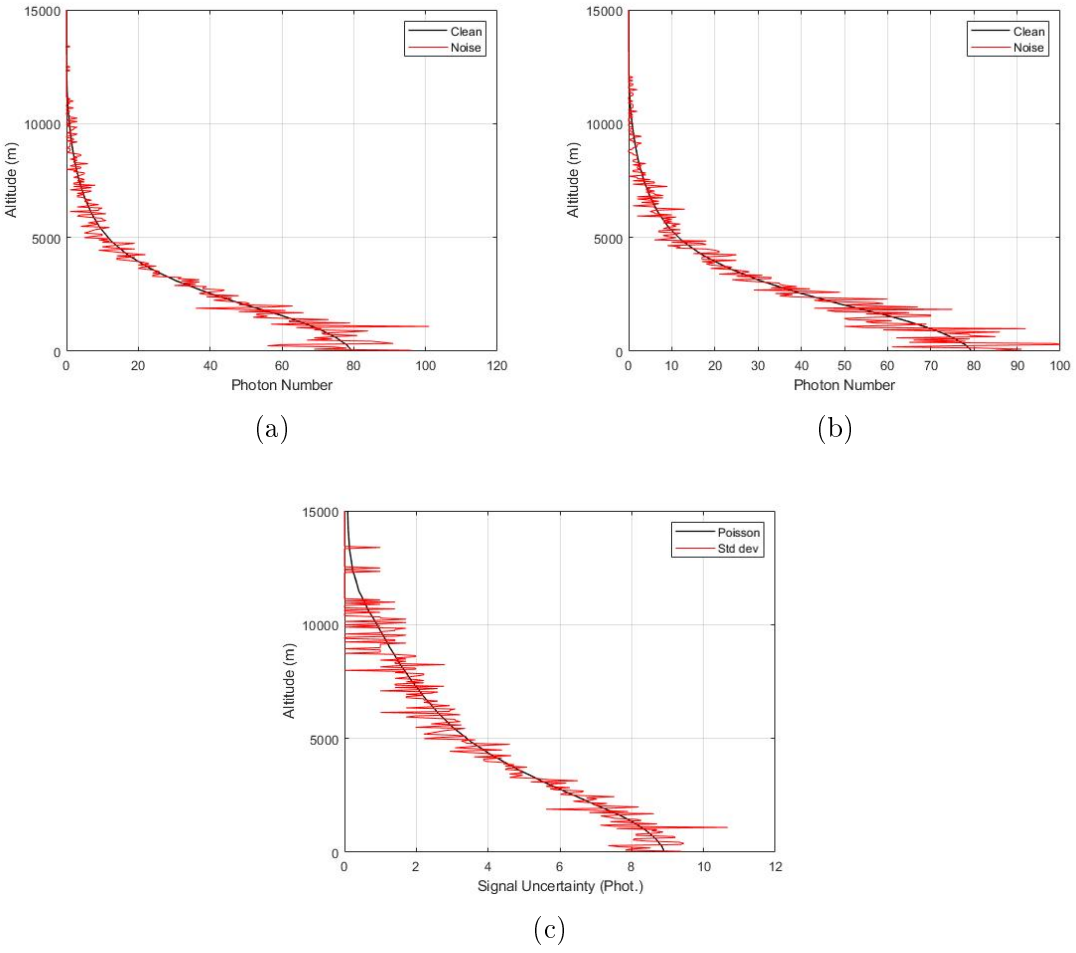



Figure 24: Signal with shot-noise calculated as (a) the sum of perturbed profiles and (b) perturbing the total profile. The statistical uncertainty (c) is calculated using the Poisson statistics and the standard deviation

bility to exploit the shots "between" discontinuous clouds, especially in case of space-born measurements. Usually, information about clouds includes a vertically resolved cloud fraction, i.e., the fraction of sky covered by clouds specified layer by layer, as well as the in-cloud optical thickness for ice and liquid clouds. Conversely, no information is usually given about the location of the clouds within the measurement or model grid. In order to properly account for the discontinuous presence of clouds, the total signal $P_{tot}$ is determined as a weighted sum of signals with different cloud attenuation, including not attenuated, partially attenuated, and strongly attenuated lidar echoes. The weights used in the sum are taken to be proportional to the cloud fraction, which quantifies the probability to intercept clouds. Specifically:

$$P_{tot}(z) = P_{clear}(z) \cdot \left[\omega_0 + \sum_{i=1}^{2^N-1} \omega_i T^2_{cloud,i}(z)\right], \text{ with } \sum_{i=0}^{2^N-1} \omega_i = 1 \tag{2.61}$$

where $N$ is the number of cloud layers. If no clouds are present ($N = 0$), the only weight is $\omega_0 = 1$ and the total signal coincides with the signal in clear sky conditions $P_{clear}$. If $N$ cloud layers are present, the number of possible paths for the beam are $2^N$ (Figure 25), which represent all the possible path combinations through the clouds. Every path is associated to a weight obtained as the product of the probability to intercept a single cloud layer.

Cloud layers are identified analyzing the cloud fraction profiles and estimating the top and base of the separate cloud layers. In the simulations, cloud fractions smaller than 1% are discarded. The probability of the laser pulse to hit a cloud layer is then taken as the maximum cloud fraction at each profile level. Finally, the total cloud transmissivity along the different paths is calculated from the cloud optical depth $\tau$ as:

$$T_{cloud}(z) = \exp\left(-\int_z^{z_{TOA}} \tau(t)dt\right) \tag{2.62}$$

Figure 26 illustrates the attenuation due to two separate cloud layers centered approximately at $7\,km$ and $0.5\,km$ with a maximum cloud fraction of 0.34 and 0.70 and a maximum optical thickness at $355\,nm$ of 0.25 and 6.55, respectively. The two layers generate four possible paths, whose respective weights are 0.198 (clear sky), 0.102 (only layer 1), 0.462 (only layer 2) and 0.238 (layer 1 and 2). The actual signal, obtained as the weighted sum of the attenuated signals along the different optical paths, is characterized by an overall attenuation of 35 % by the upper layer and by 80 % by the lower layer.

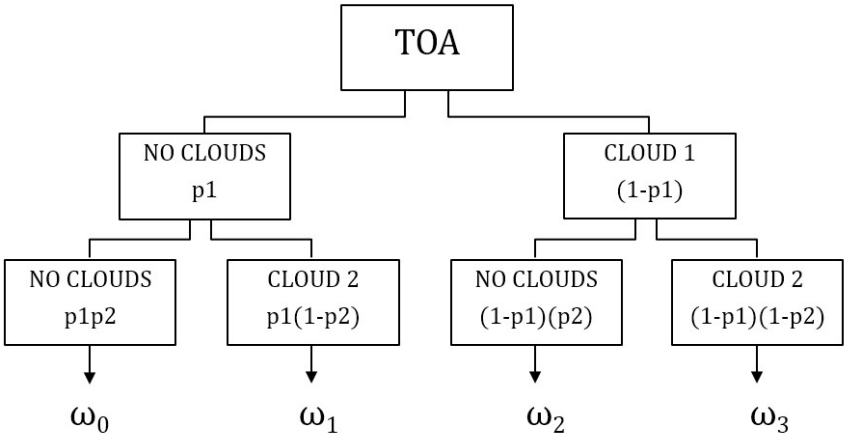



Figure 25: Decision tree with all possible paths considering two cloud layers ($N = 2$). The beam can avoid all cloud layers, intercept only one cloud layer (cloud 1 or cloud 2) or intercept both cloud layers. Every branch of the tree has a probability $p_i$ corresponding to the cloud layer probability being intercepted. The weights $\omega_i$ are calculated as the products of the probability to intercept the single cloud layer.

It is possible to estimate the amount of signal after the cloud attenuation as a function of the optical thickness and cloud coverage. A 30 % reduction in signal intensity is found for optical thicknesses up to 0.2 with a completely covered sky (cloud fraction = 1). The same values can be found for higher values of cloud optical thickness, if the cloud fraction is smaller than 30%. Indeed, the attenuation factor reveals a clear dependence on the optical thickness only up to $\tau \leq 1.6$ (Figure 27). For larger values of $\tau$, the signal is completely extinct by the clouds, so the percentage of the received signal depends only on the weight of the free-cloud path.

Regarding the simulation of the elastic signal in presence of clouds (Figure 28), the lidar equation is modified by adding the cloud reflectivity calculated through equation (2.43):

$$P_{\lambda_0}(z) = P_0 \cdot \eta\tau_{\lambda_0} \cdot \frac{c\Delta t}{2} \cdot \frac{A_{tel}}{z^2} \cdot [\beta_{mol,\lambda_0}(z) + \beta_{par,\lambda_0}(z) + R_{cloud}(z)] \cdot T^2_{\lambda_0}(z) \quad (2.63)$$

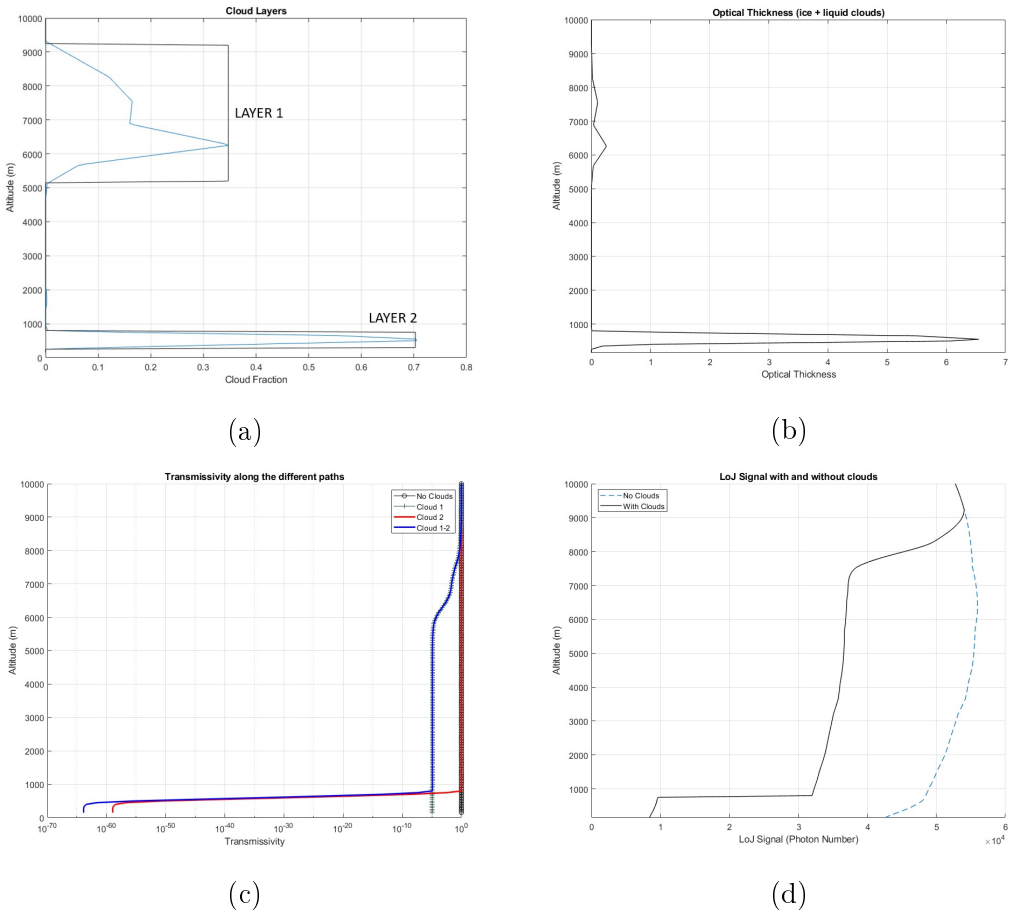



Figure 26: Different cloud layers identified using the cloud fraction (a), with the corresponding optical thickness (b). The cloud transmissivity is calculated for the four possible paths (c) and is used to attenuate the signal (d).

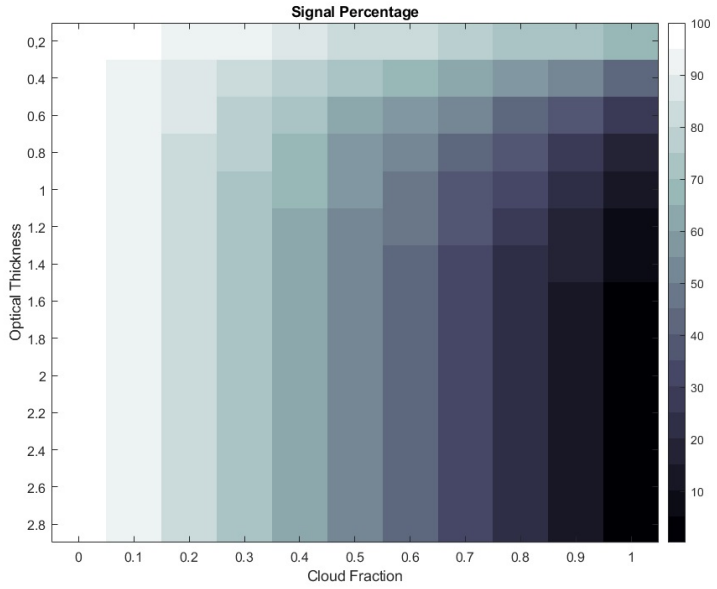



Figure 27: Signal intensity of the attenuated signal $P_{cloud}$ compared to the non-attenuated signal $P_{clear}$ for different cloud fraction and optical thickness values

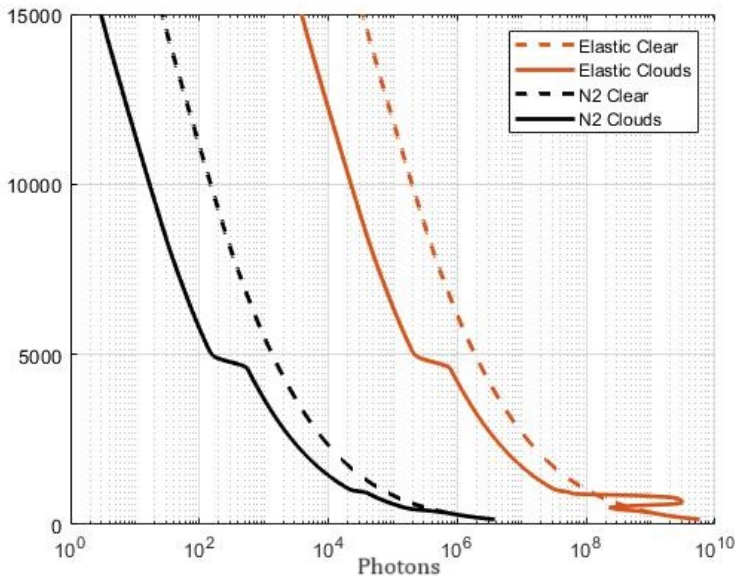



Figure 28: Ground-based simulation of $N_2$ and elastic signal in clear sky conditions and with clouds.

## 2.6 Retrieval Module

The retrieval module allows to analyze the lidar signals and apply consolidated techniques to determine the vertical profiles of water vapor mixing ratio, temperature and ancillary products, such as backscatter and extinction coefficients. It can be used indifferently with real signals or simulated signals obtained with the forward module. A vertical averaging is first performed to smooth the signals. In case of a real signal acquired in photon-counting it is possible to apply a dead-time correction. In fact in case of detector with non-paralyzable behavior, the probability of measuring an event during the dead-time $\tau$ is given by an exponential distribution. Therefore, if $P(z)$ is the measured photon count rate in a time interval $\Delta t$, the corrected signal will be:

$$P'(z) = \frac{P(z)}{1 - \frac{\tau}{\Delta t} P(z)} \tag{2.64}$$

A subtraction scheme is then applied to remove the solar background contribution from the synthetic signals. The background is estimated by averaging the last bins of the lidar echo, where the signal should be negligible and only background radiation contributes to the measured signal.
The water vapour mixing ratio profile is obtained from the ratio of the roto-vibrational $H_2O$ signal to the temperature-independent reference signal:

$$\mathrm{WVMR}(z) = c \cdot \frac{P_{H_2O}(z)}{P_{ref}(z)} \cdot \frac{T_{ref}(z)}{T_{H_2O}(z)} \tag{2.65}$$

The reference signal can be the roto-vibrational $N_2$ signal or a weighted sum of HiJ and LoJ rotational signals obtained through the expression 2.22. The calibration constant is usually determined by a comparison with co-located measurements, e.g., from radiosoundings. Theoretically, it can be estimated as:

$$c = \frac{\eta_{ref}}{\eta_{H_2O}} \cdot \frac{(d\sigma/d\Omega)_{ref}}{(d\sigma/d\Omega)_{H_2O}} \cdot \frac{n_{ref}}{n_{air}} \tag{2.66}$$

The statistical uncertainty affecting water vapour mixing ratio measurements can be determined through error propagation, considering the uncertainties affecting the Raman signals $P_{H_2O}(z)$ and $P_{Ref}(z)$. The statistical uncertainty is determined as the standard deviation of the single signal profiles from the mean profile along a horizontal path corresponding to the horizontal resolution of the measurements,

as described in Section 2.4.

$$\frac{\Delta \mathrm{WVMR}(z)}{\mathrm{WVMR}(z)} = 100 \cdot \sqrt{\left(\frac{\Delta P_{H_2O}(z)}{P_{H_2O}(z)}\right)^2 + \left(\frac{\Delta P_{ref}(z)}{P_{ref}(z)}\right)^2} \tag{2.67}$$

The atmospheric temperature profile is obtained from the ratio $R(z)$ of LoJ to HiJ rotational Raman signals. Depending on the number of parameters used to fit the ratio $R$, it can be calculated as:

$$T(z) = \frac{b}{\ln R(z) - a} \tag{2.68}$$

or

$$T(z) = \left[ -\frac{b}{2c} - \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c} + \frac{\ln R(z)}{c}} \right]^{-1} \tag{2.69}$$

The statistical uncertainty affecting temperature measurements can be again determined through error propagation, considering the uncertainties affecting the considered rotational Raman signals $P_{HiJ}(z)$ and $P_{LoJ}(z)$:

$$\Delta T(z) = R(z) \cdot \frac{\partial T(z)}{\partial R} \cdot \sqrt{\left(\frac{\Delta P_{LoJ}(z)}{P_{LoJ}(z)}\right)^2 + \left(\frac{\Delta P_{HiJ}(z)}{P_{HiJ}(z)}\right)^2} \tag{2.70}$$

The systematic error (relative bias) for both water vapour mixing ratio and temperature is calculated by comparing the mean retrieved profile with the real profile, i.e., the profile used as input for the simulator or the radiosounding profile:

$$\mathrm{BIAS}(z) = \frac{x_{real}(z) - x_{ret}(z)}{x_{real}(z)} \tag{2.71}$$

The particle backscatter coefficient is calculated from the elastic signal and a reference signal as:

$$\beta_{par,\lambda_0}(z) = \beta_{mol,\lambda_0}(z) \left[ \frac{P_{\lambda_0}(z)}{k \cdot P_{ref}(z)} - 1 \right] =: \beta_{mol,\lambda_0}(z)[R(z) - 1] \tag{2.72}$$

where $k$ is constant obtained by normalizing the elastic signal to the reference signal at a reference height in which $\beta_{par} \ll \beta_{mol}$ [25]. The statistical uncertainty

is estimated as:

$$\frac{\Delta\beta_{par,\lambda_0}(z)}{\beta_{par,\lambda_0}(z)} = 100 \cdot \frac{R(z)}{R(z)-1}\sqrt{\left(\frac{\Delta P_{\lambda_0}(z)}{P_{\lambda_0}(z)}\right)^2 + \left(\frac{\Delta P_{ref}(z)}{P_{ref}(z)}\right)^2} \tag{2.73}$$

In alternative, it is possible to use the Klett's algorithm [26, 27], with which it is possible to estimate the total backscatter coefficient $\beta = \beta_{mol} + \beta_{par}$ using only the elastic signal, with the assumption $\alpha_{par} = \mathrm{LR} \cdot \beta_{par}$ :

$$\beta(z) = \frac{\exp(S'(z) - S'(z_{ref}))}{\beta^{-1}(z_{ref}) - 2\mathrm{LR}\int_z^{z_{ref}}(S'(r) - S'(z_{ref}))dr} \tag{2.74}$$

where $LR$ is the lidar ratio, $z_{ref}$ is a reference height in which $\beta(z_{ref}) = \beta_{mol}(z_{ref})$ and the term $S' - S'_0$ is an auxiliary variable calculated as:

$$S'(z) - S'_0 = S(z) - S(z_{ref}) + 2\left(\frac{8\pi}{3} - \mathrm{LR}\right)\int_z^{z_{ref}} \beta_{mol}(r)dr \tag{2.75}$$

and $S(z) = \ln[P_{\lambda_0}(z) \cdot z^2]$ is the logarithm of the range corrected signal.
The particle extinction can be obtained from a reference signal as [28]:

$$\alpha_{par,\lambda_0}(z) = \frac{1}{2}\frac{d}{dz}\ln\left[\frac{n(z)}{P_{ref}(z)\cdot z^2}\right] - \alpha_{mol,\lambda_0}(z) \tag{2.76}$$

with the statistical uncertainty estimated as:

$$\Delta\alpha_{par,\lambda_0}(z) = \frac{1}{\sqrt{2}\Delta z}\frac{\Delta P_{ref}(z)}{P_{ref}(z)} \tag{2.77}$$

# Chapter 3

# Application of the End-to-End Simulator

The great flexibility of the end-to-end simulator and the capability to adapt its applications to different scenarios made it a central tool for different projects. Specifically, on the space-borne side it was used for the project "As-Atlas", in support of the space lidar mission ATLAS, now ATLANTIS, and is currently in use in support of the space lidar mission "Caligola". Moreover it was used in support of the design and development of the two ground-based lidars, "Concerning" and "Marco", developed at the University of Basilicata.
Simulations and results for the projects mentioned above, are discussed in this chapter.

## 3.1 Atlas

The Atmospheric Thermodynamic LidAr in Space (ATLAS) is a mission concept proposed to the European Space Agency in the frame of "Earth Explorer-11 Mission Ideas" Call with the aim to develop the first Raman Lidar in space capable to measure simultaneously atmospheric temperature and water vapour mixing ratio profiles with high temporal and spatial resolutions. A study in support of the project was funded and supported by the Italian Space Agency (As-Atlas: Activities in Support of Atlas).
Accurate measurements of water vapor mixing ratio and atmospheric temperature profiles in the lower troposphere are essential for the understanding of the water

and energy cycles, as well as the prediction of extreme events, that nowadays still show huge deficiencies on all temporal and spatial scales [29], and would have a great impact in many research areas.
The vertical structure of atmospheric humidity is highly variable in space and time, with a strong vertical gradient, especially in the tropics. Moreover, temperature and humidity profiles presents gradients and inversions that are not well retrieved by passive remote sensing systems, characterized by coarse vertical resolution ($>$ 1 km) and intrinsic limitations in the inversion of the radiative transfer equation. Water vapor is the most important atmospheric greenhouse and it has a fundamental role in the study of radiative transfer and energy budget and therefore in the study of the earth system model. Moreover, water vapour and temperature profiles have a strong impact on the long-wave down-welling radiation.
Accurate temperature profiles can be used to derive the planetary boundary layer (PBL) depth. Especially in daytime, the water vapour flux divergence becomes very important so that both the surface flux and the entrainment flux must be determined.
One of the most important aspects is that the availability of high vertical resolution profiles is limited mainly to the global radiosonde network, with discrete locations and typically daily temporal resolution. A space Raman lidar will bring entirely new, globally-sampled information at high vertical resolution in under-observed parts of the atmosphere, that could be used in data assimilation to provide initial conditions and therefore better forecasting.

### 3.1.1 Atlas Experimental setup

The space-borne Raman Lidar considered in ATLAS collects six lidar signals: the water vapour roto-vibrational Raman signal, the high- and low-quantum number $O_2$-$N_2$ rotational Raman signals both in the anti-Stokes and Stokes branches and the elastic backscattered signal at 354.7 nm. The laser is a frequency-tripled diode pumped Nd:YAG laser, with an average power of 200 W (1 J per single-pulse with a repetition rate of 200 Hz). The receiver involves a telescope with a primary mirror diameter of $2\,m$ in Cassegrain configuration, with a field-of-view of $25\,\mu$rad (FWHM) and a receiving unit for the collection and detection of the lidar signals. The spectral selection is carried out based on the use of interference filters defined to minimize the uncertainties.
The selected orbit for the satellite is a sun synchronous polar orbit with a local time descending node at 06:00 (dawn-dusk orbit) and an orbital height of 450 km. This kind of orbit guarantees a sun zenith angle on the sub-satellite point close to 90°, which allows to reduce the background radiation, and optimal illumination conditions for power generation and on-board thermal management.

| | | | |
|---|---|---|---|
| Laser | Nd:YAG @354.7 nm | | |
| Pulse Energy | 1 J | | |
| Repetition rate | 200 Hz | | |
| Power | 200 W | | |
| Telescope diameter | 2 m | | |
| FOV | 25 $\mu$rad | | |
| Quantum efficiency | 0.85 | | |
| Filter peak transmission | 0.80 | | |
| Filter CWL and FWHM | H2O | 407 nm | 0.3 nm |
| | Elastic | 354.7 nm | 0.5 nm |
| | LoJ AS | 354.36 nm | 0.3 nm |
| | HiJ AS | 353.29 nm | 0.5 nm |
| | LoJ S | 355.30 nm | 0.3 nm |
| | HiJ S | 356.25 nm | 0.5 nm |
| Satellite height | 450 km | | |
| Ascending node | 06:00-18:00 | | |
| Orbit inclination | 97.2° | | |
| Vertical resolution | 200 m | | |
| Horizontal resolution | 50 km | | |
| Temporal resolution | 7.5 s | | |

Table 3: Atlas system specifications

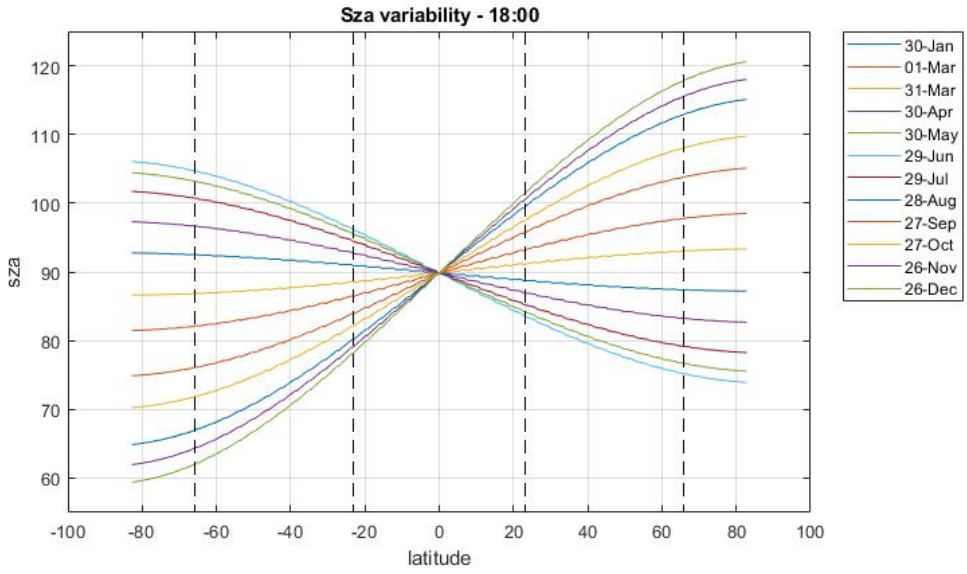



Figure 29: Solar zenith angle calculated for different day of the year, as a function of latitude

### 3.1.2 Simulations with standard atmospheric models

To assess the performance of the space-borne lidar Atlas in standard conditions, several simulations have been carried out, using as input data the thermodynamic profiles from different standard atmospheric models. Specifically, the Mid-Latitude summer and winter, Tropical and Sub-Arctic summer and winter were used. To consider the different amount of solar background depending on latitude and time of the year, a preliminary study has been performed to estimate the variability of solar zenith angle along the orbit (Figure 29). It should be noticed that the solar zenith angle is not perfectly symmetrical on the two hemisphere due to the inclination of the orbit. For each latitude belt, simulation were performed considering the maximum, mean and minimum solar zenith angle, corresponding to low, mean and high background conditions (Table 4).

Figures 30 and 31 show the statistical uncertainties affecting the simulated water vapor mixing ratio (%) and temperature ($K$) profile measurements, obtained using different atmospheric models. The tropical simulation shows very promising results for all background conditions, with an uncertainty below 10 % up to 5 km and 8 km for mean and low background respectively and below to 20% up to 4 km with high background. At mid latitudes, the performance in mean background conditions is still good in the summer, with a small increase of the error ($<$ 10% up to 7 km with low background, $<$20% up to 5 km and 4 km with mean and high background). During winter the amount of water vapor is reduced and the zenith

| Latitude | Max | Mean | Min |
|---|---|---|---|
| Tropical | 101 | 89.9 | 78.4 |
| MidLat (Summer) | 105 | 89.9 | 75.2 |
| MidLat (Winter) | 118 | 89.9 | 62.1 |
| Sub-Artctic (Summer) | 106 | 90.1 | 73.9 |
| Sub-Artctic (Winter) | 120 | 89.9 | 59.4 |

Table 4: Solar zenith angle calculated along the orbit for the different latitude belts and day of the year

angle has a higher variability, so performances are still acceptable only for low and mean background, with errors below 10% up to 5 km and and below 20 % up to 3 km respectively. For sub-arctic simulations, results are good during the summer season with uncertainties below 10% up to 7 km for low background and 4 km for mean background, while with high background is below 20 % up to 3 km.
Temperature measurements always show very promising results, with uncertainties always below 0.8 K up to 10 km for all background conditions and latitudes, decreasing to 0.6 K up to 15 km when considering mean and low background.
The observational requirements that needs to be fulfilled by a satellite according to the World Meteorological Organization [30] are random uncertainties below 20% for water vapor and 1 K for temperature, values that are reached by Atlas in almost all the simulations.

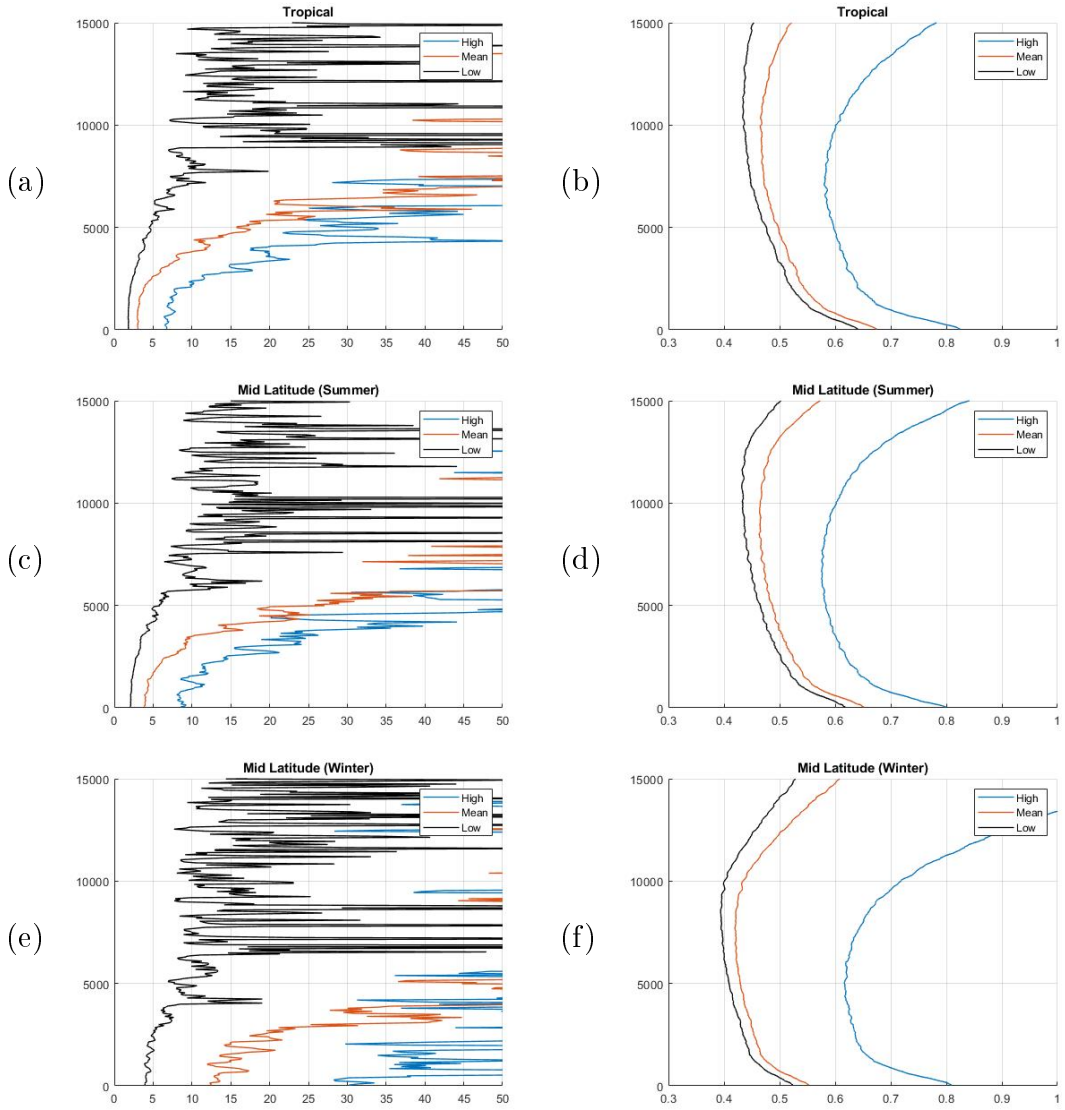



Figure 30: Statistical uncertainties for water vapor mixing ratio (%) and temperature (K) obtained using Tropical (a-b), Mid Latitude summer (c-d) and winter (e-f) atmospheric models

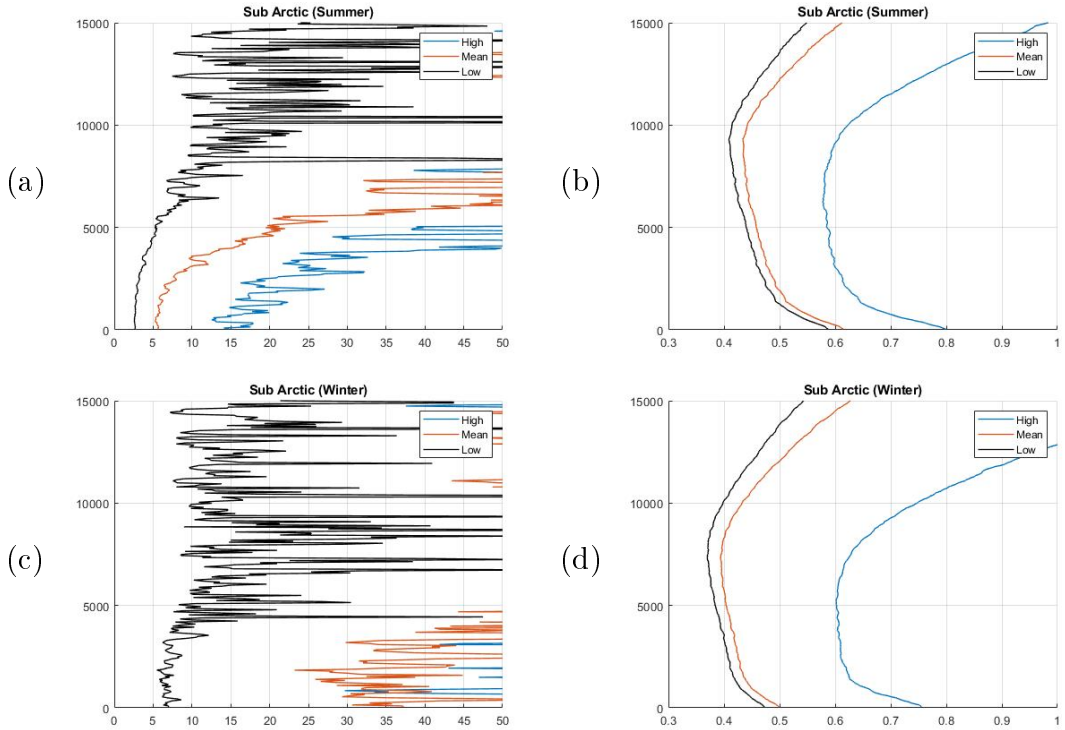



Figure 31: Statistical uncertainties for water vapor mixing ratio (%) and temperature (K) obtained using Sub-Arctic summer (a-b) and winter (c-d) atmospheric models

### 3.1.3 Performance dependence on technical specifications

To understand how to maximize the system performances, several sensitivity studies have been performed to estimate the results with variable technical specifications.

The selection of the orbit is one of the most important factors. For Atlas/Atlantis, a dawn-dusk orbit has been selected, in order to minimize the contribution of solar background and have a beta angle always close to 90°, which guarantees optimal illumination conditions for power generation and on-board thermal management. Moreover, a daily overpass at 6:00 and 18:00 (local time) allows to measure thermodynamic profiles before and after the development of PBL, which are very important for data assimilation. However, the orbital height is a point of discussion, since a lower orbit allows better performances, but increases orbital perturbations and energy consumption. Figure 32 shows the percentage variability from the signal intensities calculated at 450 km, which is the nominal value

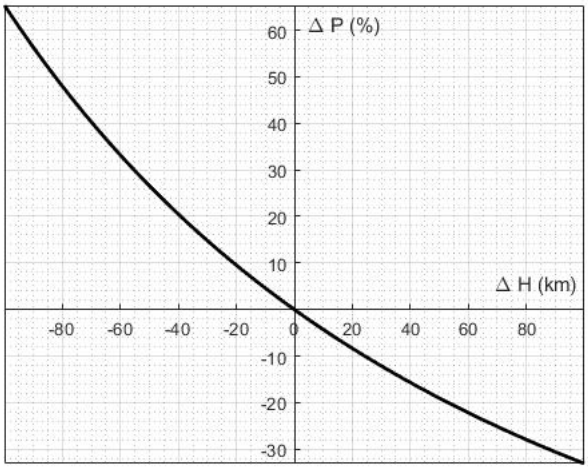



Figure 32: Percentage variation of signal intensities with respect to orbital height variation

for obit height used in simulations. For example, the lowering of the orbit to 400 km ($\Delta H = -50$ km) would increase the signals by 25 %, while solar contribution remains the same, resulting in a reduction of statistical uncertainty of about 10-20 %, depending on the background.

Another parameter that can be varied is the telescope diameter. In Atlas, a 2 m diameter has been considered, which is bigger than the one used in Aeolus (1.5 m). Recently, ESA demonstrated the possibility to develop space-qualified telescopes with a diameter up to 2.2 m [31]. A larger telescope would increase the signal intensity (Figure 33), but also the collected solar background. Increasing the diameter from 2 m to 2.2 m, would increase the intensity of 21 % and reduce the error of 9 %.

The laser source specifications are another critic point. In Atlas, a laser system with 200 W in the UV and single pulse energy of 1 J is considered, which is a concrete possibility considering the rapid development of technology for laser power amplifier. Also a less challenging laser source can be considered, with a power of 120 W and 0.6 J per single pulse. The signal intensities is obviously directly proportional to the laser power, but with the same power it is possible to consider different combinations of energy and repetition rate (Figure 34). Since the collected background increases with higher repetition frequencies, in general it is better to increase the pulse energy.

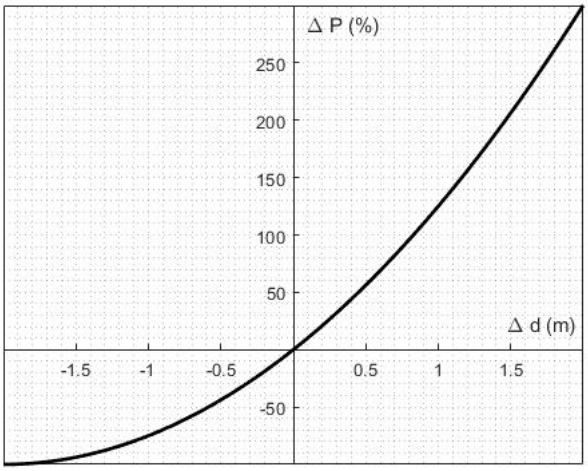
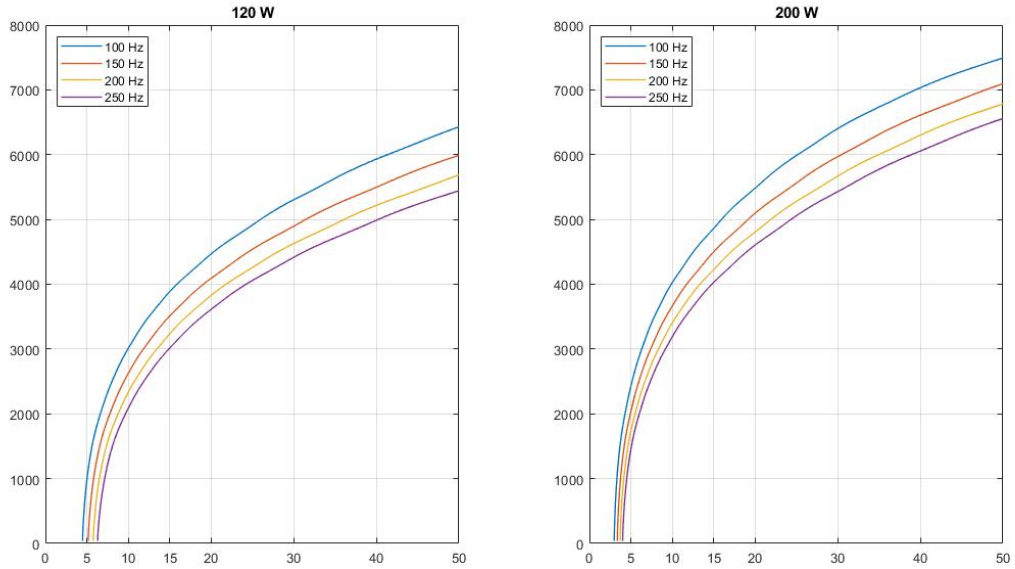
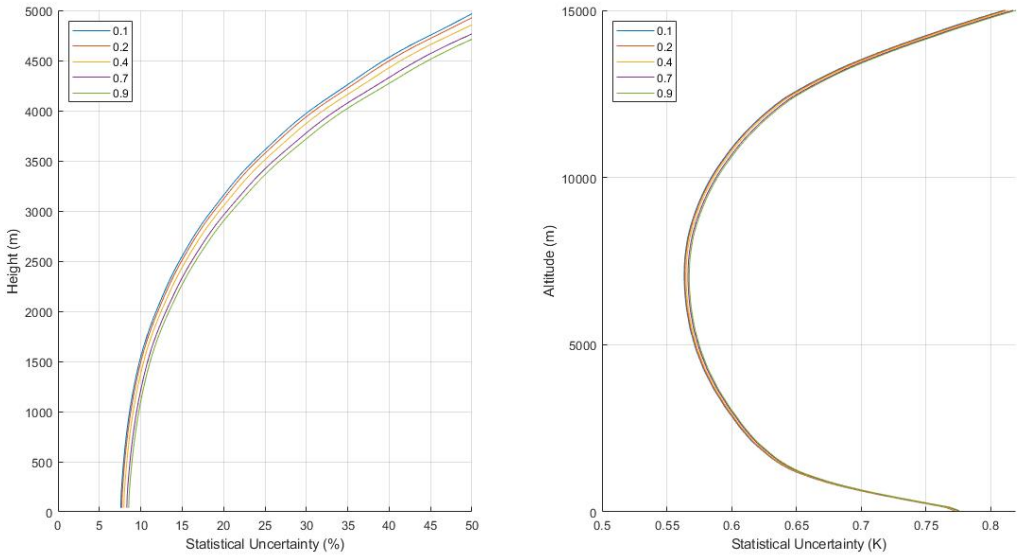



Figure 33: Percentage variation of signal intensity as a function of the telescope diameter variation




(a) (b)

Figure 34: Water vapor mixing ratio statistical uncertainty (%), obtained considering a laser source of (a) 120 W and (b) 200 W, with different repetition rate (100, 150, 200, 250 Hz). Simulations are based on thermodynamic profiles from Mid Latitude Summer atmospheric model, with a mean solar background



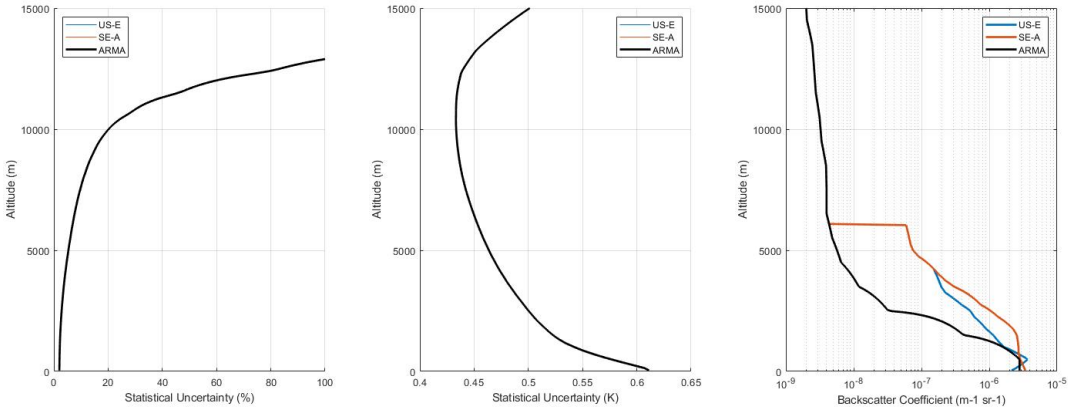

(a) (b)

Figure 35: Statistical uncertainty for water vapor and temperature for different values or surface albedo

### 3.1.4 Variability of environmental conditions

Some sensitivity studies have been carried out to evaluate the influence of environmental conditions on the lidar system's performances. Considering the standard experimental setup, several simulations were performed with different values of surface albedo: sea surface (0.1), forest and trees ($\approx 0.2$), desert (0.4), melting snow ($\approx 0.7$) and fresh snow (0.9). The average surface albedo is approximately 0.35.
As explained in Section 2.3.2, the surface contribution to the background is considered null for solar zenith angles higher than 90°, therefore for the simulations, a zenith angle of 80° was considered, using thermodynamic profiles from Mid Latitude Summer atmospheric model. Results show that the variability of surface albedo has a minimal influence on the performance, which is almost the same for all the considered values. Indeed, the surface contribution to the background is always smaller than the atmospheric contribution.

To assess the performances in relation to the amount of aerosol, different profiles of the backscatter coefficient were considered: the ARMA Aerosol Model and two nocturnal profiles from the climatological analysis of Calipso obtained from a six-year mean and related to eastern US (31-41°N, 95-75°W, July 2006-2011) and



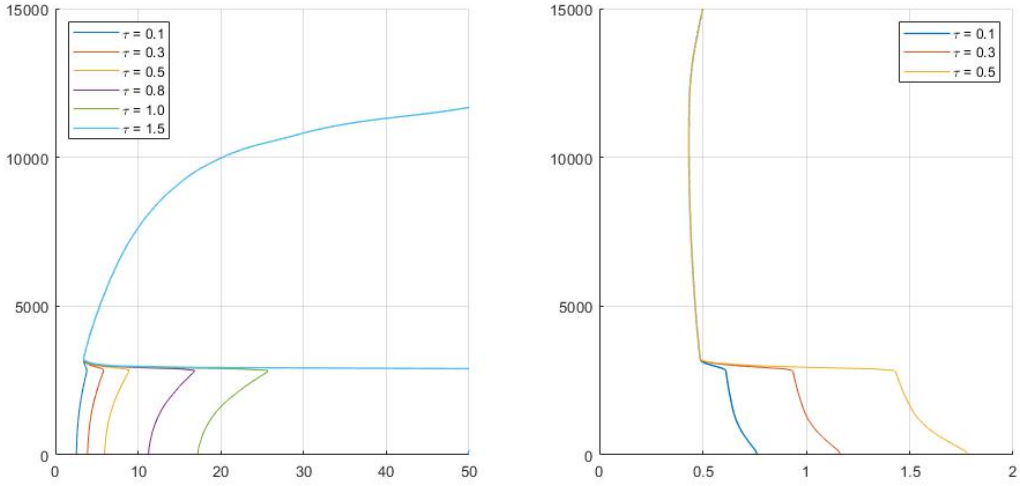

(a) (b) (c)

Figure 36: Statistical uncertainty for water vapor (a) and temperature (b) with different aerosol models (c)

South-East Asia (1-19°N, 90-110°E, August 2006-2011). In this case, simulations were performed without background, considering a night scenery. Figure 36 reveals that the variations are completely irrelevant.

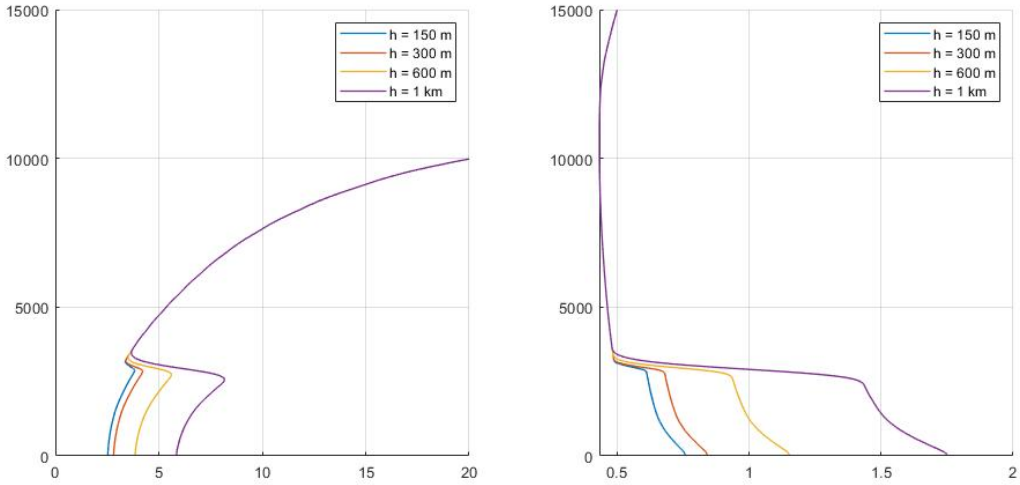
A factor that has a strong influence on the performance assessment is the presence of clouds. The attenuation of the signal depends on several factor, such as the optical thickness of the cloud and its height extension, but also the spatial distribution. For Atlas/Atlantis, a time resolution of 7.5 s is considered, which for a satellite moving at 7 km/s, corresponds to a horizontal resolution of 50 km. The total signal measured during this time period is a combination of single echo signals where some may have intercepted clouds, but others may not. Therefore the cloud fraction, i.e., the portion of the grid that actually contains clouds, is another factor to be considered. To assess how these factors influence the performance, a single gaussian-shaped cloud centered at 3 km was included in the simulations. At first, a vertical extent of 150 m was considered, with a variable optical thickness from 0.1 to 1.5. Considering a night-time simulation, the statistical uncertainty of water vapor mixing ratio keeps below 10 % up to $\tau = 0.7$ and increases to 25 % for $\tau = 1$, while temperature uncertainty increases to 1.2 K for $\tau = 0.3$ and exceeds 1.5 K for higher optical thickness values (Figure 37). Subsequently, a fixed optical thickness of 0.1 was considered, with a variable vertical extent from 150 m to 1 km (Figure 38). In this case, the water vapor uncertainty increases from 2.5 % to 6 % and temperature uncertainty from 0.7 K to 1.8 K. In both cases, the error increases exponentially.

To test the performance in a more realistic way, the presence of clouds has




(a) (b)

Figure 37: Statistical uncertainty for water vapor (a) and temperature (b) considering a cloud at 3 km, a vertical extent of 150 m and with different values of optical thickness




(a) (b)

Figure 38: Statistical uncertainty for water vapor (a) and temperature (b) considering a cloud at 3 km, optical thickness $\tau = 0.1$ and different vertical extent values

been studied also using profiles of optical thickness and cloud fraction extracted from the GEOS-5 dataset. Specifically, two profiles were selected, with different characteristics.
The first one (1 January 2006, 06:13 UTC - 5.04°N, 1.21°W) shows two cloud formations, the first one centered at 6.5 km with a small optical thickness of $\tau = 0.1$ and a cloud fraction of 35 % and a second one with a higher optical thickness ($\tau = 6.5$) centered at 500 m and covering 70 %. The solar zenith angle is 113°, so no background is considered.
The second profile (1 January 2006, 07:33 UTC - 8.07°S, 2.33°W) shows three cloud formations: a combination of liquid and ice clouds centered at 10 km with a optical thickness of 3.1 and 9.2 respectively, but with a very small spatial extension (cloud fraction $\approx 1\%$), a second cloud layer covers 25 % of the grid with $\tau = 25$ centered at 6 km and finally a third cloud has a cloud fraction of 80 %, centered at 800 m and with $\tau = 2.5$. In this case the solar zenith angle is 90°, therefore the simulation includes the atmospheric contribution to the background.
Figure 39 shows input and retrieved water vapor and atmospheric temperature profiles, both obtained from a simulation with and without clouds. It is clear that the cloud located at 6.5 km has limited influence of the results, with an increase of the statistical uncertainty of about 2% and 0.1 K for water vapor and temperature respectively, while the bias is the same in both simulations. The second cloud layer determines a larger increase, with the statistical uncertainty rising from 3% to 5.5% and from 0.6 K to 1.25 K. The bias increases from 5% 10% and from 1 K to 2 K.
Figure 40 shows results obtained using the second selected profile. The higher cloud formation has no visible effect on the system performance, since the statistical uncertainty is too high also in the clear-sky simulation, but the small cloud fraction ensures that the performance below is still satisfying. The presence of the cloud at 6 km determines an increase of the statistical error from 6& to 8% and from 0.5 K to 0.6 K, while the bias remains around 10% and 1 K. The error increases again within the third cloud layer, with the uncertainty going from 3% to 15% and from 0.65 K to 1.8 K, and the bias within clouds rising to 15-20 % and 5 K.

Clouds contribute also to the background noise for sun zenith angles smaller than 90°, as explained in section 2.3.3. Besides the direct contribution due to cloud reflectivity, the presence of clouds modifies the atmospheric background, since the solar radiation is attenuated as well as the lidar signals, while surface background is considered only where no clouds are intercepted. This means that, in some cases, the performance can even improve above the cloud. As an example, Figure 41 shows the statistical uncertainty (without shot noise) obtained using input data from Geos-5 (1 January 2006 06:42 UTC, 17.8°N, 171°E) in which there is a single

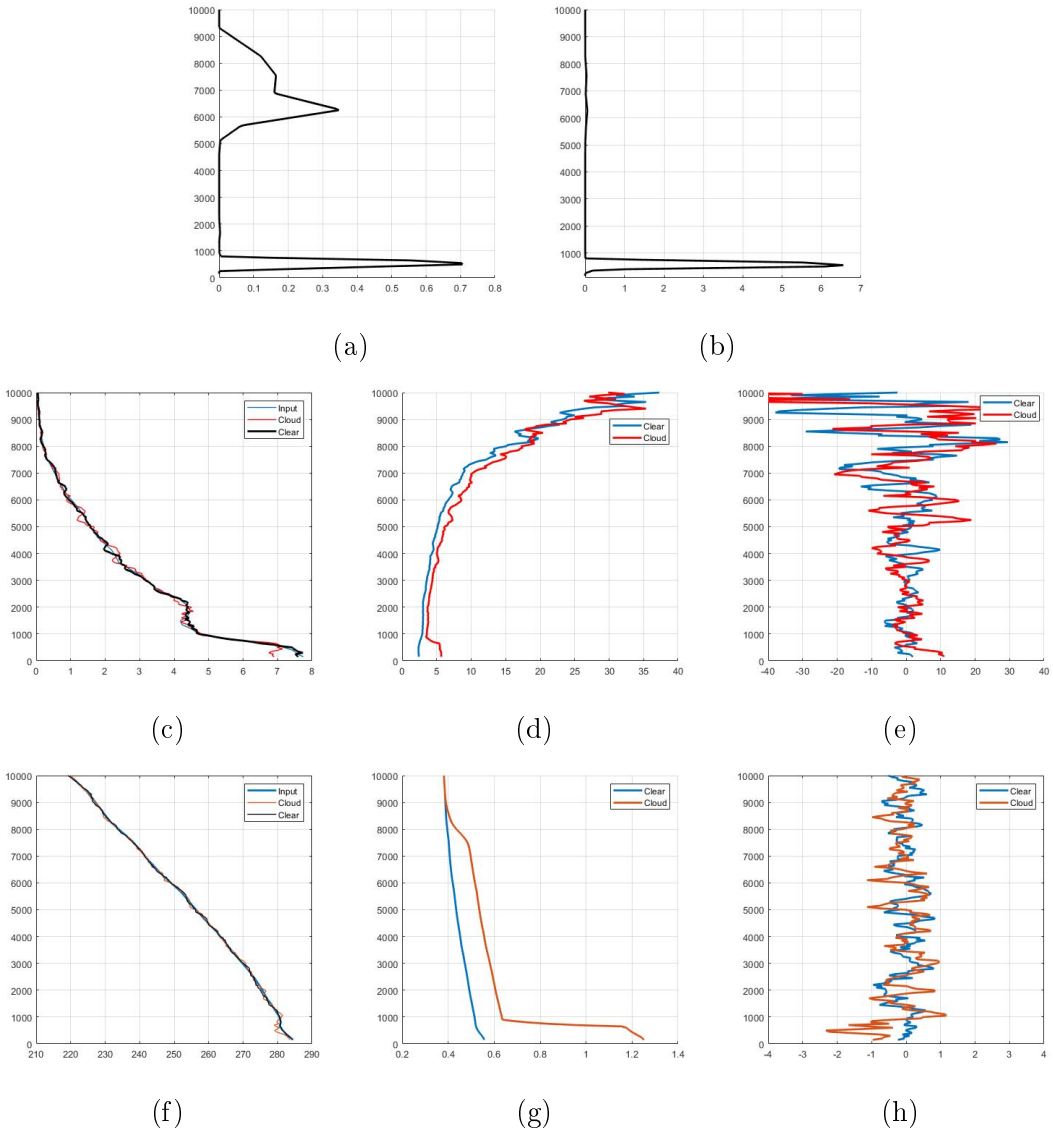



Figure 39: (a) Cloud fraction and (b) optical thickness profiles. Input and retrieved profiles of water vapor (c) and atmospheric temperature (f) together with their statistical uncertainty (d-g) and bias (e-h), with and without clouds

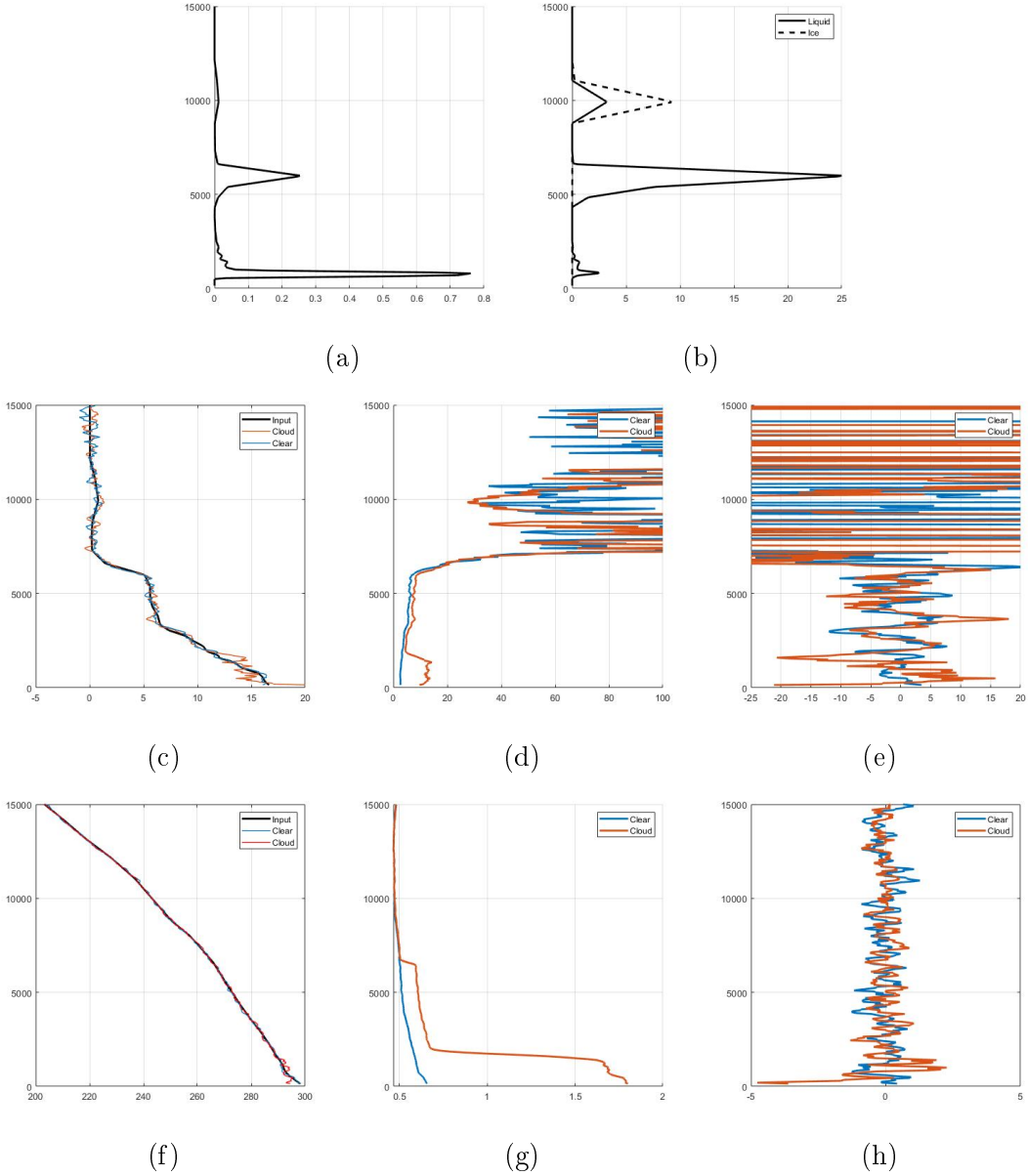
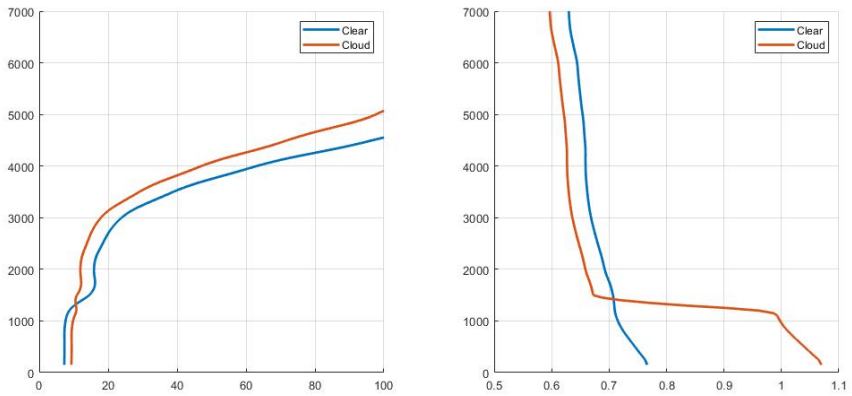



Figure 40: (a) Cloud fraction and (b) optical thickness profiles. Input and retrieved profiles of water vapor (c) and atmospheric temperature (f) together with their statistical uncertainty (d-g) and bias (e-h), with and without clouds




(a) (b)

Figure 41: Statistical error for water vapor (a) and temperature (b) with and without clouds. The presence of clouds can reduce the background and therefore the uncertainty.

cloud centered at 1 km with an optical thickness of 8.3 and covering 40% of the grid. The attenuation of the background due to the presence of clouds decreases the statistical uncertainty of 5 percentage point and 0.05 K.

### 3.1.5 Orbit simulations

In addition to the analysis of single profiles, simulations were also performed along entire orbits, to study the variability of performances for different latitudes and environmental conditions. Specifically a sun-synchronous dawn-dusk orbit was considered, with an altitude of 450 km and an eccentricity of 97.21° (Figure 42).

The input data, provided by NASA-GSFC, are extracted from the GEOS-5 Nature Run, Ganymed Release, a 2-year global, non-hydrostatic mesoscale simulation comprehending data for the period June 2005 through May 2007. The original dataset has a horizontal a resolution of 7 km (0.0625° × 0.0625°), 72 pressure levels up to 0,01 hPa (≈ 80 km) for each point at a given latitude and longitude and a temporal resolution of 30 minutes. For the simulations, a space-time interpolation along the orbit was considered, in order to get a vertical resolution of 50 m and a timestep of 1.0 s between successive profiles. The input data include thermodynamic and optical profiles, specifically air density ($kg/m^3$), pressure (Pa), atmospheric temperature (K), relative humidity (%), backscatter ($km^{-1}sr^{-1}$) and extinction ($km^{-1}$) coefficient at 355 nm, in-cloud optical thickness for liquid and ice clouds (a.u.) and surface albedo.

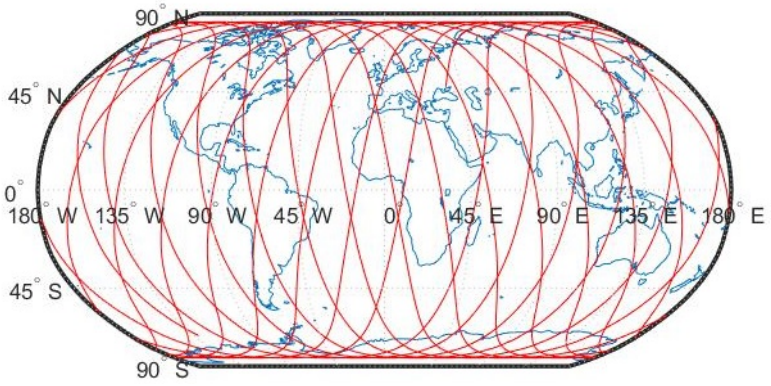



Figure 42: Dawn-dusk sub-orbital track

Three different days of the year were selected for the purpose of considering the variability of sun zenith angle, and therefore of the solar background, during the year (Figure 43): Day 0 (1 January 2006) with the largest variability (90°± 30°), Day 180 (28 June 2006) with an intermediate variability (90°± 16°) and Day 240 (27 August 2006) with a smaller variability (90°± 3°).

For each day, the simulation considers an entire orbit from 06:00 to 7:30 UTC. It should be noticed that Day 0 is a boreal winter day, therefore the southern hemisphere is located in the daylit side (Figure 44), instead Day 180 and Day 240 present the opposite situation.

This is reflected in Figure 45, which shows the estimated background at 355 nm from the three contributes - atmosphere, surface and clouds - along the orbit. The total amount of solar background is very different in the three situations: the atmospheric contribution is comparable for Day 0 and Day 180 ($\approx$ 3000 photons on average), but is reduced by about 5 times in Day 240, in which the satellite orbit keeps on the terminator, so the sun zenith angle is always near to 90°. The clouds contribution is obviously more variable, since it depends on their spatial distribution and season. The surface contribution is almost always negligible, but presents a high peak on the poles, since the ice has the higher surface albedo.

Figure 46, 47 and 48 show the statistical uncertainty and bias for water vapor mixing ratio and temperature along the orbit Day 0, 180 e 240 respectively, in a simulation including clouds. The statistical error is calculated from the standard deviation, considering an average profile obtained as a mean of 1500 shots (7.5 s, $\approx$ 50 km), while the vertical resolution is 100 m. In the orbit's dark side, the uncertainty is below 10 % up to 5 km in the tropical region, but with clear sky conditions it can reach up to 7-8 km. At mid latitudes, the strong presence of clouds worsen the average performance, which is on average characterized by uncertainties below 30 % up to 7 km and below 10 % up to 3 km in best conditions.

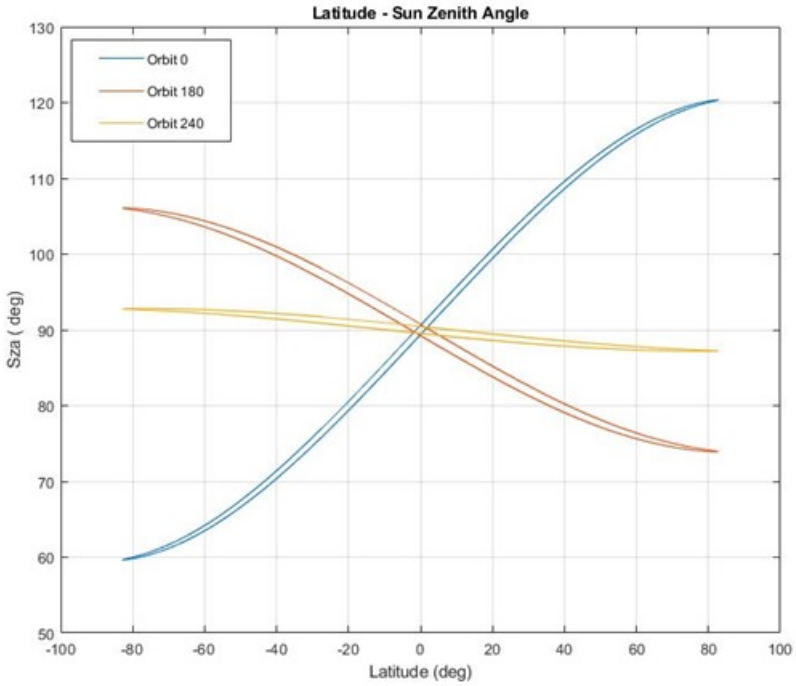



Figure 43: Variability of the solar zenith angle for the selected days

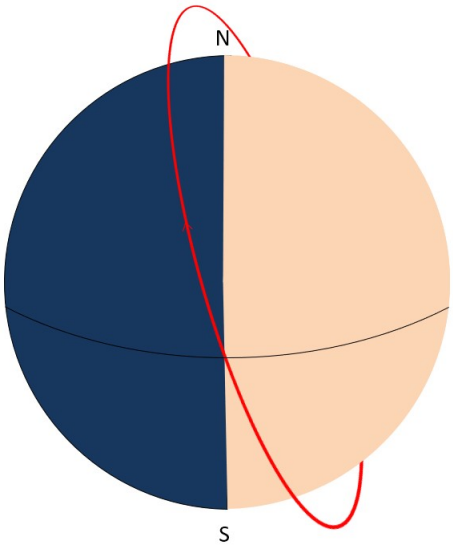
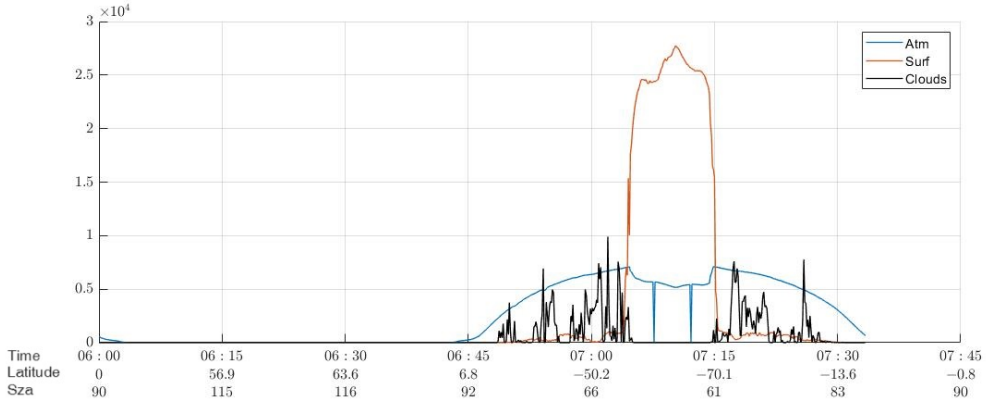
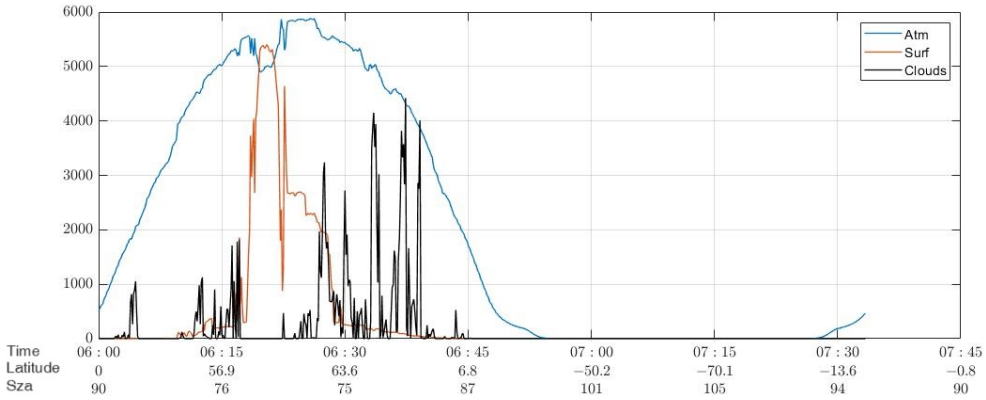


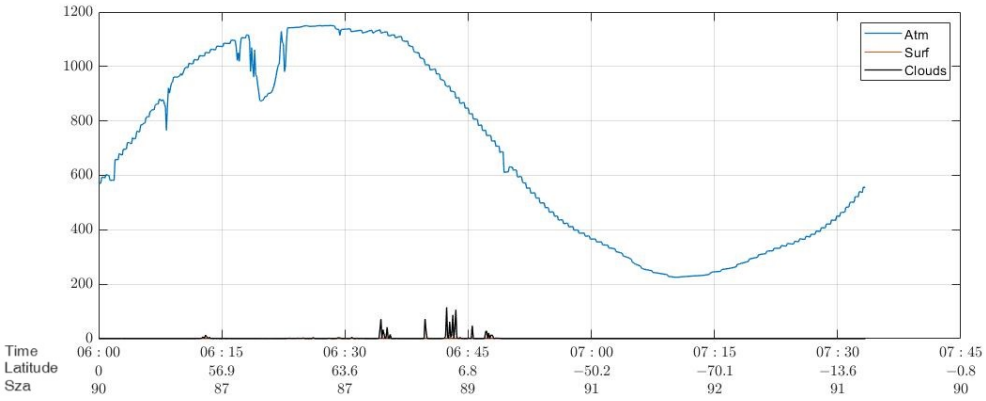

Figure 44: Schematic representation of the terminator line, that divides daylit and dark side of the planet




(a) Day 0




(b) Day 180




(c) Day 240

Figure 45: Estimated background along the orbit for different day of the year

In the sub-subarctic belt, the error is below 30 % up to 5 km. In the daylit side, the performance is is acceptable only in the tropics up to 5 km and partially in the mid latitudes (1-2 km). Panels (c-d) shows the average uncertainty calculated not including profiles with clouds. The relative bias (e-g) is on average $< 5\%$ up to 8-10 km in the darkside. In the daylit side, it is still acceptable up to 5 km. Only antarctic region is excluded since the bias exceed 20%.
For what concerns temperature, the statistical uncertainty is below 0.75 K up to 18 km with low background, while with high background it keeps $< 0.75$ K in the mid latitudes up to 15 km. The bias is always smaller than 0.5 K.

### 3.1.6 Results and further research

WMO requirements indicates for each variable the uncertainty, calculated in terms of root-mean-square-error (RMSE) from bias and statistical uncertainty combined in the root-mean square sense as [32]:

$$RMSE(z) = \sqrt{Bias^2(z) + \sigma^2(z)} \tag{3.1}$$

The uncertainty requirements are distinct in three different levels: the "threshold", that is the minim requirement, the "goal", an ideal requirement above witch no further improvements are required, and the "breakthrough", an intermediate level between threshold and goal which brings significant improvements for the targeted application. For specific humidity in global numerical weather prediction and real-time monitoring, the threshold is 10% with a horizontal resolution of 250 km and a vertical resolution of 1 km, the breakthrough is 5% with horizontal and vertical resolution of 50 km and 0.5 km, while the goal is 2%, 15 km and 0.3 km. The atmospheric temperature requirements indicate a threshold of 3 K, with a resolution of 500 km and 3 km; a breakthrough of 1K, 100 km and 1 km and finally a goal of 0.5 K, 15 km and 0.3 km. Figure 49 shows the percentage of water vapor mixing ratio and temperature measurements obtained from Day 0, which satisfy the observational requirements. On average, 35 % of water vapor measurements in the dark-night side reach the breakthrough level, especially in the tropics ($\approx 60\%$), but with good results also for mid latitudes (25%). Anyway, the threshold is reached on average in about 70% of the cases. A small portion of data ($5 - 10\%$) even manages to reach the goal threshold in the tropics. In the daylit side only the sub-antarctic zone exceeds the permitted error, while results are still acceptable for mid latitudes with good cloud conditions and for a 60% of tropical measurements.

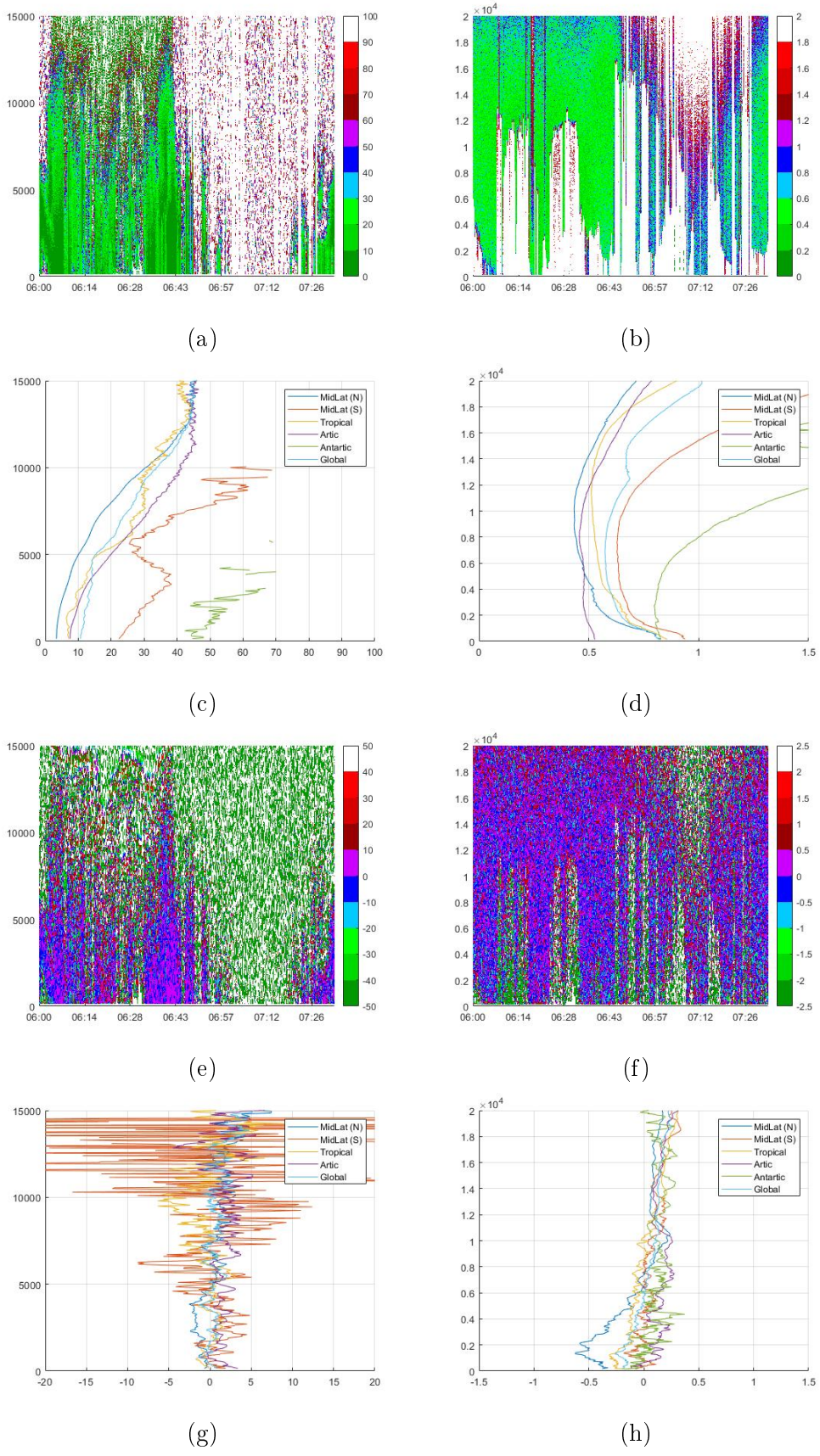



Figure 46: Performance simulation - Day 0: (a-c) Statistical uncertainty (%) and (e-g) relative bias (%) for water vapor mixing ratio; (b-d) statistical uncertainty (K) and (f-h) bias (K) for temperature

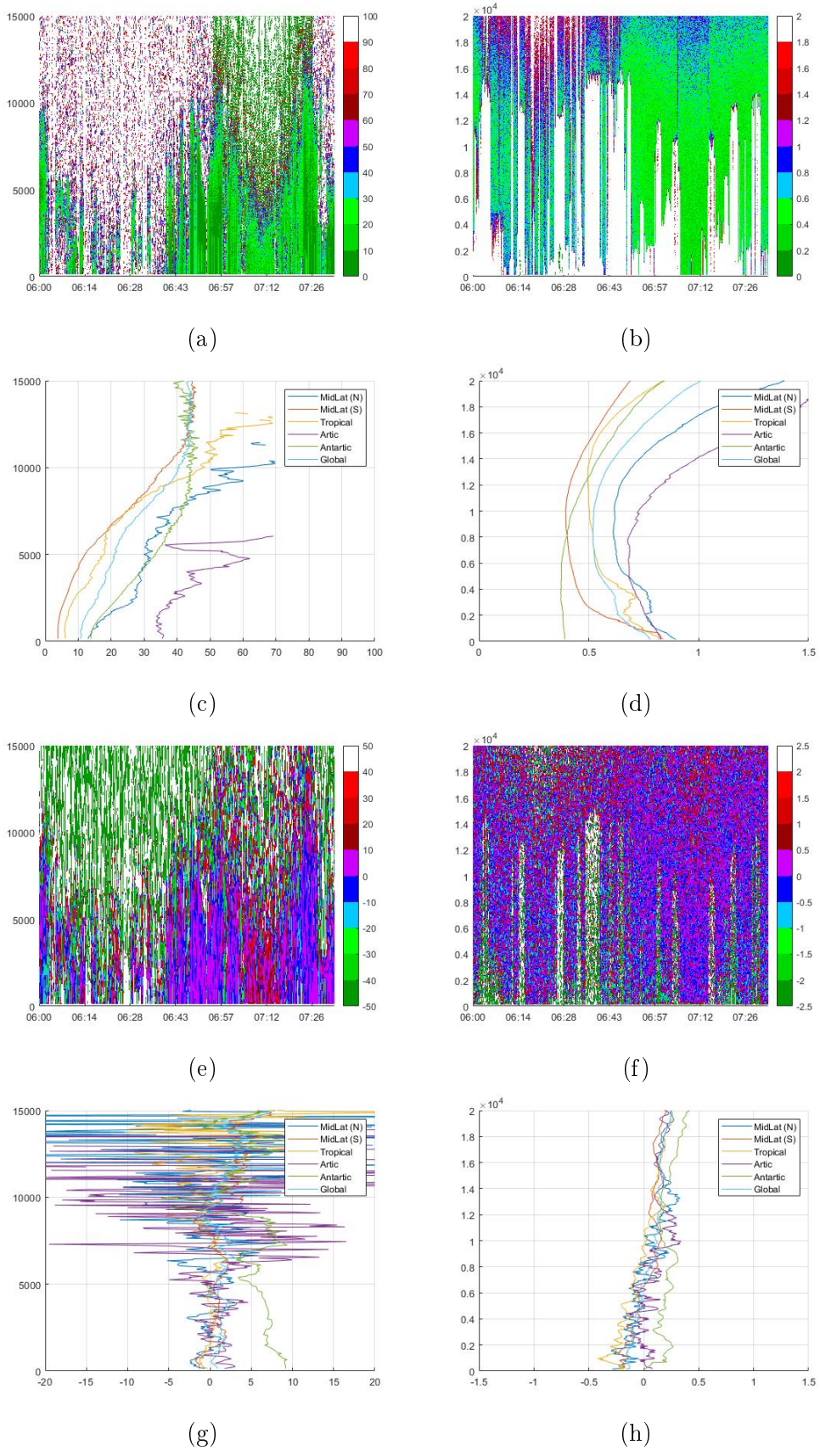



Figure 47: Performance simulation - Day 180: (a-c) Statistical uncertainty (%) and (e-g) relative bias (%) for water vapor mixing ratio; (b-d) statistical uncertainty (K) and (f-h) bias (K) for temperature

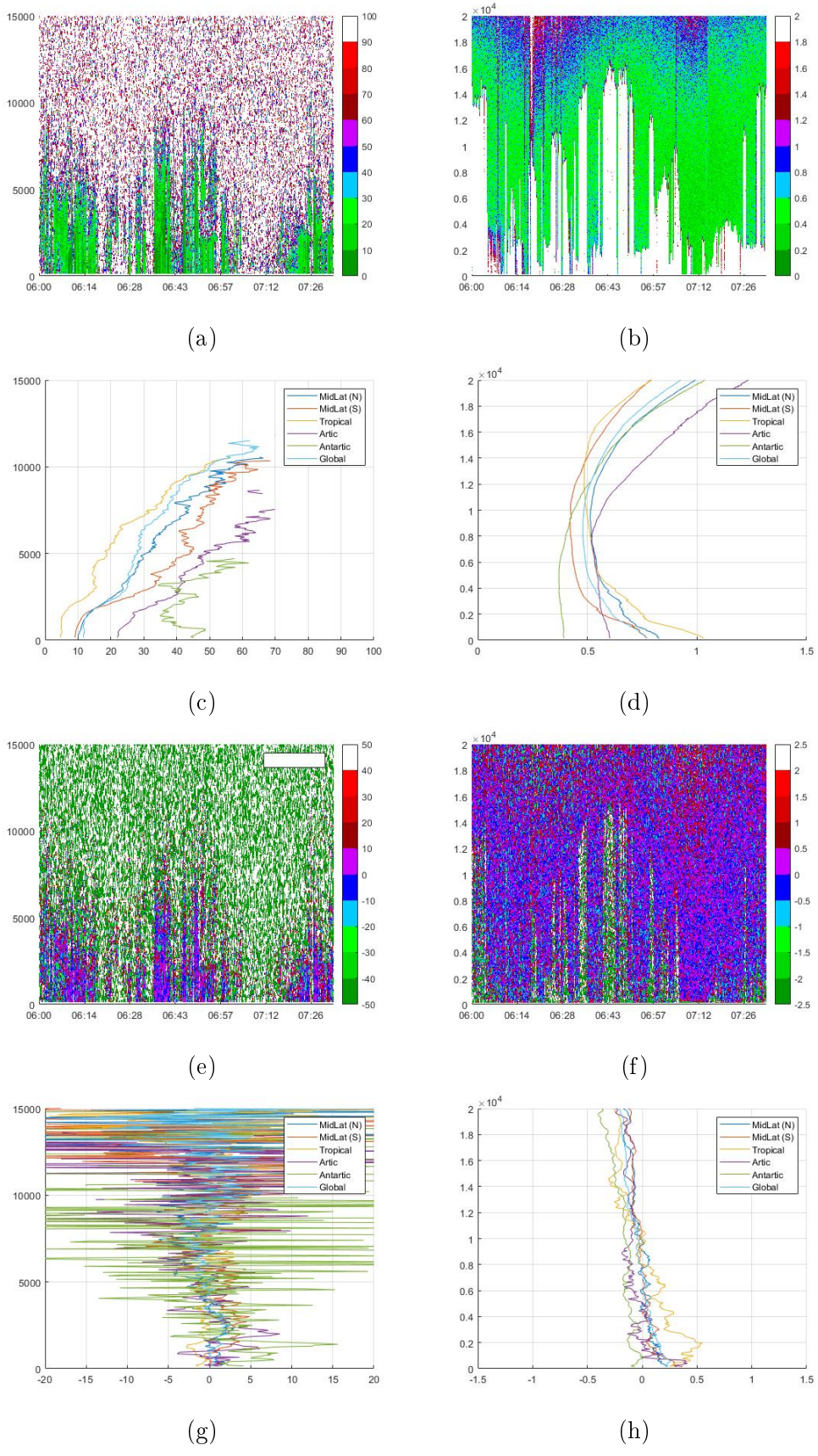
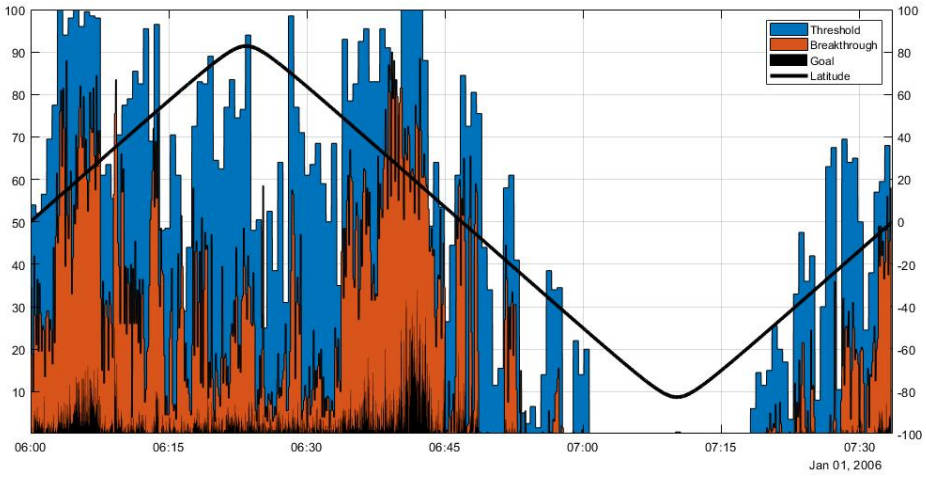
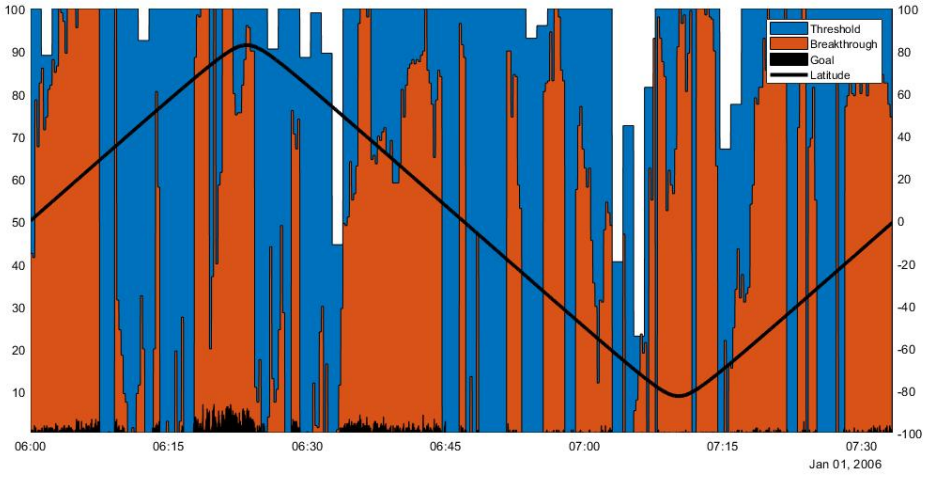



Figure 48: Performance simulation - Day 240: (a-c) Statistical uncertainty (%) and (e-g) relative bias (%) for water vapor mixing ratio; (b-d) statistical uncertainty (K) and (f-h) bias (K) for temperature




(a)




(b)

Figure 49: Percentage of data satisfying observational requirements (Day 0) for water vapor (a) and temperature (b)

For what concerns atmospheric temperature measurements, results are really encouraging: almost the totality of measurements meets the threshold, while the breakthrough is reached in 50% of cases.

All the results confirm the science impacting capabilities of this space-borne Raman lidar and are really encouraging for the continuation of the project. Specifically, the real impact that such measurements could have on forecasting, can be estimated with the development of an OSSE (Observation System Simulation Experiment): assimilating the simulated lidar signals or retrieved profiles of water vapor and temperature [33, 34] in a meteorological global model such as WRF-NOAHMP it is possible to estimate the increase of skill scores, comparing the output with and without assimilation of lidar data. For this purpose, input data for entire month (from 15 August to 15 September 2005) has been analyzed, in order to obtain simulated lidar signals and corresponding retrieved profiles.
A collaboration with the University of Hohenheim (Stuttgart - Germany) has been initiated to understand the feasibility of the OSSE and plan its realization. Specifically, the project will consist in three different simulations on WRF: a run without data assimilation, a control run with conventional observations and a final run with the assimilation of simulated lidar data in addition to the conventional observations.
As test project, a limited area model can be considered, consisting in 1600x1600x100 cells with a 2 km grid and focused on the hurricane Katrina (Figure 50). The project will subsequently be extended to a near-global domain, considering a forecast duration of three weeks and a resolution of 3 km that will require a computing time of approximately 14 million core-hours. The enormous amount of time and computational resources was not compatible with this research project, but the foundations for future collaboration were laid.

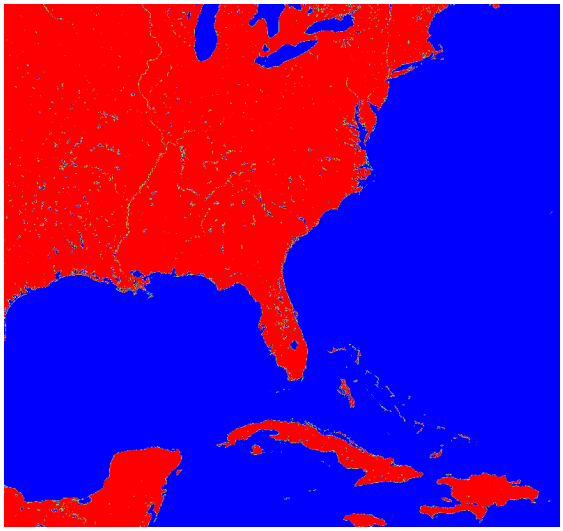


Figure 50: Local domain used for test

## 3.2 Caligola

CALIGOLA (Cloud Aerosol Lidar for Global Scale Observations of the Ocean-Land-Atmosphere System) is a mission funded by the Italian Space Agency (ASI), aimed at the development of a space-borne Raman Lidar. A Phase A study to assess the technological feasibility of the laser source and receiver system is currently underway at the Leonardo S.p.A, while scientific studies in support of the mission are conducted by the University of Basilicata. Scientific and technical studies are furthermore supported by other Italian institutions (CNR-ISMAR, CNR-IMAA), with NASA also having expressed an interest in contributing to the mission. Mission objectives include the observation of the Earth atmosphere, surface (ocean and land). Among the atmospheric objectives, the characterization of the global scale distribution of natural and anthropogenic aerosols, their radiative properties and interactions with clouds, and the measurements of ocean color, suspended particulate matter and marine chlorophyll. The expected performance of CALIGOLA has been assessed based on the application of the end-to-end simulator. Specifically, sensitivity studies have been carried out to define the technical specifications for the laser source, the telescope, the optics of transceiver, the detectors and the acquisition system.

### 3.2.1 Possible experimental setup

The space-borne Raman lidar CALIGOLA will exploit the three wavelengths emitted by a Nd:YAG laser source in the UV (354.7 nm), VIS (532 nm) and nIR (1064

nm) to obtain vertical profiles of backscatter and extinction coefficient, depolarization ratio and fluorescence. For this purpose, the receiver will collect the elastic signals at the three wavelengths, together with some reference signals obtained from appropriately selected signals. The technical specification of the lidar system are under definition, depending on results of a Phase A study conducted by Leonardo S.p.a. Actually, two possible configuration are being considered: a 50/50 splitting of elastic channels into total and polarized (p) channels or a 98/95 splitting through a polarizing beam splitter, with the acquisition of the two polarizations s and p. The best performance in terms of acquisition efficiency is obtained using the second solution, which however presents some additional difficulties for calibration. Table 5 shows a possible layout for Caligola, with the two proposed solutions, with some parameters still under investigation. For example, since the beam divergence is different at the three wavelengths, the field-of-view has been chosen as a possible trade-off, considering the possibility to capture all the signals and to reduce the received background. The hour of ascending node also depends on several mission considerations: moving the local time passage toward the dawn-dusk hour greatly reduces the solar background and improve the quality of atmospheric products. On the other side, a strong interest was shown by the scientific community for the oceanic objectives, with the possibility to improve the observation of ocean color, plankton reserves and diel vertical migration (DVM) of mesopelagic fish and these measurements can be done only considering a noon/midnight or near-noon orbit. Moreover, since Caligola will hopefully be launched in the temporal window 2030-2031, it would guarantee continuity to the data collected by Calipso, dismissed in September 2023, which had a local passage time at 13:30. It would be also beneficial to consider a synergic collaboration with other missions such as EarthCARE, which has a local passage time at 14:00 and it's planned to be launched in 2024. To overcame the degradation of atmospheric performance for noon-midnight orbits, it is possible to consider lower orbital heights (e.g. 350 km).

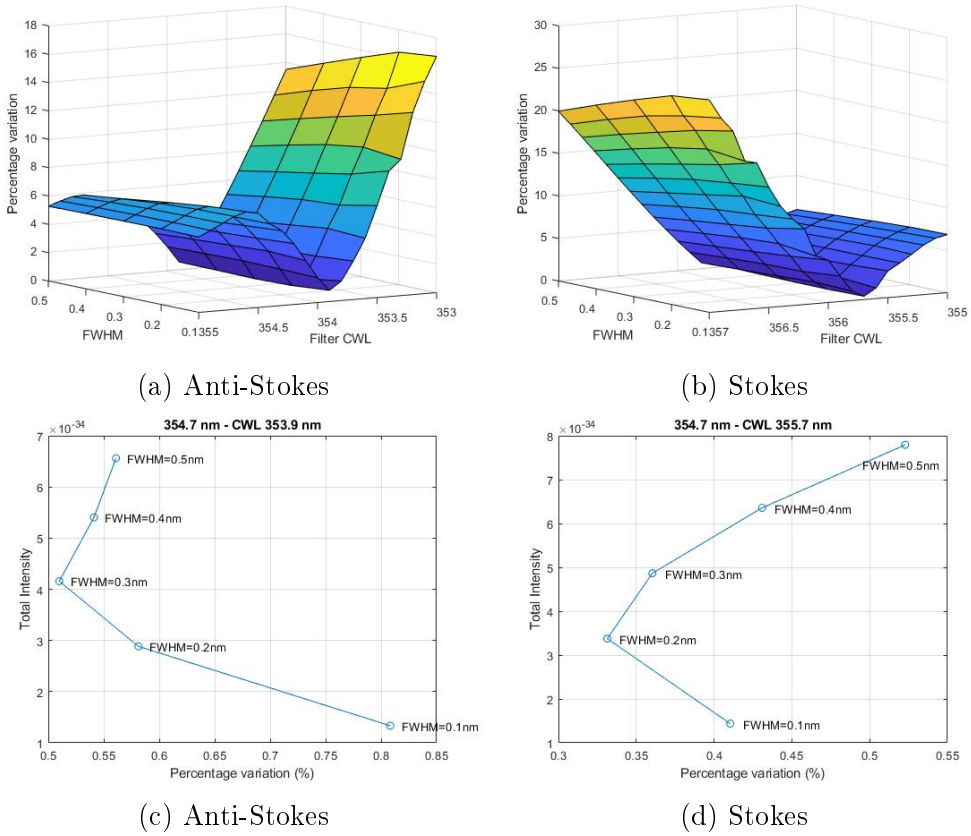
### 3.2.2 Choice of reference signal

The particle backscattering coefficient can be obtained from the ratio of the elastic signal to a reference signal, with the reference signal being useful for calibration. Both the roto-vibrational nitrogen signal or the rotational Raman echoes from nitrogen and oxygen can be used as reference signals. A sensitivity study was therefore carried out to select the rotational lines that simultaneously maximized the signal intensity and minimized its dependence from temperature. The total line intensity transmitted by an interferential filter has been calculated using equation described in 2.22, considering different combinations of central wavelength (CWL)

| Wavelength (nm) | 354.7 | 532 | 1064 |
|---|---|---|---|
| Pulse Energy (mJ) | 151 | 44 | 156 |
| Repetition rate (Hz) | 51 | | |
| Telescope diameter (m) | 1 | | |
| Telescope reflectivity (%) | 90 | 83 | 88 |
| Quantum efficiency (%) | 20 | 30 | 10 |
| Transmitter output divergence ($\mu$rad) | 20 | 33 | 53 |
| FOV ($\mu$rad) | 30* | | |
| Satellite height (km) | 400* | | |
| Hour of ascending node | 14:00* | | |

| Channel's overall efficiency | | | | |
|---|---|---|---|---|
| | Solution A | | Solution B | |
| Polarization | p | tot | p | s |
| Elastic 355 | 28 | 31 | 59 | 61 |
| Elastic 532 | 28 | 31 | 59 | 61 |
| Elastic 1064 | 37 | 88 | 73 | 80 |
| Reference 355 | 75 | | | |

Table 5: Possibile configuration for CALIGOLA. Parameters marked with (*) are modifiable based on mission considerations



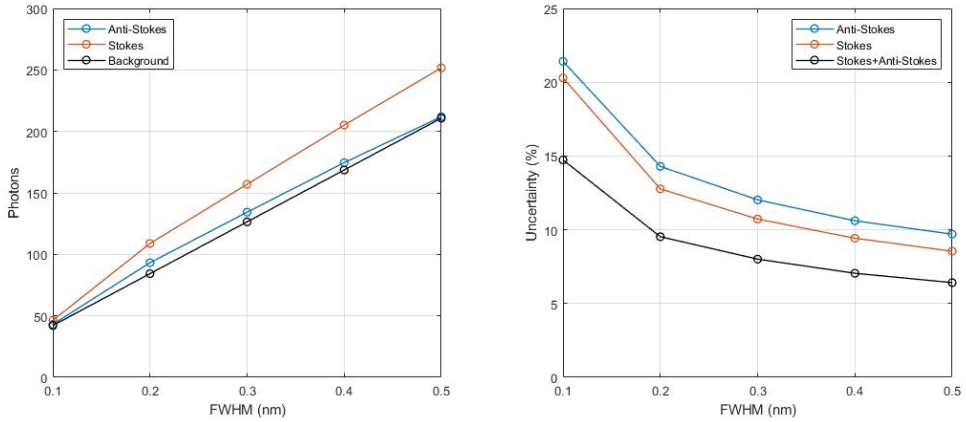

(a) Anti-Stokes (b) Stokes

(c) Anti-Stokes (d) Stokes

Figure 51: Percentage variation of total line intensity with temperature for different combinations of CWL and FWHM in the branch Anti-Stokes (a) and Stokes (b). Total line intensity and temperature sensibility for different FWHM at the selected CWL (c-d)

and filter width (FWHM), chosen so that the selected lines are far enough away from the elastic line.

Figure 51(a-b) shows the percentage variation of the transmitted line intensity calculated at two different temperature (270 K and 290 K), plotted against FWHM and CWL. Centering the filter at 353.9 nm and 355.7 nm for Anti-Stokes and Stokes branch, respectively, we obtain a signal with the minimal temperature sensitivity. In Figure 51(c-d), the total line intensity calculated with different filter widths is plotted as a function of the percentage variation. The width associated with the higher total intensity is in principle preferable, considering the slight differences in the intensity variability, but the increasing of the background noise must be taken in account. Figure 52 shows the estimated statistical error obtained considering




(a) (b)

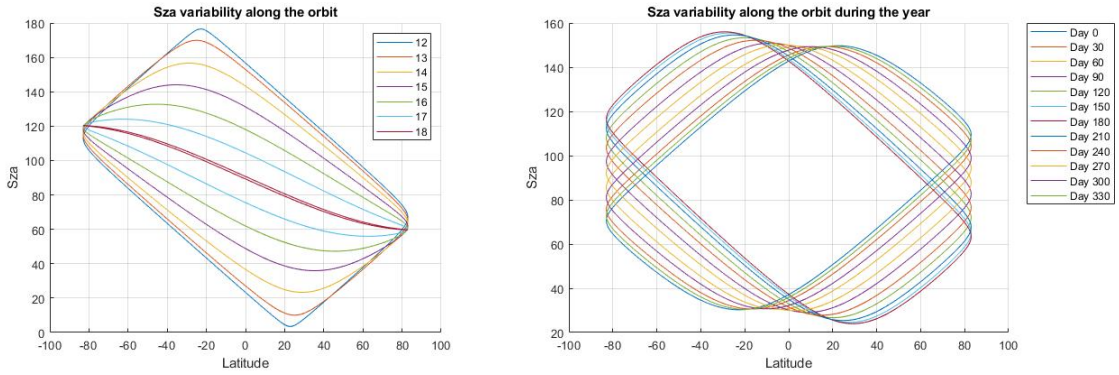
Figure 52: (a) Number of collected photons for Anti-Stokes, Stokes and background for an equatorial overpass with a local time at 14:00 UTC; (b) statistical uncertainty

as reference signal the Anti-Stokes, the Stokes or Stokes+Anti-Stokes, with an average value of the background for an equatorial overpass with a local time at 14:00 UTC. Even if the background increases linearly with the increase of the filter width, a 0.5 nm filter is preferable.

Analogous sensitivity studies have considered also for rotational signals at 532 and 1064 nm, showing that within the Anti-Stokes pure rotational branch, the minimum temperature sensitivity is reached with a filter centered at 530.1 nm and 1056.5 nm, respectively, but simulations show that the signals are too weak to be used with the specified experimental characteristics.
In the next months, a sensitivity study will be carried out to understand if it is possible also to consider rotational lines which are dependent on temperature if taken individually, but independent when added together. In this case, also temperature measurements would be possible.

### 3.2.3 Considerations on the local passage time

As already mentioned, the choice of the equator overpass time of the sun-syncronous orbit is a critical point to find the best trade-off between different mission objectives. To understand how much the local passage time influences the expected performance of Caligola, a dedicated study was conducted accounting for the variability of solar background.
Moving the local passage time produces an increase of the sun zenith angle vari-




(a) (b)

Figure 53: Sun zenith angle variability at different passage time

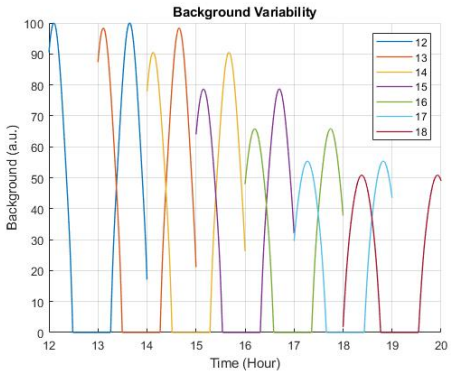



Figure 54: Normalized background for different passage time

ability, which on average changes from 60-120° for a dawn-dusk orbit to 0-180° for a noon-midnight orbit, with a variability of 25-155° for the orbit 14:00-02:00, as shown in Figure 53(a). On the other hand, the variability of the sun zenith angle during the year at a fixed hour decreases when moving to noon-midnight orbits. Figure 53(b) shows that the variability for the orbit 14:00-02:00 is ±6

Different values of sun zenith angles correspond to a different amount of solar background, which is clearly higher with overpass times near noon. Specifically, the background in a dawn-dusk orbit is about 50% of the amount estimated for a noon-midday orbit (Figure 54).

It should be mentioned that the overpass time refers only to the local hour at which the satellite cross the equator, but it has an intrinsic variability when moving to different latitudes of about ±1 hour up to a latitude of 60° N/S and ±4 hour considering also the polar regions (Figure 55). This means that there is a certain number of day and night observations in the high latitudes, that are

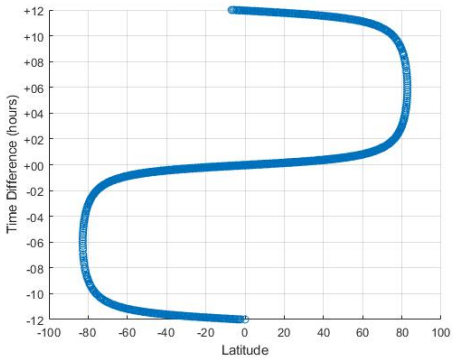



Figure 55: Difference in hour of local time at different latitudes, with respect to the local passage time on the equator

useful for oceanic objectives, also if we consider near dawn-dusk orbits. Figure 56 shows the percentage observations per day that are far enough from dawn/dusk hours, and therefore valid for oceanic purposes, when considering an equatorial local passage time at 14:00 and 18:00. A measurement is considered a "day observation" if the sun zenith angle is less than 75°, while is a "night observation" if it's higher than 105°. Of course, the dawn-dusk orbit is less efficient, since the valid observations are concentrated only in a half of the year, while the near-noon orbit has observations almost evenly distributed.

### 3.2.4 Choice of the FOV

The performance of Caligola was estimated considering different values for the field-of-view, from 20 to 60 $\mu$rad. In fact, the different output divergence among the three wavelengths makes it necessary to consider a trade-off between the possibility to collect the largest possible fraction of the lidar echoes and the need to reduce as much as possible the background. Table 6 lists the statistical uncertainty affecting the different measured signals, considering the experimental configuration of Caligola with the splitting of elastic signals into parallel (p) and cross (s) components (Solution B). The elastic signal at 355 has a divergence of 20 $\mu rad$, the beam is always entirely intercepted by the telescope field of view and the background noise increases with increasing FOVs, without any increase of the lidar echo intensity. However, both the components are strong enough to reach good results for all the considered cases. The elastic signal at 532 nm has a divergence of 33$\mu rad$, so it's slightly cut with the 30$\mu rad$ FOV. The p component is again strong enough, while the s component is weaker and has a statistical uncertainty below 30% only up to a FOV of 40$\mu rad$, thanks to the reduced background.

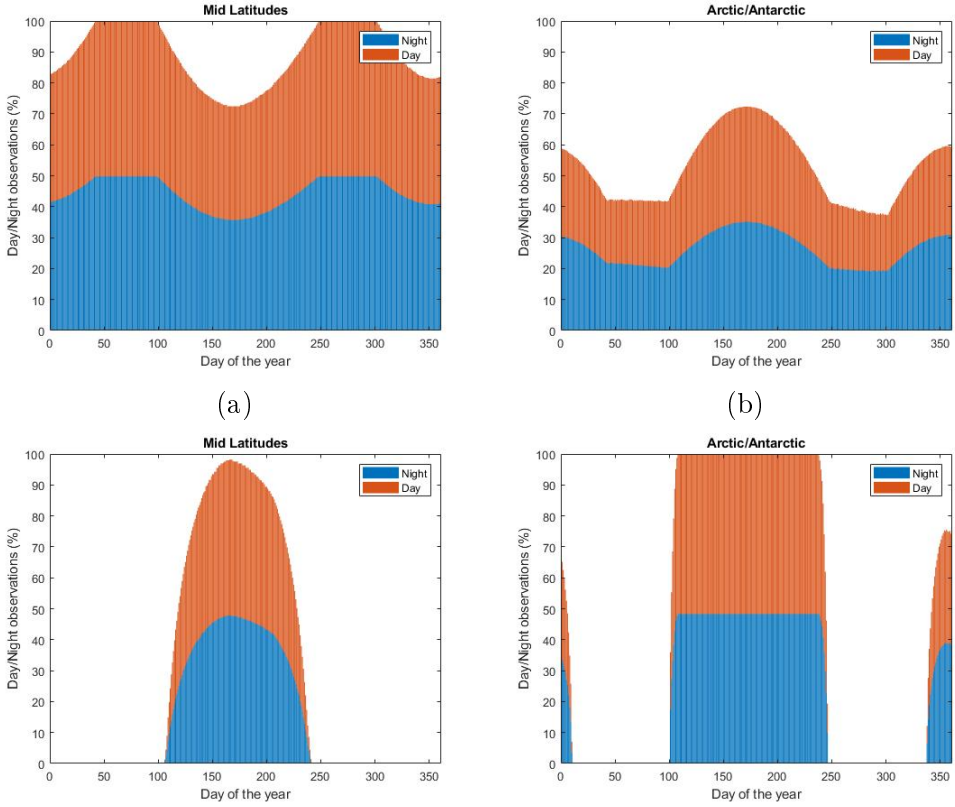



Figure 56: Number of morning/night observations considering a local passage time at 14:00 (a-b) or 18:00 (c-d)

| FOV / Ch. | $20\mu rad$ | $30\mu rad$ | $40\mu rad$ | $50\mu rad$ | $60\mu rad$ | Best |
|---|---|---|---|---|---|---|
| 355 (p) | Err $<$ 10 % up to 45 km | | | | | All |
| 355 (s) | Err $<$ 15 % up to 10-15 km | | | | | All |
| 532 (p) | Err $<$ 15 % up to 28-30 km | | | | | All |
| 532 (s) | $<$ 35% up to 2.5 km | $<$ 30% up to 2.5 km | $<$ 30% up to 2.5 km | $<$ 40% up to 2.5 km | $<$ 50% up to 2.5 km | 20-40 |
| 1064 (p) | $<$ 20% up to 18 km | $<$ 20% up to 24 km | $<$ 20% up to 28 km | $<$ 20% up to 30 km | $<$ 20% up to 30 km | 30-60 |
| 1064 (s) | $<$ 30% up to 2.5 km | $<$ 30% up to 3 km | $<$ 30% up to 4 km | $<$ 30% up to 5 km | $<$ 30% up to 5 km | 30-60 |
| Rot (S+AS) | $<$ 10-15% up to 10 km | $<$ 10-20% up to 10 km | $<$ 10-30% up to 10 km | $<$ 10-35% up to 10 km | $<$ 10-40% up to 10 km | 20-40 |
| Rot (AS) | $<$ 12-22% up to 10 km | $<$ 35% up to 10 km | $<$ 40% up to 10 km | $<$ 40% (only h18-16) | $<$ 40% (only h18) | 20-30 |

Table 6: Statistical uncertainty calculated considering different FOVs

The signal at 1064 nm is the only one which has better performance with larger FOVs, since its divergence is 53 $\mu rad$ and the solar background is usually less at this wavelength. The most problematic channel is the rotational signal used as reference signal. Using only the Anti-Stokes branch gives good results only using small FOVs. With larger values, the performance are acceptable if a dawn-dusk orbit is considered, so that the solar background is automatically reduced. On the other hand, the solution with the Stokes+Anti-Stokes reference signal gives good results, which are obviously better with small FOVs.
Therefore, taking into account all the signals, the best trade-off is found for a 30 $\mu rad$ FOV, for which the background is small enough and only the 1064 nm is slightly penalized.
Similar results were found using the split into total and p component (Solution A), with, again, the 30 $\mu rad$ FOV being the best compromise.

### 3.2.5 Performance estimation

The performance estimation of Caligola was simulated considering the experimental parameters listed in table 5 with a 30 $\mu$rad FOV and the split between parallel and cross component. Analogous simulations have been performed using the total/parallel split, but the decreased efficiency lead to support the other solution. The signal uncertainty was calculated assuming a Poisson statistics and considering different values of the solar background, calculated using sun zenith angles from 10 (bright day) to 120 (dark night). For the background contribution, the degree of polarization of skylight was taken into account. In fact, assuming a Rayleigh sky model, the light is more polarized in the direction p, the more the zenith angle is near 90° (Figure 57), with an average maximum of 80%. This implies that for the estimation of background in direction p, the skylight is entirely transmitted by the polarizing filter for a percentage equal to the polarization degree. The complementary percentage, which represents the unpolarized portion, is transmitted by 50% for the Malus law. Vice versa, for the cross signals the amount of polarized radiation is completely removed, while the unpolarized portion is transmitted by 50%.
The particle backscatter coefficient profile used for the simulation is extracted from the ARMA aerosol model and showed in Figure 58, considering the total coefficient and the two polarizations. It should be noticed that the cross component of the molecular coefficient is usually very small (0.5 %), while for the particle coefficient an average depolarization ratio was used (30 %). This implies that, for this simulation, the intensity of the cross signal decreases above 5 km, where the particle coefficient drops from $10^{-6}$ to $10^{-9} m^{-1} sr^{-1}$. The intensity of the cross signal has another reduction above 15 km and is very weak above 30 km. For this reason, the statistical uncertainty for the cross components is reported only up to this height.
Figure 59 shows the received signals in terms of photon numbers, together with their statistical uncertainty (%). The elastic signal at 355 nm shows very good results in both day and night conditions, with a statistical error smaller than 10% up to 50 km for the p component and up to 5 km for the s component, which rises to 20 % up to 18-20 km, depending on the background.
The 532 nm parallel signal has an uncertainty less than 10% up to 30 km in both day and night simulations, above this height the background slightly degrades the performances. The cross signal is weaker at this wavelength, showing an uncertainty $<$ 20% up to 5 km, while up to 15 km results are acceptable in low background conditions, but degrades very quickly in day simulations.
For the elastic signal at 1064 nm, the background has less influence, so there is less difference between day and night simulations. The parallel component has an error $<$20% up to 25 km and $<$30% up to 30 km, while for the cross component is

<30% only up to 4 km.
The simulations of the rotational signals confirm the need to consider both the Stokes and Anti-Stokes branch to increase the total intensity. With low background, the uncertainty is below 20% up to 25 km, while with high background is below 20% up to 10-15 km.
Propagating the error through equation 2.73, it is possible to estimate the statistical uncertainty of the backscatter coefficient at 355 nm. This quantity is dependent not only on the uncertainty affecting the elastic signals, but also on the uncertainty affecting the rotational reference signal. To achieve better results, a coarser spatial resolution of 150 km, corresponding to a time integration of approximately 20 s, is considered. The total particle backscatter coefficient is calculated using the total elastic signal, obtained from the sum of parallel and cross-polarized elastic signals, while the parallel and cross backscatter coefficients are derived from the single polarized signals. It should be noticed that the uncertainty associated to the total backscatter, i.e., the sum of molecular and particle contribution, depends on the uncertainty of the elastic and reference signals, but when considering only the particle backscatter, the error is amplified of a factor $R/(R-1) = \beta_{tot}/\beta_{par}$. Specifically, considering an average particle depolarization ratio of $\delta_{par} = 0.30$, the ratio $\beta^{\parallel}_{tot}/\beta^{\parallel}_{par} \approx 4$, while $\beta^{\perp}_{tot}/\beta^{\perp}_{par} \approx 1$. Simulations shows that the statistical uncertainty is about 20% in night-time conditions for parallel and total components (Figure 60a-c), which rises respectively up to 40 and 50 % with higher background. The cross component (Figure 60b) shows better results, keeping the uncertainty below 10-15 % up to 2 km. Consequently, the depolarization ratio (d) has an uncertainty of 30-60 %, based on the background conditions. The rotational signal can be used to retrieve also extinction coefficient. Considering a horizontal resolution of 200 km and a vertical resolution of 500 m, it is possible to obtain an uncertainty of about 35 % during the night (e-f). Figure 60 shows the statistical uncertainty on parallel, cross and total particle backscatter coefficient and on depolarization ratio, obtained as a sum of the two uncertainties.
The obtained results are encouraging about the capability of Caligola. Further studies will be considered to estimate also the contribution of lidar echoes from the underwater. Moreover, dedicated studies are ongoing to understand if it is possible to introduce two new channels at 460 and 680 nm, dedicated to measure the fluorescence from marine chlorophyll and atmospheric aerosols, for the characterization of ocean primary production and aerosol typing [35, 36] .

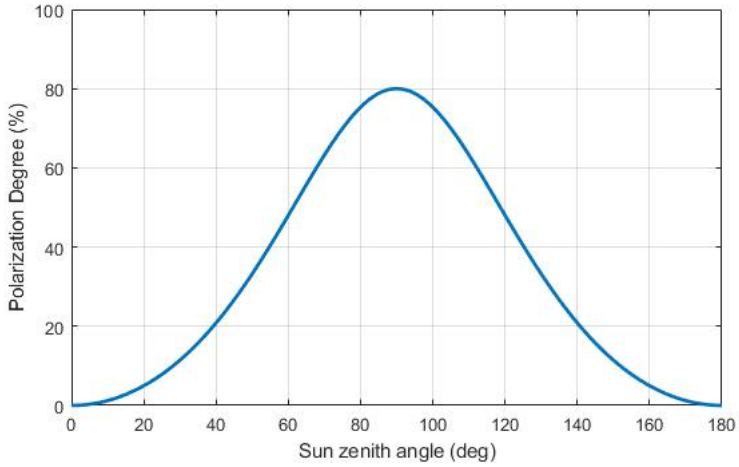



Figure 57: Polarization degree as a function of the sun zenith angle

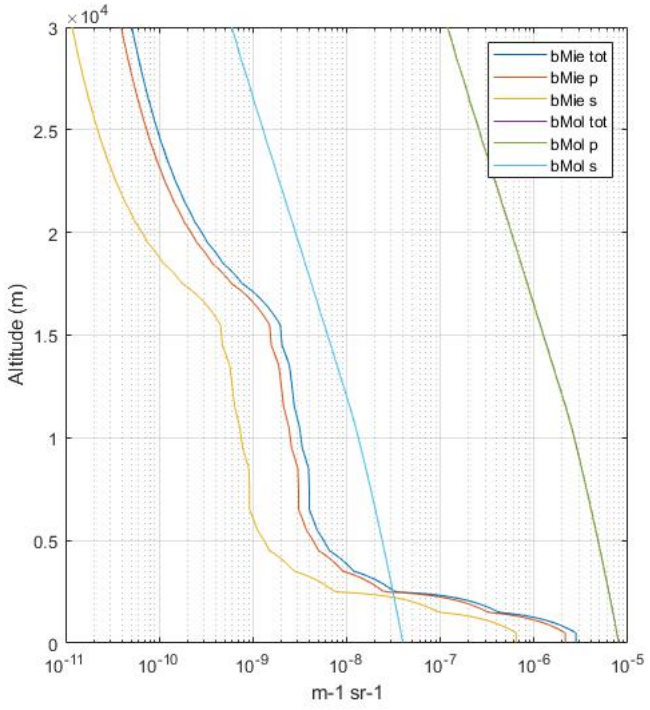



Figure 58: Particle and molecular backscatter coefficient profiles, considering the parallel and cross components

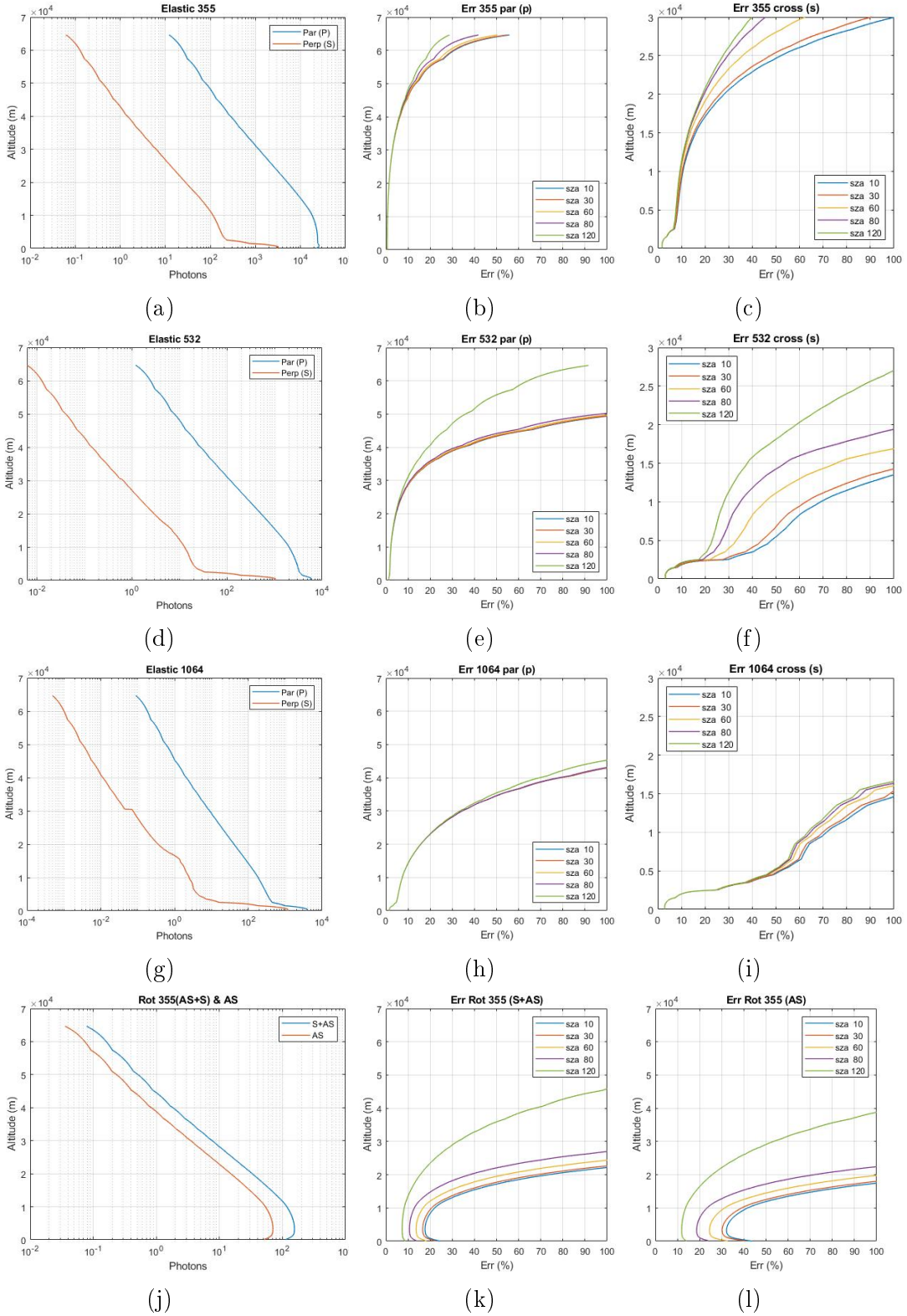



Figure 59: Performance estimation of Caligola: elastic signals (parallel (p) and cross (s) component) at 355 (a), 532 (d) and 1064 (g), together with the statistical uncertainty (%) considering different background values. As reference signal was considered the Anti-Stokes and the sum of Stokes and Anti-Stokes rotational signal (j).

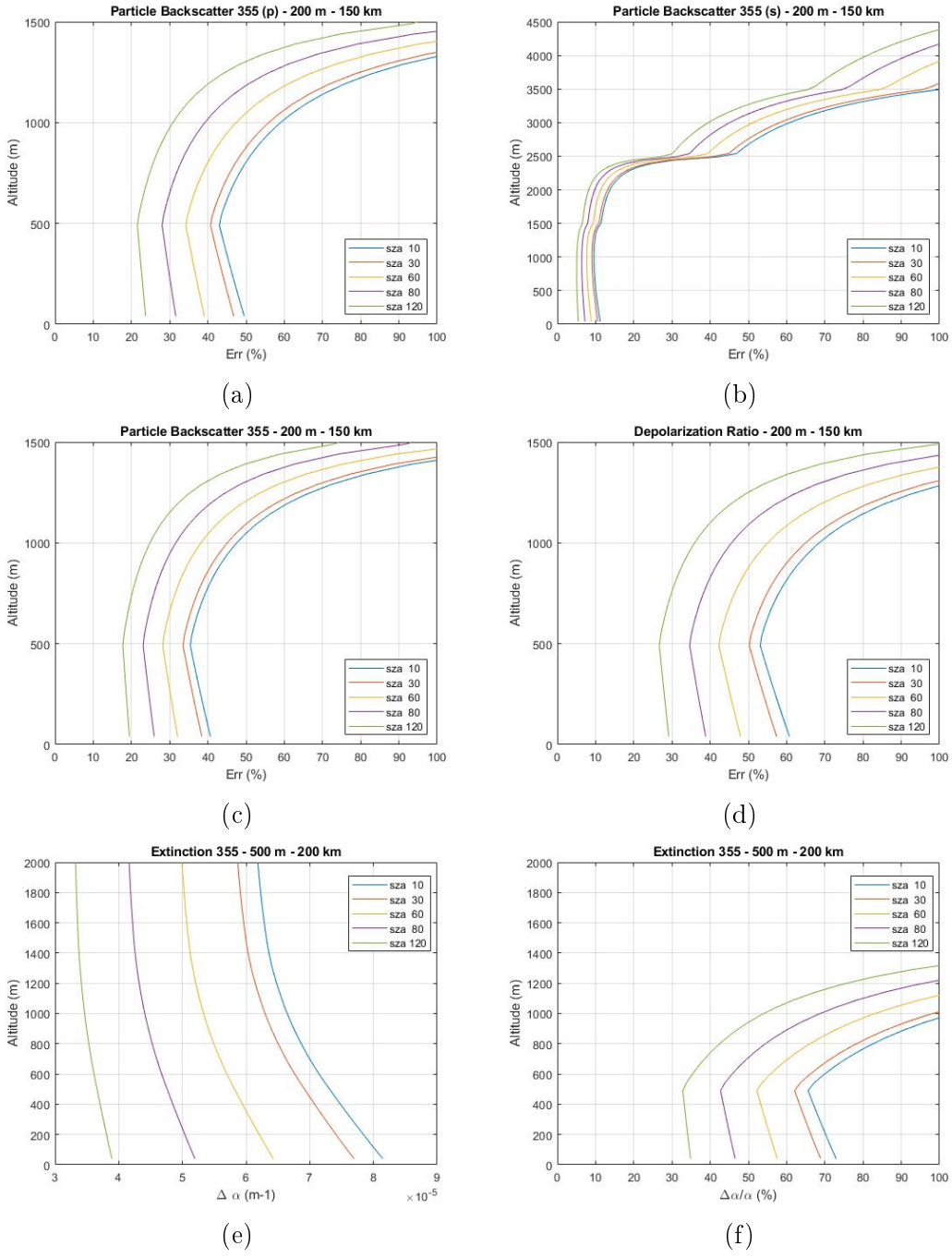



Figure 60: Statistical uncertainty (%) of parallel (a), cross (b) and total (c) particle backscatter coefficient together with the depolarization ratio (c).
(e) Absolute (m-1) and (f) relative (%) statistical uncertainty of extinction.

# Conclusions

The numerous applications showed in this work are a compelling demonstration of the enormous capabilities of end-to-end performance models. Thanks to a deep theoretical background and rigorous testing phase, the tool has demonstrated its efficacy in providing accurate and reliable assessments of Raman Lidar systems under various conditions. The simulator's ability to emulate the interactions between laser pulses and atmospheric constituents and to reproduce realistic atmospheric scenarios and various environmental conditions represents a valuable resource to optimize instrument configurations, estimate its performances and enhance the overall effectiveness of Raman Lidar deployments. Furthermore, the integration of user-friendly interfaces and visualization capabilities ensures a great usability and make it a fast and practical instrument in the development and decision-making phases.

The simulator tool has proven its utility in several projects, especially Atlas and Caligola, focused on the development of space-borne Raman Lidars. Dedicated studies have been carried out to obtain simulations tailored to meet specific mission requirements, for example the elaboration of a model for a better estimation of solar background, especially for lidars deployed on dawn-dusk orbits. The great flexibility of the developed performance simulators was also suited to realize sensitivity studies, e.g., for the selection and optimization of Lidar channels, to find the best trade-off between different experimental configurations and necessities, to select orbital parameters etc.

The modular configuration allows also an expansion of the simulation capabilities, based on incoming requests and challenges. In the next future, the simulator will include the estimation of fluorescence backscatter signals and roto-vibrational $CO_2$ signal; moreover, it will be integrated with a model for the estimation of marine echoes' contribution. Further objectives include the possibility to realize an OSSE (Observing System Simulation Experiment) to evaluate the impact of observations from a space-borne Raman Lidar on the quality and accuracy of forecasts, assimilating the simulated signals obtained through the simulator in a meteorological global model. While only briefly reported in the present dissertation, an important

activity was carried out in the frame of my PhD activities in the preparation of an adequate dataset to be assimilated in a modified version of the mesoscale model WRF (Weather Research and Forecasting). The inherent technical challenges and the significant computational time required make this project quite demanding and was consequently not finalized during the PhD activities. However, the groundwork for this project has already been laid during this period, and it is hoped that it can be realized in the future.
With the potential it holds, the simulator could be useful in many other projects and developments, contributing in the advancement of scientific research and environmental understanding.

# Bibliography


[1] S. H. Melfi, "Remote sensing for air quality management," in *Laser Monitoring of the Atmosphere* (E. D. Hinkley, ed.), pp. 9–28, Springer Berlin Heidelberg, 1976.

[2] R. T. H. Collis and P. B. Russell, "Lidar measurement of particles and gases by elastic backscattering and differential absorption," in *Laser Monitoring of the Atmosphere* (E. D. Hinkley, ed.), pp. 71–151, Springer Berlin Heidelberg, 1976.

[3] U. Wandinger, "Raman lidar," in *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere* (C. Weitkamp, ed.), pp. 241–271, Springer New York.

[4] Licel, "Transient recorder - data sheet." https://licel.com/transdat.htm. Accessed: 2023-10-05.

[5] P. Di Girolamo, N. Franco, M. Di Paolantonio, D. Summa, and D. Dionisi, "Atmospheric thermodynamic profiling through the use of a micro-pulse raman lidar system: Introducing the compact raman lidar marco," *Sensors*, vol. 23, no. 19, 2023.

[6] C. Flamant, P. Chazette, O. Caumont, P. Di Girolamo, A. Behrendt, M. Sicard, J. Totems, D. Lange, N. Fourrié, P. Brousseau, C. Augros, A. Baron, M. Cacciani, A. Comerón, B. De Rosa, V. Ducrocq, P. Genau, L. Labatut, C. Muñoz-Porcar, A. Rodríguez-Gómez, D. Summa, R. Thundathil, and V. Wulfmeyer, "A network of water vapor raman lidars for improving heavy precipitation forecasting in southern france: introducing the WaLiNeAs initiative," *Bulletin of Atmospheric Science and Technology*, vol. 2, no. 1, p. 10, 2021.

[7] D. Summa, P. Di Girolamo, N. Franco, I. Gandolfi, M. Di Paolantonio, M. Rosoldi, F. Madonna, A. Giunta, and D. Dionisi, "Latent flow measurement by Wind Lidar and Raman Lidar during WaLiNeas campaign," in *EGU*

*General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, pp. EGU–5753, May 2023.

[8] P. Di Girolamo, A. Scoccione, M. Cacciani, D. Summa, B. De Rosa, and J. H. Schween, "Clear-air lidar dark band," *Atmospheric Chemistry and Physics*, vol. 18, no. 7, pp. 4885–4896, 2018.

[9] World Meteorological Organization, *Guide to meteorological instruments and methods of observation*. World Meteorological Organization, 2008. OCLC: 928941505.

[10] M. Capderou, "Motion of orbit, earth and sun," in *Handbook of Satellite Orbits: From Kepler to GPS*, pp. 245–300, Springer International Publishing, 2014.

[11] ESA, "ARMA reference model of the atmosphere." Technical Report APP-FP/99–11239/AC/ac (European Space Agency).

[12] H. Inaba, "Detection of atoms and molecules by raman scattering and resonance fluorescence," in *Laser Monitoring of the Atmosphere* (E. D. Hinkley, ed.), Topics in Applied Physics, pp. 153–236, Springer, 1976.

[13] V. Griaznov, I. Veselovskii, P. Girolamo, M. Korenskiy, and S. Donato, "Spatial distribution of doubly scattered polarized laser radiation in the focal plane of a lidar receiver," *Applied optics*, vol. 46, pp. 6821–30, 10 2007.

[14] C. M. Penney, R. L. S. Peters, and M. Lapp, "Absolute rotational raman cross sections for n2, o2, and co2," *Journal of the Optical Society of America*, vol. 64, pp. 712–716, May 1974.

[15] A. Behrendt and T. Nakamura, "Calculation of the calibration constant of polarization lidar and its dependency on atmospheric temperature," *Opt. Express*, vol. 10, pp. 805–817, Aug 2002.

[16] P. Di Girolamo, A. Behrendt, and V. Wulfmeyer, "Spaceborne profiling of atmospheric temperature and particle extinction with pure rotational raman lidar and of relative humidity in combination with differential absorption lidar: performance simulations," *Applied Optics*, vol. 45, no. 11, pp. 2474–2494, 2006. Publisher: Optical Society of America.

[17] A. A. Kokhanovsky, "Simple approximate formula for the reflection function of a homogeneous, semi-infinite turbid medium," *Journal of the Optical Society of America*, vol. 19, no. 5, pp. 957–960, 2002. Publisher: Optical Society of America.

[18] A. Kokhanovsky, "Optical properties of terrestrial clouds," *Earth-Science Reviews*, vol. 64, no. 3, pp. 189–241, 2004.

[19] H. Van de Hulst, *Multiple Light Scattering: Tables, Formulas, and Applications*. Elsevier, 1980. Google-Books-ID: xeAkNn1ZBIwC.

[20] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*. John Wiley & Sons, 2008. Google-Books-ID: ib3EMXXIRXUC.

[21] L. C. Henyey and J. L. Greenstein, "Diffuse radiation in the galaxy," *The Astrophysical Journal*, vol. 93, p. 70, 1941.

[22] C. Bohren and E. Clothiaux, *Polarization: The Hidden Variable*, ch. 7, pp. 345–395. John Wiley and Sons, Ltd, 2006.

[23] Z. Liu, W. Hunt, M. Vaughan, C. Hostetler, M. McGill, K. Powell, D. Winker, and Y. Hu, "Estimating random errors due to shot noise in backscatter lidar observations," *Applied Optics*, vol. 45, no. 18, p. 4437, 2006.

[24] P. Di Girolamo, A. Behrendt, C. Kiemle, V. Wulfmeyer, H. Bauer, D. Summa, A. Dörnbrack, and G. Ehret, "Simulation of satellite water vapour lidar measurements: Performance assessment under real atmospheric conditions," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1552–1568, 2008.

[25] P. Di Girolamo, A. Behrendt, and V. Wulfmeyer, "Space-borne profiling of atmospheric thermodynamic variables with raman lidar: performance simulations," *Optics Express*, vol. 26, no. 7, pp. 8125–8161, 2018. Publisher: Optical Society of America.

[26] J. D. Klett, "Lidar inversion with variable backscatter/extinction ratios," *Applied Optics*, vol. 24, no. 11, p. 1638, 1985.

[27] F. G. Fernald, B. M. Herman, and J. A. Reagan, "Determination of aerosol height distributions by lidar," *Journal of Applied Meteorology and Climatology*, vol. 11, no. 3, pp. 482–489, 1972. Publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology.

[28] A. Ansmann, M. Riebesell, and C. Weitkamp, "Measurement of atmospheric aerosol extinction profiles with a raman lidar," *Optics Letters*, vol. 15, no. 13, pp. 746–748, 1990. Publisher: Optical Society of America.

[29] V. Wulfmeyer, R. M. Hardesty, D. D. Turner, A. Behrendt, M. P. Cadeddu, P. Di Girolamo, P. Schlüssel, J. Van Baelen, and F. Zus, "A review of the

remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles," *Reviews of Geophysics*, vol. 53, no. 3, pp. 819–895, 2015.

[30] WMO, *Commission for Basic Systems OPAG on Integrated Observing Systems - Implementation/coordination Team on Integrated Observing Systems (ICT-IOS), Ninth session Final report.* WMO, rev. 2 ed., 2016.

[31] ESA, "Design and breadboarding of a wide filed-of-view telescope for sentinel 2ng." TDE PRogram element EXPRO PLUS, 1-11263.

[32] WMO, "Oscar - list of all requirements." https://space.oscar.wmo.int/requirements.

[33] R. Thundathil, T. Schwitalla, A. Behrendt, S. K. Muppa, S. Adam, and V. Wulfmeyer, "Assimilation of lidar water vapour mixing ratio and temperature profiles into a convection-permitting model," *Journal of the Meteorological Society of Japan. Ser. II*, vol. 98, no. 5, pp. 959–986, 2020. Publisher: Meteorological Society of Japan.

[34] R. Thundathil, T. Schwitalla, A. Behrendt, and V. Wulfmeyer, "Impact of assimilating lidar water vapour and temperature profiles with a hybrid ensemble transform kalman filter: Three-dimensional variational analysis on the convection-permitting scale," *Quarterly Journal of the Royal Meteorological Society*, vol. 147, no. 741, pp. 4163–4185, 2021.

[35] I. Veselovskii, Q. Hu, P. Goloub, T. Podvin, M. Korenskiy, O. Pujol, O. Dubovik, and A. Lopatin, "Combined use of mie–raman and fluorescence lidar observations for improving aerosol characterization: feasibility experiment," *Atmospheric Measurement Techniques*, vol. 13, no. 12, pp. 6691–6701, 2020.

[36] I. Veselovskii, Q. Hu, P. Goloub, T. Podvin, M. Choël, N. Visez, and M. Korenskiy, "Mie–raman–fluorescence lidar observations of aerosols during pollen season in the north of france," *Atmospheric Measurement Techniques*, vol. 14, pp. 4773–4786, 07 2021.

# Publication List

[1] P. Di Girolamo, N. Franco, M. Paolantonio, S. Donato, and D. Dionisi, "Atmospheric thermodynamic profiling through the use of a micro-pulse raman lidar system: Introducing the compact raman lidar marco," *Sensors*, vol. 23, p. 8262, 10 2023.

[2] D. Summa, G. Vivone, N. Franco, G. D'Amico, B. De Rosa, and P. Girolamo, "Atmospheric boundary layer height: Inter-comparison of different estimation approaches using the raman lidar as benchmark," *Remote Sensing*, vol. 15, p. 1381, 02 2023.

[3] P. Di Girolamo, B. De Rosa, S. Donato, N. Franco, and I. Veselovskii, "Measurements of aerosol size and microphysical properties: A comparison between raman lidar and airborne sensors," *Journal of Geophysical Research: Atmospheres*, vol. 127, 07 2022.

[4] D. Summa, F. Madonna, N. Franco, B. De Rosa, and P. Girolamo, "Intercomparison of atmospheric boundary layer (abl) height estimates from different profiling sensors and models in the framework of hymex-sop1," *Atmospheric Measurement Techniques*, vol. 15, pp. 4153–4170, 07 2022.

[5] D. Whiteman, P. Girolamo, A. Behrendt, V. Wulfmeyer, and N. Franco, "Statistical analysis of simulated spaceborne thermodynamics lidar measurements in the planetary boundary layer," *Frontiers in Remote Sensing*, vol. 3, p. 810032, 03 2022.

# Abstract in international conferences

[1] N. Franco, P. Di Girolamo, D. Summa, M. Di Paolantonio, , and D. Dionisi, "Preliminary studies and performance simulations in support of the mission "caligola"," EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-16695, https://doi.org/10.5194/egusphere-egu23-16695, 2023.

[2] P. Di Girolamo, N. Franco, M. Di Paolantonio, D. Summa, D. Dionisi, A. Di Bernardino, A. M. Iannarelli, and T. Di Iorio, "Several months of continuous operation of two thermodynamic raman lidars in the frame of walineas," EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-7775, https://doi.org/10.5194/egusphere-egu23-7775, 2023.

[3] D. Dionisi, S. Bucci, C. Cesarini, S. Colella, D. D'Alimonte, L. Di Ciolo, P. Di Girolamo, M. Di Paolantonio, N. Franco, G. Gostinicchi, T. Kajiyama, G. L. Liberti, E. Organelli, and R. Santoleri Ocean color through satellite lidars: the COLOR project, EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-16196, https://doi.org/10.5194/egusphere-egu23-16196, 2023.

[4] M. Di Paolantonio, P. Di Girolamo, D. Dionisi, A. Di Bernardino, T. Di Iorio, N. Franco, G. Giuliano, A. M. Iannarelli, G. L. Liberti, and D. Summa, "Performance simulation and preliminary measurements of a raman lidar for the retrieval of co2 atmospheric profiles," EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-16149, https://doi.org/10.5194/egusphere-egu23-16149, 2023.

[5] N. Franco, P. Di Girolamo, A. Behrendt, V. Wulfmeyer, A. Comeron, D. Summa, and D. Whiteman, "Performance simulation of a space-borne raman lidar for atlas," 30th International Laser Radar Conference (ILRC-30), online, Big Sky Conference, Montana (USA), 26 Jun-01 Jul 2022.

[6] P. Di Girolamo, D. Dionisi, M. Cacciani, A. Di Bernardino, N. Franco, D. Summa, M. Di Paolantonio, A. Iannarelli, and T. Di Iorio, "Compact raman lidar for atmospheric co2 and thermodynamic profiling - concerning," 30th International Laser Radar Conference (ILRC-30), online, Big Sky Conference, Montana (USA), 26 Jun-01 Jul 2022.

[7] P. Di Girolamo, M. Bouin, C. Flamant, D. Summa, B. De Rosa, and N. Franco, "The role of dry layers and cold pools in the activation of mesoscale convective systems: A characterization study based on the combined use of raman lidar and dial measurements and meso-nh model simulations," 30th International Laser Radar Conference (ILRC-30), online, Big Sky Conference, Montana (USA), 26 Jun-01 Jul 2022.

[8] P. Di Girolamo, A. Cosentino, F. Longo, N. Franco, D. Dionisi, D. Summa, S. Lolli, E. Suetta, A. Perna, and S. Zoffoli, "Introducing the cloud aerosol lidar for global scale observations of the ocean-land-atmosphere system – caligola," 30th International Laser Radar Conference (ILRC-30), online, Big Sky Conference, Montana (USA), 26 Jun-01 Jul 2022.

[9] D. Dionisi, S. Bucci, C. Cesarini, S. Colella, D. D'Alimonte, L. Di Ciolo, P. Di Girolamo, M. Di Paolantonio, N. Franco, G. Giuliano, G. Gostinicchi, T. Kajiyama, G. Liberti, E. Organelli, and R. Santoleri, "First results of the color (cdom-proxy retrieval from aeolus observations) project," 30th International Laser Radar Conference (ILRC-30), online, Big Sky Conference, Montana (USA), 26 Jun-01 Jul 2022.

[10] D. Summa, P. Di Girolamo, N. Franco, G. D'Amico, B. De Rosa, and G. Vivone, "Abl height different estimation by lidar in the frame of hymex sop1 campaign," 30th International Laser Radar Conference (ILRC-30), online, Big Sky Conference, Montana (USA), 26 Jun-01 Jul 2022.

[11] N. Franco, P. Di Girolamo, A. Behrendt, V. Wulfmeyer, A. Comeron, D. Summa, B. De Rosa, S. Tailhades, R. Janovsky, and B. Sang, "Thermodynamic profiling of the atmosphere with a spaceborne raman lidar for atlas," Living Planet Symposium, Bonn, Germany, 23-27 May 2022.

[12] N. Franco, P. Di Girolamo, D. Summa, B. De Rosa, A. Behrendt, , and V. Wulfmeyer, "Performance assessment of the space-borne raman lidar atlas – atmospheric thermodynamic lidar in space," EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-11817, https://doi.org/10.5194/egusphere-egu22-11817, 2022.

[13] N. Franco, P. Di Girolamo, D. Summa, B. De Rosa, A. Behrendt, A. Comeron, and V. Wulfmeyer, "Sviluppo di un simulatore end-to-end per sistema lidar raman satellitare," Aisam - Milano, 2022.

[14] N. Franco, P. Di Girolamo, D. Summa, B. De Rosa, B. A., and A. Comerón, "End-to-end simulator of a space-borne raman lidar for the thermodynamic profiling of the atmosphere," EGU General Assembly 2021, online, 19–30 Apr 2021, EGU21-10789, https://doi.org/10.5194/egusphere-egu21-10789.

[15] N. Franco, P. Di Girolamo, D. Summa, B. De Rosa, A. Behrendt, A. Comeron, and V. Wulfmeyer, "End-to-end simulation of a space-borne raman lidar for the thermodynamic profiling of the atmosphere," European Lidar Conference (ELC), online, 16-18 Nov 2021, Granada.