ALMA MATER STUDIORUM -
UNIVERSITÀ DI BOLOGNA

Dottorato di Ricerca in Geofisica
Ciclo XXXIII

SETTORE CONCORSUALE: 02/C1
SETTORE SCIENTIFICO DISCIPLINARE: FIS/06

The FORUM-EE9 mission: potentiality of the FIR measurements of the Outgoing Longwave Radiation for the estimate of geophysical variables

Presentata da: Alessio Di Roma

Supervisore:
Dott.ssa Bianca Maria Dinelli
Co-Supervisore:
Prof. Tiziano Maestri

Coordinatore Dottorato:
Prof.ssa Nadia Pinardi

Esame finale anno 2021

Contents

List of Figures

List of Tables

Abstract

The Far-infrared Outgoing Radiation Understanding and Monitoring (FORUM) mission has been officially selected in September 2019 to be the Earth-Explorer 9 mission of the European Space Agency. The main instrument on board of FORUM is a Fourier Transform Interferometer measuring the outgoing longwave radiation from 100 to 1600 cm^{-1}, designed to fill the observational gap in the far infrared range (below 600 cm^{-1}) and to provide for the first time the spectrally resolved full emission spectrum of the outgoing longwave radiation of the Earth and the atmosphere in synergy with the Metop-SG - IASI-NG instrument. The far infrared spectral region contains a large part of the information on climatological key variables, such as the water vapour concentration in the Upper Troposphere / Lower Stratosphere (UTLS), the surface emissivity below 600 cm^{-1}, and the clouds, strongly driving the Earth Radiation Budget. This work investigates the information content of the radiances that will be acquired by FORUM in a large set of atmospheric scenarios mostly in clear sky, representative of the scenes that will be sounded along the polar orbit, making use of different Radiative Transfer Models, like the GBB-NADIR and the KLIMA codes. First, the sensitivity of the acquired signal in the FIR is studied by analysing the derivatives of the spectrum with respect to the water vapour concentration and the surface emissivity. The precision on the water vapour, surface emissivity and atmospheric temperature estimated from the inversion of the FORUM measurements has been evaluated with the simulation of the retrieval procedure using the Optimal Estimation Method. We then have calculated the systematic contribution due to several different sources of systematic errors on these geophysical variables. The uncertainties affecting the water vapour spectroscopic data in the FIR and the continuum parametrization have been identified to be the most important error sources on the retrieved products, being responsible for systematic errors on the UTLS water vapour and FIR surface emissivity larger than 1% and 0.01 respectively. Therefore, we assessed the consistency of recent spectroscopic data in the FIR range by analysing the agreement between the simulations performed using different spectroscopic databases and the REFIR-PAD measurements on board a stratospheric balloon over Teresina (Brazil). Finally, the measurements of the Tropospheric Airborne Fourier Transform Spectrometer (TAFTS) on board an aircraft have been analysed to retrieve the water vapour from 0 to 9 km and the quality of the obtained products has been estimated studying the

consistency of the retrieved profiles with those measured by in-situ dropsondes and with the forecast of the Numerical Weather Prediction model, determining also the information gain on the water vapour VMR in the upper troposphere due to the measurements in the FIR range.

Chapter 1

Introduction

In this Chapter, the scientific background at the basis of the EE9-FORUM mission will be presented after a full overview of the Earth Explorer missions to which FORUM is going to take part. Then, the main goals of the EE9-FORUM mission will be listed and discussed, and finally an overview of this thesis work will be reported.

1.1 The Earth Explorer Missions - Understanding our planet

The major environmental issue that we are nowadays facing is the climate change, which is determined not only by the natural variability of the Earth-atmosphere system, but also by the large-scale impact that a growing global population and the exploitation of natural resources is enforcing on the environment. It is therefore crucial to study more about our planet in order to better predict the effects of change and respond to the issues that the humankind will face in the forthcoming decades. Answering to the different questions addressed by the scientific community, the Earth Explorers (EE) missions of the European Space Agency (ESA) have been and are designed and selected through user-driven processes and launched starting from 2009. Each Explorer is a small satellite research mission dedicated to some specific aspects of our Earth environment dealing with the atmosphere, biosphere, hydrosphere, cryosphere and the Earth's interior with the overall emphasis on learning more about the interactions between these components and the impact that human activity is having on natural Earth processes. To date, the following missions have been launched in the framework of the Earth Explorer program:

- **EE1-GOCE** The Gravity field and steady-state Ocean Circulation Explorer mission was launched on the 17^{th} March 2009 and provided the data to accurately determine global and regional models of Earth's gravity and geoid in order to supply information for ocean circulation, physics of Earth's interior, geodesy, and sea-level change. The mission ended in November 2013

- **EE2-SMOS** Launched on 2^{nd} November 2009, the Soil Moisture and Ocean Salinity (SMOS) mission is making global observations of soil moisture and ocean salinity. By consistently mapping these two variables, SMOS is advancing our understanding of the exchange processes between Earth's surface and atmosphere, helping to improve weather and climate models.

- **EE3-CryoSat** The Cryosat mission should have been the first mission taking part to the EE program, however the Cryosat-1 satellite was lost in 2005 due to the malfunctioning of the Russian SS-19 Rockot launch vehicle. Then, the Cryosat-2 satellite was successfully launched on 8^{th} April 2010. The CryoSat mission is monitoring centimetre-scale changes in the thickness of ice floating in the oceans and in the thickness of the Greenland and Antarctica ice sheets. Together with satellite information on ice extent, these measurements are showing how the volume of Earth's ice is changing and leading to a better understanding of the relationship between ice and climate.

- **EE4-Swarm** Launched on 22^{nd} November 2013, Swarm is a constellation of three satellites monitoring precisely the magnetic signals that stem from the magnetosphere, ionosphere, Earth's core, mantle, crust and the oceans.

- **EE5-Aeolus** Launched in 2018, the prime aim of the Aeolus mission is to make novel advances in global wind-profile observation in order to improve weather forecasting. Aeolus is expected to pave the way for future operational meteorological satellites dedicated to Earth's wind field measurements.

In the meanwhile, the preparatory phases of the future missions that will take part to the Earth Explorer program is going on. Namely, the EE missions to be launched in the next years are the following:

- **EE6-EarthCARE** The Earth Clouds Aerosols and Radiation Explorer (EarthCARE) mission, whose launch is scheduled in 2022, is a European-Japanese mission aiming to improve the representation and understanding of Earth's radiative balance in climate and numerical weather forecast models.

- **EE7-Biomass** The Biomass mission will provide crucial information about the state of our forests and how they are changing. The data will be used to further our knowledge of the role forests play in the carbon cycle.

- **EE8-FLEX** The Fluorescence Explorer (FLEX) will map vegetation fluorescence to quantify photosynthetic activity, improving our understanding of the carbon exchange between plants and the atmosphere and how photosynthesis affects the carbon and water cycles. It will also lead to better insight into plant health and stress.

- **EE9-FORUM** The Far-infrared Outgoing Radiation Understanding and Monitoring (FORUM) mission will provide new insight into the planet radiation budget and how it is controlled, measuring the outgoing thermal radiation from 100 to 1600 cm^{-1} with a Fourier Transform Interferometer. The launch is scheduled in the 2025-2026 time frame.
- **EE10** The EE10 has been not selected yet. The Stereoid, Daedalus and G-Class candidate missions were chosen in September 2018 to enter pre-feasibility study and compete to be the 10^{th} Earth Explorer mission.

The advantages offered by the observation from a satellite platform of the Earth and atmosphere thermal emission spectrum for the understanding of different problems involving the mechanisms and feedbacks between geophysical variables and radiation driving the climate will be explained. In particular, we will describe the contribution that the future mission will provide to science with the observation of the outgoing radiation spectrum from 100 to 600 cm^{-1}.

1.2 Scientific Background

The thermal emission of the Earth and of the atmosphere system escaping to space is called the Outgoing Longwave Radiation (OLR). The emission spectrum covers mainly two wide subranges of the infrared radiation, namely the Medium or Thermal InfraRed (MIR or TIR, from 600 to 4000 cm^{-1} - from 16.6 to 2.5 μm in wavelength) and the Far InfraRed range (FIR, from 100 to 600 cm^{-1} - from 100 to 16.6 μm in wavelength). The FIR spectral range of the OLR has been identified as a very important regulator of the terrestrial radiative balance, that is strongly driven by different key geophysical variables. Model simulations suggest that around 50% of the total clear-sky long-wave cooling of the Earth to space occurs within the far infrared spectral region (Brindley and Harries (1998); Kiehl and Trenberth (1997)). An even larger fraction of this cooling is predicted to occur in the FIR under all-sky conditions (Collins and Mlynczak, 2001), because the presence of clouds causes lower emitting temperatures, and thus a shift to longer wavelengths of the peak of the black-body emission. Many previous studies have outlined the importance of the FIR rotational band of water vapour and its influence on climate (Brindley and Harries, 1998), in particular the FIR is strongly sensitive to mid-upper level tropospheric humidity producing a peak in the cooling rate of the atmosphere (Clough and Iacono, 1995). Other molecular species as O_3, N_2O, and HNO_3 have pure rotational or low-energy vibrational transitions in the FIR, but these are insignificant in terms of radiative energy transfer compared with water vapour. In fact, its contribution on the total atmospheric optical depth from 100 to 600 cm^{-1} is 3-5 order of magnitude higher than that provided by other absorbing species (Harries et al., 2008). Huang and Ramaswamy (2009) have shown that accurate and continuous measurements of the outgoing longwave radiation extending to the FIR region represent a valuable tool to monitor climate changes and to validate climate models. Furthermore, cirrus clouds play an important role in the determination of the Earth's energy budget. They may have both a warming and cooling effects, since they modulate both the incoming solar radiation and the outgoing long-wave emission. The relative importance of the two effects depends on the cloud optical and micro-physical properties (Cox et al. (2010), Lubin et al. (1998)). Moreover, the FIR component of the spectrum is very sensitive to the particle habit and the size of ice crystals (Baran (2007), Baran (2009)), which makes it very important to characterize the cirrus radiative properties (Palchetti et al., 2015). The radiative properties of the polar surface could have relevant effects on the OLR budget even though the FIR range is largely dominated by the water vapour absorption, namely the surface emissivity for wavelength higher than 15 μm has a direct impact on the OLR and infrared cooling rates where the column precipitable water vapour (PWV) is less than 1 mm (Feldman et al., 2014).

1.3 The FORUM mission

The literature reported in Section1.2 suggests several mechanisms to be further investigated, linking a wide set of climatological key parameters to the radiation escaping to space in a spectral range containing a large fraction of the Earth and atmosphere emission that has never been measured from space in its entirety. In fact, despite its importance, the technical difficulties associated with achieving the necessary instrument signal to noise ratio across the FIR region had never been overcame in the past. The previous space missions sampled only part of the FIR range (extending to wavelengths of 25μm), having neither the necessary lifetime nor the accuracy to provide quantitative measurements at the required level for climate change monitoring, therefore there have been only a few spectral measurements and datasets collected in this region (Harries et al., 2008). To close this gap, the FORUM mission has been proposed and selected to be the EE9 mission of ESA. FORUM will deliver an improved understanding of the climate system by supplying, for the first time, a complete characterization of the Earth's OLR spectrum. The FORUM mission was officially selected in September 2019 to be the 9^{th} Earth Explorer as a result of the suggestion of the Advisor Committee for Earth Observation (ACEO) to ESA at the end of the user-consultation meeting in July 2019 in Cambridge (UK). The mission will fly on a polar low orbit measuring a large part of the outgoing Earth and atmosphere thermal radiation covering the whole globe.

Measuring the FIR part of the Earth thermal emission spectrum is very important, since most part of the Earth's outgoing energy occurs in this spectral region. Furthermore, different climatological key parameters strongly affect the FIR radiation, making this spectral region of primary importance for the driving of the Earth Radiation Balance. In this respect, the forthcoming EE9-FORUM has been defined as a “new satellite to understand how Earth is losing its cool”, a definition enforced by the ESA’s Director of Earth Observation Programs Josef Aschbacher, declaring “FORUM will measure, for the first time, the far infrared part of the electromagnetic spectrum from space, thus allowing us to better understand the energy balance of our planet. FORUM will bring great benefits to climate science”.

In particular, the FORUM mission will deliver the entire OLR spectrum by flying in loose formation with the Meteorological Operational Satellite, Second Generation (Metop-SG) carrying on board the Infrared Atmospheric Sounding Instrument - Next Generation (IASI-NG) performing the measurement of the outgoing thermal radiation between 645 and 2710 cm^{-1}. In this way, FORUM together with IASI-NG will deliver a truly unique dataset of the Earth entire emission spectrum from 100 to 2760 cm^{-1} (from 3.62 to 100 μm in wavelength) that will help the scientific community to understand the links between key underlying physical processes driving climate change, their spectral signatures, the Greenhouse Effect, and the overall Earth Radiation Budget. An example of synthetic spectra - a spectrum simulated in clear sky conditions (red line) and one in presence of a cirrus cloud (light grey line) - as observed by the FORUM and IASI-NG combination are shown in Figure1.1, demonstrating that the thermal emission of the Earth and atmosphere system is almost totally contained in the signal measured by the two instruments.

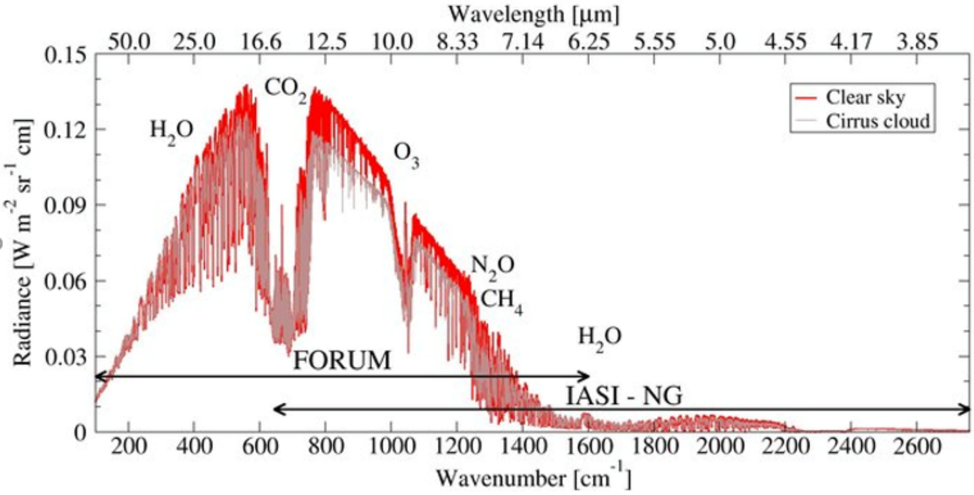


Figure1.1: Spectra calculated in the same atmospheric scenario in clear sky conditions (red line) and with cirrus cloud (light grey line) in the spectral range covered by the FORUM+IASI-NG instruments (100-2710 cm^{-1}). On the figure the spectral ranges covered by FORUM alone and IASI-NG alone are also shown. The major absorption bands of the main absorbing species are also marked

The nominal FORUM lifetime is 5 years, a time lapse that will enable to perform measurements covering different seasons and capture inter-annual variability and periodic atmospheric oscillations. A flight configuration in loose formation with IASI-NG requires a sun-synchronous orbit at an altitude of about 817 km, the one anticipated for Metop-SG, precluding the possibility to study the diurnal variability of the OLR, nevertheless it will provide a consistent set of observations for climatological assessments. The data acquired by FORUM will be organized into different processing levels. Namely, the Level-1c of the FORUM mission will provide the measured radiances at global scale that will be the basis for the future climate studies. Furthermore, they could become a benchmark for the future releases of the water vapour spectroscopic data, that are actually affected by high uncertainties in the FIR. In particular, the accurate measurements could be exploited to test the quality of water vapour spectroscopic data in the FIR, as just performed in different ground-based measurements campaigns such as SHEBA (Surface Heat Budget of the Arctic Ocean, Tobin et al. (1999)), RHUBC-I (Radiative Heating in Underexplored Bands Campaign, Delamere et al. (2010)), RHUBC-II (Turner et al., 2012) and ECOWAR (Earth Cooling by Water Vapour Radiation, Serio et al. (2008)). The Level-1c radiances will be also the starting point for the derivation of the Level 2 Products, including the main parameters to characterise the observed atmospheric state. The Level 2 Products of the FORUM mission and the related required uncertainties are reported in Table1.1.

Table1.1: Level 2 Products of the FORUM-EE9 mission with the required uncertainties

Level 2 Products	**Products Definition and Uncertainty Requirement**
All-sky broadband spectral flux (Irradiance [$nW/cm^2\ cm^{-1}$])	Estimated from the FORUM Level-1c FIR radiance extended with the Level-1c from IASI-NG. Error on flux consistent to the errors affecting the acquired radiance
Water vapour profile [ppmV]	Vertical profiles of water vapour concentration with 15 % uncertainty at 2 km vertical resolution
Surface emissivity	0.01 in the 300–600 cm^{-1} spectral range for polar region on 50 cm^{-1} spectral grid
Ice water path (IWP[g/m^2])	20 g/m^2 uncertainty on cloud ice content
Particle size parameter [μm]	20 % uncertainty on dimension

The potentiality of the FORUM-EE9 mission in the achievement of some of the aforementioned targets is discussed in the present thesis.

1.4 Overview of this thesis work

Many studies on the potentiality of the FORUM mission and the exploitation of the FIR radiation have been performed during the Phases A and B1 of the FORUM mission. They have been focused mainly to the link between the estimated outgoing radiance and flux and different geophysical key variables and to the estimate of these variables from the inversion of the radiance acquired by FORUM, with the main purpose to correct the biases actually affecting climate models investigating the main processes driving the climatic system. Part of the work performed in the last three years to study the potentiality of the FORUM measurements from space is presented in this thesis, pointing out the problems arising when the challenging FIR region is sounded and studied.

This thesis will focus principally on the investigation of the information content of the FORUM signal on the products specified in the list reported in Section1.3, data that are recognised as the main drivers of the OLR sounded by FORUM especially in the FIR range. To perform these studies, a reliable model simulating the radiative transfer in the atmosphere and the FORUM instrumental response to the incoming radiation is needed, therefore the FORUM main instruments and the Radiative Transfer Models (or Forward Models) used for simulations will be presented in Chapters 2 and 3 respectively. Then, the information content of FORUM measurements will be fully investigated in Chapter 4 through sensitivity tests and simulations of the retrievals of the geophysical variables in different atmospheric scenarios. Moreover, the study of the effect of different instrumental noise values on the uncertainty of the estimated variables and the assessment of the systematic errors affecting the Level 2 products due to several systematic error source will be shown in Chapter 5. These tests have shown that a large part of the systematic contribution to the Level 1 and Level 2 products is originated by the high uncertainties affecting the available water vapour spectroscopic data in the FIR range. The uncertainty on the spectroscopic data can be assessed by investigating the existing differences among the different spectroscopic databases, their effects on the simulations in the FIR and on the consistency between simulations and real measurements. In this context, the water vapour spectroscopic data contained in some of the most updated databases will be compared in Chapter 6. Then the simulation performed with different spectroscopic data will be compared to real FIR measurements performed by the REFIR-PAD (Radiation Explorer in the Far-Infrared) instrument (a prototype of FORUM) from a stratospheric balloon. Finally, in Chapter 7 the airborne measurements from the TAFTS (Tropospheric Airborne Fourier Transform Spectrometer) interferometer, being a prototype of the FORUM interferometer, will be inverted to obtain the water vapour estimate and information content, in order to confirm and check the information on water vapour in the upper troposphere due to the FIR radiation exploitation theoretically estimated.

Chapter 2

The FORUM payload

In this chapter the main instruments on board of the future EE9-FORUM mission, namely the FORUM Fourier Transform Spectrometer and the FORUM Embedded Imager, will be presented. Since the first is the main instrument of the mission, a brief description of the principle of operation of this instrument will be reported. Then, the characteristics of the instruments of the FORUM-EE9 satellite will be described in details. Finally, the particular measurement technique adopted by the FORUM-EE9 satellite will be described.

2.1 The Fourier Transform Interferometer

The Fourier Transform Interferometer (FTI) is an optical instrument where the radiation coming from an external source is divided into two parts that are later recombined generating an interference figure depending on the optical path difference of the split beams. The interference figure - called interferogram - contains detailed spectral information of the incoming signal, since the spectrum in frequency can be directly evaluated from the Fourier Transform (FT) of the interferogram. Optical spectra were first measured with the use of an interferometer by Michelson in the 19^{th} century. The utility of the FT technique was limited at the time, first by the lack of sufficiently quantitative detectors, and second by the difficulty of performing the necessary Fourier transforms. With the advent of electro-optical detectors and computers, those barriers were removed and many Fourier Transform Spectrometer (FTS) instruments have been flown successfully in space from 1962. The main advantage offered by the Fourier Transform Spectrometer consists in the measurements of the radiance on a large number of spectral channels with relatively high spectral resolution (generally finer than $\Delta\nu$=1 cm^{-1}) using a single detector. The simplest type of FTS instrument is based on the Michelson interferometer design, whose sketch is shown in Figure2.1.

The incoming radiation (assumed monochromatic) is collimated and divided into two beams by the beamsplitter. The beam in each leg of the interferometer is reflected back toward the

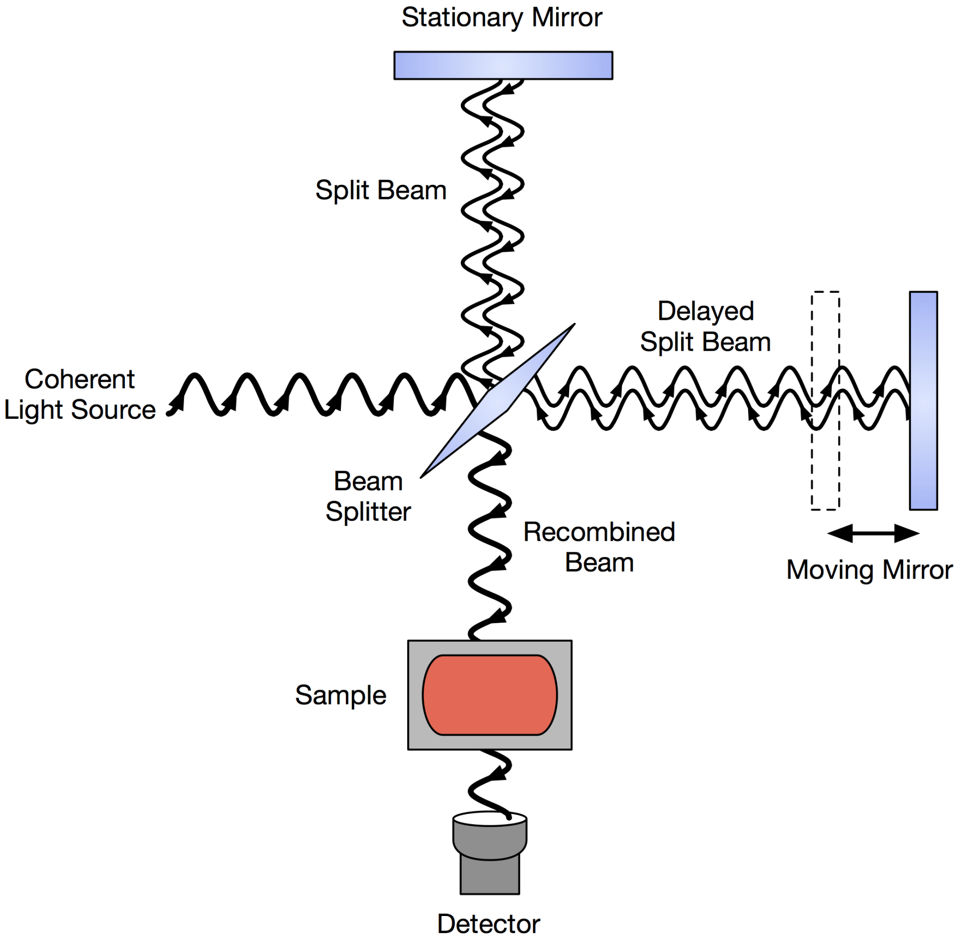


Figure2.1: Sketch of the optical device of the Michelson interferometer

beamsplitter by a mirror, where the two beams are recombined. The beamsplitter again divides the two beams into two parts, one part returning to the input source and the other going to the detector. In each direction, the beams from the two interferometer legs recombine or interfere, so the intensity of the light focused by the lens on the detector is modulated according to the optical phase difference of the two recombined beams (related to their optical path difference), equal to ϕ=2$\pi\Delta$x/λ where λ is the wavelength of the radiation, and Δx is the optical path difference (OPD) of the two legs in the interferometer. In operation, one or both of the interferometer mirrors are moved, so Δx is varied. The intensity recorded at the detector, for a monochromatic source, is then given by

$$I(\nu,\Delta x) = \frac{1}{2} I_0 [1 + cos(2\pi\nu\Delta x)] \tag{2.1}$$

where ν=1/λ is the wavenumber and I_0 the intensity at zero path difference. In reality, the input of the interferometer is a spectral radiance B_ν that is not monochromatic, therefore the total signal acquired by the detector as a function of the ODP s(Δx) is

$$s(\Delta x) = 2\int_{0}^{+\infty} S_\nu[1+cos(2\pi\nu\Delta x)]d\nu \tag{2.2}$$

$$S_\nu = \frac{1}{4}A_0\Omega\tau_{0x}\eta_s R_\nu B_\nu \tag{2.3}$$

where the acquired spectral signal S_ν results from the multiplication of B_ν with the area of the optical aperture A_0, the solid angle of accepted radiation Ω, the spectral responsivity of the detector R_ν [Volts/Watt], the beamsplitter efficiency η_s and the transmission of the remainder of the optical system τ_{0x}. Hence, the S_ν signal is composed of an unmodulated part and a modulated part of average value zero called *interferogram* that can be written as

$$m(\Delta x) = \int_{-\infty}^{+\infty} S_\nu[cos(2\pi\nu\Delta x)+isin(2\pi\nu\Delta x)]d\nu = \int_{-\infty}^{+\infty} S_\nu e^{i2\pi\nu\Delta x}d\nu \tag{2.4}$$

The interferogram m(x) and the spectral signal S_ν form a Fourier Transform pair, hence the spectral signal can be reconstructed from the interferogram as

$$S_\nu = \int_{-\infty}^{+\infty} m(x)e^{-i2\pi\nu\Delta x}d(\Delta x) \tag{2.5}$$

The interferogram m(x) is what is directly measured by the detector as a function of the OPD. However, it cannot be measured for Δx from $-\infty$ to $+\infty$, since the mirror can move only on a path of finite length, and the absolute difference between the optical paths in the two interferometer legs can range only from 0 to the Maximum Optical Path Difference (MOPD) **L**. Thus, the interferogram is truncated with a truncation function

$$S_\nu^{'} = \int_{-\infty}^{+\infty} \prod(\frac{\Delta x}{L})m(x)e^{-i2\pi\nu\Delta x}d(\Delta x) \tag{2.6}$$

where $\prod$ is the truncation function having zero values for |Δx|>L. Since $S_\nu^{'}$ is the Fourier transform of a product of two functions, then it is equal to the convolution of their inverse transforms

$$S_\nu^{'} = \pi(\nu)\star S_\nu \tag{2.7}$$

where the symbol $\star$ stands for convolution. The inverse transform $\pi(\nu)$ of the truncation function is usually called Instrumental Spectral Response Function (ISRF), representing the response of the instrument to the incoming radiation. This function is directly linked to the spectral resolution of the instrument due to the finite mirror displacement. The ideal ISRF of a FTI is represented by a *sinc* function, that is the FT of a boxcar function defined from **-L** to **+L**. The *sinc* is an oscillating function that has its first zeros at $\Delta\nu$=1/2L, that is the unapodized spectral resolution of the interferometer.

However, the real interferometers have truncation functions different from the boxcar function, due to instrumental effects that results in a not flat response along the interferogram. In order to reduce these unknown effects, the basic Fourier Transform is generally modified with a corrective procedure called “apodization”, consisting in the convolution of the spectrum with an apodizing function (apodized ISRF) that is different from the *sinc* function. The effects of the apodization process are the reduction of the spectral resolution with respect to the nominal one provided by the *sinc* and the attenuation of the sidelobes that can be observed in the unapodized spectra. The sidelobes look like feet of each line in the spectrum, so the word apodization is appropriate to describe the procedure (apod or apodal means having no feet). Different functions are used to perform the apodization process, such as Gaussian or Norton-Beer functions, that represents the FT of different truncation functions. To an extreme, the interferogram is truncated with a triangle function, whose Fourier Transform is the $sinc^2$ function, sharply reducing the sidelobes affecting the unapodized spectrum and providing a reduced spectral resolution equal to $\Delta\nu$=1/L. The spectral resolution of an FTS could be improved by increasing the MOPD in the interferometer legs, however this would require a higher acquisition time for the single interferogram reducing the spatial resolution and the number of spectra measured along a single orbit. On the other hand, a low MOPD value would allow the acquisition of high number of spectra at the expenses of the spectral resolution, therefore not resolving the spectral signatures of the atmosphere.

2.2 The FORUM Interferometer

In this Section, the main characteristics of the FORUM Fourier Transform Interferometer are presented, including also information on the thermal noise and calibration errors affecting the acquired measurements.

The Fourier Transform Interferometer on board of FORUM, also known as the FORUM Sounding Instrument (FSI) will measure the Earth’s upwelling spectral radiance from 100 to 1600 cm^{-1} (from 100 to 6.25 microns in wavelength) in the nadir direction. In order to provide measurements with high spectral detail for a suitable climate monitoring at the global scale, the preliminary studies have recommended a goal requirement on the spectral resolution of the FORUM FSI equal to 0.36 cm^{-1} and a threshold requirement equal to 0.5 cm^{-1}. The difference of FORUM measured spectra from the ’real’ incoming radiation due to the instrumental effects can be visualised in Figure2.2, showing in the left panel the TOA spectrum computed at very fine spectral resolution (0.0005 cm^{-1}, representing the "true" spectrum) and, in the right panel, the corresponding FORUM unapodized (blue line) and apodized spectra (red line), simulated convolving the high resolution spectrum with a *sinc* function having $\Delta\nu$=1/2L=0.413 cm^{-1} and the apodized *sinc* with the Norton-Beer strong function respectively. We can see in the figure that the spectral response of the interferometer does not allow to resolve most part of the narrowest absorption lines, however the broadened water vapour lines in the FIR and the envelope of intense absorption bands (e.g. the CO_2 band centered at 667 cm^{-1} and the ozone band at 1050 cm^{-1}) can be clearly identified in the spectrum observed by the FORUM FSI. The apodization applied with the Norton-Beer strong function tends to remove the ripples in the unapodized spectrum and to broaden and reduce the intensity of the absorption lines, that is the result of the reduction of the spectral resolution due to the apodization itself. The explained theory on the instrumental response of the interferometer and the assumed ISRF for test reported in Figure2.2 does not take into account possible spurious effects affecting the optical device as - for instance - the effects introduced by a wedged beamsplitter or a wavenumber dependence of the instrumental response. Therefore, the used ISRF represents an ideal condition and the real ISRF could deviate from the ideal one. A model of real ISRF affected by spurious effects used to simulate the instrumental response of the FORUM FSI in the Level 2 systematic errors assessment will be presented in Chapter 5.

All the components of the optical part of the instrument and the detector used to measure the interferogram introduce a random thermal noise in the acquired interferograms, that can be expressed in terms of the Noise Equivalent Spectral Radiance (NESR). In the preliminary studies of the mission, the Goal and Threshold Values for the NESR have been established studying the values that enable to obtain accurate Level 1 and Level 2 products from the measurements. Therefore, the industrial consortia studying the design of the satellite instrument are investigating the use of different kind of beamsplitter and detectors that will allow to get a thermal noise lower than the required values. The NESRs relative to two different real detectors will be described and studied in Chapter 5.

The goal of FORUM is to provide accurate measurements of the outgoing longwave radiation, to be used as a reference for future missions. Therefore it is important to have measurements with high Absolute Radiometric Accuracy (ARA), that is the accuracy that cannot be reduced averaging a large number of spectra. The ARA directly affects the Level 1 Products (e.g. TOA radiance and integrated flux) and is linked to the accuracy with which the various on board calibration sources are known. The calibration procedure of FORUM spectra consists in the measurements of the radiation coming from calibration sources of known temperature (e.g. internal black bodies or the deep space), used to characterize the variation of the detector response over time. Differently from the NESR, the ARA introduces a systematic contribution on the measurements and can be regarded as an unknown offset of the output spectra. Hence, the ARA value should be evaluated and is required to stay below fixed values, specified as the 3σ error of the brightness temperature measured observing a blackbody with known temperatures in the range between 190 and 300 K (the expected range of temperatures measured by FORUM).

The aforementioned requirements on the spectral resolution, instrumental noise and the absolute radiometric accuracy of the FORUM FSI are summarized in Table2.1

Tabel2.1: Goal and threshold requirements on the spectral resolution, NESR and ARA of the FORUM FSI instrument

	GOAL requirements	**THRESHOLD requirements**
Spectral resolution	0.36 cm^{-1}	0.5 cm^{-1}
NESR	40 nW/(cm^2 sr cm^{-1}) in the range 200 - 800 cm^{-1} 100 nW/(cm^2 sr cm^{-1}) in the range 200 - 800 cm^{-1}	60 nW/(cm^2 sr cm^{-1}) in the range 200 - 800 cm^{-1} 200 nW/(cm^2 sr cm^{-1}) in the range 200 - 800 cm^{-1}
ARA(3σ)	$<$ 0.2 K in the range 200 - 300 cm^{-1} $<$ 0.1 K in the range 300 - 1100 cm^{-1} $<$ 0.2 K in the range 1100 - 1300 cm^{-1} $<$ 1 K elsewhere	$<$ 0.25 K in the range 200 - 1300 cm^{-1} $<$ 1 K elsewhere

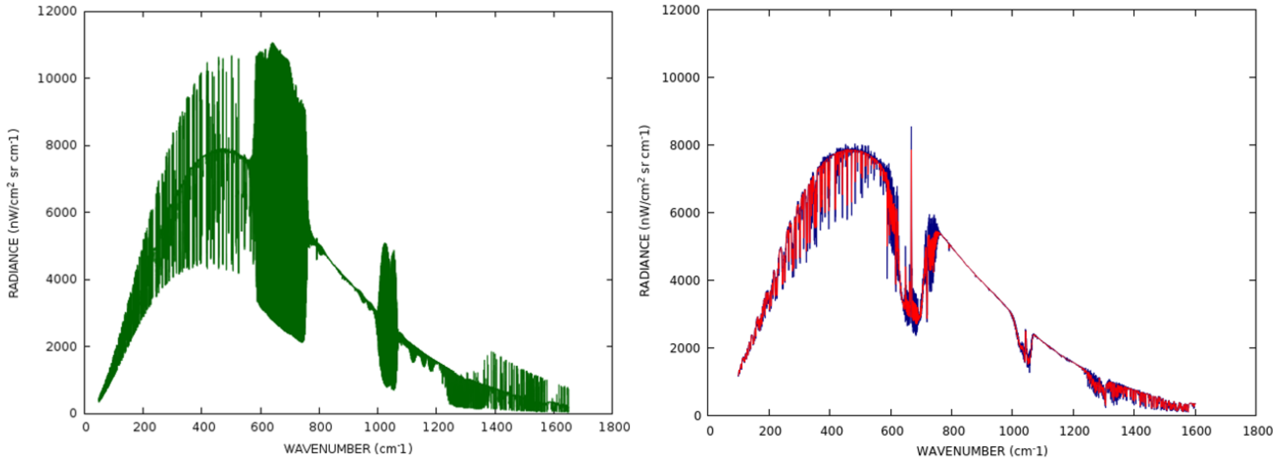


Figure2.2: Simulated TOA spectrum for a polar scenario at high spectral resolution (0.0005 cm^{-1}, left panel), compared in the right panel to the spectrum convolved with the unapodized ISRF (blue line) and the apodized ISRF (red line)

2.3 The FORUM FEI

The FORUM Embedded Imager (FEI) will observe the same scene of the FSI in order to assess the scene homogeneity (i.e. the presence of scattered clouds or the boundaries between two different surface types like ice and ocean). The FEI is a camera estimating the brightness temperature (BT) field with high spatial resolution (600 m x 600 m) measuring the incoming radiation within a spectral band centered at 10.5 μm (952 cm^{-1}), that is a spectral region with high atmospheric transmittance where numerous studies (Inoue (1985), Parol et al. (1991), Ackerman et al. (1995)) have already shown the sensitivity to both surface and cloud property variations. Differently from the FSI instrument, the Imager will not provide spectrally resolved measurements in order to get measurements with high Signal-to-Noise-Ratio (SNR) in each spatial sample of the image. In fact, the signal is determined by the spectrally integrated radiation acquired by each detector pixel that covers a very limited surface area, hence the instrument needs to measure the outgoing radiance in a broad spectral band to provide a high quality image with fine spatial resolution. The precision of the FEI is affected by the thermal noise affecting the detectors acquiring the single image pixels, being characterised by the Noise Equivalent Delta Temperature (NeDT), defined as the minimum identifiable temperature difference between any two spatial samples of the FEI image in a homogeneous scene. The NeDT should be small enough to enable the identification of inhomogeneities in the observed scenes, namely it has to be lower than 0.3K (Goal requirement) or 0.8K (Threshold requirement) at the reference temperature 210 K. At temperatures different from the reference one, the requirements will be scaled by the Radiometric Scaling Function (RSF). However, the radiometric accuracy is not an issue for the FEI if just used for monitoring scene homogeneity since only relative changes are important in this case.

2.4 FORUM Measurement Technique

The spatial resolution of the atmospheric scenarios measured by the FORUM FSI and FEI depends on the Field-Of-View (FOV) of the instruments, represented by a circle with diameter equal to 15 km for the interferometer and a square of 36 km x 36 km for the Imager on the Earth surface. The FOV of the imager should always include the FOV of the FSI (see Figure2.3). During a single measurement, while the satellite is moving along the orbit, both the instruments should observe the same scene along the satellite track in a direction close to the nadir one, integrating the incoming signal. Therefore, the FORUM measurement technique is based on the adjustment of the instruments line of sight during the whole time interval of the measurement, known as the “step and stare” measurement technique (shown in Figure2.3). Each ground pixel of the FSI is sounded with a line of sight ranging within an angle of $\pm 3.5^\circ$ around the nadir direction, allowing the acquisition of nearly constant outgoing signal from unchanged atmospheric scenarios during the "step and stare" scan, since the horizontal smearing of the tropospheric air mass sounded by the instruments will be less than 1 km for an assumed satellite altitude of 800 km, that is much lower than the FSI and FEI FOV. The ground distance between two adjacent ground pixels will be around 100 km, comparable to the horizontal spatial resolution of current global circulation models. The instrument will not perform measurements across-track, however the polar orbit will guarantee a high coverage of the polar regions, that are of great interest for the monitoring of the water vapour content and the surface radiative properties in the FIR. We have seen in the introduction that both these quantities have relevant effects on the outgoing radiation and on the radiative balance in the the polar regions. This will make FORUM a suitable mission to support for climate models providing accurate Level 1 and Level 2 Products to which they can be tested.

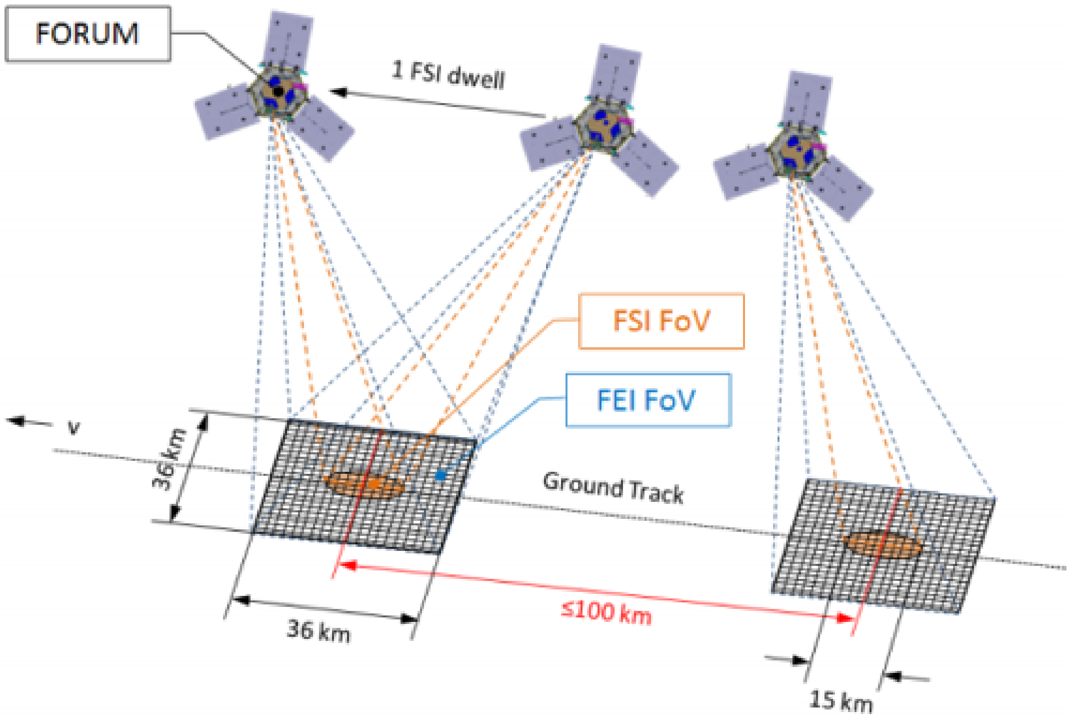


Figure2.3: Snapshot of the "step and stare" technique, showing the FSI and FEI Field Of Views and the distance between two adjacent pixels on the along-track of the satellite

Chapter 3

Forward Models for FORUM

In order to study the expected performances of the FORUM mission we need to have a tool to simulate the FORUM measurements. This tool is made of a code reproducing the radiative processes along the entire line of sight through the atmosphere determining the radiation collected by the instrument. The radiative processes can be reproduced by solving the Radiative Transfer Equation (RTE) after the calculation of the absorption, emission and scattering processes of the crossed medium. The RTE for the infrared signal propagating along a selected line of sight (LOS) through the atmospheric medium modelled with plane parallel layers can be written as:

$$\mu \frac{dI(\tau,\mu,\phi)}{d\tau} = I(\tau,\mu,\phi) - S(\tau,\mu,\phi) \tag{3.1}$$

where I is the diffuse specific intensity at τ in a cone of unit solid angle along the LOS, μ is the cosine of the zenithal angle θ, ϕ is the azimuth angle, τ is the optical depth, that is related to the optical path, and S is the "source function". The general solution of the Eq.3.1 in presence of a not constant source function S is

$$I(\tau_0,\mu,\phi) = I(0,\mu,\phi)e^{-\tau_0} + \frac{1}{\mu}\int_0^{\tau_0} S(\tau,\mu,\phi)e^{-\frac{(\tau_0-\tau)}{\mu}}d\tau \tag{3.2}$$

where τ_0 is the optical depth of the total atmospheric column, I(τ_0,μ,ϕ) is the intensity at τ_0 coming from the direction (μ,ϕ) and $I(0,\mu,\phi)$ is the radiation coming from the background at τ=0. Since the effect of the solar radiation can be neglected in the spectral region of FORUM (FIR and TIR) the source function S is the sum of the thermal emission Q_{Therm}, represented by the Planck function B(T(τ)) at temperature T when the atmosphere is in local thermodynamic equilibrium (LTE), and of the radiation scattered into the direction (μ,ϕ) from all the other directions ($\mu^{'}$,$\phi^{'}$).

$$S(\tau,\mu,\phi) = \frac{\omega(\tau)}{4\pi}\int_{0}^{2\pi}\mathrm{d}\phi'\int_{-1}^{1}P(\tau,\mu,\phi;\mu',\phi')I(\tau,\mu',\phi')\mathrm{d}\mu' + Q_{\mathrm{Therm}}(T(\tau)) \tag{3.3}$$

$$Q_{Therm}(T(\tau)) = (1-\omega(\tau))B(T(\tau)) \tag{3.4}$$

$$B(T(\tau)) = \frac{2h\nu^3}{c^2}\cdot\frac{1}{e^{\frac{h\nu}{K_BT}}-1} \tag{3.5}$$

where ω is the Single Scattering Albedo (SSA), representing the fraction of the radiation of an incident beam reflected toward the observer by an infinitesimal volume inside the medium and P is the scattering Phase Function, describing how much of the radiation coming from a given direction is scattered into another direction, h is the Planck constant, c the speed of light, K$_B$ the Boltzmann constant and ν the wavenumber. The atmospheric optical properties defined in the Eq.3.2 depend on the physical and chemical status of the atmospheric layer crossed by the LOS (i.e. temperature and pressure and Volume Mixing Ratios (VMRs) of the atmospheric gases, the particles type, the mass concentration and size distribution). In order to solve the Radiative Transfer problem, these quantities have to be supplied with the spectroscopic properties of the absorbing gases, for which the parameters describing each absorption line for a proper modelling of the absorption in the spectral range of interest are provided by a spectroscopic database. Moreover, we need to define the surface temperature and radiative properties, since the nadir line of sight (μ=1, ϕ=0°) intercepts the Earth surface, with the I(0,μ,ϕ) term in the Eq.3.2 representing the radiation both emitted and reflected by the surface to the zenith

The atmospheric optical depth in Eq. accounting for the gaseous absorption can be defined as

$$\tau(\nu,l) = (\sum_{m=1}^{N_G} K(\nu,l,m)\cdot X_{l,m})\eta\,\delta z \tag{3.6}$$

with N_G the number of the gaseous species in the atmosphere, $X_{l,m}$ the Volume Mixing Ratio of species m and layer l, η the layer air density and Δz the optical path. K(ν,l,m) is the absorption cross section [cm^2/molecule] of the gaseous species m at wavenumber ν in the layer l characterised by a given temperature and pressure, namely

$$K(\nu,l,m) = \sum_{j=1}^{N(m)} S_j(l)\cdot L_j(\nu,l) \tag{3.7}$$

where N(m) is the number of the absorption lines of the gaseous species m determining a contribution to the absorption at the wavenumber ν, $S_j(l)$ is the line strength of the jth line depending on the temperature of layer l and $L_j(\nu,l)$ is the spectral line shape of the jth line at layer *l*, depending on temperature and pressure.

The ω and P single scattering properties of the atmosphere can be calculated starting from their refractive index (real and imaginary part) and the particles size distribution and mass concentration using the Lorentz-Mie theory (as the one described by Mishchenko et al. (2002)) or can be extracted from external Look-Up-Tables containing pre-calculated optical properties of different particles distribution being representative of the scenarios that could be observed.

Different Radiative Transfer Models (RTMs) have been developed and used to simulate the outgoing signal in the spectral range of FORUM - like the Geofit Broad-Band NADIR code, and the one inside the KLIMA code - that will be described in detail in Sections 3.1 and 3.5 with an accurate explanation of the method used for the line shapes and line strengths calculation to compute the atmospheric optical depth. Moreover, RTMs have been embedded into retrieval codes used to study FORUM Level 2 performances, like the IAC retrieval code, used in this work to estimate the water vapour concentration from the Tropospheric Airborne Fourier Transform Spectrometer (TAFTS) measurements (see Chapter 7). This code is briefly described in Section3.6. All these RTMs codes calculate the TOA (Top of the Atmosphere) spectra at high spectral resolution, then the effects due to the instrumental response on the acquired signal are reproduced through the convolution procedures of the high resolution spectra with the ISRF and the Field Of View functions of the instrument, processes reported in Sections 3.3 and 3.4 respectively. Finally all the sources providing the needed ancillary data for scenarios definition that have been exploited for FORUM simulation will be listed and explained in Section 3.7.

3.1 The Geofit Broad-Band NADIR code

The Geofit Broad-Band NADIR code (GBB-NADIR) is a Line-By-Line Radiative Transfer Model computing the extinction properties of the atmosphere in the most general case of a stratified and plane parallel atmosphere both absorbing and scattering the radiation. The "Line-By-Line" denomination is due to the explicit calculation of the contribution given by each absorption line selected from a spectroscopic database on the computed optical depths. In the GBB-NADIR code, the calculation is performed dividing the whole simulated spectral range into contiguous small spectral intervals called MicroWindows (MWs) in order to reduce the high computing resources required by memorising all the molecular transition properties in a large spectral region. The GBB-NADIR code is based on the line by line code developed by Ridolfi et al. (2000) for the prototype of ESA MIPAS Level 2 analysis system. The output of the GBB-NADIR code consists of the optical depths due to the gaseous absorption and the scattering properties ω and P for each atmospheric layer, that can be used as input data for different Radiative Transfer Equation solvers like DISORT (Stamnes et al., 1988) and LIDORT (Spurr, 2008) or a fast RTE (Radiative Transfer Equation) solver developed in this thesis for TOA spectra computation in clear sky conditions. Most part of the physics at the basis of the GBB-NADIR code is described in the MIPAS Level2 ATBD (Algorithm Theoretical Baseline Document) available at the webpage *http://www.ifac.cnr.it/retrieval/documents/atbd_issue_4.1.pdf*, describing the Forward and Inverse Model used for vertical profiles of temperature and trace gases estimates from ENVISAT-MIPAS (Michelson Interferometer for Passive Atmospheric Sounding). The original GBB code was developed for limb measurements simulation, however the calculation procedures for the optical depths calculation do not depend on the viewing geometry. Therefore the nadir-adapted version contains few revisions compared to the original version of the GBB. In the next subsections several aspects of the GBB-NADIR code, dealing with the computing of the necessary quantity for the solution of Eq.3.2, will be fully described. Then, different RTE solvers that can calculate the spectra at the top of the atmosphere starting from the output data supplied by the GBB-NADIR will be described in Section3.2.

3.1.1 The GBB-NADIR spectroscopic database

The division of the spectral range in different contiguous MicroWindows requires that each MW shall get the spectroscopic data from a dedicated spectroscopic database containing the absorption lines determining a contribution to the absorption in the spectral range covered by the MW itself. The spectroscopic database of each MW is previously compiled in a different module, that performs the lines selection from one or more spectroscopic databases developed for the radiative transfer calculations, e.g. HITRAN (HIgh resolution TRANsmission molecular absorption database), GEISA (Gestion et Etude des Informations Spectroscopiques Atmospheriques) and AER. The lines selection code includes or discards a single line according to user-defined criteria that are different if the line falls inside the MW spectral range or outside the MW boundaries (within a range of 25 cm^{-1}). The criterium is based on the value of the line strength combined to the species expected average abundance, to model broadening effects and preserve computational efficiency and accuracy at the same time. The line selection code produces also a MicroWindow list file, specifying the MWs to be processed and the files containing the related spectroscopic data.

3.1.2 Input structure

The GBB-NADIR code reads all the required input data to perform the optical depths and scattering properties calculation from a settings file in ASCII format. It specifies the surface temperature, the vertical atmospheric layering, the paths of the files containing the input atmospheric profiles of pressure, temperature and Volume Mixing Ratio of the absorbing gases and the surface emissivity, the spectral resolution at which the computation is performed and the physical properties of aerosol or clouds (e.g. size distribution and particle mass concentration along the vertical) to be included in the modelled atmosphere. The spectral range to process is specified in the MicroWindow list file, also reporting the file name of each MW spectroscopic database to be read.

3.1.3 Calculation of the absorption properties

The effect of the absorption from the different gaseous species composing the atmosphere is calculated starting from the spectroscopic data contained in the spectroscopic database of the defined MWs. Each of the selected transition in the databases is described by the fundamental parameters for the correct simulation of the behaviour of the corresponding absorption line in all the possible atmospheric conditions, namely the central wavenumber of the absorption line [cm^{-1}], the intensity of the line in its centre [cm^{-1}/(molecule cm^{-2}) at 296 K standard temperature], the air and self broadening coefficients at 296 K [cm^{-1} atm^{-1}], the energy of the fundamental state [cm^{-1}], the temperature-dependence exponent for the pressure broadening effect and the air-pressure induced line shift [cm^{-1} atm^{-1}]. Therefore, the intensity and the shape of all the absorption lines are determined starting from these parameters with the methods described in the following Subsections.

Line strength calculation The line strength of each transition to be included in the optical depth calculation is provided by the officially distributed spectroscopic databases at a reference temperature (296 K). This value has to be scaled at the correct temperature of each layer using the temperature dependency of the population of the molecular energetic levels involved in the transition. The line intensity is determined with the expression Eq.3.8

$$S = S_0 \cdot F_{parts}(T) \cdot [e^{-\frac{hcE_0(T_0-T)}{K_B(T_0T)}} \cdot (1-e^{-\frac{hc\nu_0}{K_BT}})/(1-e^{-\frac{hc\nu_0}{K_BT_0}})] \tag{3.8}$$

Where S_0 is the line strength at T_0=296 K reference temperature, T is the temperature of the atmospheric layer, h the Planck constant, c the speed of light, E_0 the lower state energy of the line, K_B the Boltzmann constant and ν_0 the wavenumber of the molecular transition. F_{parts} is the ratio among the Total Internal Partition Sums (TIPS) at the required temperature and at the reference temperature calculated using the tabulated coefficients from Gamache et al. (2017) covering the temperature range between 70 and 3000 K. If the tabulated coefficients for a molecular species are not available, the F_{parts} value is approximated as the ratio between the layer temperature and T_0 to the power of 1.5.

Line shape calculation Each molecule presents determined absorption and emission lines related to the absorption or emission of a photon having wavenumber ν leading to a change in the energy of the rotational, vibrational or electronic state of the molecule itself equal to

$h\nu$, where h is the Planck constant. In the measured atmospheric spectra we can observe a broadening of the absorption lines due to collisions among the molecules (linked to the pressure) and the Doppler effect due to the thermal velocity of the absorbing molecules (linked to the temperature value). The line shape deriving from the pressure broadening is typically described with a Lorentzian function

$$L_L(\nu) = \frac{1}{\pi\alpha^L_{m,l}} \cdot \frac{1}{1+(\frac{\nu-\nu_{m,l}}{\alpha^L_{m,l}})^2} \tag{3.9}$$

Where ν is the wavenumber, $\nu_{m,l}$ the central wavenumber of transition l of species m, S the temperature dependent line strength and $\alpha^L_{m,l}$ the Lorentz broadened half width [cm^{-1}]. The Doppler broadening is represented by the following expression.

$$L_D(\nu) = \frac{2\sqrt{ln2}}{\sqrt{\pi}\alpha^D_{m,l}} \cdot e^{-\frac{2\sqrt{ln2}}{\alpha^D_{m,l}}(\nu-\nu_{m,l})^2} \tag{3.10}$$

where $\alpha^D_{m,l}$ is the Doppler half width [cm^{-1}]. This can be computed using the molecular mass and the temperature of the atmospheric layer, while the Lorentz half width is usually obtained by scaling the half width defined at reference temperature T_0=296 K and pressure P_0=1013 hPa. Therefore, these are calculated with the following expression

$$\alpha^D_{m,l} = \nu_{m,l}\sqrt{\frac{2ln2K_BT}{M_mc^2}} \quad \alpha^L_{m,l} = \alpha^{L_0}_{m,l}\frac{P}{P_0}[\frac{T_0}{T}]^\gamma \tag{3.11}$$

Where K_B is the Boltzmann constant, M_m the molecular mass, P and T are the atmospheric layer pressure and temperature respectively and γ is the coefficient of temperature dependence of the Lorentz half width. The Lorentz half width $\alpha^{L_0}_{m,l}$ at reference temperature and pressure includes both foreign $\alpha^{L_0}_{m,l,f}$ and self-broadened $\alpha^{L_0}_{m,l,s}$ components, that are the broadening coefficients describing the effects of collisions among different and identical molecules respectively. These are combined to calculate the half width $\alpha^{L_0}_{m,l}$

$$\alpha^{L_0}_{m,l} = (1-X_m)\alpha^{L_0}_{m,l,f} + X_m\alpha^{L_0}_{m,l,s} \tag{3.12}$$

where X_m is the Volume Mixing Ratio (VMR) of species m. The pressure effect is also responsible for a shift of the line centre, that is computed using the following expression.

$$\nu^{'}_l = \nu_l + \gamma_{p,l}(\frac{P}{P_0}) \tag{3.13}$$

Where $\nu^{'}_l$ is the perturbed line position of transition l, ν_l is the unperturbed wavenumber, $\gamma_{p,l}$ the air-broadened pressure shift coefficient at 296 K [$cm^{-1}/atmosphere$] of line l, P is layer pressure and P_0=1013 hPa.

In general, the final line shape is a combination of the shape resulting from both the Doppler and pressure broadening, that can be represented by the Voigt function, that is a convolution between the Doppler and Lorentz line shapes. The convolution integral of the Voigt line shape is very time consuming, therefore the GBB-NADIR code estimates the Voigt function with the rational approximation (with a relative accuracy of 10^{-4}) introduced by Humlíček (1982), calculating the complex probability function:

$$W^H(z) = e^{-z^2}(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{t^2} dt) = K_W(x,y) + iL_W(x,y) \tag{3.14}$$

Where z=x+iy and K_W(x,y) is the real part representing the Voigt function. Even using this approximation, the calculation of the Voigt function is much more time consuming than the Lorentz function, then to speed up the computation the Voigt line shape is used only in the region close to the line centre where the difference between the two shapes is relatively large. Therefore, the baseline of the code is to use the Voigt function in the region surrounding the line centre and the Lorentzian shape outside. The region where the Voigt function is used is automatically decided according to the value of the combined Doppler and Lorentz width of the line. The line shape is estimated for each line whose frequency is within an interval of $\pm 25\ cm^{-1}$ from the considered spectral point and calculated in two different regions using different spectral grids, fine close to the line centre and coarse elsewhere. Wavenumbers ν_1 and ν_2, determining the transition between the coarse and fine grids are fixed at 50 times the sum of the Lorentz and the Doppler half width from the line centre. The fine grid is spaced as the chosen spectral resolution of the computation, whereas the coarse grid step is five time larger. Therefore, the line shape is computed on the fine grid close to the line centre where its behaviour is more variable with the wavenumbers, whereas the coarse grid is used far from the line centre, where its frequency dependence is less variable. The region between ν_1 and ν_2 is further divided into two regions. In the central region, included in the intervals ν_3 and ν_4, around the line centre, the algorithm by Humlíček (1982) is employed for the calculation. In the other region the line shape is assumed pure Lorentzian. Wavenumbers ν_3 and ν_4 are fixed to 30 Doppler half-widths from the line centre in the GBB-NADIR code. However, the thresholds determining wavenumbers ν_1, ν_2, ν_3 and ν_4 can be modified if higher efficiency or accuracy is needed for the line shape modelling.

Cross section and optical depth computing Once the line strengths and the line shapes are calculated for all the lines contained in a single MW, the GBB-NADIR computes the absorption cross sections solving the equation 3.7, then it computes the optical depths with the equations 3.6. The optical depths are reported on an output ASCII files.

Tabulated cross sections The spectroscopic parameters as defined in the Subsection 3.1.3 are not available for heavy and complex molecules present in the terrestrial atmospheres, such as the CFCs (chlorofluorocarbons) and HFCs (hydrofluorocarbons). The cross sections of these gaseous species are generally tabulated by the officially distributed spectroscopic databases, being reported for a set of pressure and temperature pairs. Hence, the tabulated cross sections are interpolated on P, T and wavenumber in the GBB-NADIR code in order to calculate the cross sections on the user-defined spectral grid and at temperature and pressure of the atmospheric layers.

Water vapour continuum The water vapour continuum MT_CKD (Mlawer-Tobin-Clough-Kneizys-Davies) is a parametrization of the water vapour absorption accounting for cumulative effects due to the contribution of the wings of absorption lines far from their centre implemented in the GBB-NADIR code for a correct modelling of water vapour absorption and emission. Namely, the GBB-NADIR uses the approximation described by Clough et al. (1989) and also performed in the Line-By-Line Radiative Transfer Model (LBLRTM, Clough et al. (2005)) with the removal of Lorentzian absorption at 25 cm^{-1} from the line centre and the addiction of the water vapour continuum to the water vapour cross section. The continuum parametrization is calculated throughout the following equation expressing the Van Vleck-Huber line shape formalism (Vleck and Huber, 1977):

$$K_{CONT(H2O)}(\nu) = \nu tanh(\frac{h\nu}{2K_BT})\frac{\eta}{\eta_0}(X_{H2O}C_s(\nu,T)+(1-X_{H2O})C_f(\nu,T)) \tag{3.15}$$

Where h is the Planck constant, η and η_0 are the density of the atmospheric layer and the reference density, C_s and C_f are the continuum absorption parameters for the self and the foreign broadening at the reference number density η_0. To determine the temperature dependence of the self-broadening values, exponential interpolation between the tabulated parameters $C_s(\nu$, 260K) and $C_f(\nu$, 296K) is performed. For the foreign broadening $C_f(\nu$, 296K) coefficients are used for all temperatures. The continuum absorption coefficients are tabulated every 10 cm^{-1}, therefore the interpolation in frequency is performed linearly.

3.1.4 Calculation of the scattering properties

The single scattering properties of aerosol and clouds (extinction coefficient, single scattering albedo ω (SSA) and scattering Phase Function $P(\mu,\phi)$) - accounting for radiation scattered by one localised scattering center - are calculated for each modelled layer of the atmospheric scenario including scattering bodies. The GBB-NADIR evaluates the scattering properties of spherical particles exploiting the code described by Mishchenko et al. (2002), providing the

Phase Function evaluated at user-defined scattering angles and approximated with a Legendre polynomial series truncated at a user defined order. For non-spherical particles, the GBB-NADIR code can read an external file containing the optical properties ω and P of the particular particles distribution.

3.1.5 Surface emissivity

After the definition of the user-defined spectral grid for the computation, the surface emissivity, read from the external file, is interpolated on the wavenumber grid within each defined microwindow, then the reflectivity is computed as (1-$\varepsilon(\sigma)$) assuming no transmission below the surface level for the FIR and MIR radiation. If the spectral points of the MW are not covered by the surface emissivity data, extrapolation with a constant value is performed.

3.2 RTE solvers

The high level GBB-NADIR code output (i.e. optical depths and scattering properties) are the necessary input data for different external solvers performing the RTE solution in case of atmosphere with scattering bodies. In the following Sections, the main characteristics of the RTE solvers that can be used for the simulations of FORUM measurements in presence of scattering bodies are explained.

3.2.1 GBB-NADIR RTE Internal Solver

The GBB-NADIR Internal RTE Solver is an internal routine of the GBB-NADIR code solving RTE equation for thermal outgoing radiation at the top of the atmosphere in a computationally fast way. The internal RTE solver does not perform simulations of the TOA signal in case of clouds or aerosol, since scattering effects are not taken in account. Hence, the Eq.3.2 in absence of scattering bodies can be easily solved for the source function S(τ,μ,ϕ)=B(T(τ)) generating the Schwarzchild equation solved by the internal solver, namely

$$S(\nu) = [\varepsilon(\nu)B(T_S) + R(\varepsilon(\nu))]e^{-\tau_0} + \int_0^{\tau_0} B(T)e^{-(\tau_0-\tau)}d\tau \tag{3.16}$$

where B(T_S), B(T) are the Planck function for thermal emission of surface and atmosphere, $\varepsilon(\nu)$ is the surface emissivity, $\tau(\nu)$ is the optical depth of the atmosphere taken as the vertical coordinate and $\tau_0(\nu)$ is the optical depth of the total atmospheric column. R($\varepsilon(\nu)$) is the fraction of the atmospheric thermal radiation emitted downward towards the observed point that is reflected backward to the observer. Considering a Lambertian surface albedo, it is:

$$R(\varepsilon(\nu)) = \frac{1-\varepsilon(\nu)}{\pi} \cdot \int_0^{2\pi} d\phi \int_0^1 dcos\theta \int_0^{\tau_0} B(T) e^{-\tau'} d\tau' \tag{3.17}$$

$$\tau' = \tau/\mu \tag{3.18}$$

where τ_0' and τ' are the total and infinitesimal optical depth of the slant atmospheric column, equal to the optical depth in the nadir direction divided by μ=cos(θ), where θ is the zenithal angle. Namely, the integral over the vertical coordinate, the zenith and azimuth angles expresses the total downward atmospheric emission reaching the observed point on the surface. Integration on the azimuth provides a multiplicative factor since no horizontal inhomogeneity has been assumed in the modelled atmosphere. The Planck function (used as source function) is calculated at each frequency point to model the surface and atmospheric thermal emission. B(T_λ) is calculated using the layer temperature, however the thermal emission of each layer is corrected using the layer optical depth through the empirical weighting function represented by the expression within the square brackets in the following equation

$$B'(T) = B^{up}(T^{up}) + 2[B(T) - B^{up}(T^{up})] \cdot [\frac{1}{\tau(\nu)} - \frac{e-^{\tau(\nu)}}{1-e^{-\tau(\nu)}}] \tag{3.19}$$

Where $B^{up}(T^{up})$ is the black body emission at the layer upper boundary temperature. The weighting function is shown in Figure3.1 as a function of the layer optical depth ranging from 10^{-6} to 10^6.

The function ranges from 0.5 for negligible optical depths to 0 for very large optical depths. Therefore, the more the layer is optically thick, the more B$'$(T) converges toward the thermal emission of the upper boundary of the layer $B^{up}(T^{up})$, that is physically correct. The same applies for the downward emitted radiation, with the emission approaching the lower boundary black body emission $B^{dn}T^{dn}$ for high values of $\tau(\nu)$.

$$B'(T) = B^{dn}(T^{dn}) + 2[B(T) - B^{dn}(T^{dn})] \cdot [\frac{1}{\tau(\nu)} - \frac{e-^{\tau(\nu)}}{1-e^{-\tau(\nu)}}] \tag{3.20}$$

This correction has been validated through comparison with the DISORT model introduced in Subsection3.2.2.

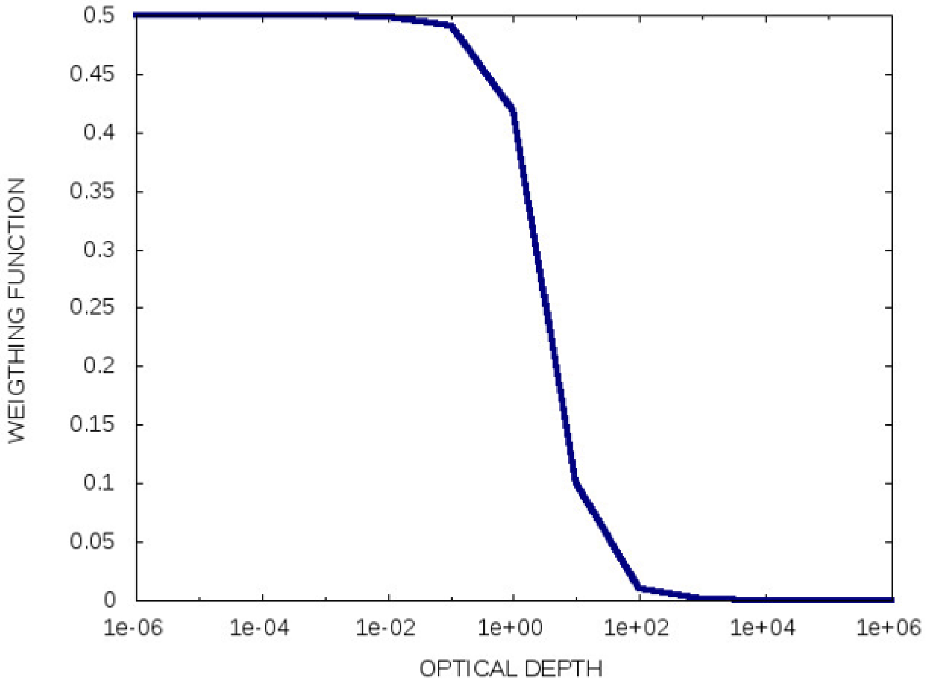


Figure3.1: Weighting function as a function of the optical depth determining the Planck emission of a single atmospheric layer

3.2.2 DISORT Radiative Transfer Model

The Discrete Ordinate Method Radiative Transfer in Scattering and Emitting Layered Media (DISORT, Stamnes et al. (1988)) code is a stand-alone Fortran module that accurately solves the radiative transfer equation in a vertically inhomogeneous plane-parallel atmosphere and has been incorporated into numerous radiative transfer codes as a separate module internally called for Multiple Scattering calculations. A single call to the DISORT routine solves the problem of the radiative transfer of monochromatic radiation, therefore it is necessary for each wavenumber to call DISORT iteratively from an interface routine passing the required input data in order to obtain an output spectrum within the desired spectral range. Atmospheric applications ranging from the UV to the radar region of the electromagnetic spectrum are possible with DISORT. The code searches for analytic solution of the RTE exploiting the discrete ordinate method by Chandrasekhar (1960), approximating the thermal emission functions by polynomials in τ and considering the atmospheric medium as made of **L** adjacent homogeneous layers in which the single scattering albedo and the phase function are assumed to be constant. The method is unconditionally stable for an arbitrarily large number of quadrature angles for the scattering and arbitrarily large optical depths. Since the solution found is analytic, the radiance can be returned at arbitrary angles and vertical levels. The surface reflection can be considered as Lambertian, specular or described by a BDRF (bidirectional reflectance) function. The DISORT code does not perform any derivative of the radiance with respect to geophysical parameters, therefore it can not be exploited as a forward model routine embedded in a numerical retrieval code for parameters estimate.

3.2.3 LIDORT Radiative Transfer Model

The Linearized Discrete Ordinate Radiative Transfer model (LIDORT, Spurr (2008)) is a stand-alone Fortran package for the analytic resolution of the full RTE in a plane-parallel atmospheric medium with both thermal and solar radiation. The binary executable program can be called by an interface routine passing the required input data, including atmospheric profiles of temperature and pressure and optic parameters, to the RTE solver. Particularly, the model needs optical properties (τ, ω, γ) for each modelled layer, respectively the total layer optical thickness for extinction, the single scattering albedo and the phase function Legendre expansion coefficients. As for the DISORT routine, LIDORT does not require any reference to the nature of the scatterer or the type of absorbing trace gases, and it calculates the monochromatic radiance field at any desired atmospheric level and viewing geometry. Differently, LIDORT can compute the analytical derivatives of computed spectrum K_ξ=$\delta I/\delta\xi$, with ξ referring to any atmospheric or surface parameter. For Jacobian output K_ξ, it is necessary to input the derivatives of properties (τ, ω, γ) with respect to the ξ parameters. The lower boundary reflection can be assumed as Lambertian, specular or described by a Bidirectional Reflectance Distribution Function (BDRF).

3.2.4 SHDOM Radiative Transfer Model

The Spherical Harmonic Discrete Ordinate Method SHDOM (Evans, 1998) enables computations of monochromatic radiation in a one, two, or three-dimensional medium for either collimated solar and/or thermal emission source of radiation. Differently from the previous presented FM, the radiative transfer is performed in a 3-D domain in which the atmospheric properties are defined on different grid points (therefore it requires extinction coefficients K[km^{-1}] instead of the optical depths). The 3-D domain can be defined as inhomogeneous on both horizontal and vertical directions, allowing accurate simulations providing radiance and integrated flux for scenarios with spatially variable atmospheric and surface properties. The source functions in the radiative transfer equation are calculated through spherical harmonics expansion in order to correctly simulate the angular behaviour of the thermal emission and scattering processes.

3.3 Convolution with the ISRF

The GBB-NADIR model described in Section 3.1 provides the simulated high resolution spectra of the radiation escaping to the outer space within the selected spectral range for a considered atmospheric scenario. To reproduce FORUM measurements, the high resolution spectra have to be convolved with the defined ISRF function simulating the interferometer response to the incoming radiation. The obtained unapodized spectrum can then be apodized in order to smooth the signal from the oscillations introduced by the *sinc* function side lobes. Therefore, the FORUM apodized synthetic measurements are obtained as the result of the following steps:

- The simulated high resolution spectral radiance reaching the instrument is convolved by the assumed ISRF. Since the oscillations of the unapodized ISRF tends to zero for large distances from the main maximum, the high resolution spectra have been calculated for a spectral interval exceeding the FORUM spectral range boundaries by 50 cm^{-1} (namely from 50 to 1650 cm^{-1})
- The spectrum is finally apodized with the Norton-Beer strong function (Norton and Beer, 1976).

On the basis of the requirements on the spectral resolution shown in Table2.1, the ISRF assumed for the FORUM FSI simulation in the performed tests of the preliminary Phase A and B1 of the mission is a *sinc* function with 1/2L=0.413 cm^{-1} spectral resolution apodized with the Norton-Beer strong function.

A different model of ISRF will be introduced in Chapter 5, being used for the evaluation of the systematic errors affecting the FORUM Level 2 Products due to the ISRF knowledge.

3.4 Convolution with the instrumental FOV

As shown in Chapter 2, the FORUM FSI will sound the underlying scenes with a circular Field of View (FOV) having a diameter of 15 km. The rigorous calculation of the signal acquired by the interferometer would require the calculation of the TOA spectra at different pencil beams in off-nadir directions within the aperture angle of the instrument, in order to take into account the behaviour of the instrumental FOV and the possible inhomogeneity of the scene. Then the convolution between the incoming signal from different directions and the FOV function should be performed. In the performed tests, we consider the angular instrumental FOV function to be a box car and the sounding area to be an homogeneous scenario. Therefore, the signal convolved with the instrumental FOV is practically the same of the simulated spectrum at the nadir direction, therefore this last will be considered to be representative for the FORUM FSI observation.

3.5 KLIMA Forward and Inverse Model

The KLIMA (Kyoto protocoL Informed Management of the Adaptation) Radiative Transfer Forward and Inverse Model (Del Bianco et al. (2014)) is a standalone exploitable ForTRAN code for both radiative transfer simulations and retrievals of geophysical variables through the inversion of the measured spectra using the Optimal Estimation Method (see Appendix A). The KLIMA model exploits the same input structure and routines for optical depth calculation as the GBB-NADIR code, making use of a very similar approach for the division of the spectral calculation into contiguous microwindows, then it calculates the RTE solution in case of not scattering with an internal fast solver for each spectral grid point. In order to provide the Jacobian matrix - containing the derivatives of the spectrum with respect to the retrieved parameters - required for the inversion procedure, the RTE solver computes also the numerical derivatives of the spectrum with respect to the atmospheric temperature, skin temperature and surface emissivity, whereas the derivatives with respect to the VMR of the absorbing species are performed analytically within the radiative transfer calculation. The Jacobian calculation is always performed when the Forward Model is run, therefore KLIMA can be used to obtain the Jacobians relative to the atmospheric input scenario on the defined input altitude grid. KLIMA requires an input ISRF function for high resolution spectra and Jacobians convolution. The instrumental spectrum, NESR and calibration error and the uncertainties on the a-priori state of the atmospheric scenario are required too for the definition of the instrumental and a-priori Variance-Covariance Matrices (VCM) when KLIMA is used as Inverse Model. The KLIMA model performs the convolution procedure with the ISRF internally defined providing both high resolution and convolved spectra.

3.6 IAC Retrieval Code

The IAC (Istituto per le Applicazioni del Calcolo) retrieval code is a standalone module performing the retrieval of different geophysical parameters from the measurement with the Optimal Estimation Method (see Appendix A) using the LBLRTM (Line-By-Line Radiative Transfer Model, freely available on *http://rtweb.aer.com/line_param_frame.html*) as Forward Model. The IAC retrieval module has been developed within the activity for the development of the FORUM end2end simulator. The IAC retrieval code has been exploited in Chapter 7 to analyse the TAFTS spectra at nadir acquired from aircraft, from which the water vapour profiles in the troposphere have been estimated.

3.7 Ancillary data for the simulation of FORUM spectra

In this Section, all the ancillary data necessary as input data for the Forward Model computations are listed and described. The data are divided into two different Subsections, the first describing the input data for the definition of the atmospheric scenarios to be simulated and the second describing the input spectroscopic data necessary for the line-by-line calculation performed by both GBB-NADIR and KLIMA.

3.7.1 Atmospheric Data

The preliminary studies to investigate the FORUM and FIR radiation potentiality need to be performed in a wide set of different atmospheric scenarios, in order to simulate the TOA radiance for all the extreme atmospheric conditions that FORUM will sound along its polar orbit. The following paragraphs explain the atmospheric datasets considered for the tests performed in the present work.

IG2 Database One of the most comprehensive atmospheric profiles database that had been used for FORUM sensitivity tests (described in Chapter 4) is the IG2 database, a climatological set of atmospheric profiles used as reference for the analysis of MIPAS/ENVISAT spectra (Remedios et al. (2007)). The database includes altitude distributions of pressure, temperature and 36 species (all gases whose spectral signature is in the 700-2500 cm^{-1} range) on a common altitude grid, from 0 to 120 km at 1 km step. The IG2 climatological database contains day and night profiles for 6 latitudinal bands, namely Polar North (90:65), Mid-Latitude North (65:20), Equatorial North (20:0), Equatorial South (0:-20), Mid-Latitude South (-20:-65), Polar South (-65:-90), 4 seasons and 12 years (2002-2013). The last version of the IG2 database (v. 5.1) has been used for the sensitivity tests, because it takes into account vertical profiles of some minor gases like PAN (peroxyacetilnitrate), phosgene ($COCl_2$) and propane (C_3H_8) recently added to the atmospheric spectroscopic databases. Given the completeness of the IG2 database, it can be used as a self-standing atmospheric dataset for the simulations of FORUM. Moreover, it can be exploited to complete the atmospheric scenarios defined with other datasets not containing a comprehensive set of the trace gases VMR or defined on a vertical grid not extended from surface to the top of the atmosphere.

ECMWF ERA-Interim reanalysis dataset The ERA-Interim (Dee et al., 2011) is a global atmospheric reanalysis produced by ECMWF. This reanalysis covers the time period from 1^{st} January 1979 to 31^{st} August 2019 with four analyses per day referred to the synoptic hours 00:00, 06:00, 12:00, and 18:00 UTC (Universal Time Coordinate). Gridded data products include a wide set of surface parameters and upper-air parameters covering the troposphere and stratosphere, including vertical profiles of pressure, temperature, water vapour and ozone, the surface temperature (both over land e sea) and some cloud data. The VMRs of species different from water vapour and ozone shall be supplied by the corresponding IG2 profiles in season and latitude. The ERA-Interim can provide data on a very fine 0.175° x 0.175° horizontal grid with the altitude grid defined from the surface up to 1 hPa on 37 pressure levels, providing a vertical resolution in the UT/LS (Upper Troposphere / Lower Stratosphere) generally ranging between 0.7 and 1.2 km. If needed, the profiles from DPD are extended at higher altitudes with the IG2 data, up to 120 km as a maximum.

ECMWF 83 Diverse Profiles Dataset The ECMWF (European Center for Medium Range Weather Forecast) 83 Diverse Profile Dataset (DPD) is an atmospheric dataset developed by Matricardi (2008) and exploited as training set for the RTTOV radiative transfer code (currently at version 12, Saunders et al. (2018)) and includes 83 different clear sky scenarios. This dataset has been extracted from a wider Diverse Profile Dataset containing a representative collection of 25000 different atmospheric profiles (Chevallier et al., 2006). The DPD dataset contains vertical profiles on 101 pressure levels for 83 different atmospheric scenarios from the ECMWF reanalysis between 2006 and 2007 evenly covering the Earth surface, as shown in the Figure3.2.

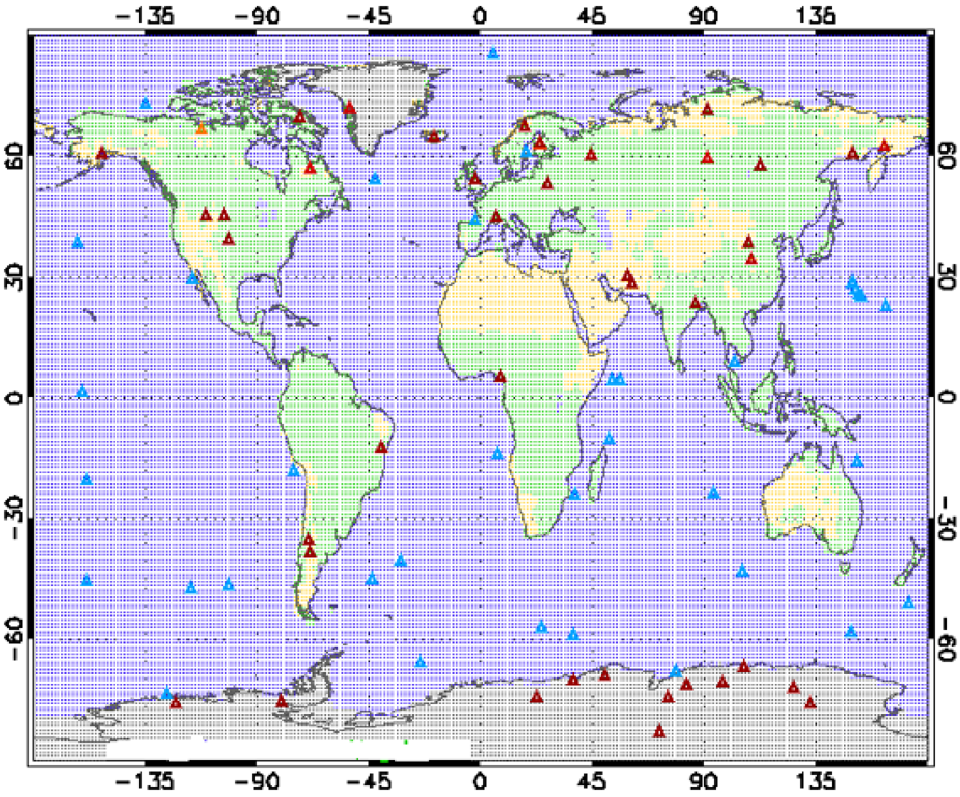


Figure3.2: Geolocations of the 83 DPD scenarios, the light blue triangles indicate the scenarios above water surface, the red ones indicate scenarios over land

Vertical profiles of temperature, pressure and some of the most relevant absorbing gases in the mid infrared range (namely H_2O, CO_2, O_3, CO and CH_4) are provided for each scenario. The VMR of the other minor species are supplied by the corresponding IG2 profiles in season and latitude. Since the highest atmospheric level is located at 80 km above sea level, if needed the profiles from DPD are extended at higher altitudes with the IG2 data, up to 120 km as a maximum.

The 21 ice clouds scenarios dataset In order to simulate the TOA radiance in cloudy conditions, a set of 21 ice cloud scenarios randomly selected out of the DPD by Laurent Labonnote from LOA (Laboratoire d'Optique Atmosphérique - Lille, France) has been used. The scenarios are referred to different geolocations, seasons and cloud optical depths including subvisible to thick clouds, as summarized in Table3.1.

Table3.1: Description of the ice cloud scenarios considered for tests with clouds. For each scenario identifier (first column), the table shows the geolocation (Latitude and Longitude), the total Ice Water Path (IWP), the ice optical depth (OD), the surface temperature (Tsurf), and the date (year, month and day)

Scenario	LAT(°)	LONG(°)	IWP(g/m^2)	OD	Tsurf(K)	Date(YYYY-MM-DD)
1	-82.30	43.00	<0.01	<0.01	213.97	2014-6-1
2	32.68	53.00	0.01	<0.01	279.61	2013-10-10
3	-77.24	-62.40	0.03	<0.01	253.88	2013-11-1
4	67.26	176.84	0.05	<0.01	252.34	2014-2-10
5	59.39	-124.53	0.27	0.01	273.25	2014-4-20
6	50.39	26.04	0.48	0.02	285.20	2014-5-1
7	-31.28	27.04	0.50	0.02	276.16	2014-6-20
8	0.63	27.00	0.87	0.03	296.94	2014-2-10
9	-33.24	112.92	12.51	0.49	290.20	2013-11-1
10	20.45	8.00	1.46	0.06	290.53	2014-1-1
11	48.99	89.38	16.46	0.65	291.76	2014-7-20
12	-68.53	147.00	2.11	0.08	253.99	2014-3-10
13	80.90	-147.50	24.08	0.95	245.08	2014-2-20
14	30.57	-8.00	25.29	1.00	287.27	2013-12-10
15	30.43	-10.24	3.48	0.13	288.04	2014-3-20
16	-68.25	-70.00	3.67	0.14	272.07	2014-2-1
17	52.50	89.32	41.24	1.63	262.66	2014-1-1
18	38.87	58.31	59.40	2.34	307.64	2013-9-20
19	-33.24	-137.78	7.50	0.29	290.71	2014-5-1
20	-1.76	111.94	8.90	0.34	297.70	2014-2-1
21	38.73	59.02	9.89	0.38	292.54	2014-5-20

Gaseous species not present in the IG2 database The VMR profiles of some of the molecules showing weak absorbing signatures in the FORUM spectral range are not available in the IG2 and DPD databases, therefore their profiles have been extracted from both models and literature. Vertical volume mixing ratio profiles of OH, HF, HCl, HBr, HI and H_2CO for a standard atmosphere are extracted from the ancillary data of the LBLRTM (Line-By-Line Radiative Transfer Model) code, whereas profiles of H_2S, HCOOH, HO_2, HOBr, C_2H_4, CH_3OH, CH_3CN, SO_3 and SF_5CF_3 molecules are taken from literature: H_2S from McElroy et al. (1980), HCOOH from Millet et al. (2015), HO_2 from Carli and Carlotti (1992), HOBr from Johnson et al. (1995), C_2H_4 from Aikin et al. (1982), CH_3OH from Cady-Pereira et al. (2012), CH_3CN from Brasseur et al. (1983), SO_3 from Reiner and Arnold (1997) and SF_5CF_3 from Sturges et al. (2000). When the VMR profiles found in literature are available just as plots, then the WebPlotDigitizer tool (*http://arohatgi.info/WebPlotDigitizer/*) is used to retrieve numerical data from the profiles.

Surface parameters The spectral surface emissivity is taken from Huang et al. (2016) database providing emissivity in the whole thermal range from 100 cm^{-1} to 3000 cm^{-1} for 16 different surface types, including desert, green covered and frozen surfaces. Skin temperature is constrained to be equal to the atmospheric temperature at the lowest atmospheric level when a scenario from the DPD is considered. For The IG2 scenarios, the seasonal zonal mean skin temperature for each latitudinal band is calculated from the ERA-Interim data and considered as representative for the climatological skin temperature. The surface parameters coupled to each atmospheric dataset selected to simulate FORUM are summarized in Table3.2.

Table3.2: Surface parameters used for each atmospheric dataset selected for the simulation of the FORUM FSI measurements

Atmospheric Dataset	**Surface temperature**	**Surface emissivity**
IG2	Seasonal zonal average from the ERA-Interim data	1.0 (no albedo)
DPD	Equal to the atmospheric temperature at 0 m	Vegetation, Snow, Desert, Ocean from Huang et al. (2016), depending on geolocation
Ice Clouds Scenarios	Equal to the atmospheric temperature at 0 m	1.0 (no albedo)
Era-Interim	Supplied by the dataset	Determined by geolocation (emissivity from Huang et al. (2016))

3.7.2 Spectroscopic data

The spectroscopic data are fundamental auxiliary data to properly simulate the gaseous absorption of the atmosphere, generally provided by different officially released spectroscopic databases. Each database reports the spectroscopic properties of each absorption line necessary for the calculation of the cross section, as reported in Subsection3.1.3. For complex and heavy molecules the spectroscopic data are available as pre-calculated cross sections that are tabulated at different pressure and temperature pairs. A brief description of the spectroscopic databases from which the spectroscopic data for the GBB and KLIMA models are extracted is presented in the following.

The HITRAN database The HITRAN (High Resolution TRANsmission) database is one of the most used spectroscopic databases for Earth′s atmosphere observations. It was developed in 1973 by the Air Force Cambridge Research Laboratories (McClatchey et al. (1973), its name was AFCRL database, it became AFGL later on and was finally changed into HITRAN in 1986) and was originally composed by spectroscopic data for 7 molecular species present in the Earth’s atmosphere. The HITRAN database has been continuously updated with new releases every 4-6 years, so the latest versions contain the spectroscopic data of the majority of the molecules composing the terrestrial atmosphere. In the following Subparagraphs, we describe the versions of the HITRAN database that have been exploited by the GBB-NADIR and KLIMA codes for the tests performed in this thesis.

HITRAN 2008 The HITRAN2008 database (Rothman et al., 2009) contains line-by-line spectroscopic parameters for 42 molecules, including many of their isotopologues. With respect to the previous version HITRAN2004, the 2008 database contains new entries in the line parameter files for gaseous species like methyl bromide (CH_3Br), methyl cyanide (CH_3CN) and tetrafluoromethane (CF_4). Anyway, the database contains a lower number of transition lines respect to the following versions, limited by the use of a cut-off in the absorption intensity (based on the sensitivity of the instruments that observe absorption over extreme terrestrial atmospheric path lengths).

HITRAN 2012 The HITRAN2012 compilation, described in Rothman et al. (2013), is freely accessible on the internet at the website *http://hitran.org/data-index/*, and includes line-by-line spectroscopic parameters for 47 molecular species with 120 isotopologues and infrared absorption cross-sections, generally representing the absorption by large polyatomic molecules for which spectral line parameters are incomplete or unavailable. The spectral coverage of the line-by-line part of HITRAN2012 ranges from the microwave region through the visible. Both line-by-line and absorption cross sections files show significant additions to those of

HITRAN2008.

HITRAN 2016 The HITRAN2016 compilation is the latest version of the HITRAN spectroscopic database. It has introduced some important innovation with respect to the 2012 version. The line-by-line part and the absorption cross sections part of the new edition have now been stored into an underlying relational database structure as opposed to the long-standing fixed-length ASCII record format adopted until HITRAN2012. The new HITRAN is greatly extended in terms of accuracy, spectral coverage and additional absorption phenomena. Moreover, molecules, isotopologues (for carbon dioxide, up to 12 different isotopologues are considered in the compilation), and new gases have been added in order to address the problems of analysing data on planetary atmospheres beyond the Earth. Line-by-line data of $COCl_2$ (phosgene) have been added. Moreover experimental IR cross-sections for almost 300 additional molecules important in different areas of atmospheric science have been added to the database (Gordon et al., 2017).

HITRAN 2020 To date, the version 2020 of the HITRAN database has not been officially released yet, however the spectroscopic parameters of the first isotope of the water vapour in the FIR range have been supplied by Jonathan Tennyson from UCL (University College London - London, UK). In the present work, the HITRAN2020 database - when mentioned - is referred to a hybrid database containing the new line parameters for the first isotope of the water vapour from 100 to 700 cm^{-1} and the line parameters from HITRAN2016 for all the other species.

GEISA 2015 The GEISA database (Gestion et Etude des Informations Spectroscopiques Atmospheriques: Management and Study of Atmospheric Spectroscopic Information) has been developed and maintained by the ARA/ABC(t) group at LMD (Laboratoire de Meteorologie Dinamique) since 1974. GEISA is constantly evolving, taking into account the best available spectroscopic data. The GEISA-2015 (Jacquinet-Husson et al., 2016) is freely available on the website *http://cds-spri.ipsl.upmc.fr/etherTypo/?id=1729&L=0*, it contains a line parameter database with 52 molecular species (118 isotopologues) and transitions in the spectral range from 10^{-6} to 35877.031 cm^{-1}, representing 5.067.351 entries. Among the previously existing molecules, 20 molecular species and 7 new isotopologues have been updated to the GEISA2015 database.

AER-LBLRTMv3.5 The version 3.5 of the AER spectroscopic database has been developed for the LBLRTMv12.6 radiative transfer model, and it contains spectroscopic data ranging from 0 to 25232 cm^{-1}. The line parameter files are available both as total line list and as separate individual line lists for each molecule. The database was built starting from HITRAN2012, with exceptions for H_2O, CO_2, CH_4 and O_2. Specifically, for the data between 240 and 510

cm^{-1}, the air-broadened half-widths of 59 H_2O lines were modified based on the analysis of REFIR-PAD (Palchetti et al., 2015) measurements taken at Cerro Toco (Chile), as part of the RHUBC-II campaign, and a re-analysis of the AERI-ER measurements (for wavenumbers greater than 400 cm^{-1}) from the ARM NSA site (first analysed in Delamere et al. (2010)), as reported on AER-LBLRTM website *http://rtweb.aer.com/line_param_frame.html*. Versions v3.6 and v3.7 of the AER database have been released in the last years, however these do not present any relevant difference in the line parameters with respect to the v3.5 within the spectral range of FORUM.

Water vapour continuum parametrisation The water vapour continuum parametrisation described in paragraph3.1.3 provides extra-absorption contribution to that provided by the line-by-line absorption with slow frequency dependence, with the aim to obtain the agreement between measurements and simulations. This has been released over the years in different versions using the same MT_CKD formalism with different parameters for the self and the foreign broadening leading to different magnitude of the absorption supplied by the continuum (see *https://github.com/AER-RC/mt-ckd/wiki*). The MT_CKDv2.5, v3.0 and v3.2 versions of the water vapour continuum have been implemented in the GBB-NADIR and KLIMA codes. However, each continuum version needs to be consistent with the exploited spectroscopic data, since the correction applied on the cross sections has to fix the Lorentzian part of the absorption far from the line centre depending on the pressure broadening coefficient of the most intense lines tabulated in the database. This point will be explained more in detail in Chapter 6, in which the consistency between real measurements in the FIR and the simulations developed with different spectroscopic databases and continuum versions will be fully investigated.

Chapter 4

Sensitivity Tests

In this Chapter, the tests performed to study the sensitivity of the TOA FIR radiation and of the FORUM measurements to the spectral signatures of the major absorbing species, the water vapour in the UTLS, and to the FIR surface emissivity will be described. First, the signatures of the gaseous absorption of different species in the FORUM spectral region in clear sky conditions are shown in Section4.1, in particular pointing out the role of the water vapour absorption in the FIR. Then, the sensitivity of the FORUM measurements in the FIR range to water vapour and surface emissivity - that are two key variables of the FORUM Level 2 products - has been qualitatively evaluated through the study of the Jacobi matrix containing the derivatives of the spectrum with respect to the parameters, since these are directly correlated to the information content of the signal. The Jacobians have been computed in the spectral region from from 100 to 600 cm^{-1} for different atmospheric scenarios, representing the average atmospheric conditions observable in a determined latitudinal band. Finally, the information content of the radiance measured by FORUM on the relevant Level 2 Products is analysed in Section4.3 performing simulated retrievals using synthetic observations for the 80 DPD clear sky scenarios.

4.1 Major absorbing species in the FORUM range

The atmospheric gaseous species with the most relevant spectral signatures in the OLR radiation acquired by the FORUM FSI are water vapour, carbon dioxide and ozone, plus some trace gases contributing to the Greenhouse Effect like N_2O and methane. The spectral signatures of these gases on the measured signal has been investigated by simulating two spectra: one including the whole set of molecules and one excluding one of these species at time, settings its VMR to zero, then computing the differences between the two simulated radiances in a reference scenario. The top panel in Figure4.1 shows the differences computed for the Mid-Latitude Boreal Spring 2013 atmosphere from the IG2 database in clear sky, with the surface temperature set to 290 K and surface emissivity equal to 1. The spectra have been calculated from 100 to

1600 cm^{-1} at a spectral resolution of 0.0005 cm^{-1} with the GBB+DISORT model in order to resolve every single absorption line, then the resulting radiances have been convolved with a *sinc* ISRF function with Full Width at Half Maximum (FWHM) $\Delta\nu$=0.3 cm^{-1}. Where the difference between the spectra computed in the "unperturbed" and the "perturbed" scenarios is positive, then the single species produces a decrease of the thermal radiation escaping to space, on the contrary a negative difference indicates an increase of the TOA radiation due to the thermal emission of the gaseous species. The bottom panel shows the comparison between the convolved spectrum in greenish line calculated with the contribution of all the absorbing species and the one in purple line obtained by setting the atmospheric optical depth to zero, that shows the thermal emission of the Earth surface that would escape to space if no atmosphere was present.

All the considered gaseous species show a strong contribution in the absorption of the upward thermal radiation emitted by the surface and the atmosphere. In particular, we can state that the considered species determines most part of the atmospheric absorption by comparing the differences between the spectra reported in the bottom panel of Figure4.1 to the magnitude of the absorption due to each species in the top panel. The comparison shows clearly the high absorption provided by the carbon dioxide (green line) and the ozone (blue line) between 600 and 750 cm^{-1} and from 1000 to 1100 cm^{-1} respectively, with each of these species providing a reduction of the radiation escaping to space of the order of thousands of nW/(cm^2 sr cm^{-1}). The methane (black line) shows a similar contribution due to the absorption band centered at 1300 cm^{-1}, while the N_2O (purple line) shows minor spectral signatures from 600 to 1300 cm^{-1}.

However, the major contribution on the absorption of the radiation is provided by the water vapour (red line in the top panel), showing the highest spectral signature from 100 to 1600 cm^{-1} a part from the region where the absorption is dominated by the CO_2 and O_3 absorption. The water vapour is the gaseous species responsible for almost the total gaseous absorption in the FIR range, as just demonstrated by Harries et al. (2008). Therefore, in clear sky conditions, the atmospheric FIR transmittance strongly depends on the water vapour VMR. This suggests a high sensitivity of the FORUM measurements in the FIR to the water vapour concentration.This will be investigated in the next Section4.2 for different atmospheric scenarios.

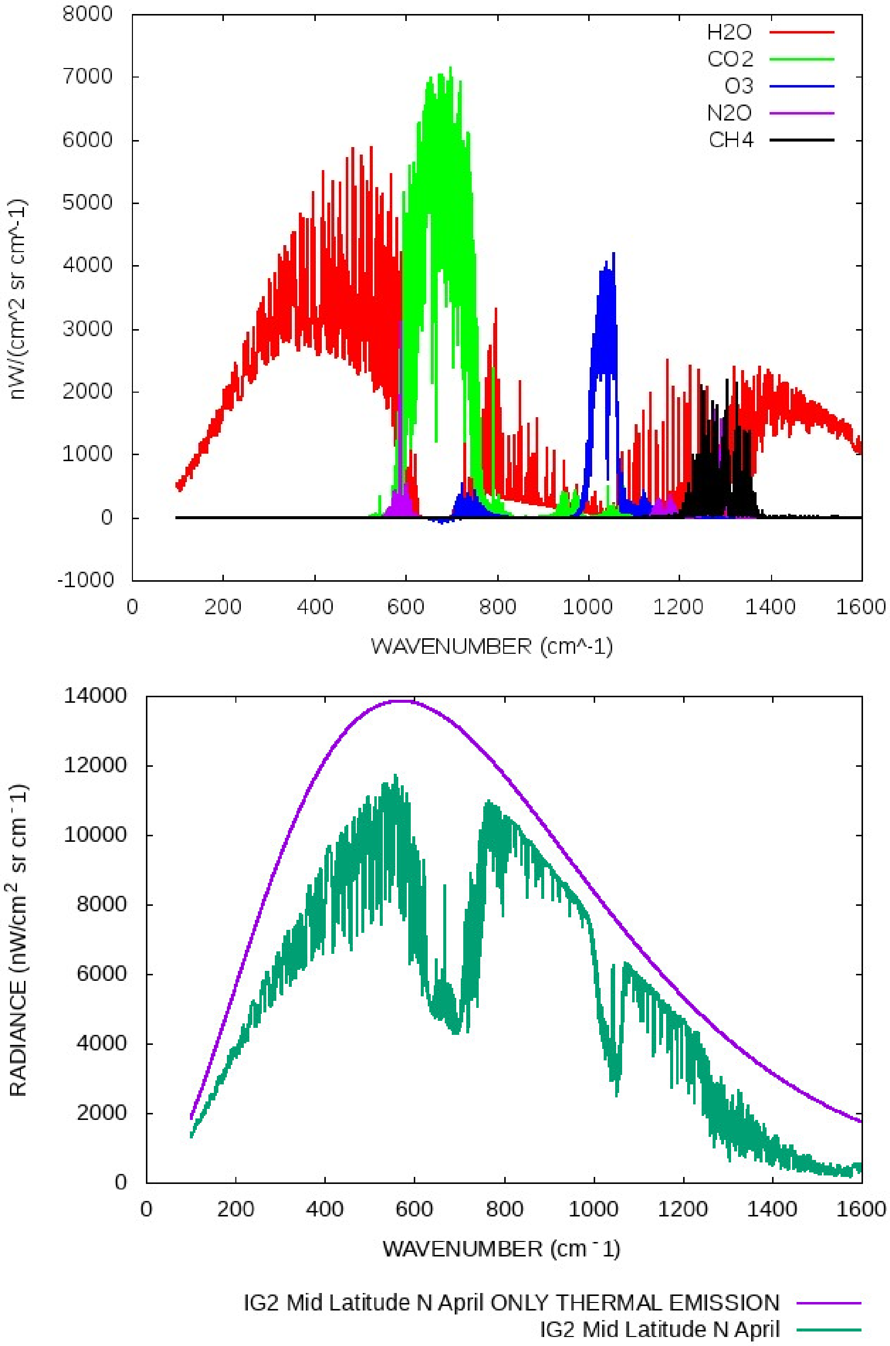


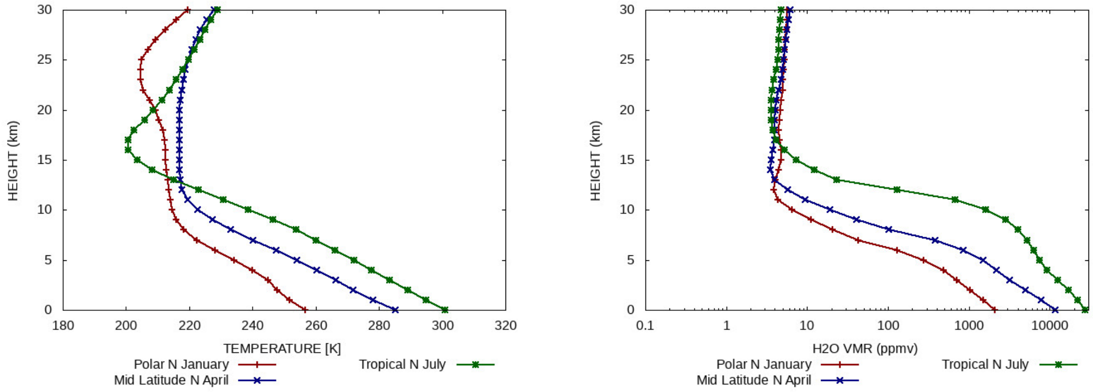
Figure4.1: Top panel: differences between the spectrum simulated with the IG2 Mid-Latitude Boreal 2013 Spring atmosphere and the spectra simulated in the same scenario setting to 0 the VMR of the H_2O (red line), CO_2 (green line), O_3 (blue line), N_2O (purple line) and CH_4 (black line). Bottom panel: difference between the spectrum observed in the IG2 Mid-Latitude Boreal 2013 Spring atmosphere and the spectrum calculated with τ equal to zero

4.2 Water vapour and surface emissivity Jacobians

The Jacobians from 100 to 600 cm^{-1} with respect to the water vapour VMR and the surface emissivity have been examined for different climatological atmospheres selected from the IG2 database and the results will be reported in the next Subsections. The analysis aims to provide insight in the sensitivity of the FORUM acquired signal in the FIR for scenarios representative of the mean atmospheric conditions that can be observed in different part of the globe. Therefore, the Jacobians have been calculated at very high spectral resolution (0.0005 cm^{-1}) with the KLIMA RTM and then have been convolved with the FORUM ISRF function described in Section3.3. The scenarios considered are:

- the Polar North Winter atmosphere of the year 2013, with surface temperature set to 2772.5 K
- the Mid-Latitude North Spring atmosphere of the year 2013, with surface temperature set to 290 K
- the Tropical North Summer atmosphere of the year 2013, with surface temperature set to 300.93 K

The used water vapour and temperature profiles from 0 to 30 km of the Polar scenario (red lines), Mid Latitude scenario (blue lines) and Tropical scenario (green lines) are reported in Figure4.2. The Total Column of Water Vapour (TCWV) in the selected scenarios ranges from 3 Kg/m^2 in the Polar atmosphere to 50 Kg/m^2 in the Tropical atmosphere, range of values that includes the TCWV values of most part of the atmospheric scenes that will be observed by FORUM over the whole globe.



(a) Temperature

(b) Water vapour VMR

Figure4.2: Temperature (left panel) and water vapour (right panel) atmospheric profiles for the selected Polar (red dotted lines), Mid-Latitude (blue dotted lines) and Tropical (green dotted lines) scenarios

4.2.1 Sensitivity to water vapour

The sensitivity of the FORUM measurements in the FIR to the water vapour concentration is studied through the examination of the Jacobi matrix. The physical meaning of the water vapour Jacobian (in nW/(cm^2 sr cm^{-1} ppmv)) can be interpreted as the variation in the measured signal due to a 1 ppm increase at defined atmospheric level, being negative if the increment of the VMR value causes a reduction of the acquired signal or positive if it leads to an increase of the acquired signal. Therefore the sensitivity to the atmospheric water vapour VMR can be evaluated with the absolute value of the Jacobian in the selected atmospheric scenarios, shown by the maps in Figure4.3 as a function of the altitude (from 0 to 25 km) and the wavenumber. The light grey colour corresponds to an absolute value of the Jacobian lower than 10^{-3} nW/(cm^2 sr cm^{-1} ppmv).

The highest absolute values of the Jacobian are observed from 7 to 12 km above the surface below 400 cm^{-1} in the Polar scenario (top panel of Figure4.3). The water vapour Jacobians calculated in the Mid-Latitude (central panel of Figure4.3) and Tropical (bottom panel of Figure4.3) scenarios show a similar sensitivity peak that is shifted upward to between 10 and 15 km in the Mid-Latitude atmosphere and between 13 and 17 km in the Tropical scenario, becoming more intense as the water vapour of the analysed atmospheric scenario increases. Comparing the maps of Figure 4.3 with the plots of Figure 4.2, we can see that the altitude at which the maximum sensitivity occurs is close to the tropopause height, identified by the inversion observable in the temperature profiles in the left panel of Figure4.2. This confirms that the FIR range is sensitive to the UTLS humidity, and the sensitivity is larger in the Tropical scenarios. The absolute values of the Jacobians for altitudes lower than 5 km are almost negligible for the Tropical scenario as the intense absorption due to the very high water vapour content totally masks out the signal coming from the lower troposphere. Differently, the tropospheric water vapour concentration weakly contributes to the signal acquired by the FORUM FSI in the Polar scenario, with non negligible values of the Jacobian at surface level for wavenumbers higher than 300 cm^{-1}. This is due to the higher transparency in the FIR related to the low water vapour content. This suggest that part of the FIR radiation escaping to space could be supplied from the surface and that the FIR radiation can contain information on the surface properties in the Polar regions. This point will be investigated in the next Subsection4.2.2

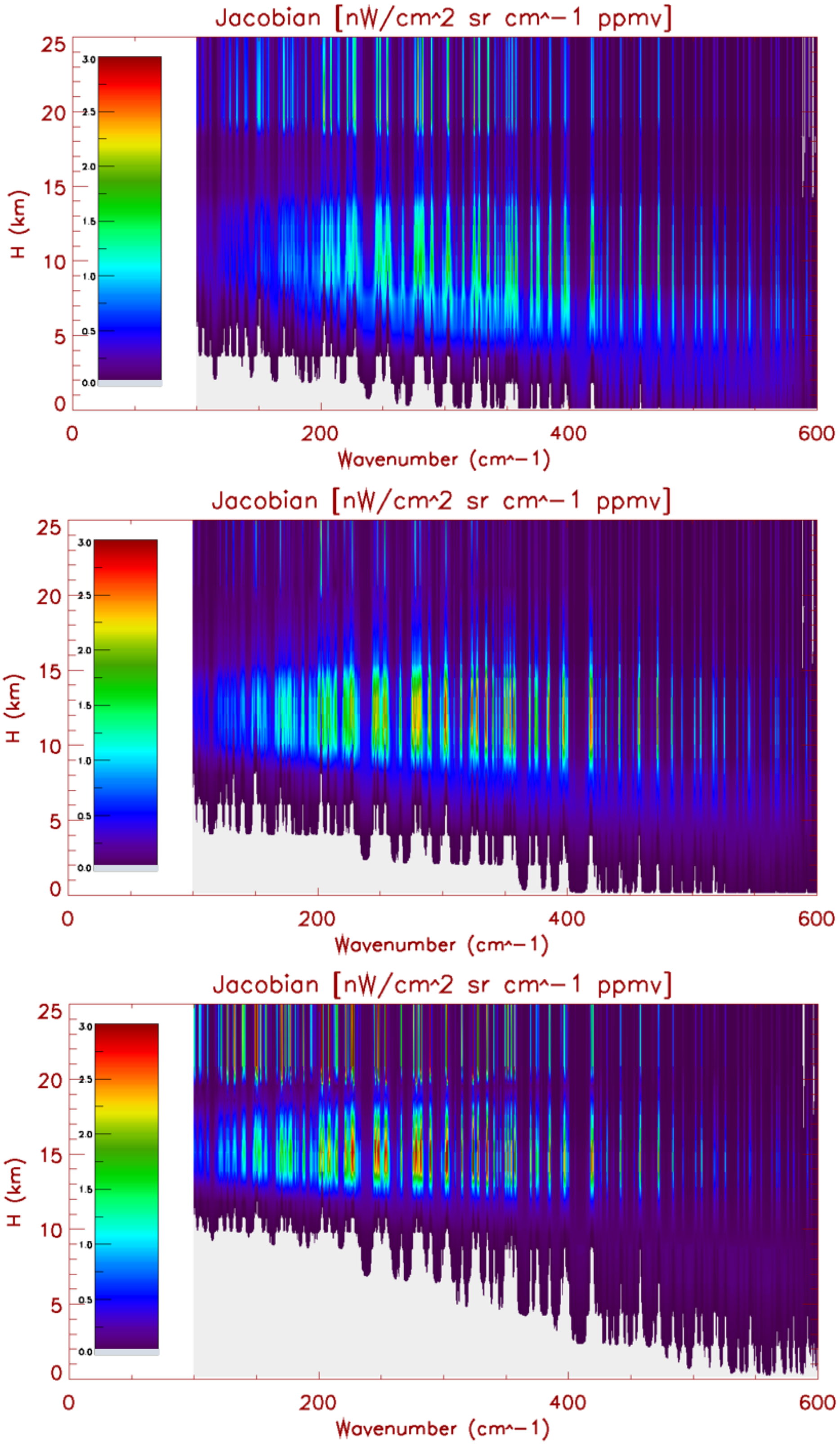


Figure4.3: Water vapour Jacobian with respect to water vapour VMR from 100 to 600 cm^{-1} for the selected Polar (top panel), Mid-Latitude (central panel) and Tropical (bottom panel) scenarios. The values range from 0 to 3 nW/(cm^2 sr cm^{-1} as shown by the colour palette, the absolute values lower than 10^{-3} nW/(cm^2 sr cm^{-1}) are reported in grey

4.2.2 Sensitivity to surface emissivity in the FIR

The analysis performed on the Jacobian with respect to water vapour VMR reported in the previous Subsection4.2.1 suggests that the FIR measurements in warm and humid atmospheric scenes at mid-latitude and tropical latitudes are not sensitive to the surface and lower troposphere properties. On the other hand, we expect a higher sensitivity to the surface properties in the Polar regions. From the analytical point of view the atmospheric transmittance strongly affects the Jacobian with respect to the surface emissivity, that we can obtain by deriving both sides of Eq.3.16 with respect to $\varepsilon(\nu)$

$$\frac{dS(\nu)}{d\varepsilon(\nu)} = [B(T_S) - \frac{1}{\pi}\int_0^{2\pi} d\phi \int_0^1 dcos(\theta) \int_0^{\tau_0} B(T)e^{-\tau^{'}} d\tau^{'}]T_{TOT} \tag{4.1}$$

$$\tau^{'}(\nu, z) = \tau(\nu, z)/\mu \tag{4.2}$$

where T$_{TOT}$ is the transmittance of the total atmospheric path and μ=cos(θ). The sensitivity of the TOA radiance to the surface properties depends principally on the optical properties of the atmosphere and does not show any dependence on the surface type. The Jacobians for the surface emissivity calculated for the three selected IG2 scenarios, are shown in Figure4.4, reporting, in the top panel, the Polar surface emissivity Jacobian, the Mid-Latitude Jacobian in the central panel and the Tropical Jacobian in the bottom panel.

The sensitivity to the surface emissivity of the FORUM acquired signal in the average Polar scenarios is large for wavenumbers higher than 400 cm^{-1}, with Jacobians values higher than 1000 nW/(cm^2 sr cm^{-1}) far from the absorption lines, while it is strongly reduced in the Mid-Latitude scenario with values lower than 100 nW/(cm^2 sr cm^{-1}) for wavenumbers higher than 500 cm^{-1} and almost negligible for lower frequencies. For the very high water vapour content typical of the Tropical scenarios, no information can be observed on the surface emissivity, as the surface optical properties are totally masked out by the very intense water vapour absorption due to the large TCWV.

A further test to investigate more in detail the sensitivity to surface properties in Polar regions can be performed by comparing the Jacobian calculated in the Polar climatological scenario and the one calculated in an extremely dry scenario typical - for instance - of the conditions that can be observed on the Antarctic Plateau during the Austral Winter. Therefore, the surface emissivity Jacobian has been calculated reducing by a factor of 10 the water vapour VMR along the atmospheric column for the IG2 Polar scenario. The used water vapour profiles are compared in Figure4.5 (blue line for the water vapour profile in the "dry" scenario, green line for the water vapour in the climatological scenario), while the Jacobians are shown in the left panels of Figure4.6 compared to the corresponding transmittance profiles (right panels of Figure4.6) of the atmosphere. The Jacobians and the atmospheric transmittances in the climatological IG2 scenario are reported in red, while the ones in the "dry" Polar scenario with the dark yellow line.

We see in Figure 4.6 that the FIR is not sensitive to the surface properties below 400 cm^{-1} in the climatological scenario. However, part of the radiation from 250 to 400 cm^{-1} coming from the surface could escape up to the TOA when the water vapour content is strongly reduced. We also see that in that case the sensitivity from 400 to 600 cm^{-1} is sharply enhanced with respect to the climatological scenario. This confirms that the FIR sensitivity to the surface emissivity depends almost totally on the water vapour content that is the main contributor to the atmospheric transmittance in that spectral region. This point is enforced by the comparison between the surface emissivity Jacobians and the atmospheric transmittance of the corresponding scenario, showing very similar spectral shape.

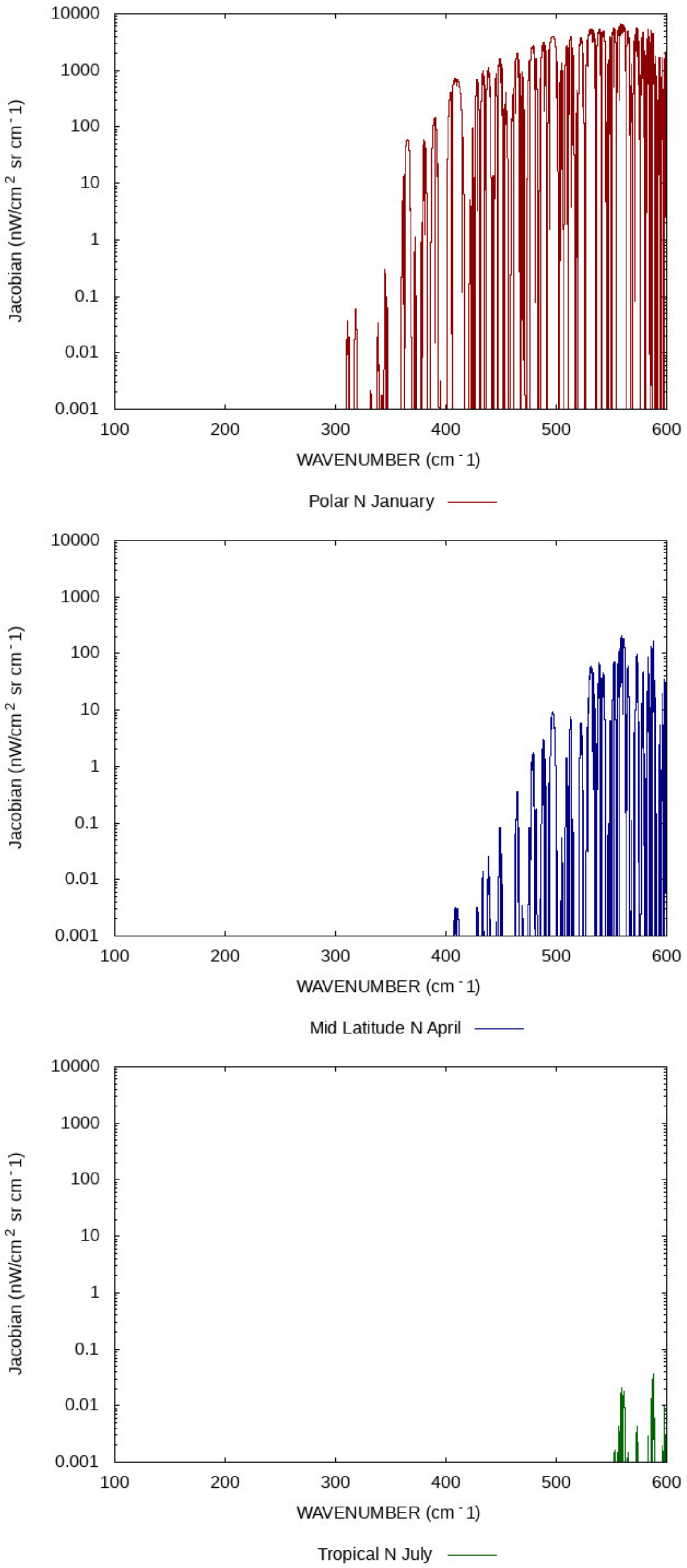


Figure4.4: Jacobian respect to the surface emissivity from 100 to 600 cm^{-1} for the selected Polar (top panel, red line), Mid-Latitude (central panel, blue line) and Tropical scenarios (bottom panel, green line). The Jacobian values are reported in logarithmic scale

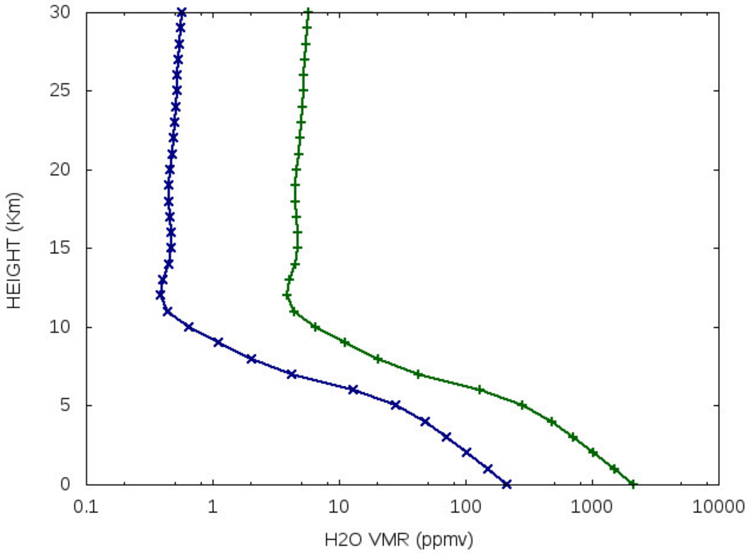
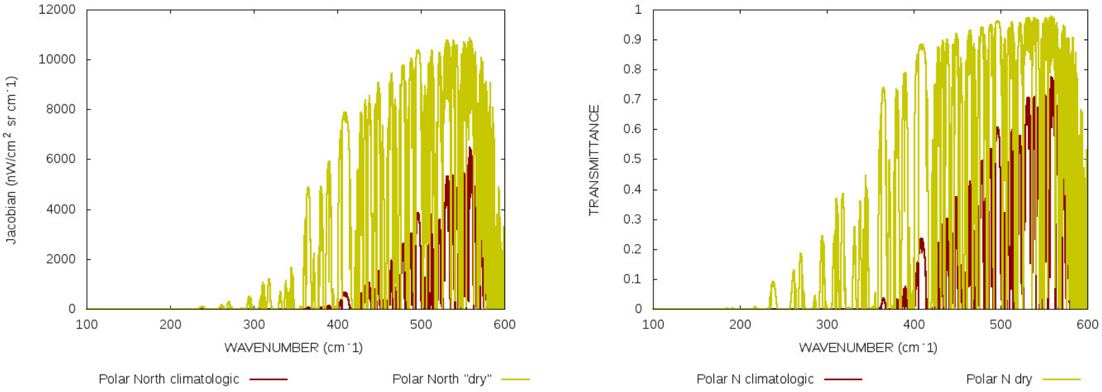


Figure4.5: Water vapour profiles of the climatological Polar IG2 scenario (green line) and the "dry" polar scenario (blue line)



(a) Surface emissivity Jacobian

(b) Atmospheric transmittance

Figure4.6: Jacobian with respect to the surface emissivity (left panel) for the climatological IG2 Polar scenario (red line) and for the "dry" Polar scenario (dark yellow line) and the atmospheric transmittance (right panel) for the climatological IG2 Polar scenario (red line) and for the "dry" Polar scenario (dark yellow line) from 100 to 600 cm^{-1}. The figures show clear evidence of the high correlation between the sensitivity to the surface properties and the atmospheric opacity

4.3 Simulated Retrievals

In this Section the results obtained performing retrieval simulations are reported. These tests have been performed in order to study the information content of the FORUM measurements in different atmospheric scenarios. The retrievals have been performed on synthetic observations using the KLIMA code that is based on the Optimal Estimation Method using the Gauss-Newton iterative algorithm (see Appendix A). The synthetic observations have been simulated for the 80 DPD clear sky scenarios using the GBB-NADIR code for the optical depths calculation and the LIDORT solver for the calculation of the high resolution radiances. The instrumental response is simulated using the ISRF function described in Section3.3. The retrieval tests have been performed with the following characteristics:

- The retrieval analysis is performed using the KLIMA code
- A multi-target retrieval approach is used. The retrieval state vector includes the surface temperature and surface emissivity, and the vertical distributions of water vapour VMR, ozone VMR, atmospheric temperature.
- The emissivity is retrieved every 50 cm^{-1} from 100 to 1600 cm^{-1}, the vertical retrieval grid for the atmospheric parameter is defined with vertical steps of 2 km from surface to 8 km, 1 km steps from 8 to 20 km and 5 km steps from 20 km up to the TOA set to 80 km.
- the a-priori values for H_2, O_3 and T(z) are set to the values used by the UK MetOffice for the assimilation of the IASI radiances in the NWP (Numerical Weather Prediction) Model, and are reported in Table4.1. A correlation length of 5 km for the atmospheric variables is used. The assumed a-priori errors on surface emissivity and surface temperature are respectively 0.1 and 0.2 K.
- the assumed instrumental noise of the FORUM FSI is the goal noise as reported in Table2.1
- the Initial Guess of all the parameters is set equal to the sum of the true values and the a-priori errors

To study the performances of the retrievals we first examine the convergence of the performed retrieval tests. As reported in Appendix A, the Optimal Estimation Method is a procedure estimating the state vector through the minimization of the χ^2. The reduced chi-square values (see Eq.8.12), showing the agreement between the simulations using the retrieved state vector and the synthetic observations, are reported in Figure4.7 for the initial (red dots) and final (green dots) steps of each retrieval test. We can observe that each retrieval test starts from chi-square values of the order of 10^2-10^3, and when the convergence is achieved the chi-square

Table4.1: A-priori errors defined for atmospheric water vapour, ozone and temperature

Height (km)	Water vapour (%)	Ozone (%)	Temperature (K)
120.0	10.0	10.0	1.4
64.6	10.0	10.0	1.4
32.1	10.0	10.0	0.6
15.8	10.0	10.0	0.6
8.0	30.0	10.0	0.5
2.6	30.0	10.0	0.5
0.0	3.0	10.0	0.5
Correlation length (km)	5.0	5.0	5.0

values are close to 1 for most part of the scenarios, indicating the consistency between the synthetic observations and the simulations at the end of the retrieval procedure. Therefore, the information contained in the signal allows to get a state vector correctly reproducing the observed radiances.

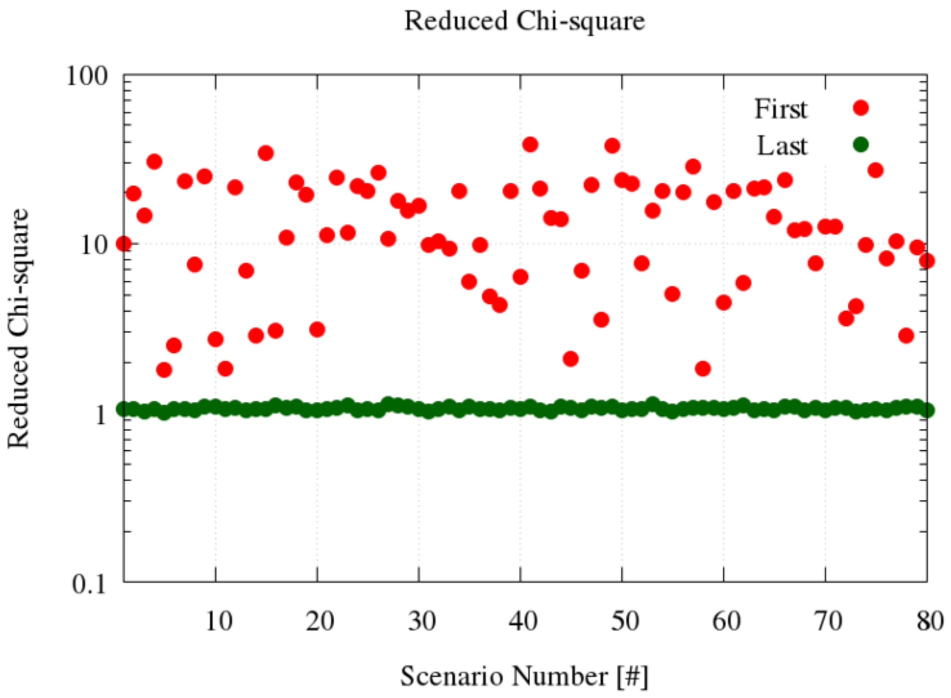


Figure4.7: Initial (red dots) and final (green dots) reduced chi-square for the 80 retrieval tests

The quantities used to assess the quality of the retrievals are the accuracy of the retrieval, defined as the difference between the “true” parameters and the retrieved ones, and the precision, given by the mapping of the measurement noise and the a-priori errors onto the retrieved parameters, and defined by the square root of the diagonal elements of the Variance Covariance Matrix of the retrieved parameters S_x (see Appendix A) The mean accuracy and precision of the retrieval tests in the 80 DPD scenarios on water vapour, temperature and surface emissivity have been calculated in different latitudinal bands (namely Polar, Mid-Latitude and Tropical). The latitudinal bands have been defined with to the following criteria:

- the Polar Band includes the DPD scenarios referred to geolocations with absolute latitude higher than 65 degrees
- the Mid-Latitude Band includes the DPD scenarios referred to geolocations with absolute latitude ranging from 20 to 65 degrees
- the Tropical Band includes the DPD scenarios referred to geolocations with absolute latitude lower than 20 degrees

The mean accuracy and precision of the retrieval tests in the different defined latitudinal bands are shown in Figure4.8. The random error in the UTLS water vapour is strongly reduced in the Tropical scenarios (solid red line in top left panel of Figure4.8) with respect to Mid-Latitude and Polar scenarios (solid green and blue lines respectively), coherently to the maximum of the Tropical water vapour Jacobian found at 15 km altitude as can be observed in the bottom panel of Figure4.3. High information content on FIR surface emissivity is present in the signal acquired in Polar scenarios, as can be observed by the low value of the random error of the FIR emissivity (solid blue line in bottom panel of Figure4.8), due to the transparency of dry atmospheric scenarios in the FIR highlighted already in the sensitivity tests, while the surface properties cannot be retrieved at the Tropics as the water vapour absorption totally masks the radiation emitted and reflected by the surface. In fact, the red solid lines in the bottom panel of Figure4.8 shows a random error of the FIR emissivity almost equal to the a-priori, pointing out that negligible information is provided by the measurement in this atmospheric scenario. High accuracy can be generally observed for the retrieved surface emissivity, with low values of the differences between the “true” and the retrieved state vector (dashed lines in the bottom panel of Figure4.8). The accuracy obtained for water vapour (dashed lines in the top left panel of Figure4.8) is good in the UTLS region, while is very low in the troposphere. Temperature accuracy (dashed lines in the top right panel of Figure4.8) in the UTLS is of the same order of magnitude of the precision, while is worse for the tropical bands (red dashed line). On the other hand, the accuracy of the retrieved products can be degraded by systematic effects due to uncertainties related to the approximations used in the Forward Model and to reproduce instrumental effects or to not-retrieved variables contributing to the observed radiance. These effects cannot be quantified with the retrieval tests performed on synthetic observations since they are not reproduced in the Forward Model simulations used to generate FORUM observations. A full analysis of the errors on the FORUM Level 2 Products due to a wide set of different systematic error sources will be presented in Chapter 5.

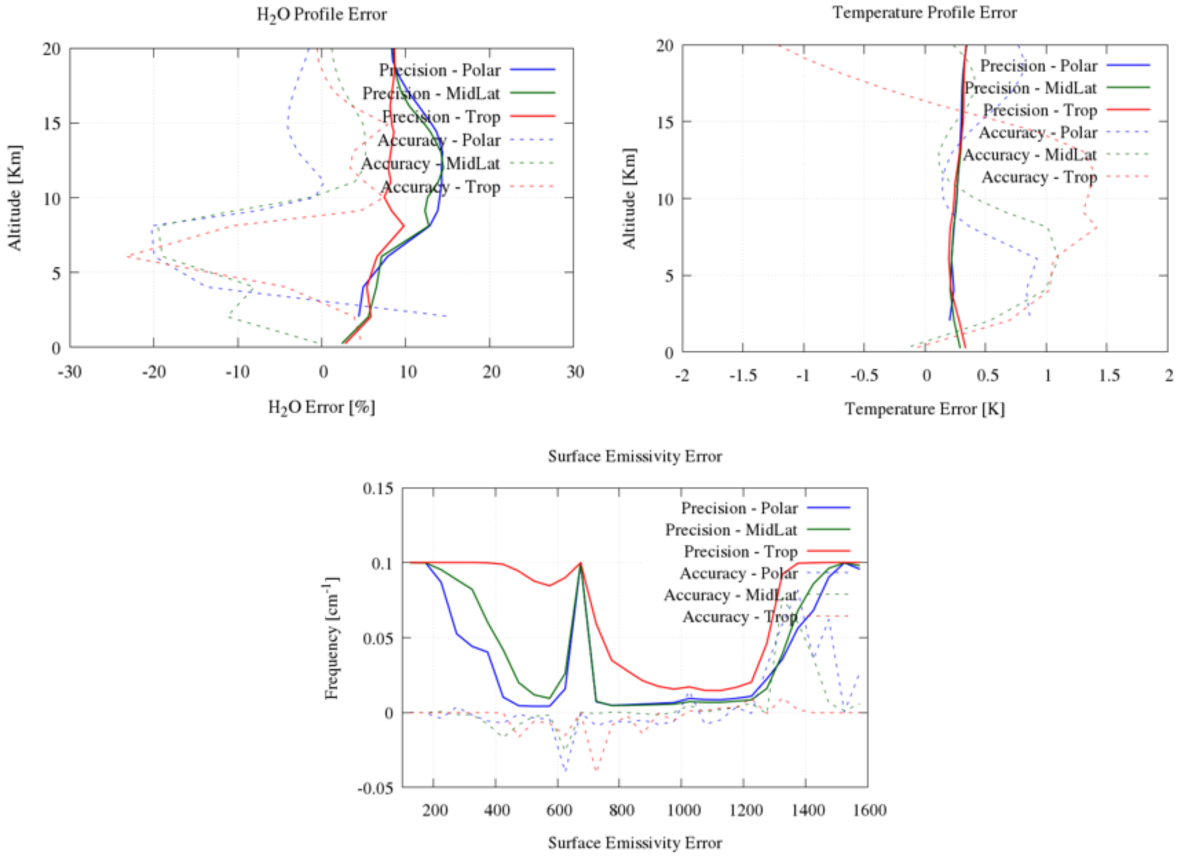


Figure4.8: Mean accuracy (dashed lines) and precision (solid lines) of the retrieved water vapour (top left panel), temperature (top right panel) and surface emissivity (bottom panel) for the Polar (blue), Mid-Latitude (green) and Tropical (red) bands

Chapter 5

Errors on the Level 2 Products

We have investigated in Chapter 4 both the precision and the accuracy of the FORUM Level 2 key products, namely the water vapour, and temperature profiles and the surface emissivity in different atmospheric conditions. Their precision depends on the instrumental noise that can be significantly different from the goal and threshold requirements defined in Chapter 2. In the Section5.1 we will calculate the precision attainable by FORUM (through the evaluation of the Variance-Covariance Matrix - VCM - of the retrieved variables as reported in Eq.8.10 of Appendix A) using two different real NESR models related to different types of detectors. The obtained values (that we will call random errors of the retrieved parameters) will be compared to those obtained using the threshold and goal NESRs reported in Table2.1. The accuracy of the Level 2 products achieved with the inversion of remote sensing measurements could be degraded by several systematic errors originating from different sources, that are generally related to the uncertainties affecting a wide set of parameters assumed as known in the Level 2 analysis. These errors have to be estimated to understand the real performances of the FORUM instrument. In Section5.2 the method used to evaluate the effect of the systematic errors on FORUM Level 2 products is briefly explained, the systematic error sources considered in the analysis are then reported and, finally, the results are reported and discussed.

5.1 Random errors comparison study

The random error on each retrieved state vector variable depends on the sensitivity of the measured signal to the parameter itself and the measurement noise. The requirements on the Goal and Threshold NESR have been determined in the preliminary studies of the FORUM mission in order to achieve good quality Level 1 and Level 2 Products. These are reported in Table2.1 and in Figure5.1, where the values are shown with the blue and black lines. In the same figure two different NESR values, coming from models of the instrumental noise evaluated for two different detectors and supplied by the industrial consortium working at the development of the FORUM instrument, are also shown. These noise models are labelled as “106 Consortium A ” (106ca green line) and “099 Consortium A” (099ca red line). The Figure5.1 shows clearly that both the 106ca and 099ca noises are compliant to the threshold and the goal requirements for most part of the spectral range of the FORUM instrument

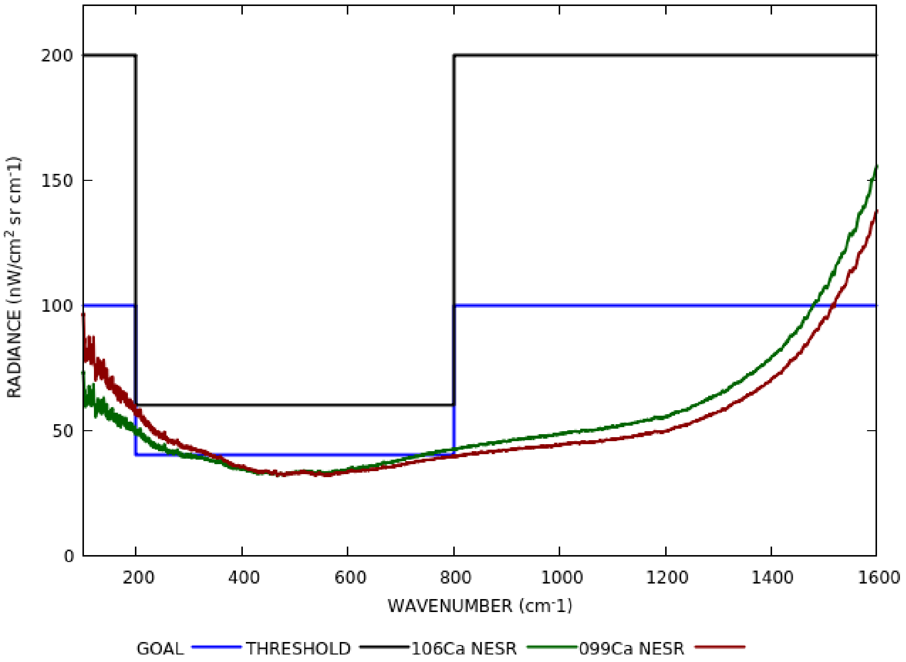


Figure5.1: 106ca (green line) and 099ca (red line) noise models compared to the goal (blue line) and threshold (black line) noise

The VCM S_x of the retrieved parameters (the square root of its diagonal elements represents the random errors on the retrieved parameters) has been calculated for the 80 DPD and 24 IG2 clear sky scenarios using the different NESR models, then the average random errors have been calculated for the same latitudinal bands described in Section4.3 and compared in order to study the effect of the different instrumental noise on the estimated uncertainties. As the GOAL noise has been considered for the retrieval tests reported in Section4.3, its value is taken as reference for the present tests. We evaluated the ratio between the random errors calculated with each retrieval configuration and the reference calculated with GOAL noise. The 106ca and 099ca noise have been obtained by the industry using updated instrumental characteristics, related to a real model of ISRF - denominated as "Consortium A ISRF" that will be described in detail Section5.3. This function has been used in the calculation of the retrieval VCM for the noise models different from the reference. The different configuration used for the present test are

- **0.413, GOAL** ⇒ Retrieval test performed with GOAL noise and ideal sinc ISRF with spectral sampling $\Delta\nu$=0.413 cm^{-1}. This is the reference random error
- **0.357, THRESHOLD** ⇒ Retrieval test performed with THRESHOLD noise and real

ISRF with spectral sampling $\Delta\nu$=0.357 cm^{-1}

- **0.357, GOAL** ⇒ Retrieval test performed with GOAL noise and real ISRF with spectral sampling $\Delta\nu$=0.357 cm^{-1}
- **0.357, 099ca** ⇒ Retrieval test performed with 099ca noise and real ISRF with spectral sampling $\Delta\nu$=0.357 cm^{-1}
- **0.357, 106ca** ⇒ Retrieval test performed with 106ca noise and real ISRF with spectral sampling $\Delta\nu$=0.357 cm^{-1}

The Figures 5.2 to 5.4 show the ratio between the random errors on the different Level 2 products calculated in the different configurations and the random errors calculated with the reference configuration. Namely, the top left panels of Figures5.2 and 5.3 are referred to the errors in the Polar band affecting water vapour and temperature respectively, the top right to the errors in the Mid-Latitude band and the bottom panel to the errors in the Tropical band. The errors on the surface emissivity have been reported in Figure5.4 for the Polar band only.

The lowest errors in the estimate of water vapour in the UTLS are obtained when using the 106ca noise (orange lines in Figure5.2), while a generally lower improvement can be observed when using the real ISRF and the GOAL noise (blue lines in Figure5.2), indicating that the retrieval error is mostly determined by the instrumental noise rather than the number of spectral points used in the analysis (directly related to the spectral sampling). As can be seen in Figure5.1 the 106ca represented by the green line is lower than the 099ca noise (red line) from 100 to 400 cm^{-1}, while it is higher above 500 cm^{-1}. From the sensitivity tests shown in Section4.2 we can expect that a more accurate estimate of the water vapour in the UTLS should be obtained using the 106ca noise. This is confirmed by the behaviour of the 106ca noise (red line in the left top panel of Figure5.2), that shows that the 106ca noise allows to have a better estimate of the water vapour in the UTLS with respect to the 099ca noise. On the other hand, the configuration using the real ISRF and the GOAL noise (blue lines in Figure5.2) provides the lowest error above 20 km, indicating that most part of the information in this altitude region is coming from the water lines in the 200-300 cm^{-1}, where the goal noise is lower than both 099ca and 106ca noise models. For temperature we see that both 099ca and 106ca noises (orange and red lines in the right top panel of Figure5.3) produce a lower noise error than the goal one. This is due to the fact that both noises are lower than the goal noise in the CO_2 band between 500 and 750 cm^{-1}, region where most part of the information on the atmospheric temperature is coming from. For the retrieved emissivity, the spectral behaviour of the error ratios is fully correlated to the relative differences among the noise behaviours shown in Figure5.1. In fact, the emissivity is retrieved with lower uncertainty with the 106ca noise for wavenumber lower than 400 cm^{-1} and with the 099ca noise from 400 to 1600 cm^{-1}, as can be observed by the comparison between the orange and the red lines in Figure5.4. However, the observed difference between the two noises

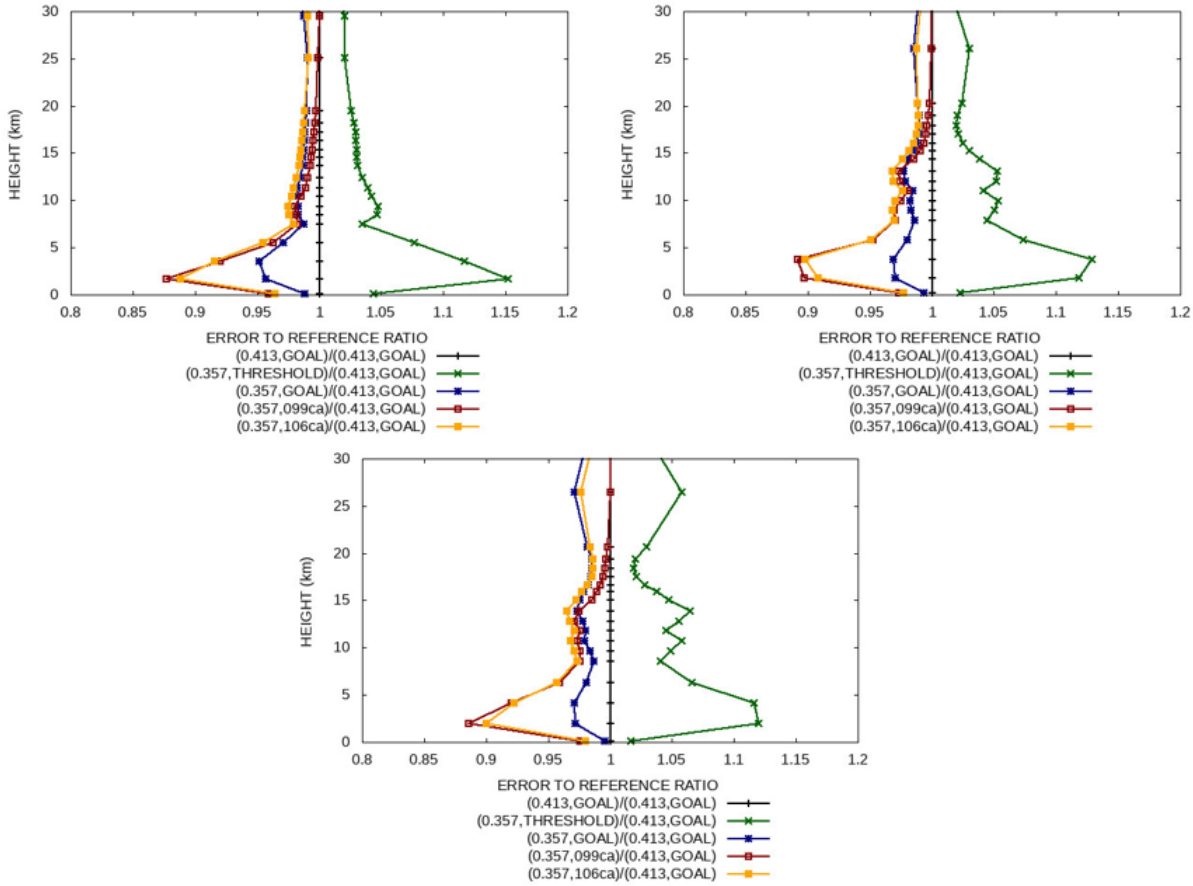


Figure5.2: Ratio between the random errors on water vapour obtained for the different configurations: the reference [0.413, GOAL] (black lines), the [0.357,THRESHOLD] (green lines), the [0.357, GOAL] (blue lines), the [0.357, 099ca] (red lines) and the [0.357, 106ca] (orange lines), shown for Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel)

in the FIR is very small, suggesting that the obtained uncertainty is highly influenced by the joint retrieval of the surface temperature and atmospheric temperature that extracts information from the same spectral regions where the sensitivity to the emissivity is larger. The highest errors on the retrieved variables are obtained - as expected - with the threshold noise (green lines in Figure5.2, 5.3 and 5.4), however the errors are slightly increased (10% in the troposphere and 5% in the UTLS) for water vapour and temperature, while the FIR surface emissivity estimate is more degraded with random errors increased by 20% with respect to the reference between 500 and 600 cm^{-1}.

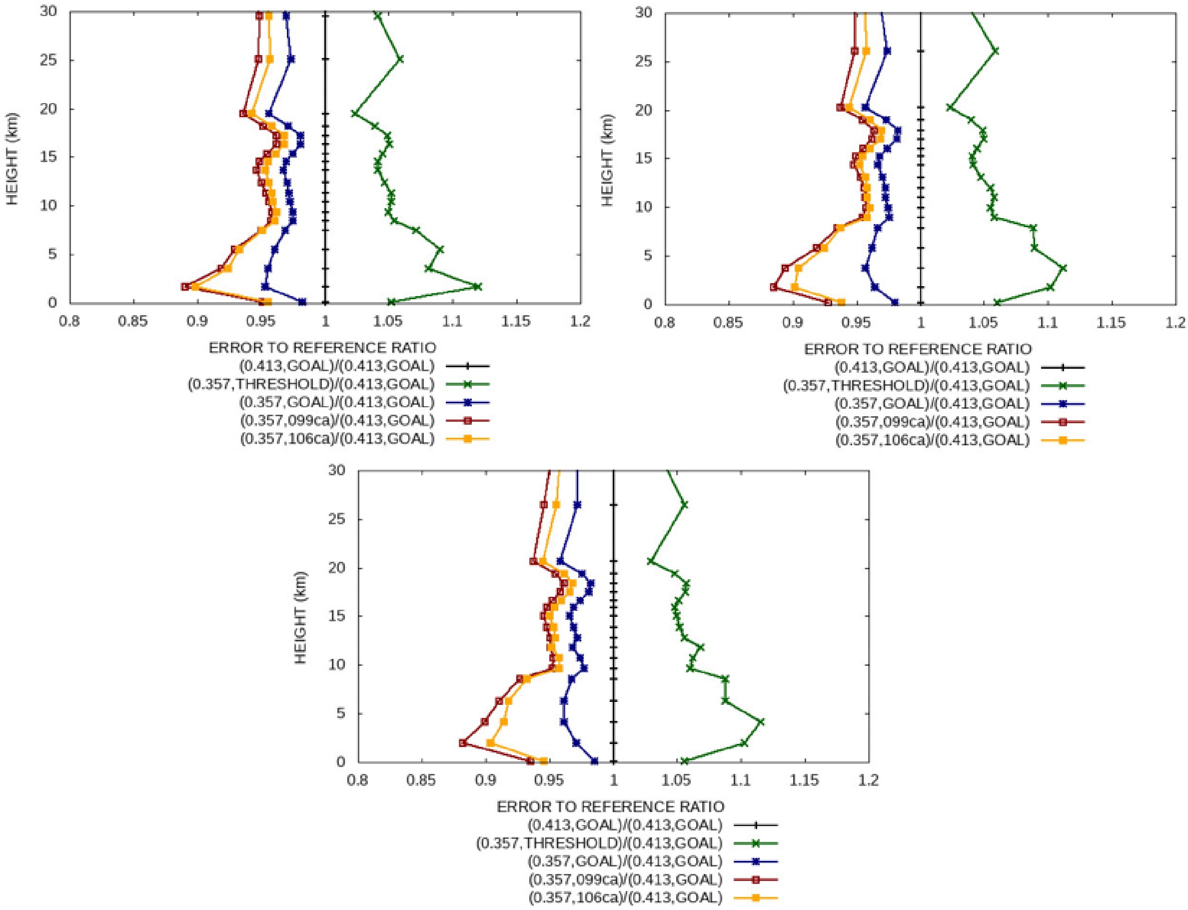


Figure5.3: Ratio between random errors on temperature achieved with different configurations: the reference [0.413, GOAL] (black lines), the [0.357,THRESHOLD] (green lines), the [0.357, GOAL] (blue lines), the [0.357, 099ca] (red lines) and the [0.357, 106ca] (orange lines), shown for Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel)

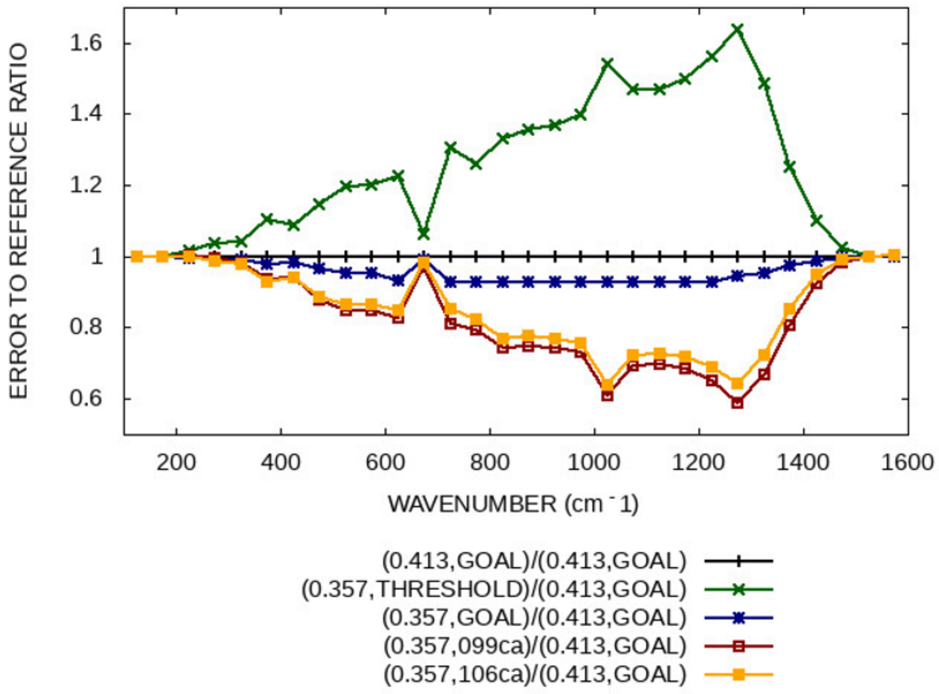


Figure5.4: Ratio between random errors on surface emissivity in the Polar band achieved with the different configurations: the reference [0.413, GOAL] (black lines), the [0.357,THRESHOLD] (green lines), the [0.357, GOAL] (blue lines), the [0.357, 099ca] (red lines) and the [0.357, 106ca] (orange lines)

5.2 Method for the systematic error assessment

In this Section the method used for the calculation of the systematic errors coming from the different error sources is discussed. The systematic errors affecting the FORUM Level 2 products are studied using the method developed by Dudhia et al. (2002) for the calculation of the systematic errors affecting the MIPAS-ENVISAT Level 2 products.

The method for the errors assessment is based on the Optimal Estimation Method exploiting the Gauss-Newton numerical algorithm (see Appendix A), that is used for the estimate of the defined state vector. In presence of a systematic error, the simulated radiances will differ from the radiances computed neglecting it by a contribution due to the considered systematic error and the new residuals $n^{'}$ will be affected by the same systematic contribution. The analysis of the observations including the systematic error effect will produce a new solution $x^{'}_{i+1}$ that can be expressed as:

$$x^{'}_{i+1} = x^{'}_{i} + (((K^T S_y^{-1} K) + S_a^{-1})^{-1} [K^T S_y^{-1} n^{'} + S_a^{-1} (x_0 - x_i)]) \tag{5.1}$$

The mapping of the systematic error source onto the state vector is then equal to the difference Δx=($x^{'}_{i+1}$-x_{i+1}). Thus, defining **G** as the gain matrix of the retrieval

$$G = ((K^T S_y^{-1} K) + S_a^{-1})^{-1} K^T S_y^{-1} \tag{5.2}$$

we have that

$$\Delta x = G \cdot (n' - n) = G \cdot \delta n \tag{5.3}$$

where the quantity δn is defined as the error spectrum, corresponding to the difference between the simulated spectra calculated in unperturbed conditions and with the single parameter perturbed with its systematic uncertainty. Thus, the systematic errors on the retrieved parameters can be estimated as the expectation value $< \cdot >$ of the errors affecting the respective elements in the retrieval state vector

$$< \delta x >_j = (\sqrt{G \cdot < \delta n \cdot \delta n^T > \cdot G^T})_{j,j} \tag{5.4}$$

where the expectation value of the product between the error spectrum vector and its transpose defines the variance-covariance matrix of the error spectrum S_n.

$$S_n = < \delta n \cdot \delta n^T > \tag{5.5}$$

The S_n matrix is determined by the spectral correlation of the systematic source, namely S(i,j)=$c_{i,j} \delta n_i \delta n_j$ with $c_{i,j}$ correlation coefficient equal to 1 if the systematic source introduces spectrally correlated error on the radiance and 0 if each spectral channel is perturbed independently from the others. The estimated systematic errors on the Level 2 products will be compared to the random errors expressed by the square root of the diagonal elements of the Variance Covariance Matrix S_x (see Appendix A), namely

$$x_E(j) = \left(\sqrt{[(K^T S_y^{-1} K) + S_a^{-1}]^{-1}}\right)_{j,j} \tag{5.6}$$

In the performed tests the retrieval state vector consists of the vertical distributions of water vapour, ozone and the atmospheric temperature, plus the skin temperature and the spectral surface emissivity. The vertical grid assumed for the atmospheric variables is the same adopted for the simulated retrievals reported in Chapter 4, namely

- 2 km step from ground level to 8 km altitude
- 1 km step from 8 km up to 20 km altitude
- 5 km step from 20 km up to the top of the atmosphere (fixed to 80 km)

and the surface emissivity is retrieved at 30 evenly spaced spectral points covering the whole FORUM spectral range from 100 to 1600 cm^{-1}. Furthermore, the relative frequency shift (RFS), corresponding to the relative shift of each spectral channel from its nominal frequency, is included in the state vector. The IWP (Ice Water Path), CTH (Cloud Top Height) and CBH (Cloud Bottom Height) will be included in the state vector in Section5.7 to evaluate the systematic errors affecting the cloud parameters as well.

The VCM of the observations $\mathbf{S}_y$ is computed using the Goal NESR, taking into account the apodization processes that correlate the errors affecting nearby spectral points. The a-priori errors on the state vector parameters are equal to those reported in Table4.1 in Chapter 4, with assumed correlation length on atmospheric H_2O, O_3 and temperature equal to 5 km. The a-priori error of the constant frequency shift is fixed to 2 ppm.

In order to achieve a complete picture of the expected systematic errors of the FORUM measurements, the systematic errors have been computed in the 24 IG2 and 80 DPD clear sky scenarios, then they have been averaged using the latitudinal bands defined in Section4.3. The Jacobi matrices containing the derivatives of the spectrum with respect to each defined parameter have been computed using the KLIMA Forward and Inverse Model, whereas the error spectra and Jacobians have been simulated with the GBB-NADIR+Internal RTE Solver at very fine spectral resolution (0.0005 cm^{-1}), then all of them have been convolved with the ISRF function of FORUM.

5.3 Systematic error sources

In this Section, the different error sources used in this work are listed, and the methods used for the calculation of the related error spectra are fully explained. The labels listed in brackets **[]** are those used in the following Sections to identify the various error sources in both figures and tables. All the systematic error sources listed below have been assumed to be fully spectrally correlated.

- **VMR of the interfering species** $\Rightarrow$ The systematic component is set to the IG2 1σ climatological variability that was estimated for the atmospheric scenarios of mid-latitude Spring, equatorial, polar Winter and polar Summer. Only the gases that show non-negligible spectral signatures in the FORUM spectral range have been used: CO_2, N_2O, CH_4, HNO_3. Each error spectrum is calculated as the difference between the spectra simulated with the unperturbed and the perturbed concentration of the considered interfering species **[CO2, N2O, CH4, HNO3]**
- **Water Vapour Continuum Model** $\Rightarrow$ The systematic errors are evaluated considering the uncertainties on the continuum parametrization reported by Mlawer et al. (2019) that are 30% in the frequency range 80-210 cm^{-1}, 20% in the frequency range 210-400 cm^{-1} and 7% in the frequency range 400-600 cm^{-1}. The differences between calculated spectra with unperturbed and perturbed continuum represent the used the error spectra **[CONT-PERC]**
- **Frequency grid** $\Rightarrow$ The error spectrum is the difference between the spectrum computed before convolution with a very fine spectral resolution 0.0005 cm^{-1} (that enables to correctly model all the line centres) and the spectrum simulated with the frequency grid used in the retrievals with the aim to reduce the CPU time (0.0025 cm^{-1}) **[HR GRID]**
- **Uncertainties of the spectroscopic parameters** $\Rightarrow$ The error spectra are calculated as the difference between spectra in unperturbed condition and applying a random perturbation to the line strengths and pressure broadening coefficients according to the error values reported in the HITRAN2016 database. The contributions of the spectroscopic uncertainties are evaluated separately for:
  - Errors due to line strengths and pressure broadening coefficients of all the transition lines **[errSPDB]**
  - Errors due to line strengths and pressure broadening coefficients of H_2O **[errSPDB_H2O]**
  - Errors due to line strengths of H_2O **[errSPDB_H2O_LS]**
  - Errors due to pressure broadening coefficients of H_2O **[errSPDB_H2O_HWHM]**

- **Absolute radiometric Accuracy (ARA)** ⇒ the error spectra have been computed using the 1σ goal threshold, corresponding to one third of the ARA requirement specified in Table2.1 **[ARA 1S]**. This error source is used as a reference in the error assessment, since it represents the minimum error that can be obtained for the retrieved products. In fact, the systematic effects due to the ARA cannot be reduced by performing the average of many spectra, differently from the error due to the measurement noise.

In addition to the above list of systematic error sources, the systematic effects due to the imperfect knowledge of the FORUM ISRF have been examined in the error assessment. The characterization of the ISRF and the considered ISRF errors are reported in next section. The results of the evaluation of the systematic errors related to the uncertainty on the ISRF knowledge are shown in Section5.4, whereas the errors due to the other error sources are presented separately in Section5.5.

A further systematic error source affecting the retrieved products is represented by the contribution on the measured radiance due to undetected clouds occupying the Field Of View of the FORUM FSI. The related systematic errors on water vapour, temperature and emissivity are shown in Section5.6

5.3.1 Characterization of a real ISRF

The ISRF could deviate from the ideal *sinc* function because in a real instrument it is unlikely that the interferogram truncation can be represented by a boxcar function (see Chapter 2). The ideal ISRF has its maximum at the central frequency of the relative spectral channel and is symmetric around that point. In reality, often the ISRF peak is centered at a different frequency. The difference between the expected and real frequency of the peak of the ISRF function is called frequency shift. In FT spectrometers often the frequency shift is proportional to the nominal frequency of the considered channel, therefore if evaluated as a relative number, obtained by dividing the spectral shift by the nominal frequency. We call this quantity Relative Frequency Shift (RFS) and we express its value in part per million (ppm). Moreover, the ISRF function of a real instrument can be asymmetric around its peak, due to the not ideal characteristics of the optical components of the instrument, like the effects due to a wedged beam-splitter. The effects due to the ISRF distortions have to be evaluated not only on the FORUM Level 2 data but also for the Level 1 ones (spectra and integrated fluxes), since they can affect the comparison of a FORUM measured spectrum with a spectrum measured by a different instrument (e.g. for an inter-calibration check). In fact, for the direct inter-comparison of two spectra (i.e. graphical comparison and/or channel-by-channel difference computation) the real frequency of each channel plus the relative ISRF of both instruments have to be known to perform a correct resampling of the two spectra on a common sampling grid. If the two spectra have different spectral resolutions and frequency grids, they can be resampled on a common grid by convolving the spectra with an apodization function broader than the broadest ISRF of the two measurements. On the contrary, if the ISRFs of the two spectra to be compared include a spectral shift and/or significant asymmetries, the spectra can still be compared only when the spectral shift and asymmetry are well characterized.

To assess the acceptable magnitude of the ISRF errors on FORUM Level 1 data, it has been required that the frequency shift and the ISRF shape should be known with a sufficiently small error, so that the difference between two simulated spectra, one at the nominal frequency grid and one using the perturbed ISRF should be equal or smaller than one tenth of the GOAL NESR (see Table2.1). Therefore, tests have been performed simulating FORUM measurements in a variety of atmospheric conditions applying to the simulated spectra an ISRF affected by errors of known quantities, and evaluating the difference in the radiances produced by this errors.

In particular, we have investigated and reported in Figure5.5 the error spectra simulated for a Tropical, a Mid-Latitude and a Polar scenario selected out of the 24 IG2 scenarios due to the relative frequency shift (top left panel), and the error due to the ISRF asymmetry (top right panel) that is analytically defined as ratio among the absolute difference between the first two minima and the main ISRF maximum. We have also reported the effect of the uncertainty on the ISRF width equal to 0.6% (bottom panel). A perturbation of 1 ppm on the RFS and 1.67% on the asymmetry have been considered.

The errors introduced on the radiance by the different components of the uncertainty on the ISRF are larger in the Tropical scenario, as shown by the relative error spectra in Figure5.5 reported in darker colours with respect to the Mid Latitude and Polar scenarios. However, a RFS equal to 1 ppm is acceptable as introduces a systematic error on the spectra that is almost totally below the NESR/10. On the contrary, a not-modelled ISRF asymmetry and the uncertainty on the ISRF width introduce differences of the order of the GOAL NESR. On the other hand, we are confident that possible asymmetry features of the FORUM ISRF could be modelled with an accuracy of 10%, producing an acceptable systematic contribution on the spectra within the (GOAL NESR)/10 for most part of the FORUM spectral range. On the basis of the performed analysis, the requirements on the ISRF asymmetry and the relative frequency shift can be summarized as in the following Table5.1

Tabel5.1: Summary of ARA and NESR requirements for the FORUM spectrum

	GOAL requirements	**THRESHOLD requirements**
Frequency shift (FSH)	< 1ppm	< 2ppm
ISRF asymmetry	< 1.67%	< 3.34%
ISRF asymmetry modelling	< 0.167%	< 0.334%

A model of a not-ideal FORUM ISRF being affected by both frequency shift and asymmetry - labelled as "Consortium A ISRF" (CA ISRF) - has been supplied by the industry to ESA in April 2020. The CA ISRF is a wavenumber dependent complex function, hence composed

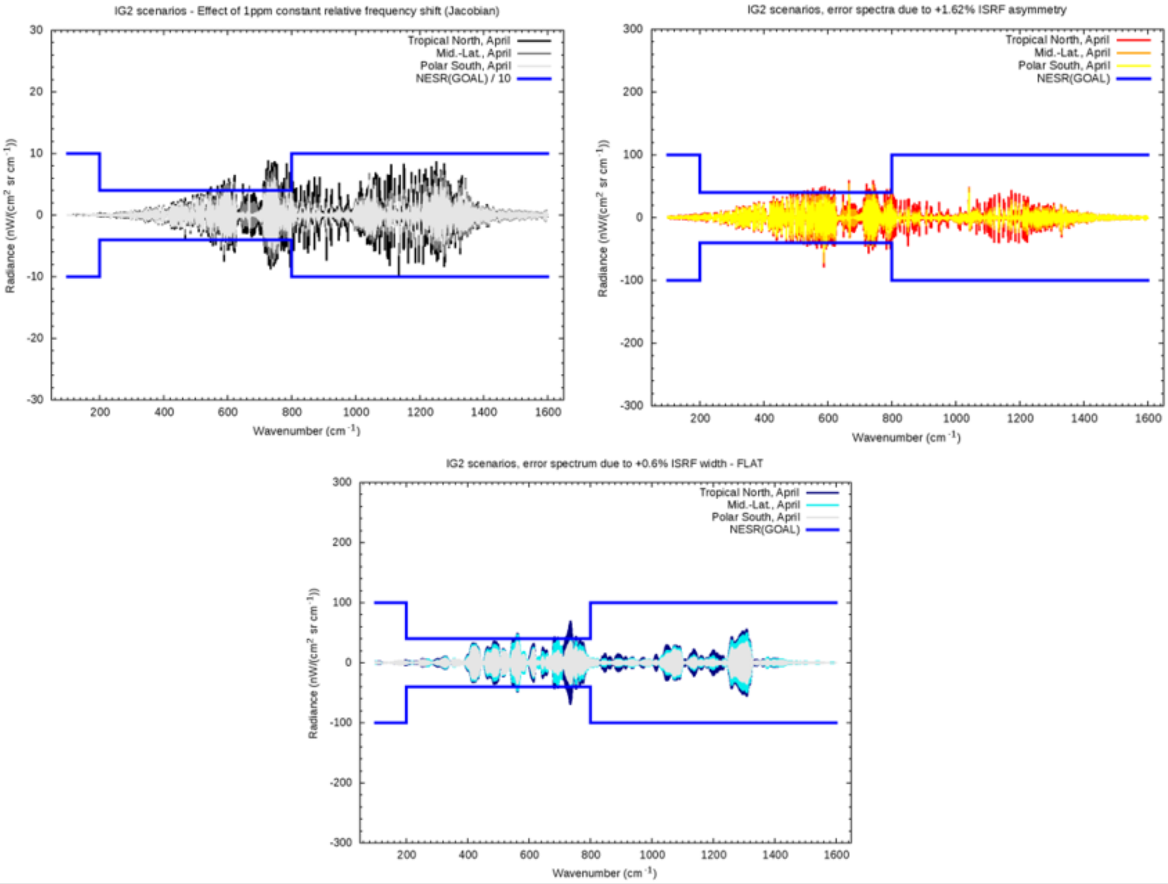


Figure5.5: Error spectra due to a constant RFS equal to 1ppm (top left panel), to a missing modelling of the ISRF asymmetry equal to 1.67% (top right panel) and to a constant 0.6% error in the ISRF width model (bottom panel), for Tropical, Mid-Latitude and Polar IG2 scenarios. Note that the error spectra referred to the asymmetry and the ISRF width are compared to the GOAL NESR, while the error spectra of the RFS are compared to one tenth of the GOAL NESR

by Real and Imaginary parts, tabulated every 10 cm^{-1} from 100 to 1600 cm^{-1}. After the calibration, it is anticipated that the FORUM ISRF function will have just a Real part, so only that has been examined and it is shown in the Figure5.6 for frequencies equal to 100 (orange), 850 (red) and 1600 (blue) cm^{-1}.

It is evident from Figure5.6 that the CA ISRF is not a symmetric function, showing also frequency dependent frequency shift, Full Width at Half Maximum (FWHM) and asymmetry. These quantities plus the distance between the main maximum and the first adjacent zero of the ISRF, that is related to the sampling of the Fourier Transform of the measured interferogram, are plotted in the Figure 5.7 for each of the provided ISRFs. The asymmetry (blue solid line in Figure5.7) meets the requirement specified in Table 5.1 as it ranges from zero to a maximum value of 1.6% at 1600 cm^{-1}, whereas the relative frequency shift (black line) is larger than the 2 ppm threshold and not constant with respect to the wavenumbers. However, one could be confident that the characteristics of the final FORUM ISRF will be less fluctuating and drifting

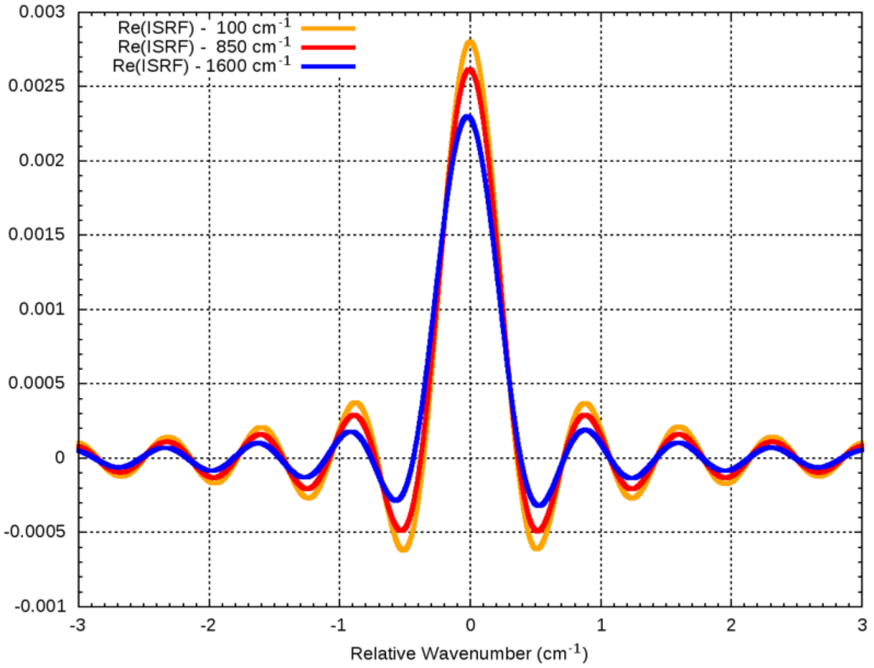


Figure5.6: Examples of the Real part of the ISRF supplied by ESA (Consortium A), for three different wavenumbers: 100 (orange), 850 (red), and 1600 cm^{-1} (blue)

than the CA ISRF reported here. The minimum value of the distance between the main ISRF maximum and the first adjacent zero (highlighted with the orange line in Figure5.7) is smaller than the 0.413 cm^{-1} assumed for the sinc ISRF, it is equal to 0.357 cm^{-1} for the CA ISRF, corresponding to a MOPD equal to 1.4 cm.

In order to better highlight the asymmetric feature affecting the CA ISRF, one can compare the amplitude of the non-symmetric part of the function to the symmetric one. In fact, every function F(x) can be represented by the sum between an even and an odd function, that are highlighted with the red and the orange lines respectively in Figure5.8 with the CA ISRF at 1600 cm^{-1} (blue line). As expected, the anti-symmetric component is very small as compared to the symmetric part and to the ISRF as well.

5.3.2 Error sources due to imperfect ISRF knowledge

The uncertainty on the knowledge of the RFS, the FWHM and asymmetry of the ISRF are considered as systematic error sources in the analysis of the errors affecting the FORUM Level 2 products. The systematic errors due to the imperfect knowledge of the FORUM ISRF are calculated for two different ISRF shapes: one using the ideal *sinc* function introduced in Chapter 2 and one using the CA ISRF shapes. For each single error source, different perturbations for the two considered cases are defined as following:

- **Uncertainty on the ISRF asymmetry modelling** ⇒ It is assumed that the ISRF asym-

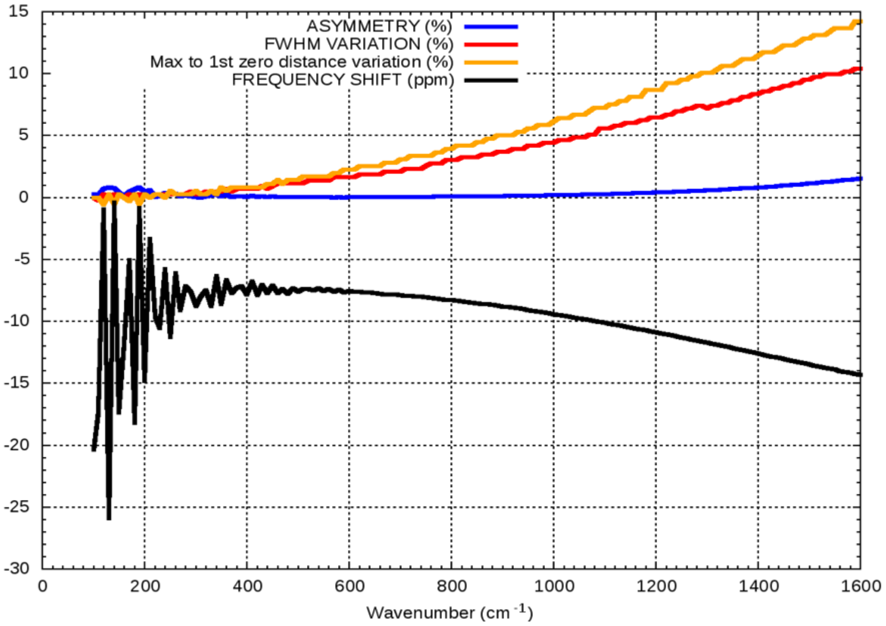


Figure5.7: Behaviour of the Consortium A ISRF asymmetry (%, blue line), FWHM (cm^{-1}, red line), spectral sampling (cm^{-1}, orange line) and RFS (ppm, black line) as a function of the wavenumber

metry can be modelled with the 10 % accuracy. For the *sinc* ISRF, the error spectrum is calculated as the difference between the spectrum convolved with the unperturbed *sinc* and the *sinc* convolved with the ISRF perturbed with the derivative of the Gaussian function providing an asymmetry equal to GOAL value in Table5.1 divided by 10 **[ASYMG]**. For the CA ISRF, the error spectrum is calculated as the difference between the spectrum convolved with the even part of the CA ISRF and the one convolved with the same perturbed with the 10% of the odd part of the function as represented in Figure5.8 **[ASY]**

- **Uncertainty on the ISRF width** ⇒ For the *sinc* ISRF, the error spectrum is calculated as the difference between the spectrum convolved with the unperturbed *sinc* and the one convolved with the *sinc* whose width is perturbed by 0.6% **[DWFLT]**. For the CA ISRF, the unperturbed or reference ISRF at each wavenumber is assumed to be the supplied ISRF at the nearest wavenumber. The perturbed ISRF is obtained by convolving the reference ISRF by a *sinc* function with a FWHM such that the resulting ISRF has a FWHM exceeding that of the reference ISRF by 0.6% **[DWD]**

- **Uncertainty due to "residual" or "uncorrected" frequency shift of the ISRF, changing randomly from spectral channel to spectral channel** ⇒ This is an error component equally defined for the *sinc* and the CA ISRFs. The error spectrum in this case is obtained by applying to the ISRF a relative frequency shift that, at each spectral channel, is taken

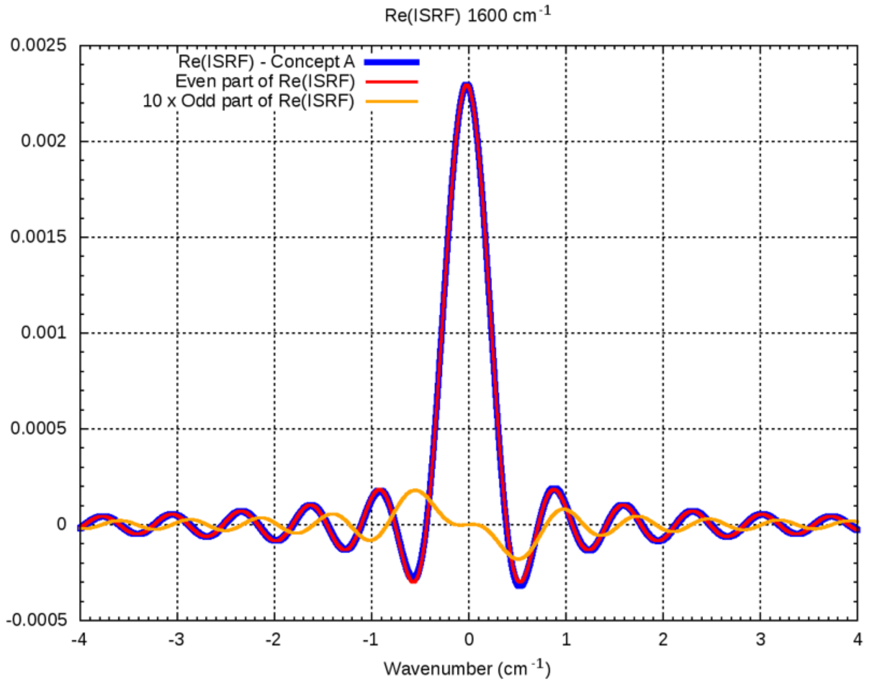


Figure5.8: Consortium A ISRF at 1600 cm^{-1} (blue), its symmetric (red) and anti-symmetric (orange) parts. For plot's clarity, the anti-symmetric part is multiplied by 10

from a pseudo random Gaussian distribution with zero mean and standard deviation equal to 1 ppm **[RFE]**

- **Specifically for the *sinc* ISRF we evaluated the effect of a not retrieved RFS equal to 2 ppm** ⇒ The error spectra have been evaluated as the Jacobian with respect to the RFS multiplied by 2 **[2ppm-RFS]**

The error sources related to the uncertainty on the ISRF width and asymmetry are considered as spectrally correlated, on the contrary no correlation is introduced in the error spectrum VCM for the RFE source. In fact, the residual unretrieved frequency shift affecting each single spectral channel should be independent from the RFS affecting the other FSI channels.

5.4 Errors due to imperfect ISRF knowledge

First the latitudinal average errors due to the different ISRF uncertainty components for the ideal ISRF *sinc* have been assessed. The Figures 5.9 and 5.10 show the mean systematic errors on water vapour and temperature calculated in the Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel), while the Figure 5.11 shows the systematic errors on surface emissivity computed in the Polar band only. The systematic errors affecting the RFS are reported in Table5.2.

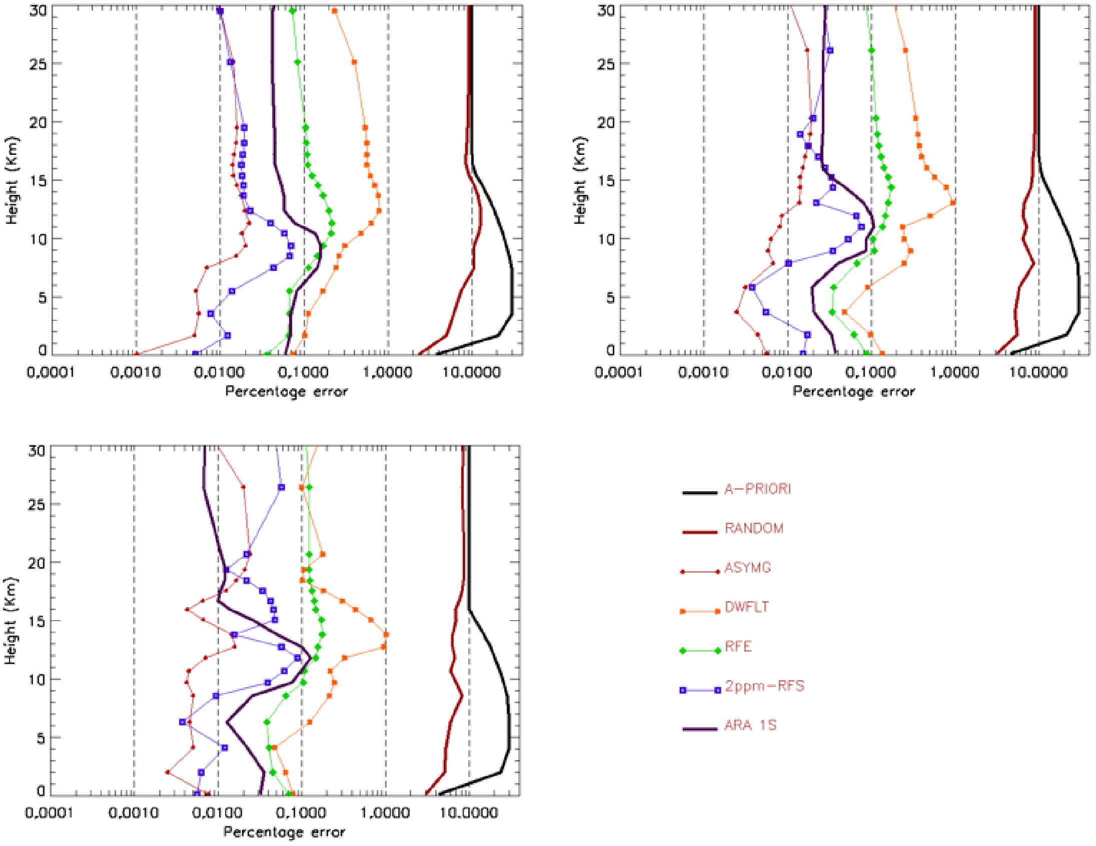


Figure5.9: Mean systematic errors on the retrieved water vapour profile due to the ISRF uncertainty components for the Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel). The considered systematic errors are the ones originating from the ASYMG (red lines), DWFLT (orange line) and RFE (green line) error sources, compared to the assumed a-priori (thick black lines), random (thick red lines) and ARA (thick purple lines) errors

The uncertainty on the ISRF width (orange lines) is the component introducing the largest errors on the retrieved parameters, with uncertainties up to 1% in water vapour and 0.1 K for temperature in the UTLS, whereas the contribution introduced by a 10% error in the ISRF asymmetry model (red lines) is negligible. The random frequency shift (RFE) is responsible for small systematic errors represented by the green lines on both temperature and water vapour,

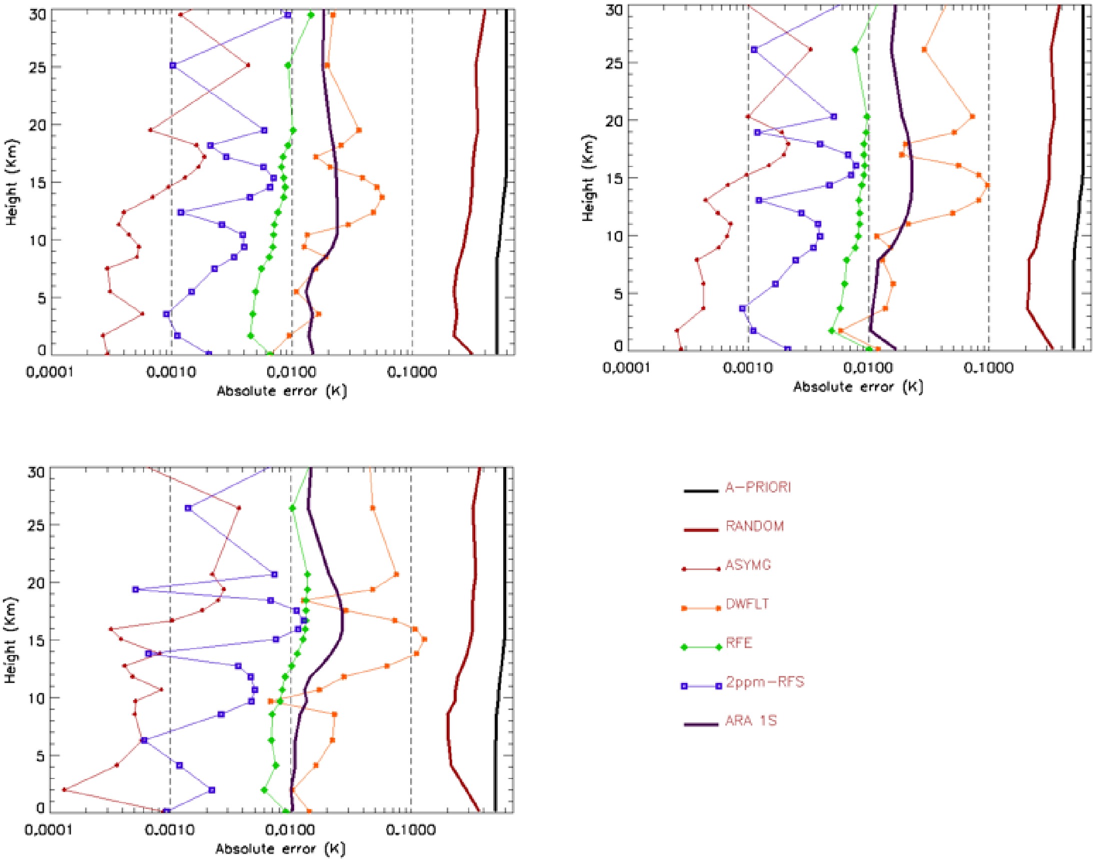


Figure5.10: Mean systematic errors on the retrieved temperature profile due to the ISRF uncertainty components for the Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel). The considered systematic errors are the ones originating from the ASYMG (red lines), DWFLT (orange line) and RFE (green line) error sources, compared to the assumed a-priori (thick black lines), random (thick red lines) and ARA (thick purple lines) errors

even though its contribution is about one order of magnitude larger than the ARA contribution (thick purple line) in the UTLS. On the contrary a flat RSE error of 2 ppm has negligible effects on the retrieved level 2 data, almost always below the ARA. All the ISRF error components generate small errors on the retrieved emissivity, since the information on the surface properties mainly comes from the sampled spectral points far from the absorption lines.

All the ISRF error components have a relevant effect on the retrieved RFS. In particular, a residual random RFS of 1 ppm rms is responsible for uncertainties comparable to the noise error. However, this will be a problem affecting the retrieval from a single FORUM spectrum, since the error due to the random component of the RFS should become negligible when many spectra are averaged. On the other hand, a not retrived RFS equal to 2 ppm introduces lower systematic errors with respect to those provided by the random residual RFS, even though the perturbation applied on the RFS is higher for the “2ppm-RFS” error source. This is due to the full spectral correlation assumed for the Forward Model error resulting from the unretrieved

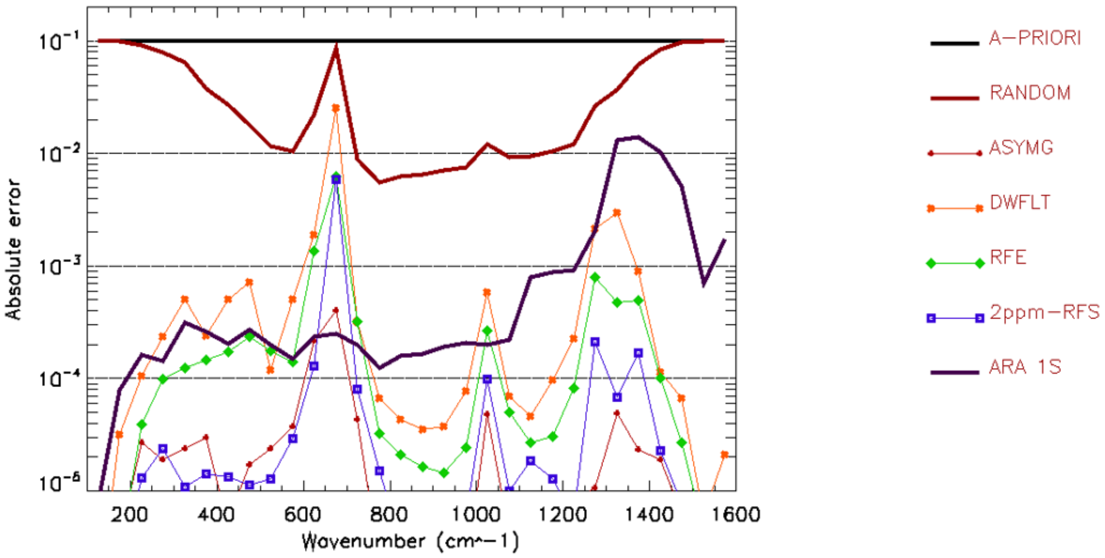


Figure5.11: Mean systematic errors on the retrieved surface emissivity profile due to the ISRF uncertainty components for the Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel). The considered systematic errors are the ones originating from the ASYMG (red lines), DWFLT (orange line) and RFE (green line) error sources, compared to the assumed a-priori (thick black lines), random (thick red lines) and ARA (thick purple lines) errors

RFS, tending to reduce the systematic effects on the estimated variables with respect to similar not correlated error source. Even though we have assumed a value equal to the 10% of the goal asymmetry requirement (see error component ASYMG defined in Section 5.3), the relative systematic contribution on the retrieved RFS is large. This is due to the arbitrary choice of introducing in the ISRF an asymmetry distortion with the shape of the first derivative of a Gaussian function. In this way the error spectrum presents the same shape of the derivative of the spectrum with respect to the RFS, generating high correlations that lead to high systematic contributions. A more realistic asymmetry distortion is provided by the Consortium A ISRF, therefore the present tests have been repeated considering the FORUM instrumental response represented by the CA ISRF and the related error sources linked to its knowledge that has been introduced in Section 5.3. The Figures 5.12 and 5.13 show the mean systematic errors on water vapour and temperature calculated in the Polar band (top left panel), Mid-Latitude band (top right panel) and Tropical band (bottom panel), while the Figure 5.14 shows the systematic errors on surface emissivity computed in the Polar band only. The Table5.3 shows the systematic errors affecting the retrieved RFS due to the uncertainty on the CA ISRF asymmetry.

The systematic error produced by the uncertainty in the CA ISRF asymmetry model is negligible, being smaller than 0.01% for water vapour (red line in Figure5.12, of the order of 0.1 mK for the atmospheric temperature (red line in Figure5.13) and of the order of 0.01 ppm for the estimated RFS. This is due to the very small amplitude of the CA ISRF odd part and to the

Table5.2: Mean error components (ppm) on the retrieved RFS due to the ISRF uncertainties, for Polar, Mid-Latitude and Tropical bands

RFS (ppm)	**POLAR**	**MID-LAT**	**TROPICAL**
ASYMG	0.442	0.484	0.486
RFE	0.161	0.174	0.172
DWFLT	0.227	0.226	0.235
ARA 1S	0.001	0.001	0.001
RANDOM	0.915	0.599	0.460
A-PRIORI	2.000	2.000	2.000

Table5.3: Mean error components (ppm) on the retrieved RFS due to the uncertainty on the CA ISRF asymmetry, for Polar, Mid-Latitude and Tropical bands

RFS (ppm)	**POLAR**	**MID-LAT**	**TROPICAL**
ASY	0.012	0.015	0.014

fact that in this case the error spectrum and the RFS Jacobian have a completely different shape and are therefore mostly uncorrelated. The error due to the uncertainty in the ISRF width is slightly reduced for the water vapour in the UTLS, as we can observe by comparing the orange lines reported in top left panels of Figures5.12 and 5.9, whereas the RFE introduces a more relevant contribution for both water vapour and temperature due to the larger number of spectral sampling points. Namely, the RFE contribution is one order of magnitude larger than the ARA contribution for water vapour and almost equal to the ARA contribution for temperature, as we can observe by comparing the green lines (RFE errors) to the purple lines representing the ARA errors. No significant variations are observed for the systematic errors affecting the surface emissivity, as shown by the comparison between the Figures 5.11 and 5.14.

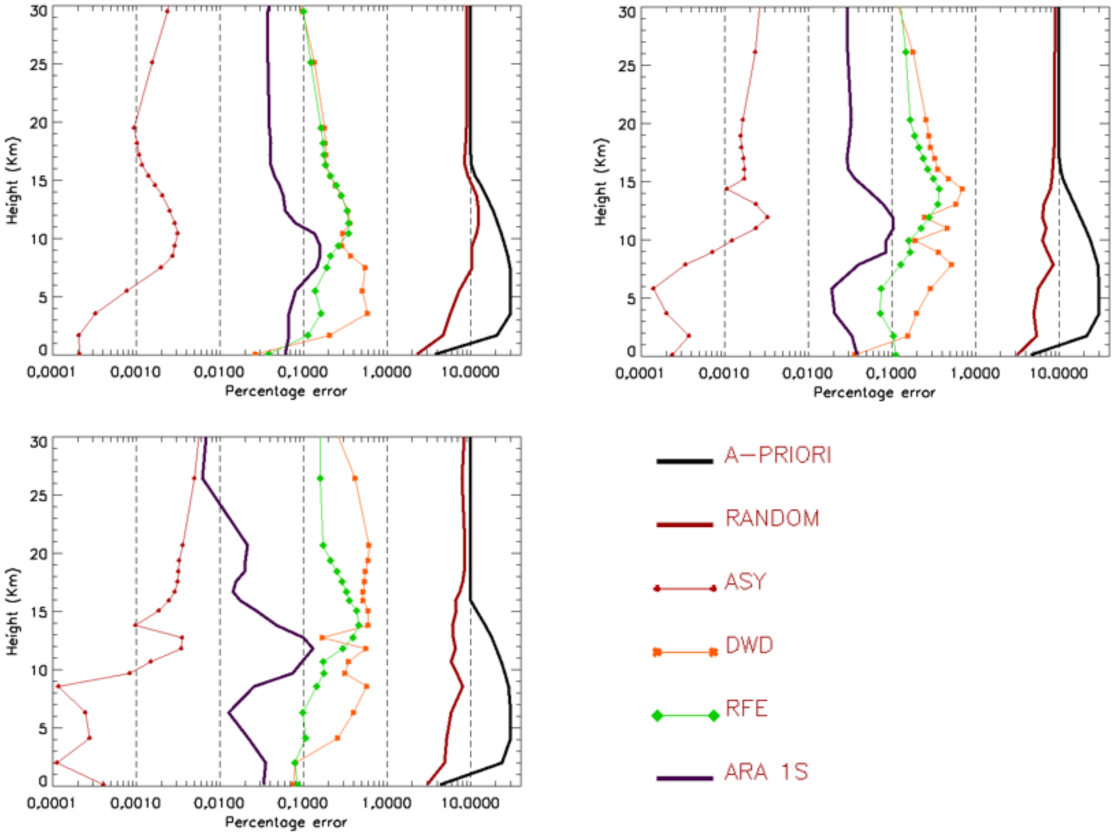


Figure5.12: Average ISRF error components on the retrieved water vapour, assuming Consortium A ISRF. The three panels refer to different latitude bands: Polar (top left), Mid-Latitude (top right) and Tropical (bottom panel). The considered systematic errors are the ones originating from the ASY (red lines), DWD (orange line) and RFE (green line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

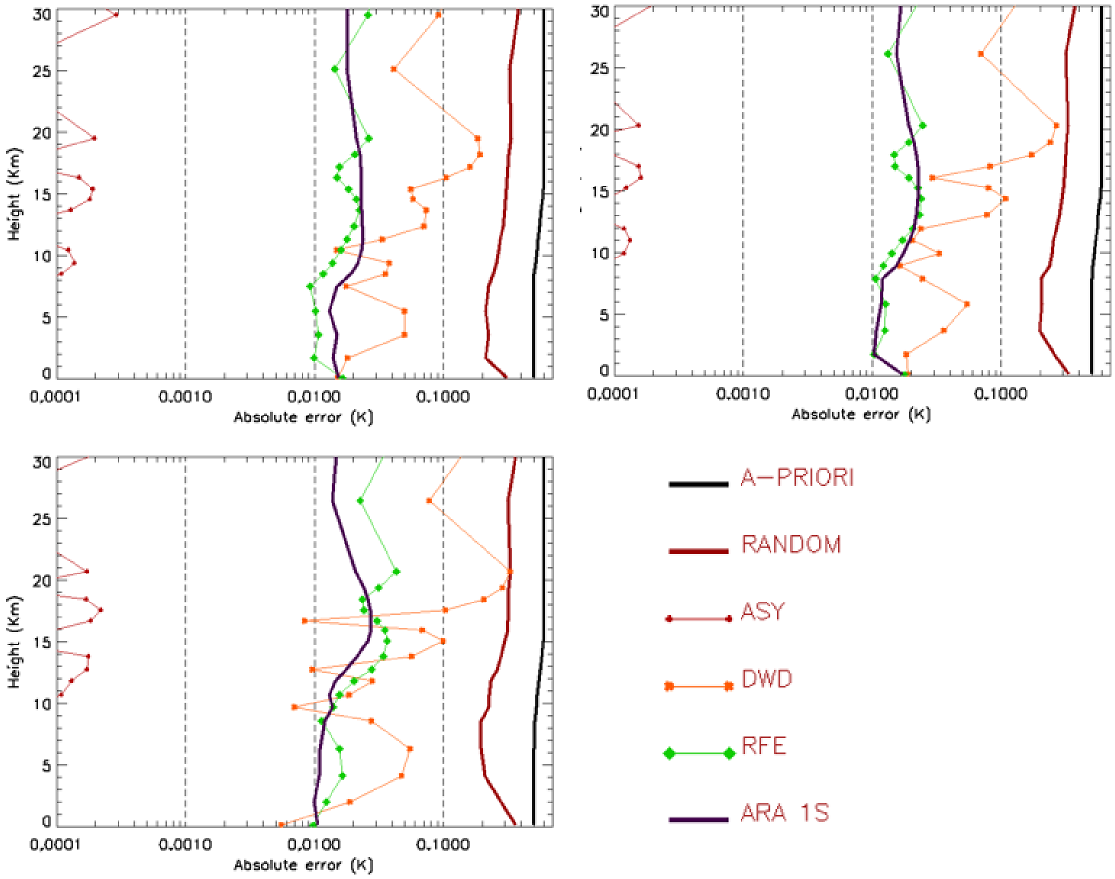


Figure5.13: Average ISRF error components on the retrieved temperature assuming Consortium A ISRF. The three panels refer to different latitude bands: Polar (top left), Mid-Latitude (top right) and Tropical (bottom panel). The considered systematic errors are the ones originating from the ASY (red lines), DWD (orange line) and RFE (green line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

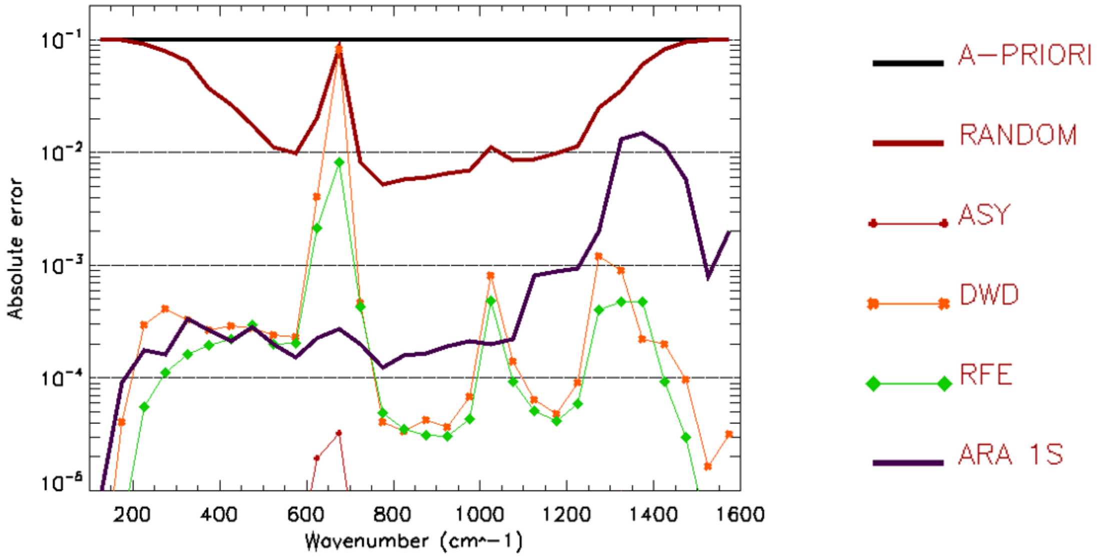


Figure5.14: Mean systematic errors on the surface emissivity due to the ISRF Consortium A uncertainty components for the Polar band. The considered systematic errors are the ones originating from the ASY (red lines), DWD (orange line) and RFE (green line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

5.5 Errors due to the other sources

In this section the other components of the Level 2 systematic errors are shown. For clarity, the errors on water vapour and temperature due to the uncertainty on the VMR of the interfering species (Figures 5.15 and 5.17) have been separated from the ones due to spectroscopy (Figures 5.16 and 5.18), with the errors in the Polar band shown in the top left panels, the ones in the Mid-Latitude band shown in the top right panels and the ones in the Tropical band shown in the bottom panel. The errors on the surface emissivity in the Polar band only are reported in Figure5.19. The instrumental response of the FORUM FSI has been simulated through the ideal *sinc* ISRF for the results shown in the present Section.

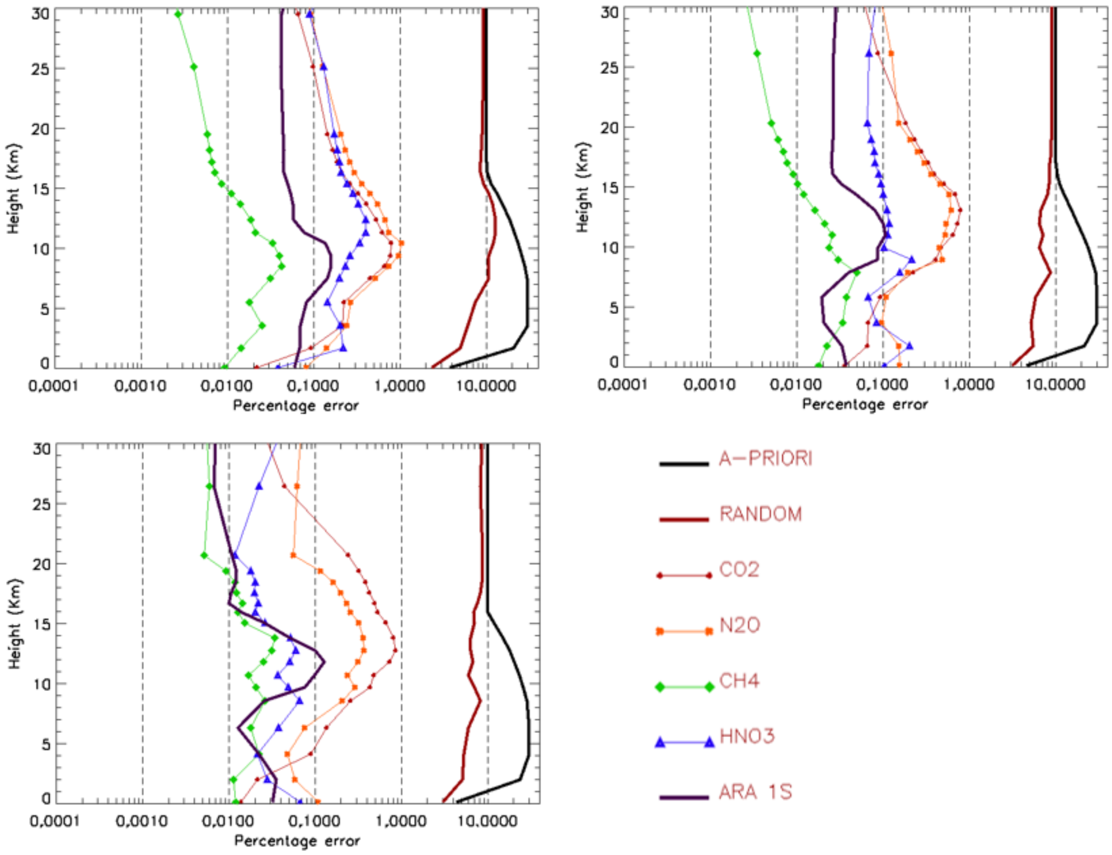


Figure5.15: Mean systematic errors on the water vapour profile due to the uncertainties in the VMR of the interfering species, for Polar (top left), Mid-Latitude (top right) and Tropical (bottom panel) bands. The considered systematic errors are the ones originating from the CO2 (red lines), N2O (orange line), CH4 (green line) and HNO3 (blue line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

First, it can be observed that the uncertainty in the continuum parametrization is the most relevant error source for tropospheric water vapour, temperature and FIR surface emissivity (purple thin lines in Figures from 5.15 to 5.19. Other spectroscopic uncertainties produce

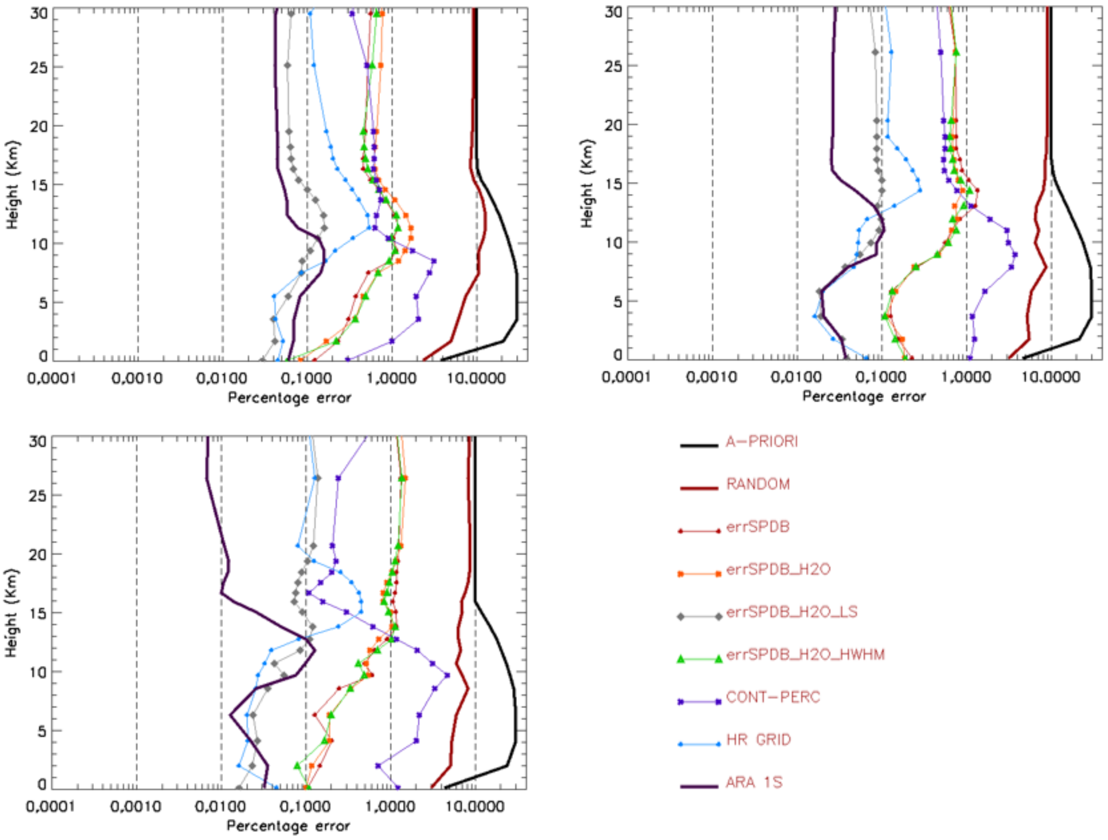


Figure5.16: Mean systematic errors on the water vapour profile due to the uncertainties in the spectroscopic data, for Polar (top left), Mid-Latitude (top right) and Tropical (bottom panel) bands. The considered systematic errors are the ones originating from the errSPDB (red lines), errSPDB_H2O (orange lines), errSPDB_H2O_LS (grey line), errSPDB_H2O_HWHM (green lines), CONT-PERC (thin purple lines) and HR GRID (sky blue) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

systematic error components of about 1% in the UTLS for water vapour, and up to 0.005 in the surface emissivity, with the main contribution due to the uncertainty affecting the pressure broadening coefficients of the water vapour absorption lines (see the green lines in Figure5.16 and bottom panel of Figure5.19). The CO_2 and N_2O variability can introduce errors ranging from 0.5% to 1% on the water vapour around the tropopause (see red and orange lines in Figure5.15, whereas the systematic contribution due to the methane (green lines in Figure5.15) variability is considerably lower. The HNO_3 variability produces generally low systematic contributions, however its variability in polar regions is responsible for the largest error component among the interfering species affecting the FIR surface emissivity with a value of 0.01 at 400 cm^{-1}, as shown by the blue line in the top panel of Figure5.19. This is due to the intense HNO_3 absorption band present in this region, and to the large HNO_3 variability at the Poles, due to the wintertime polar vortex activity. Table5.4 shows the error components affecting the retrieved RFS due to the same error sources.

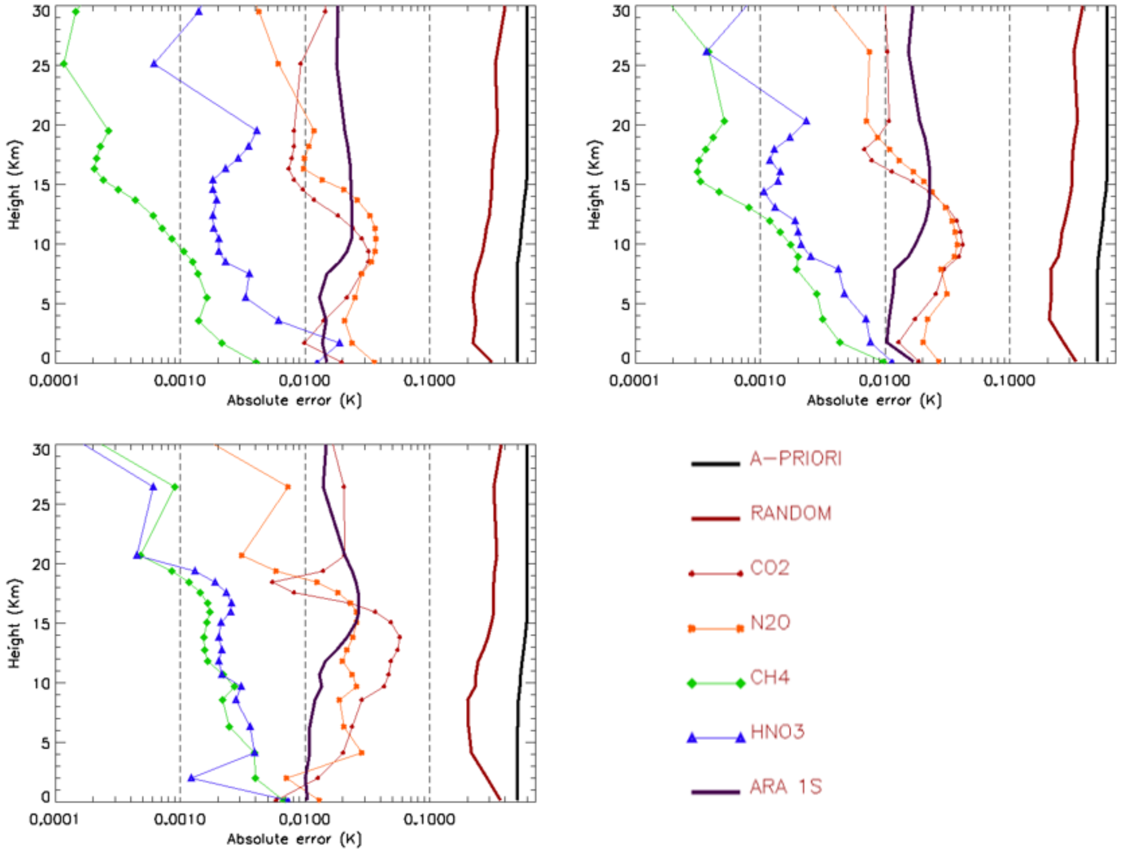


Figure5.17: Mean systematic errors on the temperature profile due to the uncertainties in the VMR of the interfering species, for Polar (top left), Mid-Latitude (top right) and Tropical (bottom panel) bands. The considered systematic errors are the ones originating from the CO2 (red lines), N2O (orange line), CH4 (green line) and HNO3 (blue line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

The variability of all the considered gases introduces low errors on the retrieved RFS, however the HNO_3 variability in the polar regions is responsible for systematic errors of about one order of magnitude larger than in the other latitudinal bands. Differently from the other analysed state vector parameters, the RFS is affected in a negligible way by the high uncertainty on the continuum parametrisation, since the information on the RFS comes from the lines centre, that are poorly affected by the continuum. On the other hand, the spectroscopic uncertainty introduces remarkable systematic errors, up to almost 0.1 ppm, with the main contribution ascribable to the absorption lines of the absorbing species showing relevant spectral signature in the FORUM radiance (H_2O, CO_2, N_2O, CH_4, O_3). However, the highest systematic contribution is introduced by the coarser spectral resolution used for the simulation of the high resolution spectra and Jacobians. This causes a degradation of the line shape around its centre, therefore the error spectra are generally highly correlated with the RFS Jacobians.

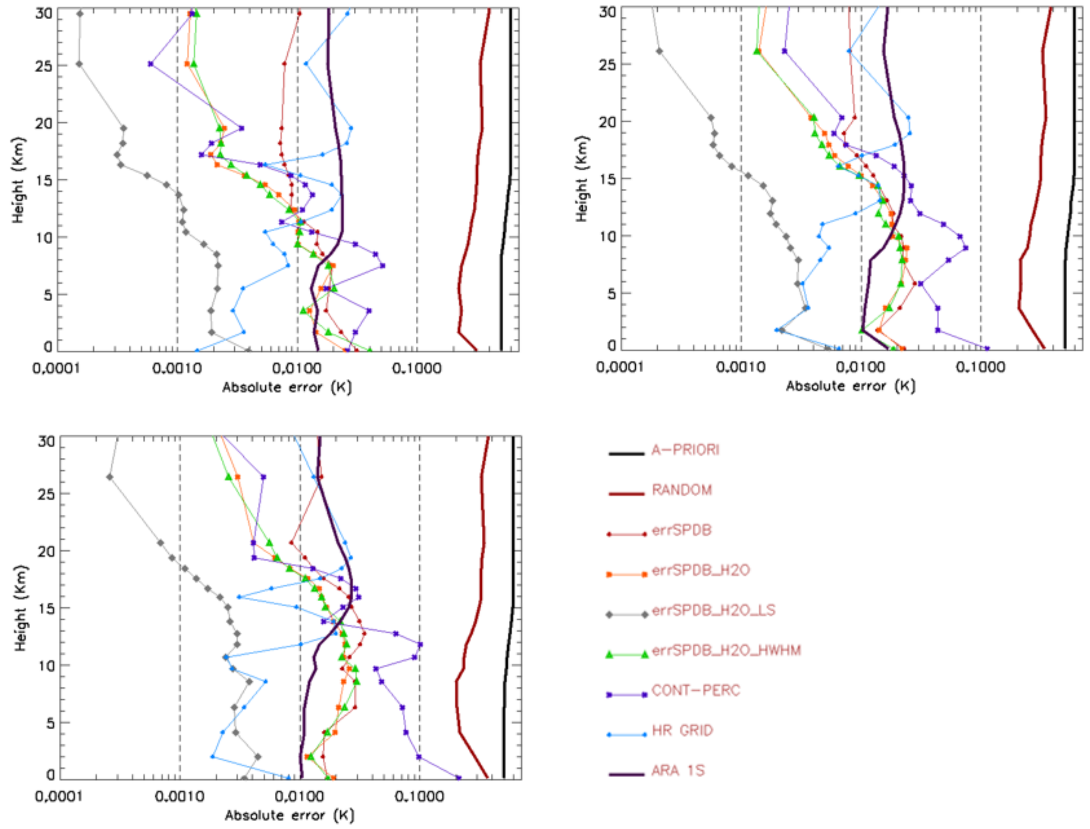


Figure5.18: Mean systematic errors on the temperature profile, due to the uncertainties in the spectroscopic data, for Polar (top left), Mid-Latitude (top right) and Tropical (bottom left) bands. The considered systematic errors are the ones originating from the errSPDB (red lines), errSPDB_H2O (orange lines), errSPDB_H2O_LS (grey line), errSPDB_H2O_HWHM (green lines), CONT-PERC (thin purple lines) and HR GRID (sky blue) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

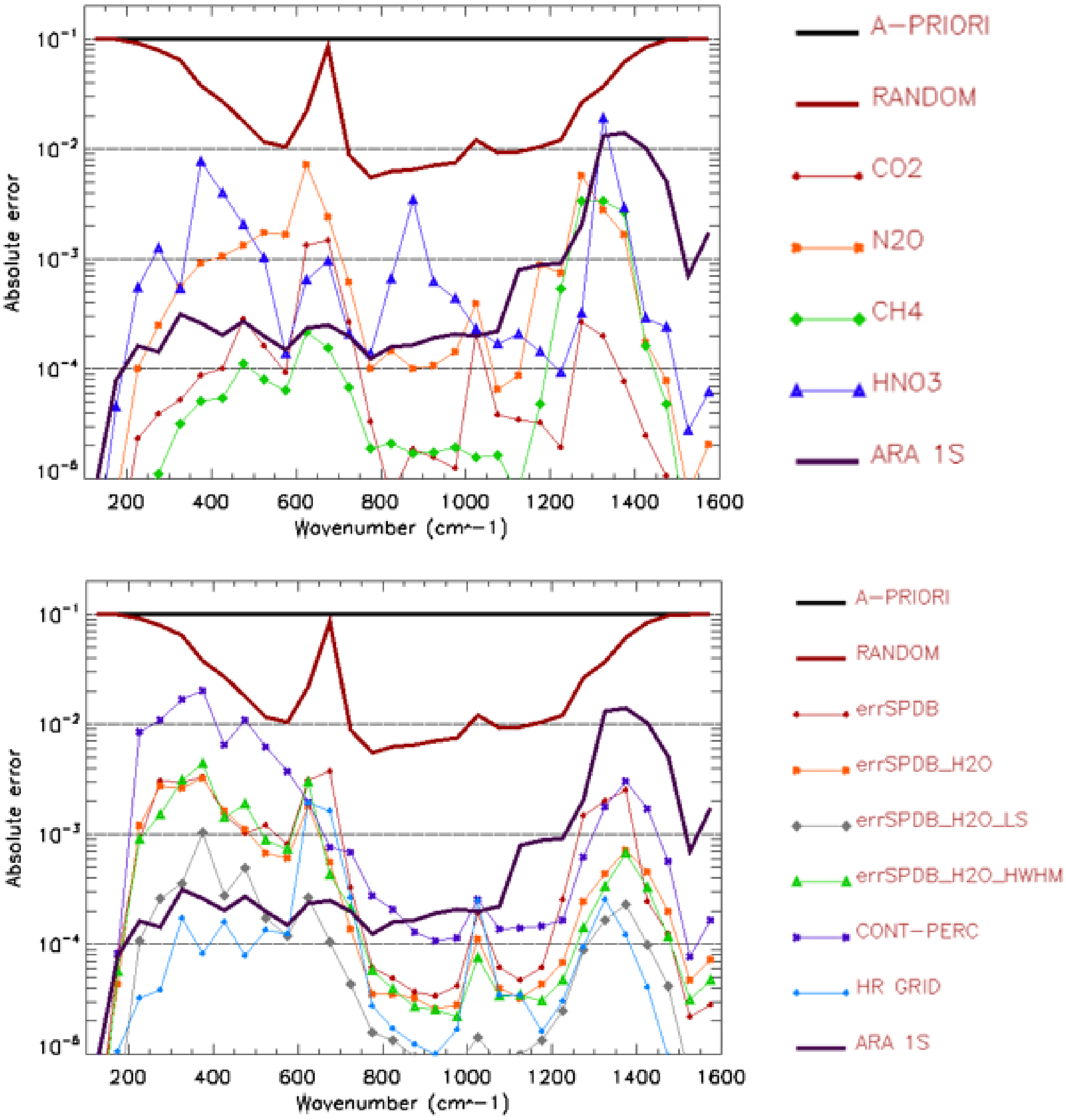


Figure5.19: Mean systematic errors for the surface emissivity in Polar regions, due to the uncertainties in the VMR of the interfering species and on the spectroscopic parameters. Top panel: The systematic errors shown are the ones originating from the CO2 (red lines), N2O (orange line), CH4 (green line) and HNO3 (blue line) error sources. Bottom panel: The systematic errors shown are the ones originating from the errSPDB (red lines), errSPDB_H2O (orange lines), errSPDB_H2O_LS (grey line), errSPDB_H2O_HWHM (green lines), CONT-PERC (thin purple lines) and HR GRID (sky blue) error sources. The systematic errors are compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

Table5.4: Mean error components (ppm) on the retrieved RFS due to the ISRF uncertainties, for Polar, Mid-Latitude and Tropical bands

RFS (ppm)	**POLAR**	**MID-LAT**	**TROPICAL**
CO2	0.002	0.001	0.002
N2O	0.026	0.011	0.004
CH4	0.002	0.003	0.001
HNO3	0.052	0.007	0.002
errSPDB	0.076	0.092	0.087
errSPDB_H2O	0.053	0.085	0.071
errSPDB_H2O_LS	0.014	0.014	0.008
errSPDB_H2O_HWHM	0.077	0.078	0.099
HR GRID	0.237	0.165	0.170
CONT-PERC	0.007	0.005	0.001
ARA 1S	0.001	0.001	0.001
RANDOM	0.915	0.599	0.460
A-PRIORI	2.000	2.000	2.000

5.6 Errors due to undetected clouds

The presence of clouds partially or entirely occupying the instrumental FOV represents a very important error source if the FORUM measurement occurs in a cloudy scenario and the cloud parameters are not included in the retrieval state vector. This can occur, for example, when the sounded atmospheric scenario is erroneously flagged as ”clear sky” by a cloud mask algorithm based on the study of the FOV homogeneity. To the date, the detection of clouds contaminating the FSI scene is based on the FSI measurements only, nonetheless the FORUM Embedded Imager (FEI) described in Chapter 2 could be also exploited to detect the homogeneity of the FORUM footprint due to cloud contamination or discontinuities of the underlying surface. Then, we will perform the analysis of the errors due to undetected clouds introducing first a criterium for the cloud detection based on the FORUM FEI measurements and then calculating the errors for specific scenarios in which the cloud detection could be not straightforward. In particular, the detection from the FEI can be difficult if the occurring cloud does not show a clear signature in the 10.5 μm Brightness Temperature field. Therefore, in order to investigate the detectability of different cloud types from the FEI measurements, a preliminary analysis on the correlation among the 10.5 μm BT difference between cloud and clear sky scenario and the cloud Ice Water Path (IWP) for the 21 ice clouds scenarios reported in Table3.1 has been performed. This is shown in the Figure5.20.

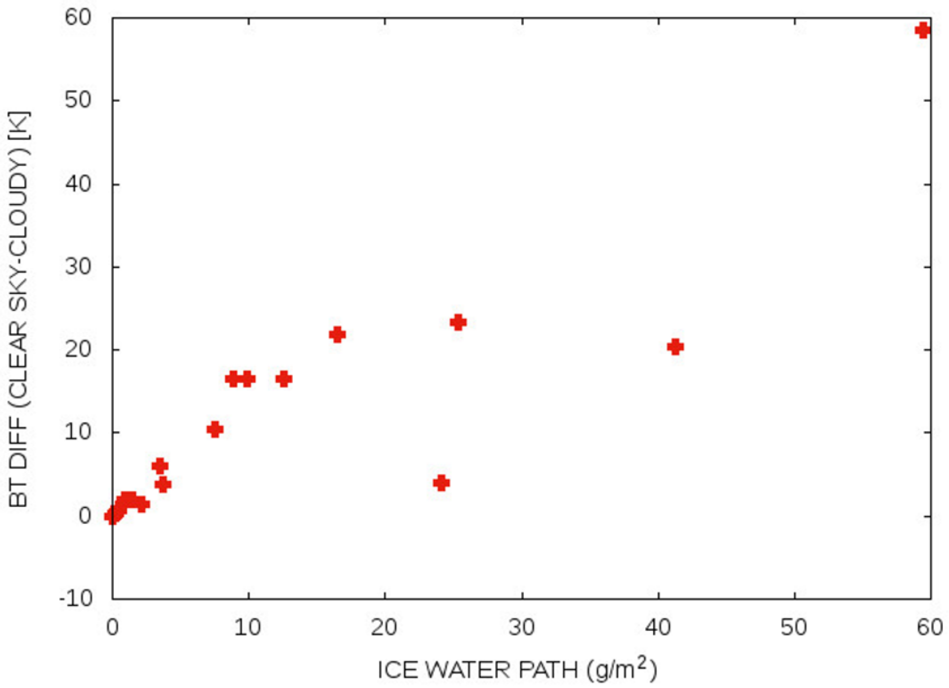


Figure5.20: Differences between the cloudy and clear sky brightness temperatures at 11 μm wavelength for the analysed scenarios

As expected, the thinnest clouds have few differences in brightness temperature from the corresponding clear sky scenarios. Generally, the BT difference between cloud and the corresponding clear sky scenario lays above the instrumental noise of the Imager, however the presence of a thin cloud in the FOV could provide inhomogeneity in the BT field not clearly distinguishable from some kinds of discontinuities due to the background properties (as coastal lines, pack ice boundaries and so forth). On the other hand, clouds covering the whole instrumental FOV do not produce inhomogeneity in the BT field, but this can be very different from the measured BT in the adjacent clear sky scenarios. Moreover, the corresponding FSI spectra can show very high differences that could not be attributable to the clear sky variability of the atmosphere in a spatial scale of 100 km. That is true mainly for thick clouds as BT difference generally increase with the IWP, however the cloud BT could result quite close to the clear sky one if the cloud top is not high and/or the surface is very cold. The preliminary analysis performed on the considered statistics can be summarized in the following two points on cloud detectability and the consequent strategy for the errors assessment taking into account the statistics shown in Figure5.20

- Clouds with IWP < 3 g/m^2 may not be detected by the Imager because of the small BT difference between the cloud and the clear part of the scenario. Therefore, the assessment is performed for 50% and 100% cloud coverage, as undetected thin clouds could produce high systematic errors if occupying a large part of the FOV. (cloud fraction α=[0.5,1]).
- Clouds with IWP > 3 g/m^2 partially occupying the instrumental FOV can be detected with high confidence by the FORUM FEI and/or on the FSI radiance themselves, nevertheless when the cloud covers the whole FOV it could not be detected if the relative BT in the FEI window is very close to the clear sky BT (i.e. in the case of very low thick clouds). On the other hand, thick clouds having strongly different BT from the clear sky ones could produce even acceptable systematic errors on the retrieved variables if occupying a small part of the FOV. Therefore the error assessment is performed for 5%, 10%, 20% and 100% cloud coverage (cloud fraction α=[0.05,0.1,0.2,1]).

In the following Subsections, we will show the systematic errors affecting the retrieved water vapour, temperature and surface emissivity for some of the 21 ice clouds scenarios defined in Section3.7. The selected scenarios are listed in Table5.5, where the IWP, the 10.5 μm BT of the clear sky and cloudy scene and the Cloud Top Height (CTH) are shown. This last and the difference between the clear sky and cloudy BTs will be considered in addition to the IWP as a benchmark to discuss the detectability of each single cloud from the FORUM FEI.

After computing the clear sky and cloudy spectra for the selected scenarios and convolving them with the FORUM ISRF function, the error spectra δn for the cloudy FOVs have been calculated with the linear combination of the convolved clear sky S$_{clear}$ and completely cloudy S$_{cloudy}$ spectra regulated by the cloud fraction α.

Table5.5: Description of the ice cloud scenarios selected for the assessment of the systematic errors due to undetected clouds. For each scenario identifier (first column), the table shows the corresponding scenario identifier in Table3.1, the total Ice Water Path (IWP), the brightness temperature (BT) measured by the FORUM FEI in clear sky and cloudy conditions respectively and the cloud top height (Z_{TOP})

Case	**Scenario (Table3.1)**	**IWP(g/m^2)**	**BT Clear(K)**	**BT Cloud(K)**	**Z_{TOP}(km)**
1	4	0.05	249.47	249.51	4.5
2	12	2.11	253.65	252.14	8.41
3	18	59.40	301.78	243.32	10.22
4	13	24.08	243.40	239.44	4.79

$$S_{pert} = \alpha S_{cloudy} + (1-\alpha) S_{clear} \tag{5.7}$$

$$\delta n = S_{pert} - S_{clear} = \alpha (S_{cloudy} - S_{clear}) \tag{5.8}$$

Case 1

For the first case we have selected the scenario 4 among the ones reported in Table5.5. Since the cloud is very thin and its perturbation on the measured BT is in the order of 10^{-2} K, it is likely that it cannot be detected by the Imager and the systematic errors are performed for 50% and 100% cloud coverage. The error spectra referring to 50% and 100% cloudy scenario, compared to the GOAL NESR of the FORUM FSI, are shown in the left panel of Figure5.21, while the cloudy and clear sky spectra are reported in the right panel. The results on water vapour, temperature and emissivity retrieval are shown in the top left panel, top right panel and bottom panel of Figure5.22 respectively.

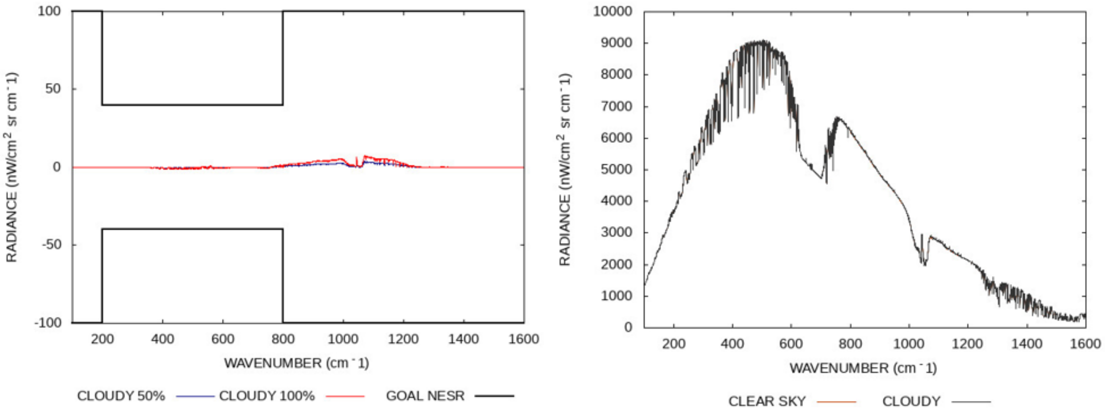


Figure5.21: Error spectra for 50% (blue line) and 100% (red line) cloudy spectra (left panel), plus the FORUM clear sky (orange line) and totally cloudy (black line) spectra (right panel). The very low difference between the cloudy and clear sky spectra in this scenario does not allow to visualize the clear sky spectrum with the adopted plot scale

The impact of this cloud on the TOA radiance is equal to few nW/cm^2 sr cm^{-1}, far below the instrumental noise, so that the cloudy and clear sky spectra cannot be distinguished in the right panel of Figure5.21. As a consequence, the impact of the related systematic errors on the retrieved quantities is negligible compared to the random error, unless for the retrieved water vapour under the cloud top, where systematic errors range between 0.1% and 1% (see blue and red lines in the top left panel of Figure5.22. Generally, ice clouds with IWP $<$ 0.1 g/m^2 do not affect retrievals in a significant way.

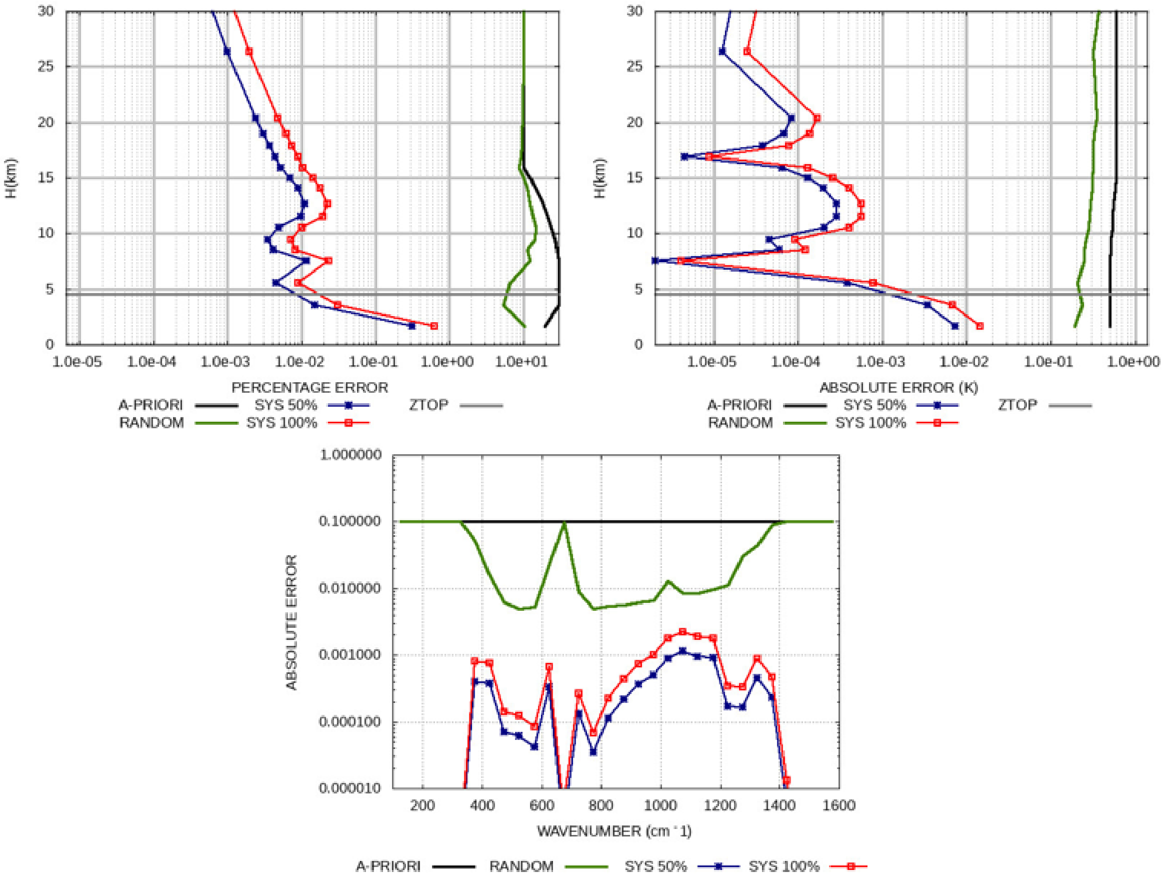


Figure5.22: Systematic errors affecting retrieval of water vapour (top left panel), temperature (top right panel) and surface emissivity (bottom panel) for the ice cloud scenario 4 in Table3.1 with cloud occupying the 50% (blue lines) and the 100% (red lines) of the instrumental FOV. The systematic errors are compared to the assumed a-priori (black line) and random (olive green lines) errors. The cloud top is highlighted with grey thick line in the top panels

Case 2

For case 2 we have selected the scenario 12 among the ones reported in Table5.5. The cloud produces inhomogeneity in the BT field far above the instrumental noise of the FORUM FEI (see Section2.3), nevertheless a similar difference could be produced by a discontinuity in the underlying surface. Therefore, the Imager may not be able to detect this cloud, therefore it is considered as ‘not detectable’ in this analysis. For this reason, the systematic error assessment is performed for 50% and 100% cloud coverage. The clear sky and cloudy spectra (orange and black lines respectively) are shown in the right panel of Figure5.23, results on water vapour, temperature and emissivity retrieval are shown in the top left panel, top right panel and bottom panel of Figure5.24 respectively.

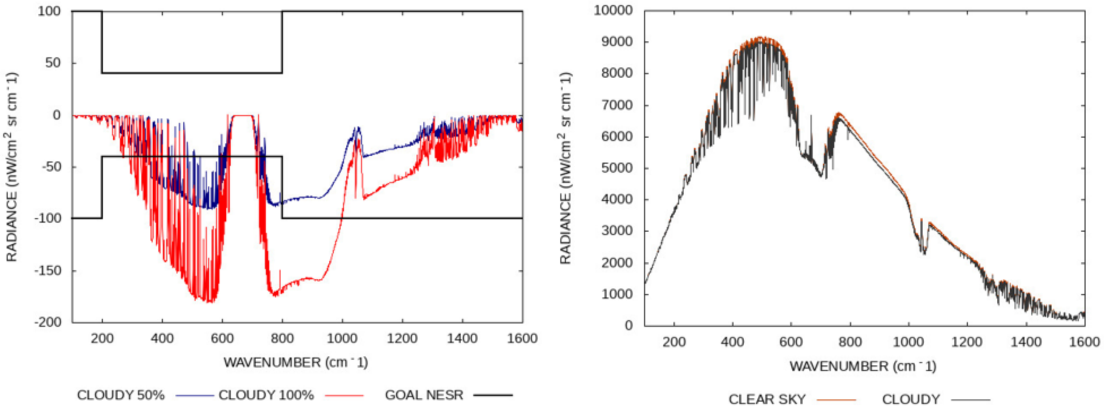


Figure5.23: Error spectra for 50% (blue line) and 100% (red line) cloudy spectra (left panel), plus the FORUM clear sky (orange line) and totally cloudy (black line) spectra (right panel)

The error spectra related to cloud coverage are larger than the FORUM instrumental noise (both in the far infrared and in the thermal atmospheric window, as can be observed in the right panel of Figure5.23), producing high systematic uncertainties on the retrieved water vapour and temperature, respectively equal to 30% and 1 K, larger than the random errors below the cloud top, whereas the errors are one order of magnitude smaller above the cloud top (see blue and red lines in the top panels of Figure5.24). The errors on the FIR surface emissivity ranges between 0.01 with cloud coverage equal to 50% to 0.02 with cloud coverage equal to 100% and are generally larger than the random error. As similar clouds are difficult to be detected, these clouds could result in having a large error on the retrieved geophysical parameters.

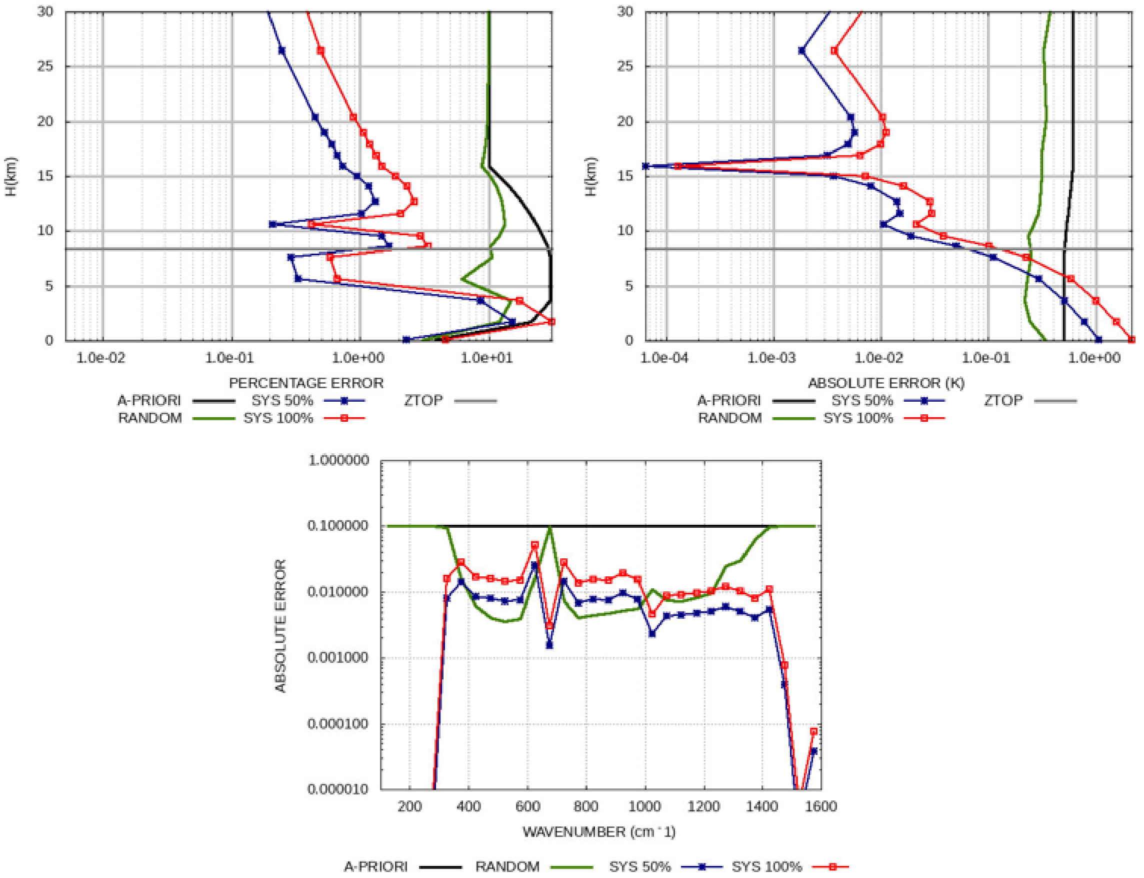


Figure5.24: Systematic errors affecting retrieval of water vapour (top left panel), temperature (top right panel) and surface emissivity (bottom panel) for the ice cloud scenario 12 in Table5.5 with cloud occupying the 50% (blue lines) and the 100% (red lines) of the instrumental FOV. The systematic errors are compared to the assumed a-priori (black line) and random (olive green lines) errors. The cloud top is highlighted with grey thick line in the top panels

Case 3

For case 3 we have selected the scenario 18 among the ones reported in Table3.1. This scenario is affected by a very thick and high cloud producing a huge BT difference, almost 60 K, inside the FORUM FEI Field Of View. The cloud can be detected for every cloudy fraction, as even the spectrum with a 5% cloud coverage produce differences in the measured radiances larger than FORUM FSI NESR. The systematic error assessment is performed for 5%, 10%, 20% and 100% cloud coverage, with the error spectra referring to the same cloud fraction shown in the left panel of Figure5.25, while the FORUM spectra in clear sky conditions and in presence of 100% cloud coverage are reported in the right panel of Figure5.25. The results on water vapour, temperature and emissivity retrieval are shown in the top left, top right and bottom panel of Figure5.26 respectively. The green lines report the errors spectra and systematic errors on the parameters due to cloud coverage equal to 5%, the light blue lines to cloud coverage equal to 10%, the dark blue lines to cloud coverage equal to 20% and the red lines to cloud coverage equal to 100%.

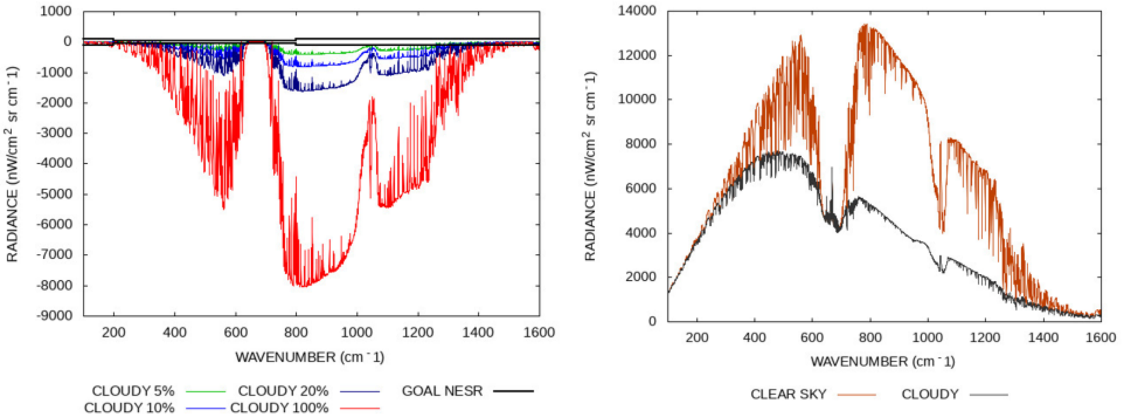


Figure5.25: Error spectra for 5% (green line), 10% (light blue line), 20% (dark blue line) and 100% (red line) cloudy spectra (left panel), plus the FORUM clear sky (orange line) and totally cloudy (black line) spectra (right panel)

If the cloud covers most of the instrumental FOV, the scenario would be flagged as ‘cloudy’ and the cloud parameters would be included in the retrieval state vector. This cloud would be detected even if covering a very small part of the scene, however the retrieval of water vapour and temperature in the UTLS (over the cloud top) could be performed with acceptable errors up to 10% cloud coverage. On the contrary, the systematic errors on the retrieved emissivity is larger than the random error for all the considered cloud fractions and for the whole spectral range of the FORUM FSI.

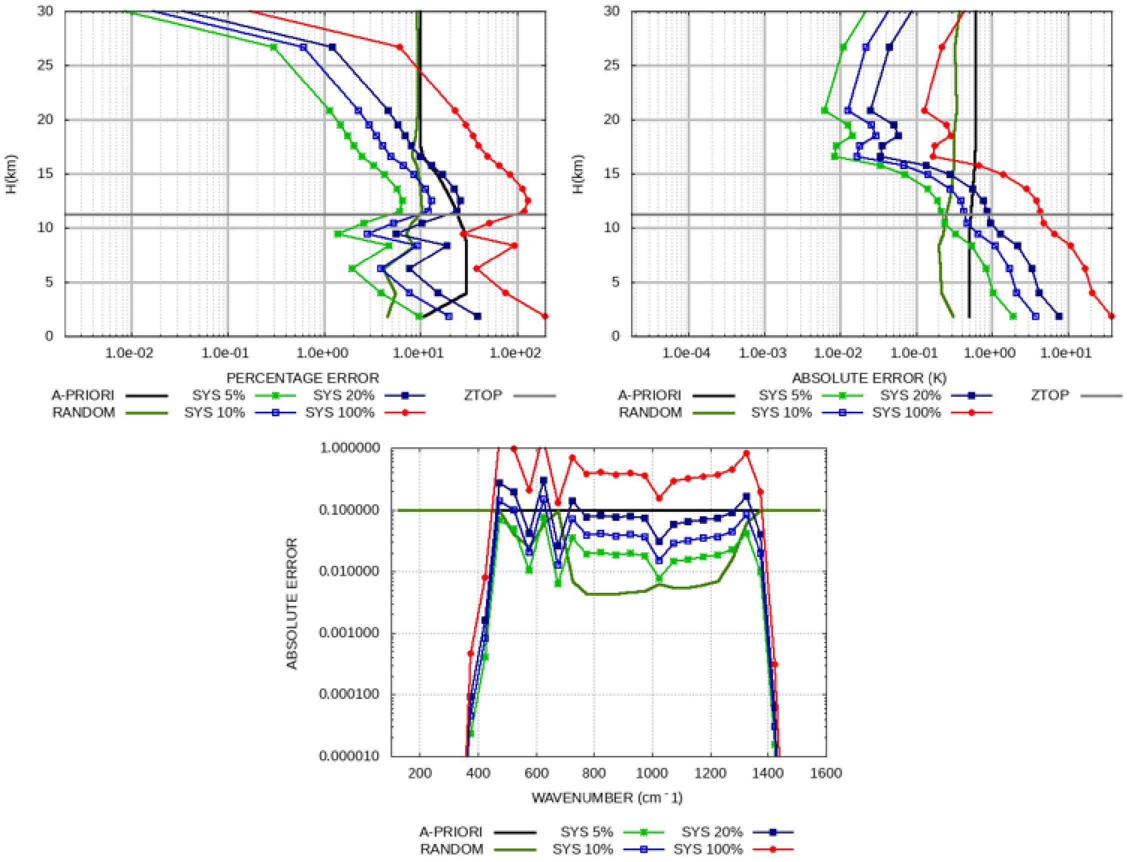


Figure5.26: Systematic errors affecting retrieval of water vapour (top left panel), temperature (top right panel) and surface emissivity (bottom panel) for the ice cloud scenario 12 in Table3.1 with cloud occupying the 5% (green lines), the 10% (light blue lines), the 20% (dark-blue lines) and the 100% (red lines) of the instrumental FOV. The systematic errors are compared to the assumed a-priori (black line) and random (olive green lines) errors. The cloud top is highlighted with grey thick line in the top panels

Case 4

For case 4 we have selected the scenario 13 among the ones reported in Table3.1. Despite of the thickness of this cloud, the BT difference with the clear sky scenario is equal to 4 K, as the cloud top altitude is below 5 km and the underlying surface is very cold (see Table3.1). If the cloud covers part of the instrumental FOV it may be detected by the Imager, the 100% cloudy case is indeed more complicated. In fact, for the cloud covering the whole FOV, the measured spectrum and BT have the same behaviour of the spectra generated by different properties of the underlying surface (skin temperature or reflectivity). For this reason, the cloud detection in this particular case could be not successful. Therefore, the systematic error assessment is performed for 5%, 10%, 20% and 100% cloud coverage. The results on water vapour, temperature and emissivity retrieval are shown in top left panel, top right panel and bottom panel of Figure5.28, with the error spectra referring to the same cloud fractions shown in the left panel of Figure5.27. The green lines report the error spectra and systematic errors on the parameters due to 5% cloud coverage, the light blue lines to 10% cloud coverage, the dark blue lines to 20% cloud coverage and the red lines to 100% cloud coverage. The FORUM FSI spectra in clear sky condition and in presence of 100% cloud coverage are shown in the right panel of Figure5.27

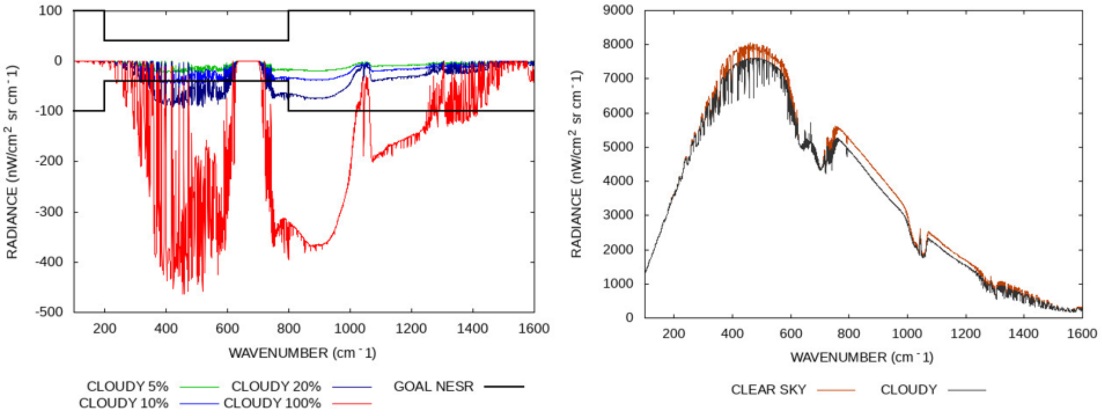


Figure5.27: Error spectra for 5% (green line), 10% (light blue line), 20% (dark blue line) and 100% (red line) cloudy spectra (left panel), plus the FORUM clear sky (orange line) and totally cloudy (black line) spectra (right panel)

The error spectra in Figure5.27 are larger than the instrumental noise both in the FIR and the MIR ranges, but not enough to indicate the clear presence of a cloud in the FOV even when the cloud fraction is 100%. Similarly to the analysed Case 3, the retrievals can be performed with high but acceptable systematic errors affecting water vapour and temperature in the UTLS region for cloud fraction equal to or below 20%. Even surface emissivity can be retrieved with systematic components in the order of 0.01 for a small cloud fraction in the FOV. However, if the cloud covers the entire FOV and it is not detected, the retrieved variables will have systematic errors even larger than the a-priori ones (see red lines in Figure5.28.)

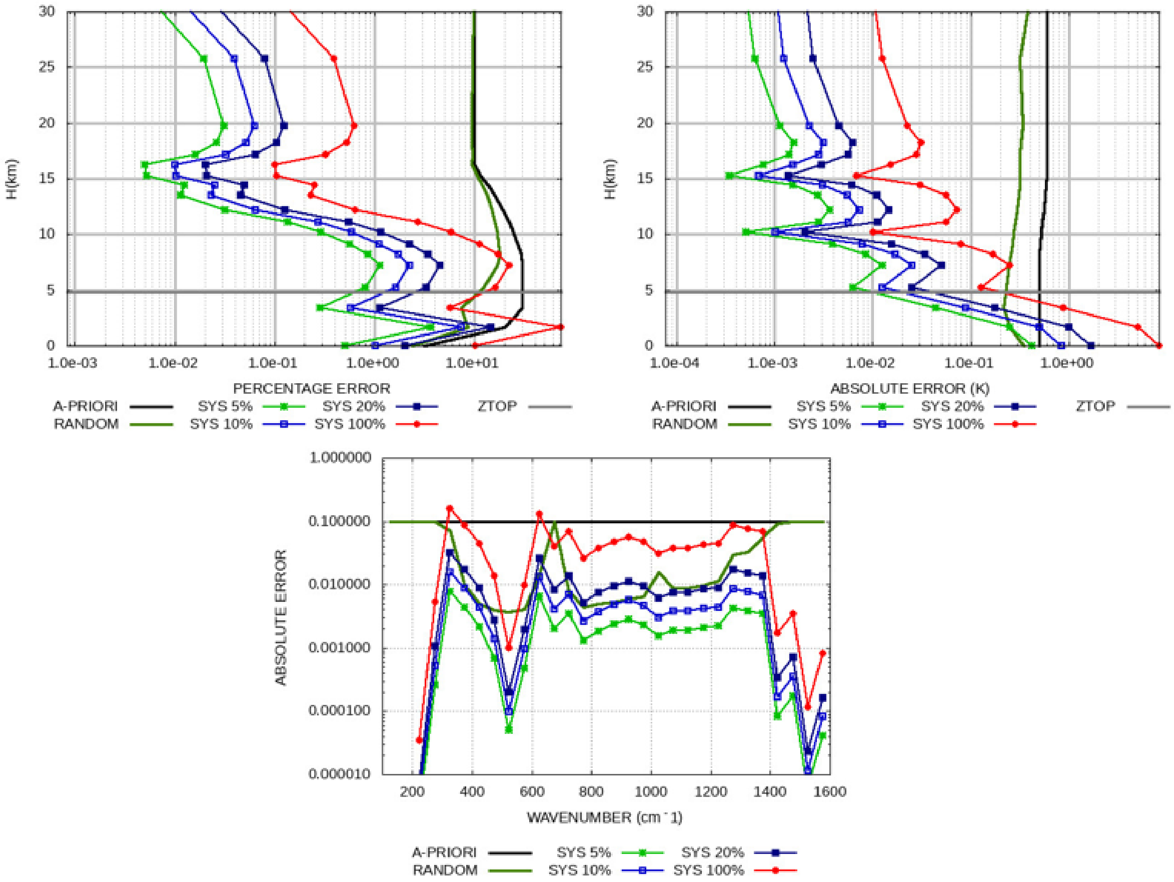


Figure5.28: Systematic errors affecting retrieval of water vapour (top left panel), temperature (top right panel) and surface emissivity (bottom panel) for the ice cloud scenario 13 in Table3.1 with cloud occupying the 5% (green lines), the 10% (light blue lines), the 20% (dark-blue lines) and the 100% (red lines) of the instrumental FOV. The systematic errors are compared to the assumed a-priori (black line) and random (olive green lines) errors. The cloud top is highlighted with grey thick line in the top panels

5.7 Errors on the cloud parameters

Here the systematic errors affecting the estimate of the cloud parameters are calculated by performing the analysis on the 21 ice clouds scenarios defined in Table3.1. First, the retrieval state vector is updated with the addiction of the following parameters

- Ice Water Path (IWP [g/m2], A-Priori error = 100%)
- Cloud Top Height (CTH [km], A-Priori error = 100 km)
- Cloud Bottom Height (CBH [km], A-Priori error = 100 km)

The specified retrieval setup has already been adopted for several tests performed during the FORUM Phase A studies and proved to be successful in the assessment of the synergistic retrievals from FORUM and IASI-NG measurements (Ridolfi et al., 2020). In the presented test, the systematic contribution due to the CO_2 and CH_4 variability and to the ISRF uncertainty components on the IWP, CTH, CBH and RFS parameters are analysed. The CA ISRF function described in Section5.3 has been considered to simulate the FORUM response, hence the error components related to the ISRF knowledge are the ASY, DWD and RFE components as defined in Section5.3. To better highlight the effect of the ice cloud on the accuracy of the Level 2 products, the various error components are plotted as a function of the IWP of the test scenarios. When the retrieved target is a profile, an average error is plotted. Namely, the average errors for H_2O and temperature profiles in the UTLS (from 7 to 16 km) and the average error on the FIR surface emissivity between 300 and 600 cm^{-1} are reported. Figure5.29 shows the average errors for the UTLS water vapour (top left panel) and temperature (top right panel), and the average FIR errors on surface emissivity (bottom left panel). Similarly, Figure5.30 shows the various retrieval error components for the RFS, IWP, CTH and CBH. For reference, the mentioned figures also show the A-Priori, the NESR and the ARA error contributions.

Figure5.29 does not show relevant differences with respect to the clear sky scenarios analysed in the previous sections, at least with respect to the different ISRF error components. These components are confirmed to provide a small contribution to the total error, except the case of the 0.6% error on the ISRF width highlighted with the blue lines. This ISRF model error may introduce errors in the UTLS water vapour and temperature up to 1% and 0.1 K, that is occasionally increased by the contrast between the cloud black body emission and the gaseous absorption occurring above the cloud top. The presence of clouds generally produces a loss of information on the surface properties, that translates in an increased random error of the emissivity, assuming larger values with the increasing IWP, as shown by the red thick line in the bottom left panel of Figure5.29. However, the random error of the FIR emissivity is significantly smaller than the a-priori error in the polar scenarios even with high IWP values. In fact, due to the low water content, the transmittance of the Polar atmospheres is preserved

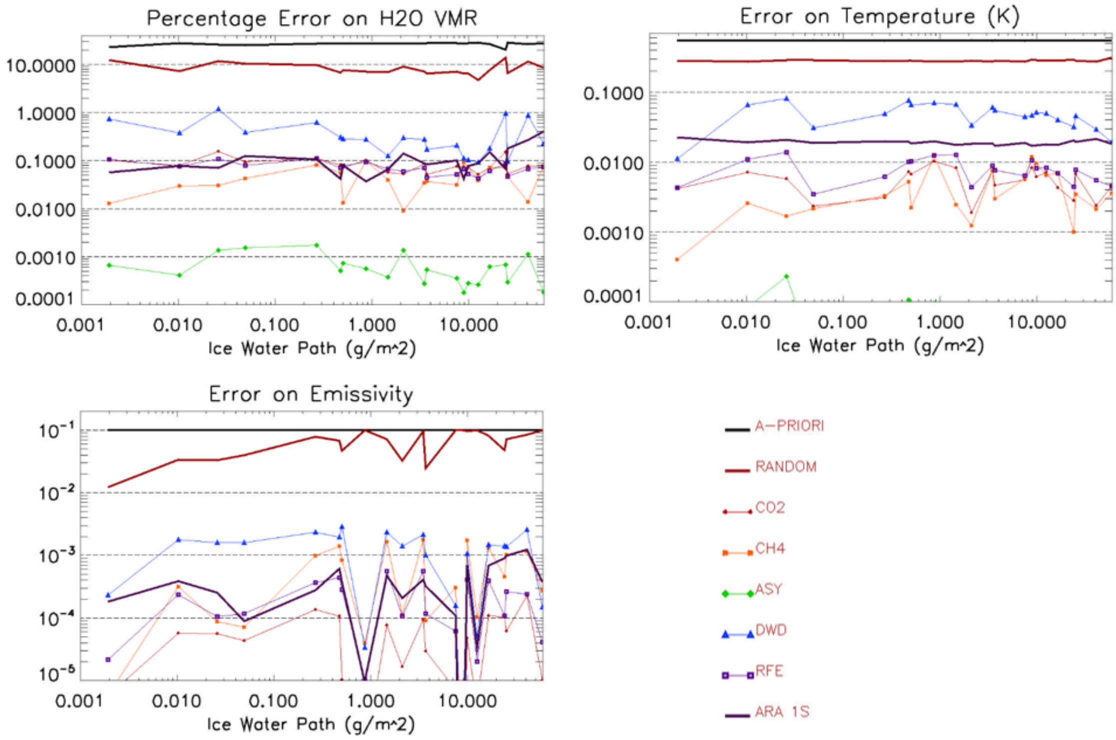


Figure5.29: Average UTLS error components affecting water vapour (top left panel) and temperature (top right panel). Average FIR (300 to 600 cm-1) error components of surface emissivity (bottom left panel). Errors are plotted versus the IWP value of the considered test cases. The considered systematic errors are the ones originating from the CO2 (red lines), CH4 (orange lines) ASY (green lines), DWD (blue line) and RFE (thin purple line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

also in the presence of the considered ice clouds. The same holds for the information of the measurements on the surface properties. The effects of CO$_2$ and methane interferences (red and orange lines respectively) are generally reduced as the cloud extinction increases. The major contribution to the RFS error is confirmed to be the RFE component (green lines), whereas the other inaccuracies introduce error components equal or smaller than 0.01 ppm. It can be observed from Figure5.30 that the errors on IWP and CTH decrease with the IWP itself, at least in the range of the IWP values encountered in the considered test cases. The errors on CBH behave similarly as those of CTH, but the dependence on IWP is less pronounced. Finally, we can conclude that the considered error components provide contributions to the error budget of the retrieved cloud parameters that are generally from one to three orders of magnitude smaller than the error due to measurement NESR.

The results for the noise errors are consistent with those published in Ridolfi et al. (2020), performing the analysis on the same atmospheric dataset. Furthermore, all the considered error components provide contributions to the error budget of the retrieved cloud parameters that are from one to three orders of magnitude smaller than the error due to measurement NESR.

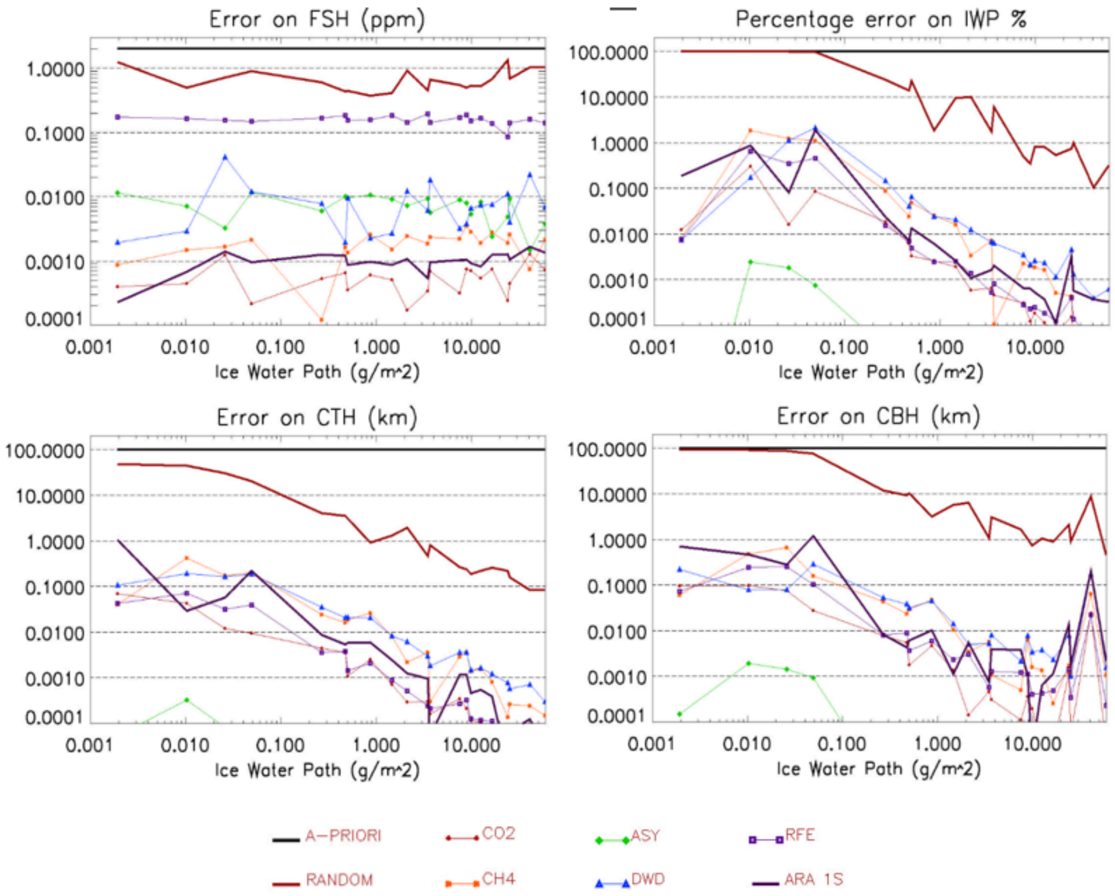


Figure5.30: Retrieval error components as a function of IWP in test cases. The various panels refer to RFS (top left panel), IWP (top right panel), CTH (bottom left panel) and CBH (bottom right panel). Errors are plotted versus the IWP value of the considered test cases. The considered systematic errors are the ones originating from the CO2 (red lines), CH4 (orange lines) ASY (green lines), DWD (blue line) and RFE (thin purple line) error sources, compared to the assumed a-priori (black thick line), random (thick red lines) and ARA (thick purple lines) errors

Chapter 6

Spectroscopic databases comparison

The spectroscopic databases are at the basis of every RT simulation. Their quality directly affects the ability of a RTM model to correctly represent the measurements. Moreover, the quality of the spectroscopic data used in the level 2 analyses of most remote sensing instruments has a direct influence in most of the retrieved data. However, the spectroscopic data may be affected by non-negligible uncertainties and each database has differences with respect to the others, e.g. reporting a different number of absorption lines or different line strengths and pressure broadening coefficients for the same absorption line, introducing uncertainties on the computed spectra and relevant systematic contributions on the retrieved products from signal inversion as observed in Chapter 5. The FIR region, in which the very strong water vapour absorption is the most important contribution to the atmospheric extinction in clear sky conditions, is particularly challenging as there is a lack of direct measurements. Therefore most of the spectroscopic data reported in the databases come from theoretical calculations obtained through the extrapolations of the molecular parameters measured in other spectral regions. Therefore, the uncertainty affecting the water vapour lines parameters is higher than the one in the MIR spectral range, with average uncertainty on the air broadening half-widths and line strengths of about 8% and 3% respectively (see the left panel of Figure6.1), despite the fact that the FIR water vapour lines are more intense, with line strengths up to 10^{-17} cm^{-1}/(molecule cm^{-2}), as clearly shown from the right panel of Figure6.1. In other spectral regions, however, the uncertainty on the water vapour lines is less important due to the interference with the gaseous absorption of the other major species (e.g. carbon dioxide, ozone and methane).

On top of that, the water vapour continuum parametrization is also affected by high uncertainty in the FIR range (Mlawer et al., 2019), mostly for wavenumbers below 400 cm^{-1}. This leads to large discrepancies in the computed FIR optical depths and TOA radiances, when different spectroscopic databases are exploited in the simulations. In the next Sections, the differences between the spectroscopic data from the different databases are investigated through the comparison between key line parameters (namely line strength and pressure broadening

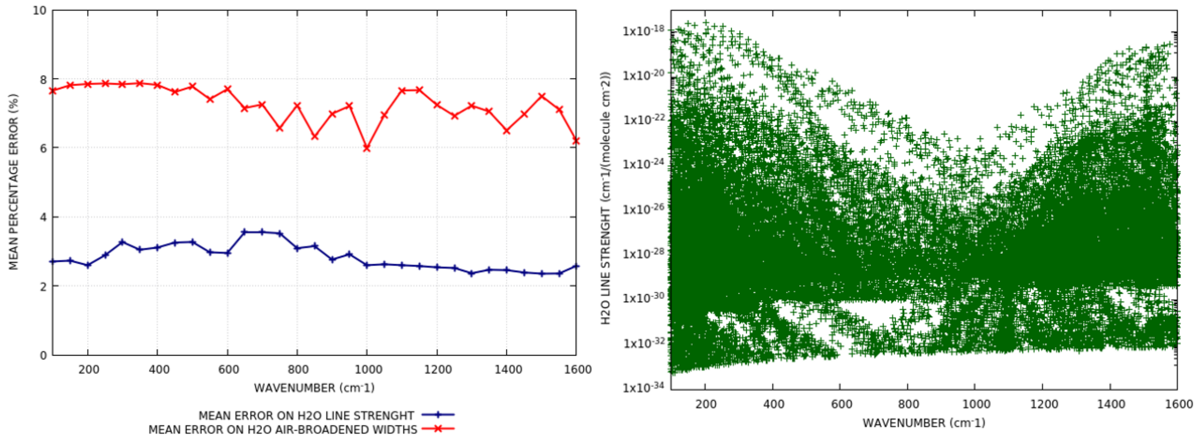


Figure6.1: Mean uncertainty of the water vapour lines line strengths (blue line) and air-broadened half widths (red line) from the HITRAN2016 database from 100 to 1600 cm^{-1}. Each value is calculated considering the water vapour lines included in spectral intervals 50 cm^{-1} wide (left panel). The line strengths of the water vapour lines in the same spectral range from HITRAN2016 are shown in the right panel

coefficient) reported in the different databases. Then two water vapour continuum models will be compared: the continuum models MT_CKDv2.5 and MT_CKDv3.0. Finally, the simulations using the different databases will be compared to the measurements of the REFIR-PAD sensor from a stratospheric balloon, in order to assess the consistency between the simulations using different spectroscopic data and the measured spectra.

6.1 Spectroscopic parameters

In this section a comparison between the line strength and pressure broadening coefficients reported in the most updated databases used by GBB-NADIR and KLIMA has been performed. Figures6.2 and 6.3 show the line strengths and the air-broadening coefficients of the water vapour lines from 50 to 650 cm^{-1} in the HITRAN2020, AERv3.5 and GEISA2015 databases, overlapped to the parameters reported in the baseline HITRAN2016 database. In each of the graphs reported in Figure6.2 and 6.3 the line strengths and air-broadening coefficients of the water vapour transitions from the reference HITRAN2016 are reported with purple crosses, while the green crosses refer to the parameters from the other analysed databases.

The examined databases show negligible differences for the line strengths of the water vapour lines in the FIR range, whereas considerable inconsistency between the pressure broadening coefficients of the HITRAN and AERv3.5 databases, reported in the top right panel of Figure6.3, can be appreciated. This is due to the different concepts on which the databases have been developed, since in the AERv3.5 release the water vapour air-broadened coefficients have been estimated on the basis of the analysis of real FIR measurements, acquired during the ARM

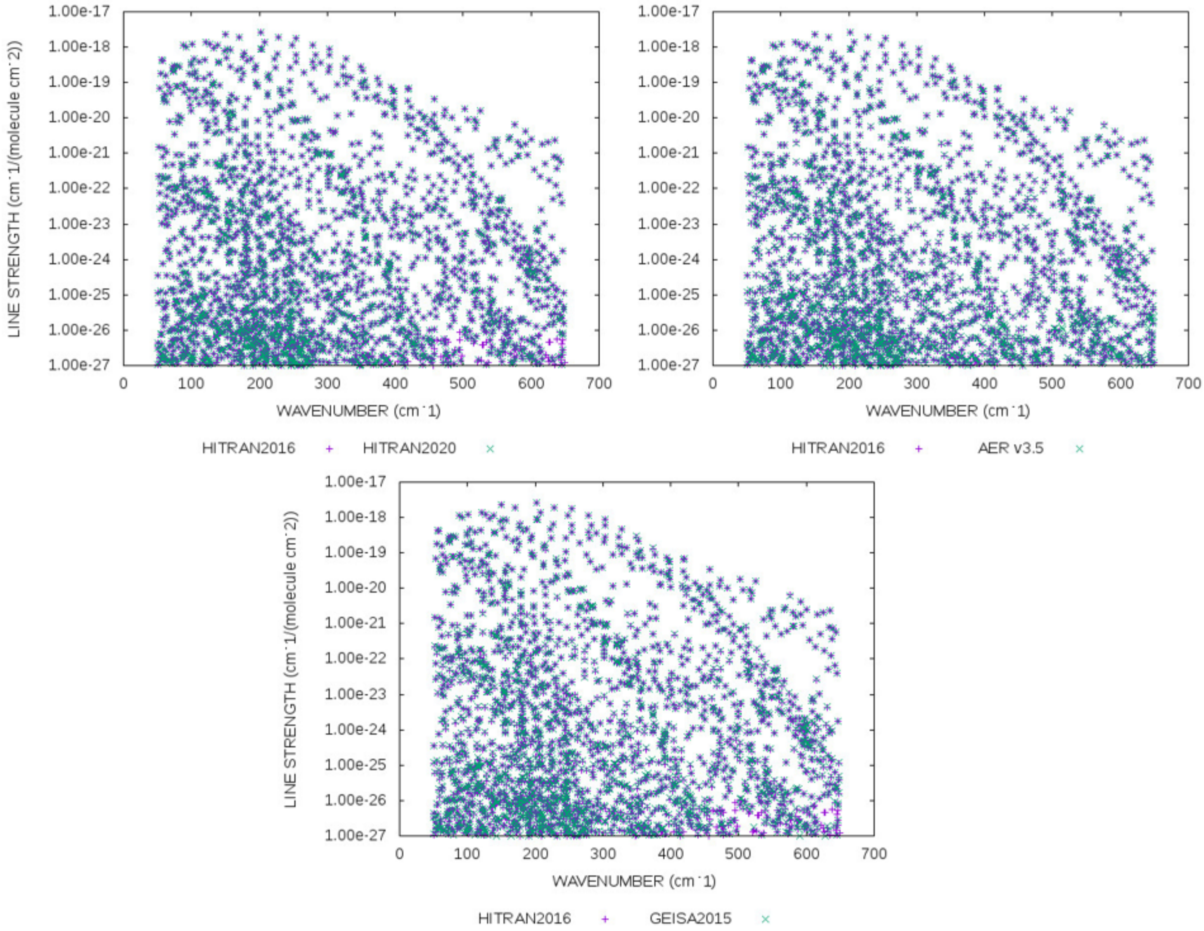


Figure6.2: Line strengths of the water vapour lines in the FIR range as a function of the wavenumber from the HITRAN2016 reference database (purple crosses) compared to the FIR water vapour line strengths from HITRAN2020 (top left panel), AERv3.5 (top right panel) and GEISA 2015 (bottom panel) reported with green crosses

RHUBC-II campaign (Palchetti et al., 2015) and a re-analysis of the AERI-ER measurements (first analysed in Delamere et al. (2010)), whereas the HITRAN databases use the water vapour air-broadening coefficients obtained with the “Diet” procedure (Gordon et al., 2007), first used in the 2008 version, with some modifications in HITRAN2012 and HITRAN2016.

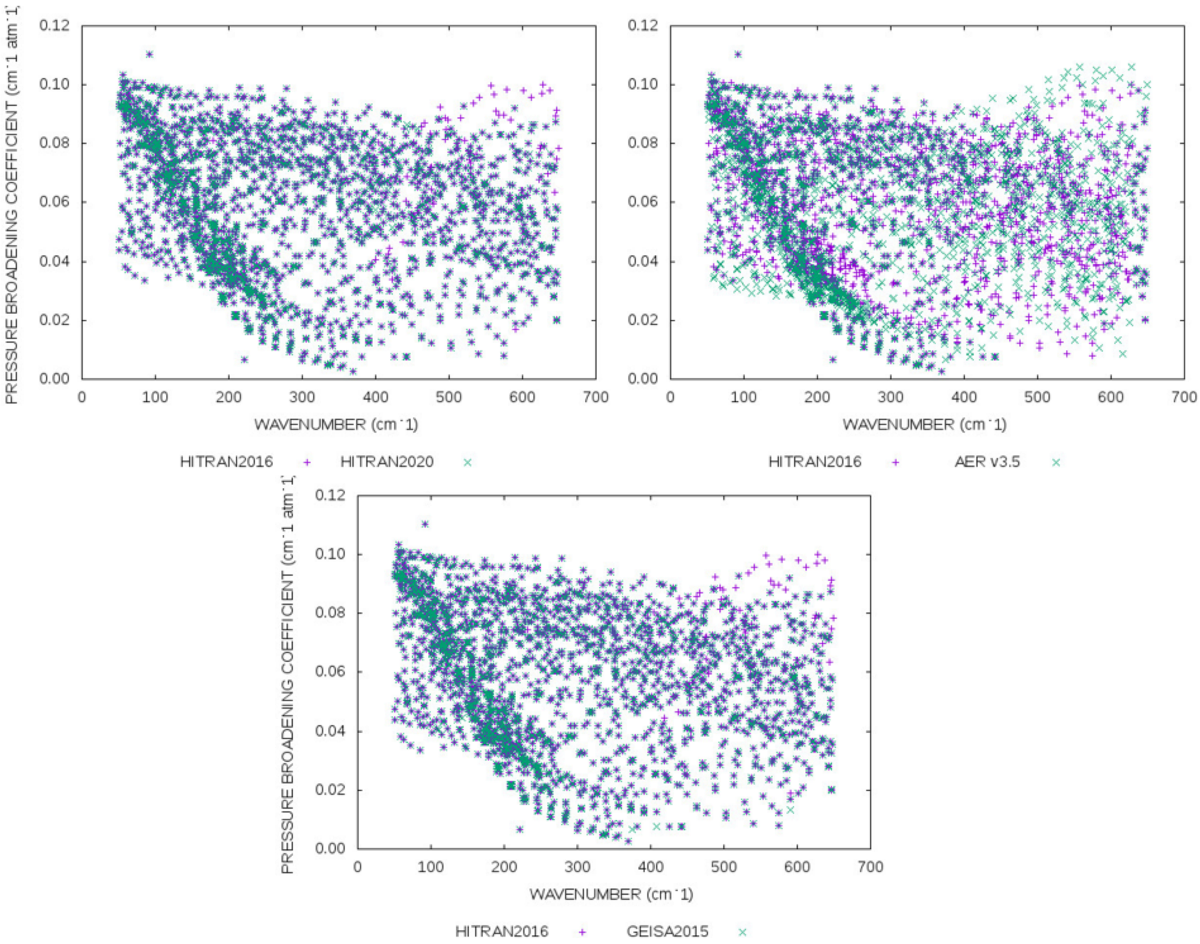


Figure6.3: Pressure broadening coefficients of the water vapour lines in the FIR range as a function of the wavenumber from the HITRAN2016 reference database (purple crosses) compared to the FIR water vapour pressure broadening coefficients from HITRAN2020 (top left panel), AERv3.5 (top right panel) and GEISA2015 (bottom panel) reported with green crosses

6.2 Water vapour continuum versions

In a radiative transfer calculation, the division between what is considered absorption due to spectral lines and that due to the continuum is somewhat arbitrary, with the only absolute requirement that the appropriate total absorption is preserved (Mlawer et al., 2012). Each version of the MT_CKD water vapour continuum has been released with the specific versions of the AER spectroscopic database and the LBLRTM (Line-By-Line Radiative Transfer Model) code, the last performing the line-by-line calculation using the Lorentzian profile for each line within 25 cm^{-1} from its centre, whereas all absorption beyond 25 cm^{-1} from the line centre is defined as continuum absorption (Clough et al., 2005). This strategy enables to reduce the computing time needed to correctly reproduce the observed spectra in a line-by-line calculation, nevertheless it needs to be consistent with the used line parameters, particularly with the coefficients involved in the pressure broadening effects, as anticipated in Chapter 3. The contribution provided on the water vapour cross section by the MT_CKDv2.5 and the MT_CKDv3.0 versions of the water vapour continuum parametrization (shown in Figure6.4 with red and green lines respectively) present relevant differences in the FIR range as the latter has been released in pairs with the AERv3.5 database, that has introduced important changes to the air-broadened coefficients of the FIR water vapour lines, and therefore changing their far wing contribution to the spectrum. Differently, the AER versions previous to the v3.5 have adopted the same water vapour air-broadening coefficients of the HITRAN database, therefore the HITRAN releases are more or less consistent with the water vapour continuum version previous to the MT_CKDv3.0. Thus, the AERv3.5 database has to be used with the MT_CKDv3.0 continuum, since these database and continuum version are coupled in the LBLRTMv12.6 code, while the previous versions of the continuum have to be coupled with the GEISA database and the different HITRAN versions, not containing the adjustment on the FIR water vapour lines included in AERv3.5. Specifically, the HITRAN2012 and HITRAN2016 versions should be coupled with the MT_CKDv2.8 version, as the coupled AER versions previous to the v3.5 take the water vapour transitions from HITRAN as the reference. The HITRAN2020 release has revised only the line position of some transitions of the first water vapour isotope, whereas the GEISA2015 database does not show relevant changes with respect to the HITRAN versions, therefore these should be coupled to the MT_CKDv2.8. However, this version does not show differences with respect to the previous MT_*CKD*v2.5 version in the FIR range, therefore the MT_*CKD*v2.5 model has been exploited for all the following simulations with the HITRAN and GEISA databases.

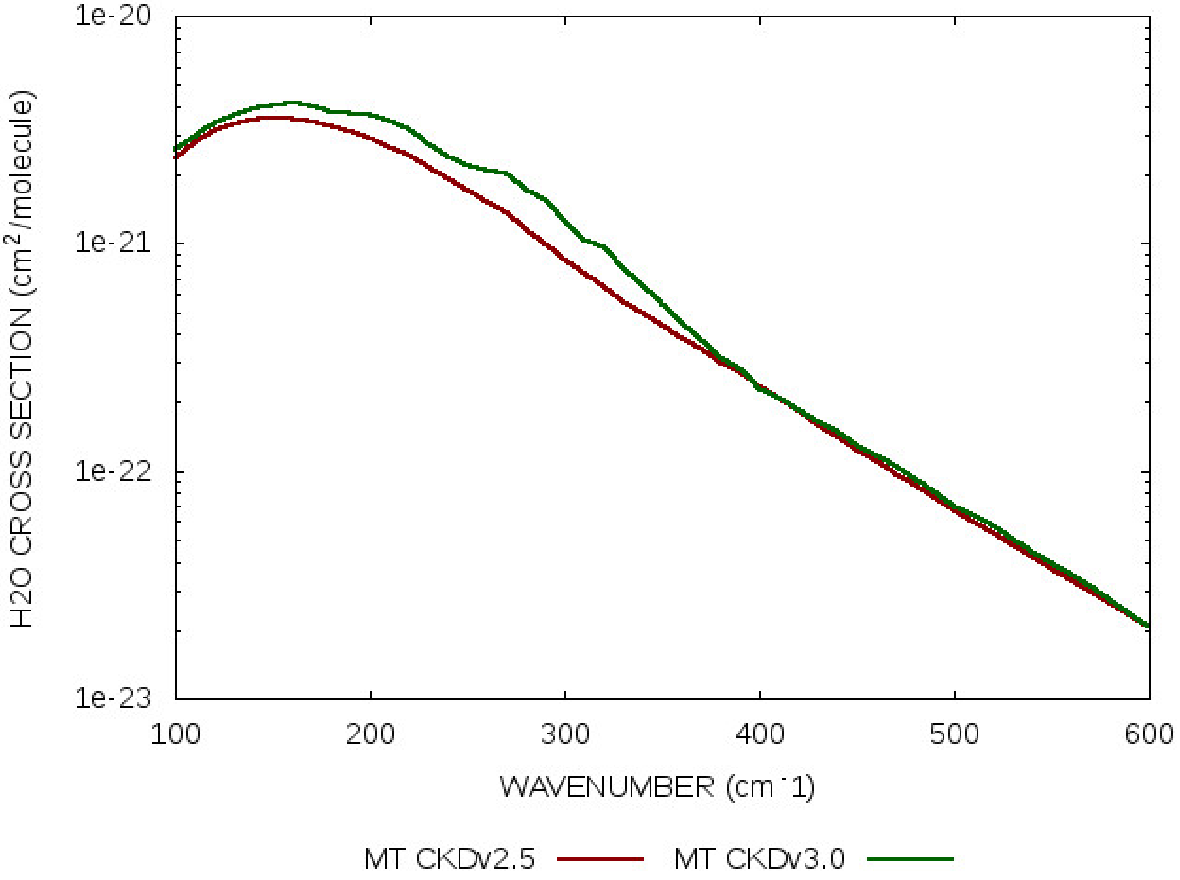


Figure6.4: Contribution to the H_2O cross section from the MT_CKDv2.5 (red line) and the MT_CKDv3.0 continuum models in the far infrared (green line). The computation has been performed using an atmospheric layer with T=285K and P=1013 hPa

6.3 Comparison with the REFIR-PAD measurements

In this Section, the comparisons between the radiance simulations performed with the different databases and the clear sky measurements of REFIR-PAD sensor above Teresina (Brazil) on June the 30^{th} 2005 are reported. First, the general characteristics of the REFIR-PAD sensor are presented, following the method used for the generation of the synthetic spectra, the method developed for their comparison with the measurements and the results are discussed.

6.3.1 The REFIR-PAD instrument and the ELBC campaign

The Radiation Explorer in the Far InfraRed, Prototype for Applications and Development (REFIR-PAD) instrument is one of the prototypes designed in the framework of the REFIR project for the study of the outgoing FIR radiation at the top of the atmosphere. REFIR-PAD is a wideband Fourier transform spectrometer optimised for stratospheric measurements from balloon-borne platforms. The instrument acquires the outgoing radiance from 100 to 1500 cm^{-1} at 0.5 cm^{-1} spectral resolution and sampling equal to 0.39 cm^{-1} and represents a precursor of the FORUM FSI instrument. REFIR-PAD was flown for the first and only time on June the 30^{th} 2005 on board the Laboratoire de Physique Moleculaire pour l'Atmosphere et l'Astrophysique (LPMAA) gondola hosting the Infrared Atmospheric Sounding Interferometer (IASI) – balloon during the Equatorial Large Balloons Campaign (ELBC) performed in Brazil in June 2005. The gondola was launched at 3:36 local time from the Timon airfield next to Teresina, located in the North-Eastern Brazil at 5° 5' S and 42° 52' W (see Figure6.5) and it has floated at the planned altitude of 34 km for about 8 hours. REFIR-PAD acquired 54 nadir spectra during the flight (10 during the ascending leg and 44 at floating altitude), the Figures6.6 and 6.7 show the flight track of the gondola and the altitude of the instrument for each measured spectrum respectively.

Figure6.8 shows the NESR and the calibration error of the REFIR-PAD instrument along the whole spectral range. REFIR-PAD shows acceptable noise values in the FIR range, while it tends to degrade in the mid infrared, due to the absorption features of the internal MyLar beamsplitter.

The Instrumental Line Shape (ILS) function of the REFIR-PAD sensor is a linear combination between a *sinc* function resulting from the finite Maximum Optical Path Difference in the interferometer legs and a $sinc^2$ contribution simulating the self-apodization effect due to the finite angular aperture of the interferometric beam (Bianchini and Palchetti, 2008).

The next Section explains the method used for the computation of the REFIR-PAD spectra generated with the different available spectroscopic databases, then the comparisons with the real measurements are shown.

Figure6.5: Geolocation of Teresina in the North-Eastern part of Brazil

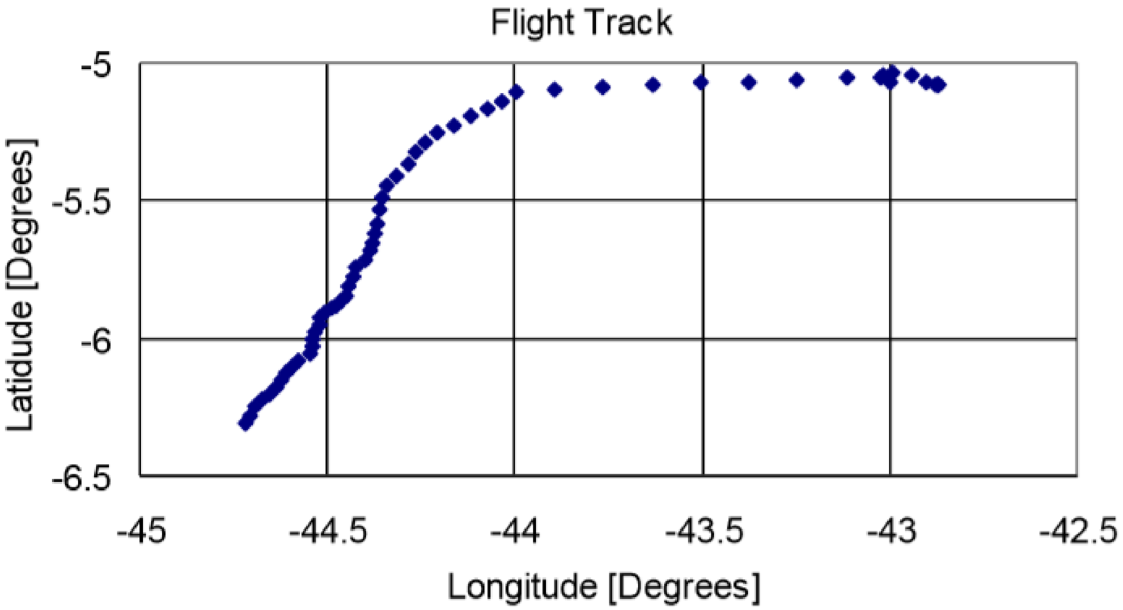


Figure6.6: Flight track of the LPMAA gondola

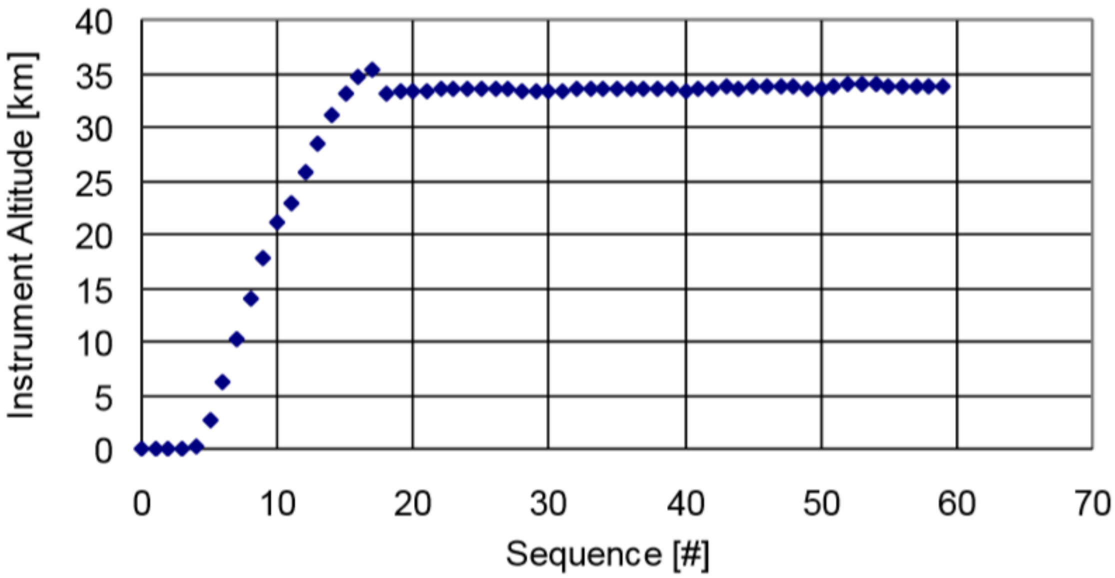


Figure6.7: The instrument altitude vs. the sequence number

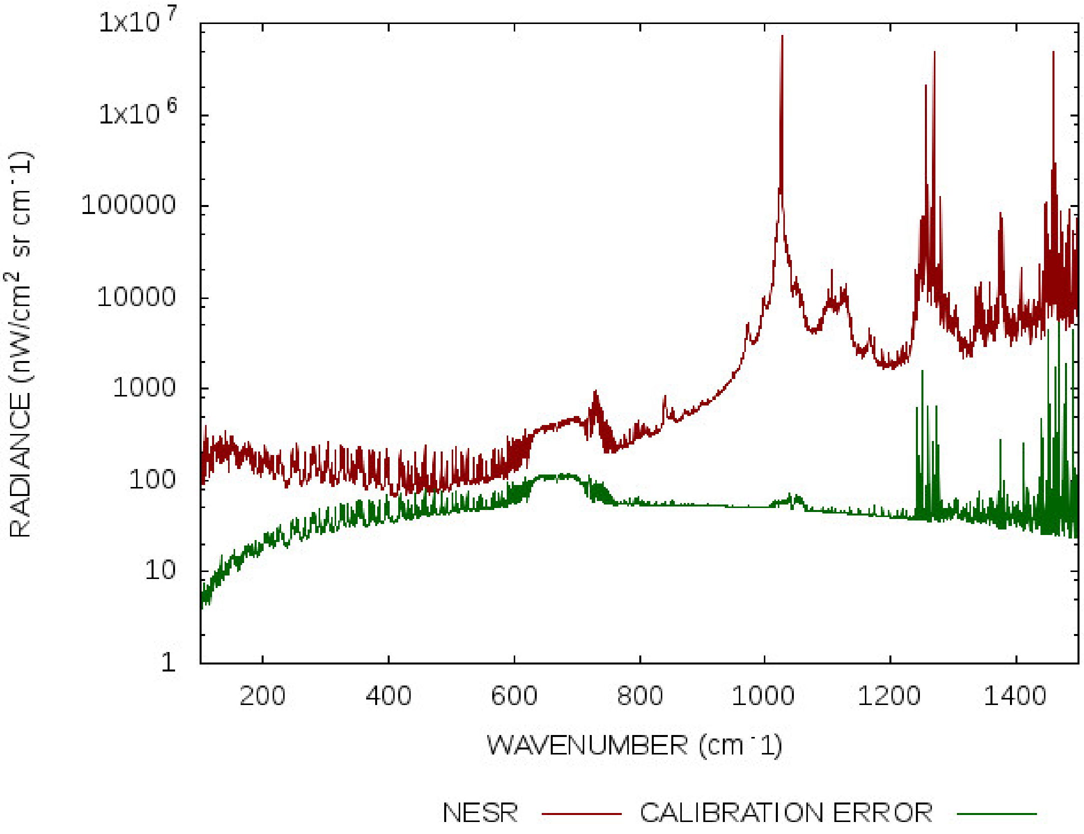


Figure6.8: REFIR-PAD NESR (red line) and calibration error (green line)

6.3.2 REFIR-PAD simulations

The high resolution spectra from 50 to 650 cm^{-1} have been computed with the GBB-NADIR model using the internal RTE solver on a fine wavenumber grid of 0.0005 cm^{-1}. The atmospheric scenarios for the simulations have been extracted from the ERA-Interim database on the finest lat/lon grid corresponding to 0.125° x 0.125° resolution, including the vertical profiles of water vapour, temperature and ozone plus the surface temperature. The underlying surface emissivity has been simulated with the tabulated emissivity of the Amazon rainforest from the database of Huang et al. (2016). For each measured spectrum, the corresponding synthetic one has been computed using the atmosphere of the grid point of the ERA-Interim dataset nearest to the geolocation of the REFIR-PAD measurement. In order to avoid high uncertainty on the atmospheric data deriving from the misalignment between the measurement time and the ERA-Interim reference time, only the spectra collected around the 12UTC have been included in the present analysis. The obtained high resolution spectra are then convolved with the ILS function specified in the previous Section to simulate the acquired spectra from REFIR-PAD. We reported in the top left panel of Figure6.9 the residual between the REFIR-PAD spectrum 35 and the simulation using the HITRAN2020 databases, in the top right panel the residual between the REFIR-PAD spectrum 35 and the simulation using the AERv3.5 and the residual between the REFIR-PAD spectrum 35 and the simulation using GEISA2015 in the bottom left panel. The residuals are compared to the NESR of REFIR-PAD and the residual between the REFIR-PAD spectrum 35 and the simulation using the HITRAN2016 database taken as the reference. Moreover, the differences among the spectrum computed with the reference HITRAN2016 and the other simulated spectra are also shown in the bottom right panel.

We can observe from both the residuals between measured and simulated spectra and the differences between the computed spectra that the simulation using the GEISA2015 database shows few differences with the HITRAN2016 spectrum, whereas the latter considerably differs from the HITRAN2020 spectrum next to some water vapour absorption lines between 400 and 500 cm^{-1} (see top left and bottom right panel of Figure6.9). However, the highest inconsistency occurs between the HITRAN2016 and the AERv3.5 spectra due to the previously discussed differences among the pressure broadening coefficients and the water vapour continuum models, producing differences between the simulations of the order of the NESR of REFIR-PAD, as shown by the green line in bottom panel of Figure6.9. The consistency of the simulations with the measurements could significantly change with the used spectroscopic data. In order to investigate the performances of the analysed databases and to study the consistency of the corresponding simulations with the measurements in different spectral regions within the FIR range, the reduced chi-square defined in Eq.8.12 of Appendix A has been calculated for spectral windows 25 cm^{-1} wide. The results of the spectral chi-square calculation are shown in the

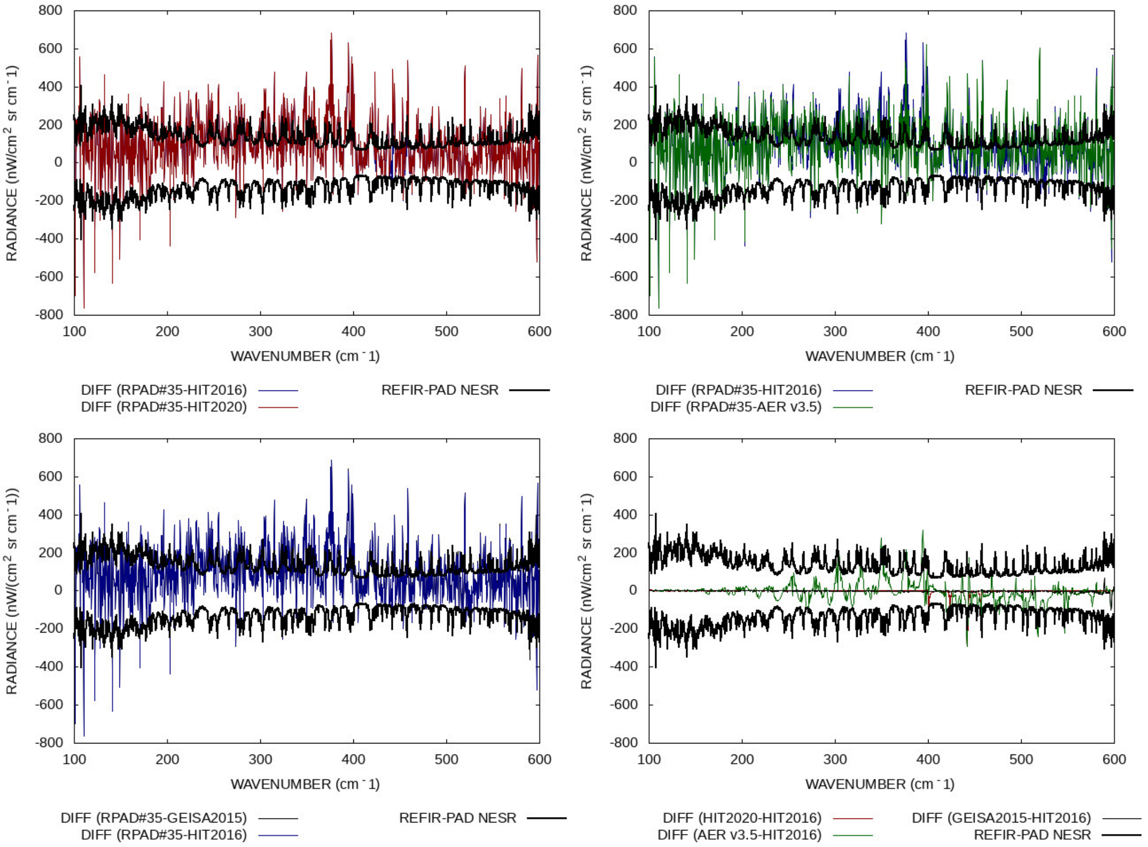


Figure6.9: Residuals between the REFIR-PAD spectrum 35 and the spectra computed with HITRAN 2020 (top left panel, red line), AERv3.5 (top right panel, green line), GEISA 2015 (bottom left panel, black thin line), compared to the residual between the observed spectrum and the spectrum computed with HITRAN 2016 (blue line) and the REFIR-PAD NESR (black thick line). In the bottom right panel the differences among the simulations and the HITRAN 2016 spectrum are reported

Figure6.10 for three different spectra acquired during the gondola flight. Namely, the chi-square for the REFIR-PAD spectrum 34, 35 and 36 are shown in the top left, top right and bottom left panels respectively, the average of the chi-square parameter for the REFIR-PAD spectra from 33 to 39 is shown in the bottom right panel.

The chi-square calculation shows a higher consistency between 300 and 400 cm^{-1} of the simulations with the AERv3.5 database (represented with green lines) with respect to the others, whereas the simulations with the HITRAN databases and GEISA2015 are more in agreement with the measurements out of that spectral range. The HITRAN2020 database (red lines) does not show improved consistency in the region corresponding to the modified water vapour lines.

However, the obtained results do not take into account possible systematic errors affecting the simulations. On the basis of the analysis performed in the previous Chapter 5, the REFIR-PAD spectra and chi-square parameter are calculated by perturbing the simulations with different systematic error sources, identified in the systematic errors assessment, in order to investigate

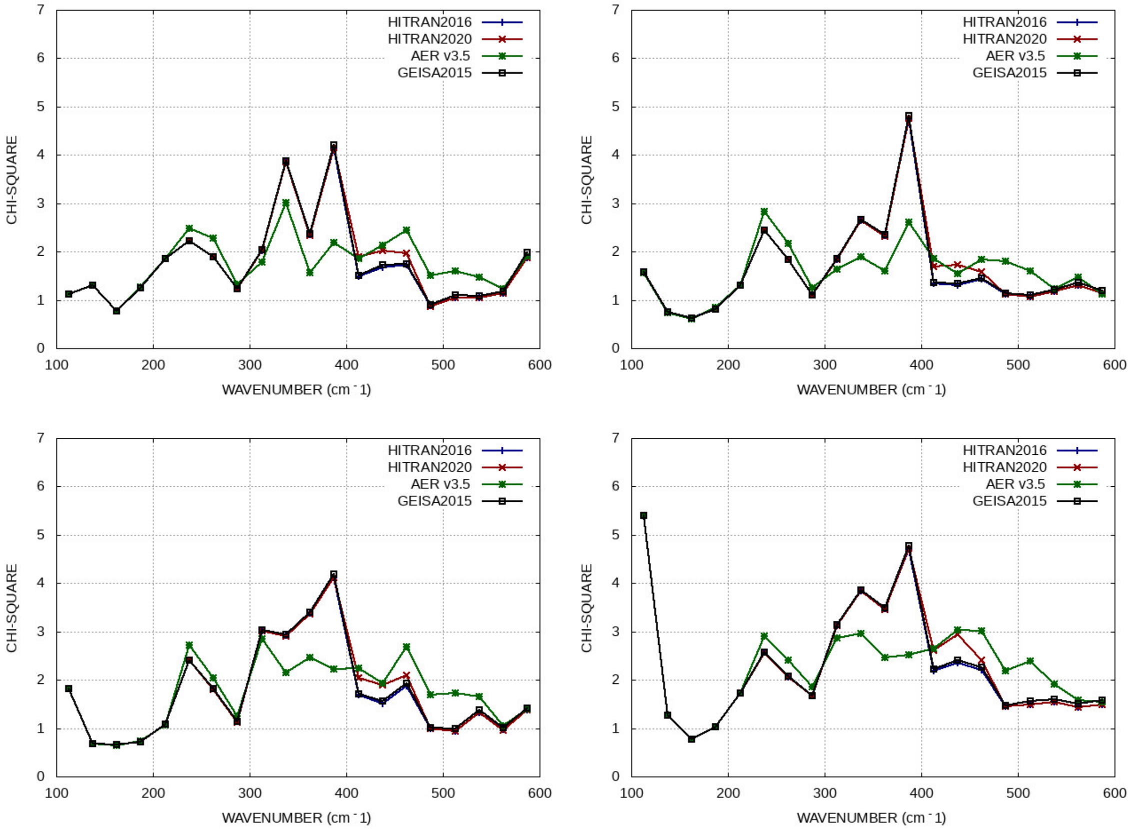


Figure6.10: Spectral chi-square calculated for the simulations with HITRAN2016 (blue line), HITRAN2020 (red line), AERv3.5 (green line) and GEISA2015 (black line), of the REFIR-PAD spectra 34 (top left panel), 35 (top right panel) and 36 (bottom left panel) using the different analysed spectroscopic databases. The mean chi-square for the REFIR-PAD spectra from 33 to 39 is reported in the bottom right panel.

the reliability of the analysis performed on the consistency between the measured spectra and the simulations. The considered systematic uncertainties and the method used to include these in the analysis are explained below:

- **Error due to uncertainties on the ERA Interim data** ⇒ the atmospheric data from the ERA-Interim database are affected by uncertainties that have been evaluated with the standard deviation of the data used to simulate the atmospheric scenarios sounded by REFIR-PAD ($H_2O(z)$, $T(z)$, O_3 and surface temperature) among the 8 grid cell points surrounding the selected atmosphere in the horizontal domain. The spectra are calculated in two different atmospheric scenarios, respectively adding and subtracting the standard deviation of each parameter to the reference scenario

- **Error due to the Forward Model approximations** ⇒ each Radiative Transfer solver uses different approximations and optimizations, being source of uncertainties in the TOA radiance calculation. The tests have been repeated with the high resolution spectra calculated with the LIDORT RTE solver.

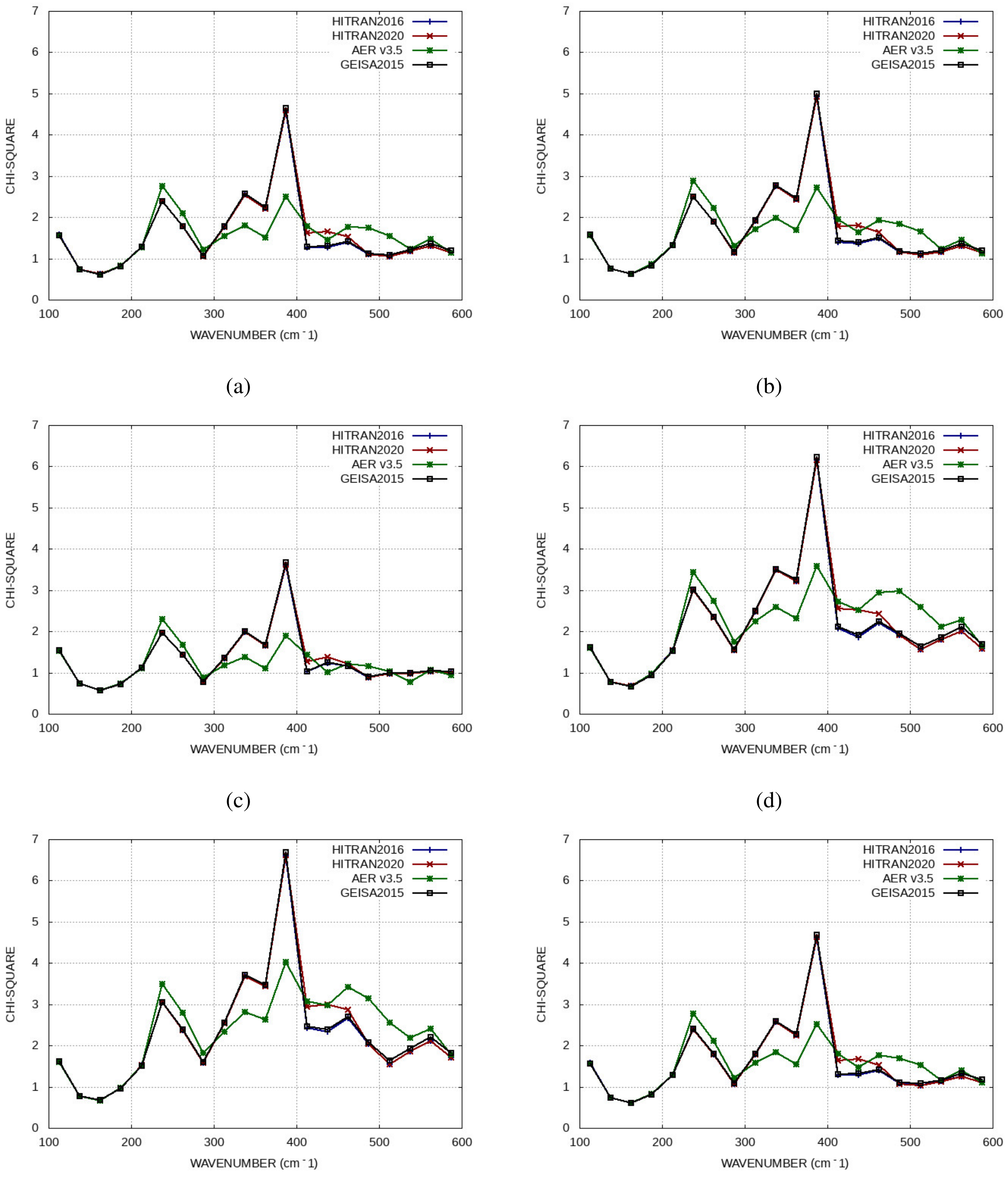
- **Error due to the uncertainty on the ILS** ⇒ the uncertainty on the ILS width is the most relevant systematic component among those relative to the ISRF knowledge on the retrieved products, as observed in the previous Chapter 5. Hence, the tests have been repeated by applying a perturbation on the MOPD **L** equal to 1%
- **Error due to the calibration accuracy of REFIR-PAD** ⇒ the tests have been repeated perturbing the REFIR-PAD measurements with the calibration error shown in Figure6.8.

The Figure6.11 shows the calculated spectral chi-square referring to the consistency between the REFIR-PAD spectrum 35 and the simulations with the different spectroscopic data, accounting the listed systematic perturbations in the analysis. Namely, panels (a) and (b) report the chi-square calculated with the 1σ standard deviation of the ERA-Interim data respectively subtracted and added to the reference atmospheric scenario. The chi-square calculated with the REFIR-PAD calibration error respectively subtracted and added to the REFIR-PAD measurement are reported in panels (c) and (d). Panel (e) shows the chi-square accounting for a perturbation on the MOPD of the instrument equal to 1% and (f) shows the chi-square square for the simulations computed by using the LIDORT model

The perturbations applied in the test reported in the panels (a) and (b) of Figure6.11 do not introduce high variations in the calculated chi-square with respect to the unperturbed case shown in the top right panel of Figure6.10. In fact, the atmospheric scenarios of the adjacent grid cell points of the used ERA-Interim reanalysis are very similar, showing a standard deviation in the water vapour VMR and atmospheric temperature of the order of 100 ppm and 0.1K respectively along most part of the atmospheric column. Moreover, the 1σ perturbation applied on the surface temperature and O_3 VMR determines negligible effects on the computed spectra, since the signal coming from the surface and the ozone absorption are almost totally masked out by the strong water vapour absorption occurring in the humid equatorial scenarios. The chi-square values computed using the LIDORT model as RTE solver are comparable to those in Figure6.10, demonstrating the agreement between the LIDORT and the Internal RTE solver of the GBB-NADIR code, generating convolved spectra with differences far below 10 nW/(cm^2 sr cm^{-1}) in the FIR range. The highest systematic contributions are introduced by the calibration error and the perturbation applied on the REFIR-PAD ILS width (subfigures (c), (d) and (e) of Figure6.11), having relevant effects on the agreement between measurement and simulations. Specifically for REFIR-PAD spectrum 35, the chi-square values are lower with respect to the unperturbed case in Figure6.10 if the calibration error is subtracted from the spectrum (test (c)), indicating a non-negligible systematic contribution introduced on the measured radiance by the calibration procedure.

However the comparison between the chi-square curves does not present relevant changes even when considering the effects of different systematic sources, since the simulations performed with AERv3.5 and the MT_CKDv3.0 water vapour continuum are still in better agreement with the observations between 300 and 400 cm^{-1}. On the contrary, the other spectroscopic databases fit more adequately the measured spectra for the remaining part of the FIR range, however the spread between the calculated chi-square is strongly reduced in test (c) of Figure6.11 for wavenumbers higher than 400 cm^{-1}, pointing out that the analysed spectroscopic databases may have similar performances in simulating real measurements between 400 and 600 cm^{-1}. Anyway, this point needs to be further investigated.

It can be concluded from the differences between the chi-squares shown in Figure6.10 and 6.11 and the differences between the computed spectra shown in Figure6.9 that the actual spectroscopic databases are not consistent between 300 and 400 cm^{-1}. In this respect, the future FORUM-EE9 mission would become a benchmark for the future releases of the water vapour spectroscopic data in the FIR by providing accurate measurements in several different scenarios over the whole globe.



(a)

(b)

(c)

(d)

(e)

(f)

Figure6.11: Spectral chi-square for the REFIR-PAD spectrum 35 calculated by subtracting and adding the 1σ standard deviation of the ERA-Interim data to the reference atmospheric scenario ((a) and (b) graphs respectively), by subtracting and adding the calibration error to the REFIR-PAD spectrum ((c) and (d) graphs respectively), by applying a perturbation equal to 1% on the MOPD (e) and calculating the TOA radiances with the LIDORT RTE Solver (f)

Chapter 7

Retrieval from TAFTS measurements

The information content of the FORUM FSI signal about different geophysical parameters, namely water vapour and surface emissivity, has been evaluated through the simulation of the retrieval in different atmospheric conditions and the calculation of the random errors due to the instrumental noise defined with the square root of the Variance Covariance Matrix of the retrieved parameters S_x (see Appendix A). The theoretical study has shown that the exploitation of the far infrared part of the spectrum leads to an information gain for the water vapour in the UTLS atmospheric region, with random errors considerably lower than the a-priori ones. In this chapter the results obtained by the analysis of the measurements collected by a prototype of the FORUM FSI instrument to obtain the water vapour vertical distribution are reported. This exercise will help in understanding if the results shown in the previous chapters are confirmed when real measurements are used.

The Tropospheric Airborne Fourier Transform Spectrometer (TAFTS) is a Martin-Puplett polarizing interferometer covering the FIR spectral region from 80 cm^{-1} to 600 $cm-1$ with a spectral resolution of 0.12 cm^{-1} (Cox et al., 2007). The instrument measured the upwelling FIR radiation on board of the FAAM-BAE (Facility for Atmospheric Airborne Measurements – British Aerospace) 146-301 ARA aircraft during the C153 flight on $03/13/2019$ above the Sea between Aberdeen (Scotland) and Newcastle (England) in the framework of the PiknMix-F field measurement campaign. TAFTS measured the outgoing radiation from 12:30 UTC to 13 UTC from an altitude of 30 kft (corresponding to 9.2 km), acquiring more than 150 spectra continuously (with intervals due to internal calibration cycles). Figure7.1 shows a single spectrum collected by TAFTS during the C153 flight and the noise and calibration error of the instrument.

TAFTS measures the outgoing spectrum in two separate spectral bands, the first "long-wave" band (80-300 cm^{-1}) uses a GeGa (Gallium doped Germanium) photoconductor detector, while the second "shortwave" band (330-600 cm^{-1}) uses a SiSb (Antimony doped Silicon) photoconductor detector (Bellisario et al., 2017), with the signal-to-noise ratio degraded between

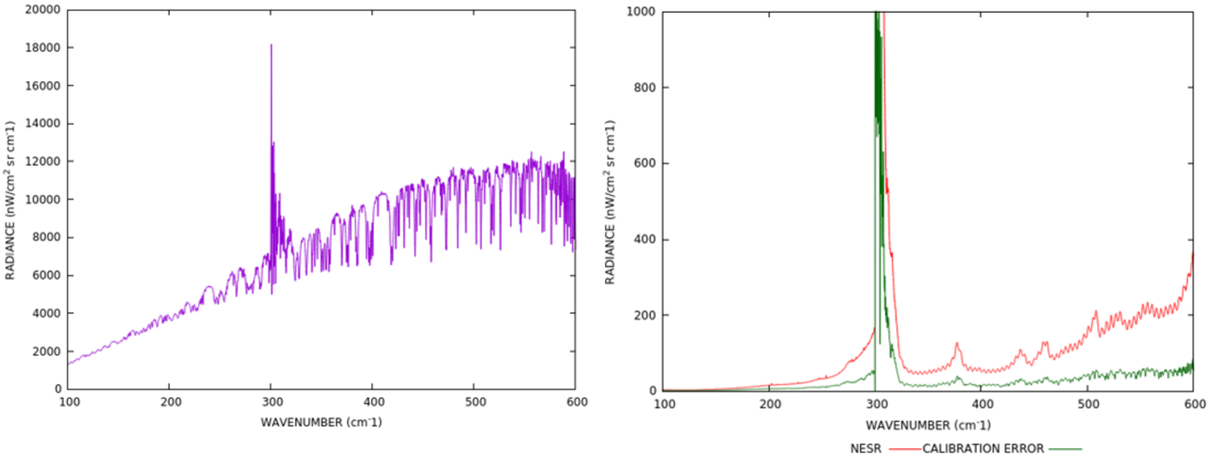


Figure7.1: Example of a single TAFTS measurement during the C153 flight (left panel) compared to the Noise and Calibration Errors(red and green lines in the right panel)

the spectral bands and close to the edges of the detectors ranges. The Signal-To-Noise Ratio is higher at lower frequencies, while it tends to degrade for wavenumbers higher than 400 cm^{-1}. During the time span in which TAFTS acquired the analysed spectra, three dropsondes (namely dropsondes 6-7-8 in the FAAM C153 flight report), measuring the temperature and water vapour profiles with an altitude grid step of about 5m, were released from the aircraft, providing very detailed atmospheric profiles that can be used as reference to assess the quality of the retrieved quantities. The comparison of the results of the retrievals of TAFTS spectra with the dropsondes measurements will enable to assess if the measured spectra contain information about the variability of the water vapour in the middle and upper troposphere. Therefore, we will show the analysis of TAFTS spectra close to the launch of the dropsondes in order to get a comparison between the estimates from the different instruments when they sound very similar atmospheric column. The dropsondes measurements occur in a time span much longer than the necessary one for the acquisition of one interferogram by TAFTS, nevertheless we could be confident that the atmospheric scenarios sounded by TAFTS and dropsondes have the same properties since the dropsondes tend to be drifted by the atmospheric flux.

7.1 Method

For the retrieval procedure, the IAC retrieval code described in Section3.6 has been exploited. The water vapour and temperature profiles are evaluated from each TAFTS spectrum on a coarse vertical grid in the lower troposphere with points fixed at surface level, 1 km, 2.5 km and 4 km, and finer up to the aircraft altitude with a vertical step of 0.5 km from 4 km up to the aircraft altitude. The a-priori atmospheric temperature - equal to the initial guess for this test - along the FAAM flight track is retrieved from the Met-Office NWP (Numerical Weather Prediction) forecast model, that provides data defined on 70 model levels from 0 to 80 km on grid points with 0.09° x 0.14° lat/lon horizontal resolution and 1h time resolution. The profiles for each scenario are determined from the 12UTC forecast run of $03/13/2019$ interpolating the predicted profiles in the grid cells closest to the location of the FAAM aircraft at the measurement time. The temperature is not a target of the retrieval tests, however it is included in the retrieval state vector in order to minimize the systematic effects originating from the uncertainties affecting the modelled profiles. Therefore, its a-priori error is fixed to a low value (0.2 K) constant with altitude, with the aim to apply only a correction to the NWP temperature profiles.

The water vapour a-priori and initial guess profiles are instead defined as much dryer profiles than the NWP ones. They have been defined by scaling the corresponding NWP profiles below 6 km by a factor equal to the altitude value in km and fixed to a constant value of 15 ppmv above 6 km. This has been done to start the retrieval iterations with a largely transparent atmosphere, since the convergence of the retrieval could be simplified when the FIR absorption lines of the water vapour at the first iteration are not saturated by the intense absorption occurring with high water vapour concentrations. Since the purpose of the test is the estimate of the water vapour VMR in the UTLS, the error on the a-priori profile has been fixed to 3% at surface level, 10% at 2 km altitude and 80% from 3 km up to the aircraft level, as low information content on the properties of the lower troposphere is expected from the inverted signal. The surface temperature and emissivity data have been taken by the NWP data and the database from Huang et al. (2016) respectively. Figure7.2 shows the comparison between the water vapour initial guess, the measured profile from the dropsonde 6 and the NWP profile referring to the location of the dropsonde 6 launch.

The water vapour from dropsondes and the NWP model, highlighted by the dotted black and green lines respectively, show high differences due to the high variability of the H_2O distribution at the fine horizontal scale measured by the dropsonde that cannot be resolved by the NWP vertical grid points.

The longwave and shortwave bands of TAFTS have both been exploited in the retrieval procedure, namely the analysed spectra range from 80 to 260 cm^{-1} and from 330 to 440 cm^{-1}. The signal from 260 to 330 cm^{-1} has been discarded from the retrieval procedure supplying no additional information because of the high NESR values. Furthermore, the wavenumbers higher than 440 cm^{-1} have been excluded, as the highest frequencies of TAFTS suffer some structural issues affecting the TAFTS optics. In fact, the incoming radiation is divided inside the interferometer by a thin film beamsplitter that is prone to the strong aircraft vibrations, so it could present different oscillation modes with amplitude of few microns introducing spurious components in the acquired interferograms. The longwave part of the signal is not affected by this problem since the signal wavelengths are much larger than the oscillation amplitude, while the spectra at higher frequencies correspond to wavelengths of close to the ones of the beamsplitter vibrations.

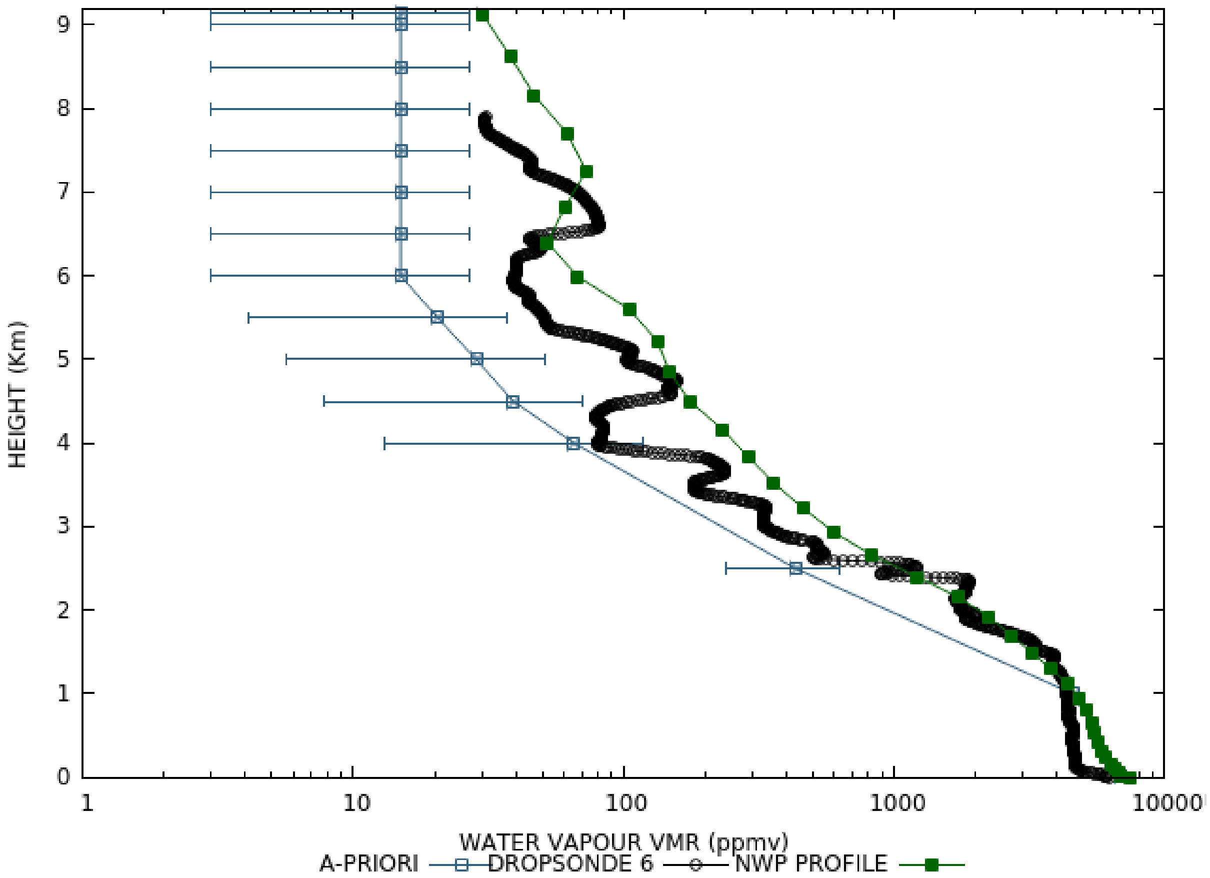


Figure7.2: Water vapour profile measured by the dropsonde 6 (black dotted line) compared with a-priori profile and the error on the a-priori (blue line) and NWP profile (green line)

7.2 Results

The Figure7.3 shows the results obtained in the analysis of three TAFTS spectra acquired close to the launch of dropsonde 6 (top left panel), 7 (top right panel) and 8 (bottom panel). In the figures the retrieved profiles (in red) are compared to the dropsondes measurements (black dotted lines), the NWP profiles for the grid cell corresponding to the dropsonde launch (green lines) and the a-priori profiles with their error bars (blue lines).

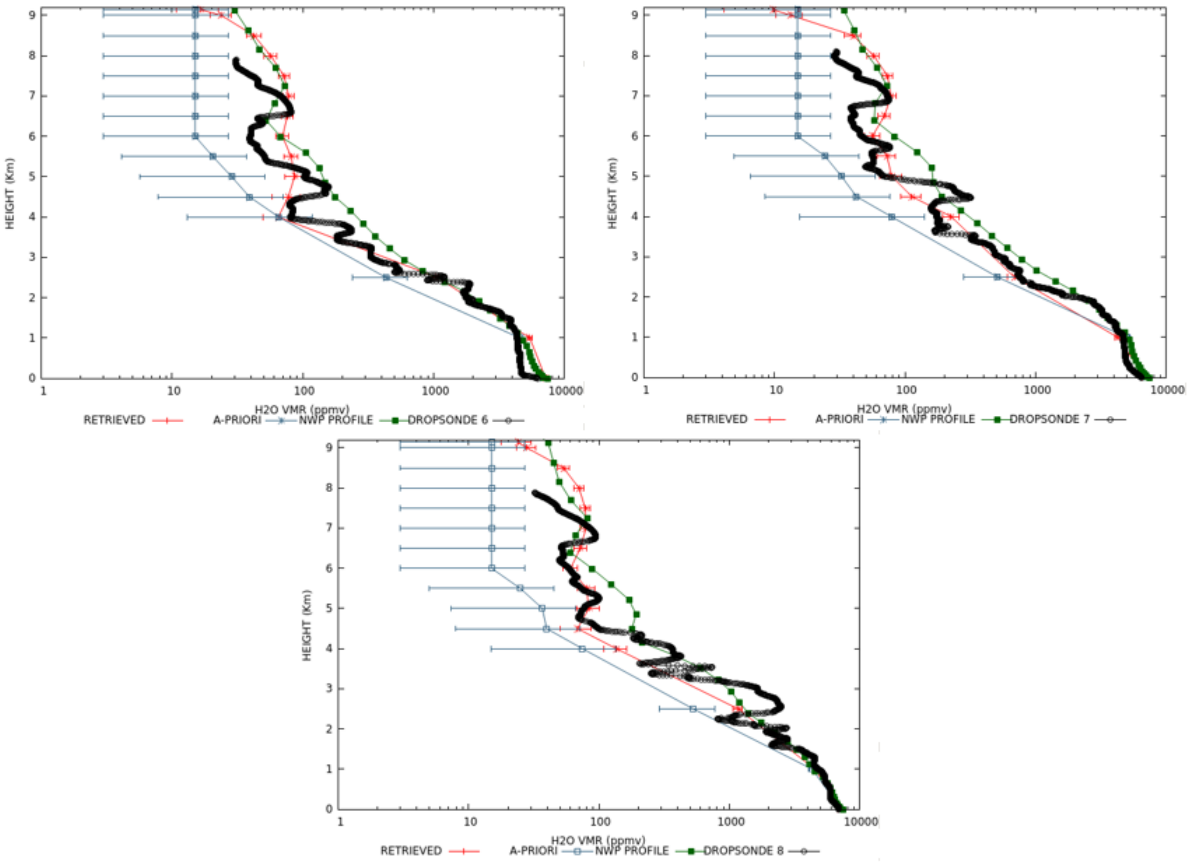


Figure7.3: Retrieved water vapour profiles (red lines) from TAFTS inverted spectra close to the launch of dropsonde 6 (top left panel), 7 (top right panel) and 8 (bottom panel) compared to the measured profiles by the dropsondes (black dotted lines) and the NWP profiles (green lines) referred to the corresponding geolocations. The a-priori profiles with their error bars for each scenario are also shown with blue lines

It can be observed that the retrieved profiles are in agreement with the dropsondes profiles, that are closer to the real state of the atmosphere than the NWP forecast, as the dropsondes can measure the atmospheric variability at a small vertical scale that is not resolved by the NWP data. However, the fine structures on the vertical observable in the dropsondes profiles are not reproduced by the retrieved profiles, despite of the fact that a fine retrieval grid has been chosen for altitudes higher than 4 km. This can be attributed to a smoothing error due to the vertical correlations of the retrieved parameters. This effect can be investigated by analysing the averaging kernel matrix and, in particular, the averaging kernel functions (see Appendix B).

The rows of the AKM - defined as the averaging kernel functions of each retrieved atmospheric parameter - are reported in the Figure7.4 for the retrieved water vapour from the analysed TAFTS spectrum close to the launch of dropsonde 7, showing the correlation between the estimated water vapour values along the vertical.

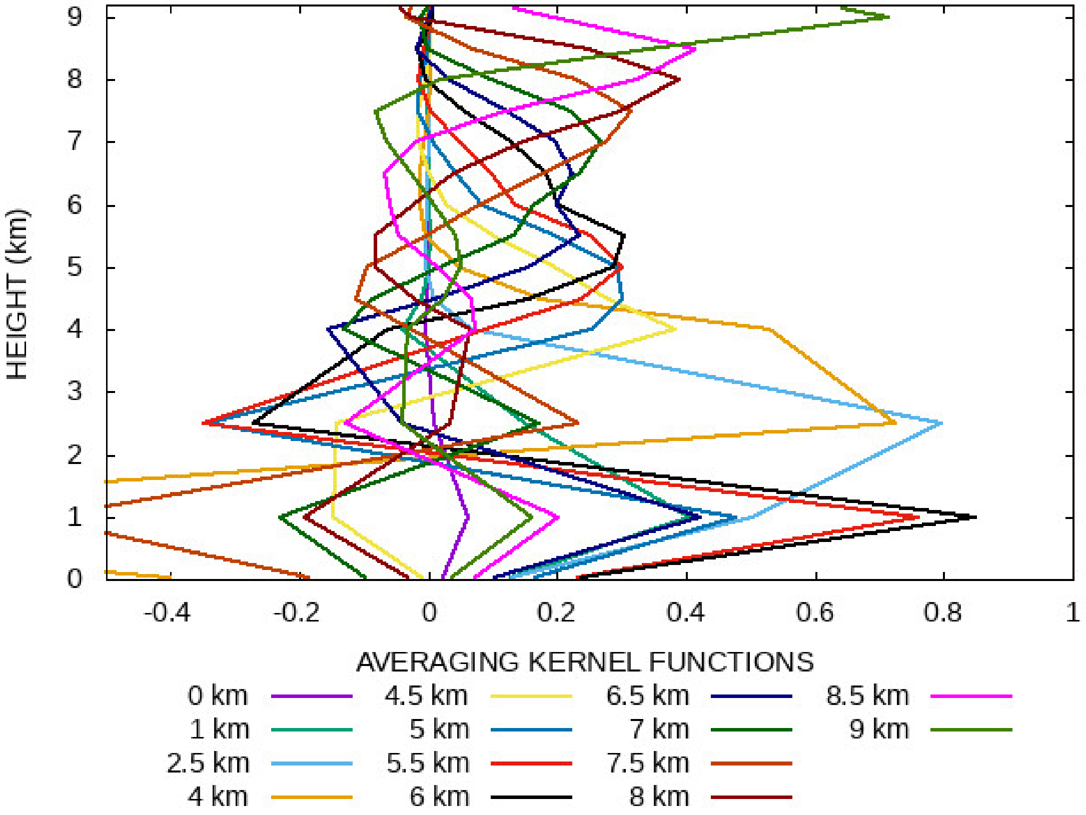


Figure7.4: Averaging kernel matrix of the retrieved water vapour profile for a single inverted TAFTS spectrum

The vertical resolution of the retrieval, given by the width of the averaging kernel functions, is larger than the separation between adjacent vertical grid points of the retrieval grid. This also indicates that the vertical resolution of the retrieval cannot be improved by adopting a finer vertical grid. This is confirmed by the number of degrees of freedom (DOF) of the retrieval (see Appendix B) for the estimated water vapour values, that is equal to 5.27 for the AKM reported in Figure7.4, value that is very similar to the ones obtained for the rest of the inverted TAFTS measurements. This number is much lower than the number of points in the used vertical grid, implying a not negligible vertical correlation between the adjacent retrieved points introducing a smoothing error on the retrieved profiles.

The convergence of the retrieval procedure can be evaluated by both visualizing the residuals between the simulated spectra and the measurements - shown for the first and the final steps of the retrieval procedures in the Figures7.5 and compared to the NESR of TAFTS - and studying the chi-square computed at the different steps of the retrieval procedure, shown in Figure7.6. We calculated a reduced chi-square dividing the analysed spectral region in small spectral intervals to check if the convergence of the simulated spectra to the observation in the different parts of the TAFTS spectral range highlights unmodelled spectral effects not allowing a correct fit of the water vapour lines on their centre only for certain regions. The spectra simulated with the retrieved profiles converge to the TAFTS measurements within the instrumental noise mainly in the longwave band and for most part of the spectral points in the shortwave band.

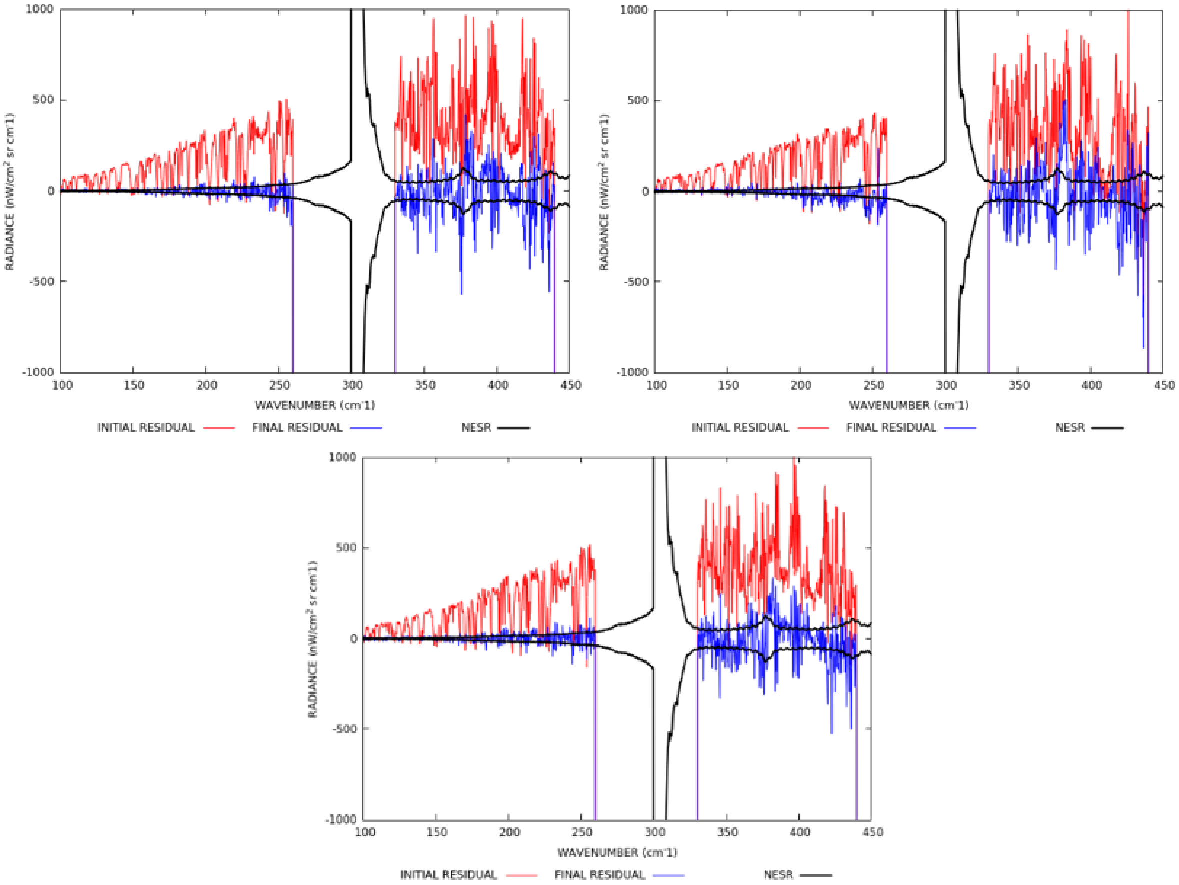


Figure7.5: Residuals between simulation and TAFTS measurements at the initial (red lines) and final (blue lines) retrieval steps for the inverted spectra close to the launch of dropsonde 6 (top left panel), 7 (top right panel) and 8 (bottom panel)

Figure 7.6 shows that the residuals are generally reduced to values within the measurement noise at the end of the retrieval procedure in the longwave band, as we can observe comparing the blue line and the black one representing the NESR of TAFTS in the Figure7.5, while some spectral points show larger differences, probably because of instrumental effects not taken into account as the presence of water vapour in the optical device and/or the uncertainty on the ISRF of TAFTS. Comparing the spectral chi-squares at the various stages of the iterative procedure we see that the spectral chi-square is monotonically reduced below 260 cm^{-1}, while the chi-square

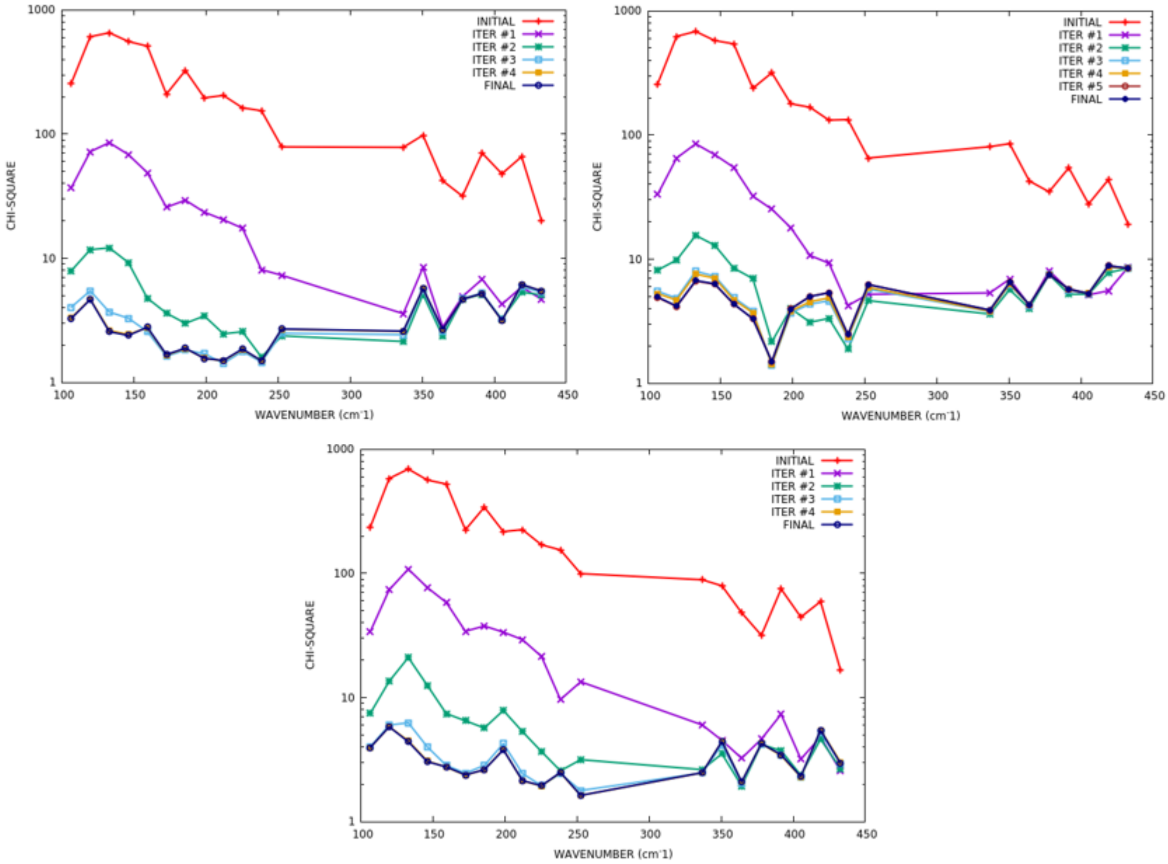


Figure7.6: Spectral chi-square at the different retrieval steps for the inverted spectra close to the launch of dropsonde 6 (top left panel), 7 (top right panel) and 8 (bottom panel)

values in the shortwave band do not decrease monotonically with the iteration steps, pointing out some problems of the forward model to fit correctly the water vapour lines. In any case, the retrieved profiles are consistent with the profiles from the dropsondes, furthermore the random errors of the retrieved parameters are generally reduced with respect to the initial a-priori errors, pointing out that some information on the humidity of the atmospheric path has been gained. This can be quantified by the information gain parameter $\mathbf{q}_j$ introduced by Dinelli et al. (2009), computed as the ratio between the **j** diagonal elements of the S_y and S_a VCMs matrices as defined in Eq.7.1. This quantifier should not be considered as an absolute quantifier, as it depends on the errors assumed for the a-priori, but can nevertheless suggest at which altitudes the retrieved values are less affected by the a-priori knowledge.

$$q_j = -0.5 \cdot log_2[\frac{S_y(j,j)}{S_a(j,j)}] \tag{7.1}$$

When there is very low information on the parameters then the random error is almost equal to the a-priori one ($q_j \sim 0$), whereas when the information is large the values of q_j become more and more large, with a value of 0.5 corresponding to a ratio between the random and a-priori error equal to $\sqrt{2}/2$. The Figure7.7 shows the information gain calculated for the retrieved water vapour profile from the inverted TAFTS measurement close to the dropsonde 7 launch

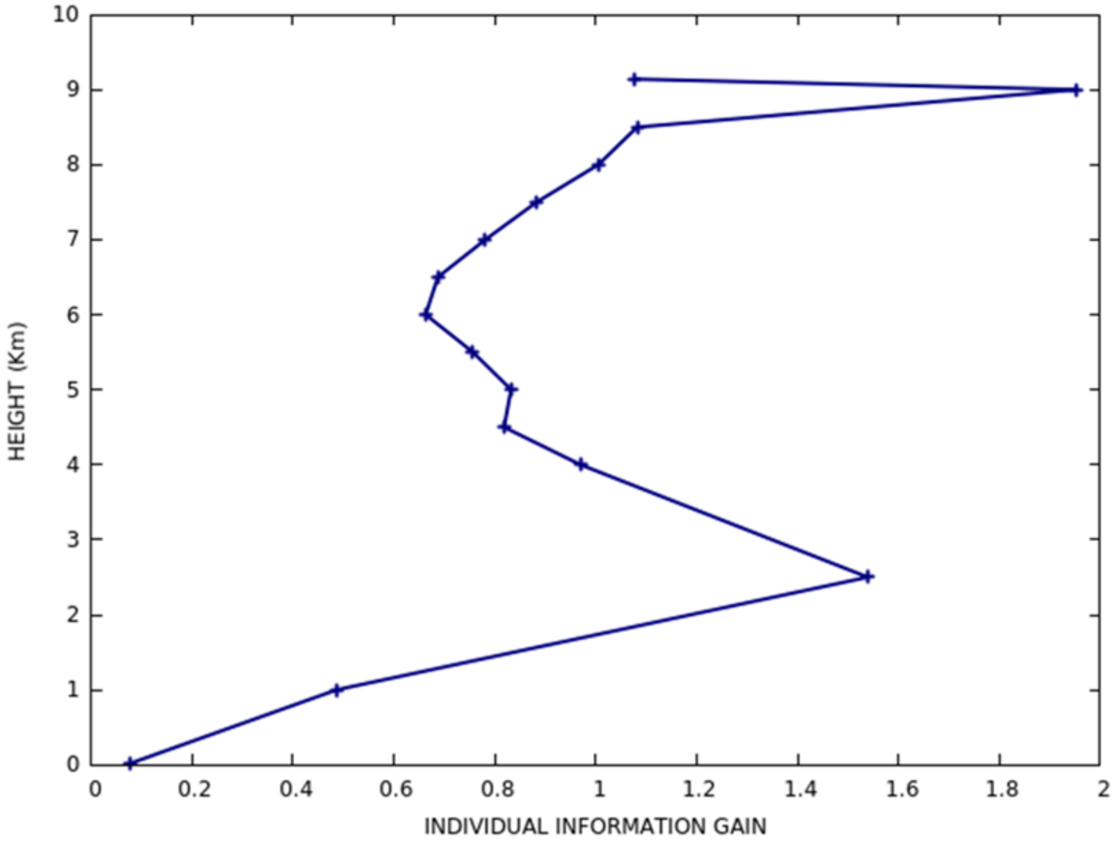


Figure7.7: Individual information gain for the water vapour VMR retrieved along the atmospheric path

It can be observed that the analysed TAFTS measurements allow to improve the water vapour estimate for altitudes higher than 1 km. The gain on the VMR knowledge is extended up to the aircraft level, however a minimum in the information gain occurs at 6 km altitude. This is due to the lack of the spectral information contained in the wavenumbers range discarded from the retrieval procedure, mainly the spectral points of the band around 300 cm^{-1}. This is confirmed by the analysis reported in Section5.1, for which the comparison between the average random errors on water vapour have been calculated in different latitudinal bands with different FORUM FSI noise types, clearly showing higher accuracy on the estimated water vapour provided by the noise with better performances below 400 cm^{-1}. In a similar way, the lack of spectral information from 260 to 330 cm^{-1} produces a minimum in the information content at 6 km, however it tends to increase up to the aircraft level due to the high signal-to-noise ratio of TAFTS in the longwave band and to the higher sensitivity of the acquired signal to the water vapour VMR above 6 km as well, as shown by the Jacobians with respect to water vapour reported in Figure4.3

Chapter 8

Conclusions

The work performed in this thesis has as main topic the assessment of the performances expected for the FORUM instrument, that will be launched in 2026. The studies reported in this thesis have played an important role in making the FORUM proposal successful in the mission selection process conducted by ESA.

The main results obtained from the analysis reported in this thesis are summarised in this section. The investigation on the sensitivity of the TOA FIR radiation and the FORUM signal to the water vapour in the UTLS and the surface emissivity, performed in Chapter 4 using both the study of the derivatives of the spectrum with respect to these parameters and the theoretical retrieval study, show that the information gain due to the contribution of the FIR is very large. Particularly, the FORUM measurements can provide accurate estimates of the water vapour at the tropopause level, mainly in the tropical scenarios, while the FIR surface emissivity can be accurately determined from space only in polar scenarios, as the higher transmittance of the atmosphere due to reduced water vapour content allows the radiation emitted from surface to escape to space. While this effect makes the FIR surface emissivity a strong driver of the ERB, it also enables the FORUM mission to get informations on it. This is of great importance, since we know that the polar contribution to the Earth energy budget can significantly change depending on the ice pack extension and snow coverage. Despite its importance, climate models rarely take the spectral variation of the polar surface emissivity into account. FORUM measurements will therefore help in an improvement of the climate models providing for the first time a direct measurement of the surface effects of the poles on the OLR.

The study of the errors obtained with different instrumental noise models confirms that the information content on the water vapour in the UTLS is strongly related to the accuracy of the measurements in the FIR, as lower errors on the retrieved H_2O above 5 km can be obtained with lower instrumental errors for wavenumbers below 500 cm^{-1}. However, the accurate estimates of the water vapour in the UTLS and the polar surface emissivity provided by FORUM will be affected by several systematic error sources. They have been identified and their contribution on the main geophysical variables has been evaluated in chapter 5. The highest systematic contributions on the water vapour and spectral emissivity in the FIR are introduced by the uncertainty affecting the water vapour spectroscopic data in the FIR and the water vapour continuum model. Particularly, the continuum model can introduce systematic errors in the estimates of the water vapour VMR of the order of the random error that is higher at the Tropics due to the higher water vapour content and absorption. For the surface emissivity as well, the water vapour continuum parametrization can introduce errors up to 0.03 between 200 and 300 cm^{-1}, while the uncertainties on the line parameters of the FIR water vapour transitions can contribute with errors up to 0.01 on the FIR emissivity. Moreover, the undetected clouds occupying the FORUM FSI Field Of View may introduce huge systematic errors on the retrieved products in particular cases where the signal emitted and scattered by the cloud is similar to that coming from the underlying surface. The other analysed systematic error sources do not introduce relevant systematic errors on the analysed Level 2 parameters. The only exception is the errors introduced by the high variability of the nitric acid VMR in the wintertime polar vortex, leading to systematic uncertainties on the retrieved emissivity at 400 cm^{-1} even up to 0.01. This can be mitigated introducing the HNO_3 VMR into the state vector. The uncertainties on the FORUM ISRF introduce overall low systematic errors; however the uncertainty on the ISRF width could produce systematic errors on water vapour and atmospheric temperature in the UTLS up to 1% and 0.1K respectively, while the errors on surface emissivity due to the imperfect knowledge of the ISRF are negligible.

Given the results on the errors affecting water vapour and FIR surface emissivity (both random and systematic), we can state that the estimates of these two variables from the FORUM measurements will satisfy the requirements on the uncertainty listed in Table1.1. However, the uncertainty on the water vapour continuum provides on both water vapour VMR and surface emissivity very high systematic errors, that could be reduced by exploiting the FORUM measurements to improve the knowledge of the water vapour spectroscopic data in the FIR.

Since the uncertainties on the FIR spectroscopic data of the water vapour represent the largest systematic error source for the water vapour and surface emissivity, some of the most updated spectroscopic database have been inter-compared (Chapter 6), then the different databases have been used to simulate the REFIR-PAD measurements above Teresina (Brazil) with the GBB-NADIR code. The simulations with the different databases in the FIR range show differences of the order of the instrumental noise, with the largest discrepancies occurring in the spectral range between 300 and 400 cm^{-1} that contains a significant part of the information on the water vapour concentration at the tropopause level. Actually the AERv3.5 database provides the best fit to the observations in this spectral range. This test could be repeated in future using FIR measurements in different atmospheric scenarios, like - for instance - the REFIR-PAD ground-based measurements of the downwelling radiation on the Antarctic Plateau and the spectra acquired by other prototypes of the FORUM FSI as FIRMOS (Far-Infrared Radiation

Mobile Observation System), during the field measurement campaign at Zugspitze (Germany) in November 2018.

Finally, the retrieval of the water vapour concentration has been performed exploiting the TAFTS measurements from the FAAM-Bae aircraft during the PiknMix-F campaign in March 2019. Even though the information content of the acquired signal is limited by technical issues due to the edges of the detectors and the aircraft vibrations, the retrieved profiles show good agreement with the water vapour profiles measured by the dropsondes launched from the aircraft above 2 km altitude. However, the fine structures on the vertical scale cannot be reproduced even though a fine vertical grid has been chosen for the performed retrieval test, as the degrees of freedom of the signal are not large enough to provide profiles affected by low smoothing error. In future work, the water vapour profiles along the flight track of the FAAM aircraft in the C153 flight could be retrieved using the joint measurements of TAFTS and the ARIES (Airborne Research Interferometer Evaluation System) interferometer also on board of the FAAM aircraft measuring the outgoing radiation from 550 to 3000 cm^{-1}. This would allow to study the retrieval performance for observations acquired in a similar way of the future measurements that will be provided by the FORUM and IASI-NG instruments. Furthermore, the retrieval exercises could be repeated when improved spectroscopic data will be available to confirm if the improvement are going in the right direction.

Appendix A: The Optimal Estimation Method

The geophysical variables estimate from the inversion of remote sensing measurements is an ill-posed problem since it can generally be satisfied by many configurations of the state vector. A mostly used expedient to reduce the instabilities of the inverse problem is to include some information on the a-priori state of the measured scenario. The Optimal Estimation Method (Rodgers, 2000) is a Bayesan approach for the solution of the inverse model, that enables to weight the information from the observations with the a-priori knowledge of the quantities to be estimated. From the Bayesian theory of probability, the measurement **y** and the state vector **x** are represented by two probability functions P(**x**) and P(**y**), determining the prior knowledge of the problem. The inverse problem consists in finding the conditional probability of the state vector **x** if the result of the measurement is **y**, that is P(**x**|**y**). This conditional probability is linked to the conditional probability P(**y**|**x**) by the Bayes theorem

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{8.1}$$

the measurement and the state vector are linked by the relation **y=f(x)+e**, where **f** is the direct problem mapping **x** onto **y** and **e** is the measurement error. Assuming that the error **e** is gaussian and assuming that its covariance is given by $\mathbf{S}_y$, the conditional probability P(y|x) can be written as

$$P(y|x) = c_n e^{-(y-f(x))^T S_y^{-1}(y-f(x))/2} \tag{8.2}$$

where c_n is a normalization constant. The a-priori probability, linked to the a-priori information on the state vector **x** can be written as

$$P(y|x) = c_m e^{-(x-x_a)^T S_a^{-1}(x-x_a)/2} \tag{8.3}$$

where c_m is a normalization constant and S_a is the covariance matrix of the a-priori state. P(y) could be calculated from P(y) = $\int$P(y|x)P(x)d^{n}x, but it is not needed since it does not depend on **x** and is just a normalization factor. Hence, the expression for P(**x**|**y**) is

$$P(y|x) = Ce^{-[(y-f(x))^T S_y^{-1}(y-f(x))+(x-x_a)^T S_a^{-1}(x-x_a)]}/2 \tag{8.4}$$

The usual way to determine the best estimate of the state vector as solution of the inverse problem is to assume the value **x** for which P(**x**|**y**) is maximum, known as the Maximum Likelihood Method or Maximum a Posteriori Solution. Hence the cost function χ^2 has to be minimised

$$\chi^2 = (x-x_a)^T S_a^{-1}(x-x_a) + (y-f(x))^T S_y^{-1}(y-f(x)) \tag{8.5}$$

Thus, the points for which $\nabla\chi^2$ is zero have to be found

$$g(x) = \nabla\chi^2 = -2K^T S_y^{-1}(y-f(x)) - 2S_a^{-1}(x-x_a) = 0 \tag{8.6}$$

where **K** is the Jacobian matrix of the direct model **f** with respect to each of the parameters. The equation $\nabla\chi^2$=0 can be solved with an iterative procedure. The zero value closest to the point x_1 of a scalar function g(x) can be determined with the Newton algorithm, whose iterative procedure at the point x_i:

- calculate the tangent at x_i
- follow the tangent till it intersects the x axis at the new point x_{i+1}

Generalized to the case where g is a vectorial quantity depending on the vector **x**, the Newton's method can be written

$$x_{i+1} = x_i - [\nabla g(x_i)]^{-1} g(x_i) \tag{8.7}$$

the gradient of g($\mathbf{x}_i$) can be determined deriving the Eq.8.6 on both sides, obtaining

$$\nabla g(x_i) = 2K^T S_y^{-1} K - 2[\nabla K^T] S_y^{-1}(y-f(x)) + 2S_a^{-1} \tag{8.8}$$

The second term in the expression above is related to the second derivatives of the function **f**. The calculation of this term requires a significant effort in terms of computational time. The approximation known as Gauss-Newton method consists in neglecting this term, which is often small. Thus, the single iterative step is

$$x_{i+1} = x_i + ((K^T S_y^{-1} K) + S_a^{-1})^{-1} [K^T S_y^{-1} n + S_a^{-1}(x_0 - x_i)] \tag{8.9}$$

where **n=y-f(x)** is the residual between observation and Forward Model simulation. The error on the retrieved parameters x_i can be expressed in form of the Variance-Covariance Matrix S_x, that is:

$$S_x = ((K^T S_y^{-1} K) + S_a^{-1})^{-1} \tag{8.10}$$

namely, the errors affecting the retrieved parameters are equal to the square root of the diagonal elements of the S_x matrix.

The main problem with the Gauss-Newton method is that it may fail convergence due to too large steps. The idea introduced by the Levemberg-Marquardt optimization was to reduce the single steps by adding a new term into Eq8.9

$$x_{i+1} = x_i + ((K^T S_y^{-1} K) + S_a^{-1} + \lambda_{LM} D)^{-1} [K^T S_y^{-1} n + S_a^{-1}(x_0 - x_i)] \tag{8.11}$$

where D is a diagonal matrix including the diagonal elements of $\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K}$, λ_{LM} is a positive parameter known as dumping factor. The strategy used with the Levemberg-Marquardt algorithm is to start from a small λ_{LM} ($\lambda_{LM} = 0.1$ for instance) and then scale it if the χ^2 is reduced in the following steps.

The inversion procedure stops when the difference between the χ^2 between two consecutive steps of the procedure is below a fixed threshold, indicating that the minimization procedure has been performed successfully. The agreement between the simulation performed with the state vector retrieved at the end of the inversion procedure and the observation can be expressed by the reduced chi-square parameter, defined as

$$\chi' = \frac{1}{N} \sum_{i=1}^{N} \frac{(y_i - f(x_i))^2}{\sigma_i^2} \tag{8.12}$$

where N is the number of the observations and σ_i is the measurement noise at the ith fitted spectral point.

Appendix B: The Averaging Kernel Matrix

A test of the sensitivity of the retrieval to the atmospheric parameters can be performed through the use of the Averaging Kernel Matrix (AKM, Rodgers (2000)). The Averaging Kernel provides a measure of the sensitivity of the retrieved state vector x to the true atmospheric state $x^{'}$, with A(i,j) = $\delta x_i / \delta x^{'}_j$, where**A** is the averaging kernel matrix, x_i is the i^{th} component of the final retrieved state x and $x^{'}_j$ is the j^{th} component of the true atmospheric state. The expression for the Averaging Kernel Matrix is:

$$A = \frac{dx}{dx^{'}} = ((K^T S_y^{-1} K) + S_a^{-1})^{-1} K^T S_y^{-1} K^{'} \tag{8.13}$$

where $K^{'}$ is the Jacobian matrix calculated with respect the variables used to represent the true atmospheric state $x^{'}$. In the real inverse problem the retrieval of the different atmospheric parameters produce not independent results, that are highlighted by the out of diagonal correlation terms in the AKM. In particular, the correlation of the i^{th} retrieved parameter in the state vector with the true atmospheric state along the vertical can be studied by analysing the corresponding averaging kernel function, defined as the i^{th} row of the AKM.

The width of the averaging kernel functions is a measure of the vertical resolution of the retrieval in the different part of the atmospheric column. A useful scalar metrics of the sensitivity of the retrieval is the number of degrees of freedom of the retrieval, defined as the trace of matrix **A**

$$nDOF = \sum_j A(j,j) \tag{8.14}$$

that corresponds to the number of uncorrelated parameters that can be estimated from the inversion procedure. For a well constrained problem the number of degrees of freedom is close to the number of retrieved parameters.

Bibliography

S.A. Ackerman, W.L. Smith, A.D. Collard, X.L. Ma, H.E. Revercomb, and R.O. Knuteson. Cirrus cloud properties derived from high spectral resolution infrared spectrometry during FIRE II. Part II: Aircraft HIS results. *Journal of the atmospheric sciences*, 52(23):4246–4263, 1995.

A.C. Aikin, J.R. Herman, E.J. Maier, and C.J. McQuillan. Atmospheric chemistry of ethane and ethylene. *Journal of Geophysical Research: Oceans*, 87(C4):3105–3118, 1982. doi: 10.1029/JC087iC04p03105.

A.J. Baran. The impact of cirrus microphysical and macrophysical properties on upwelling far-infrared spectra. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133(627): 1425–1437, 2007.

A.J. Baran. A review of the light scattering properties of cirrus. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(14-16):1239–1260, 2009.

C. Bellisario, H.E. Brindley, J.E. Murray, A.E. Last, J. Pickering, R.C. Harlow, S. Fox, C. Fox, S.M. Newman, M. Smith, D. Anderson, X. Huang, and X. Chen. Retrievals of the far infrared surface emissivity over the Greenland Plateau using the Tropospheric Airborne Fourier Transform Spectrometer (TAFTS). *Journal of Geophysical Research: Atmospheres*, 122(22):12–152, 2017.

G. Bianchini and L. Palchetti. Technical note: REFIR-PAD level 1 data analysis and performance characterization. *Atmospheric Chemistry and Physics*, 8(14):3817–3826, 2008. doi: 10.5194/acp-8-3817-2008.

S. Del Bianco, B. Carli, M. Gai, L.M. Laurenza, and U. Cortesi. XCO2 retrieved from IASI using KLIMA algorithm. *Annals of Geophysics*, 56, 2014.

G. Brasseur, E. Arijs, A. De Rudder, D. Nevejans, and J. Ingels. Acetonitrile in the atmosphere. *Geophysical Research Letters*, 10(8):725–728, 1983. doi: 10.1029/GL010i008p00725.

H.E. Brindley and J.E. Harries. The impact of far i.r. absorption on clear sky greenhouse forcing: sensitivity studies at high spectral resolution. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 60(2):151 – 180, 1998. ISSN 0022-4073. doi: https://doi.org/10.1016/S0022-4073(97)00152-0. URL `http://www.sciencedirect.com/science/article/pii/S0022407397001520`.

K.E. Cady-Pereira, M.W. Shephard, D.B. Millet, M. Luo, K.C. Wells, Y. Xiao, V.H. Payne, and J. Worden. Methanol from TES global observations: retrieval algorithm and seasonal and spatial variability. *Atmospheric Chemistry and Physics*, 12(17):8189–8203, 2012. doi: 10.5194/acp-12-8189-2012. URL `https://www.atmos-chem-phys.net/12/8189/2012/`.

B. Carli and M. Carlotti. *Far-infrared and microwave spectroscopy of the Earth's atmosphere*, pages 1–95. Academic Press, 1992.

S. Chandrasekhar. *Radiative Transfer*. Dover, New York, 1960.

F. Chevallier, S. Di Michele, and A.P. McNally. Diverse profile datasets from the ECMWF 91-level short-range forecasts. *European Centre for Medium-Range Weather Forecasts, NWP SAF Report No. NWPSAF-EC-TR-010*, 2006.

S.A. Clough and M.J. Iacono. Line-by-line calculation of atmospheric fluxes and cooling rates: 2. Application to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons. *Journal of Geophysical Research: Atmospheres*, 100(D8):16519–16535, 1995. doi: 10.1029/95JD01386. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JD01386`.

S.A. Clough, F.X. Kneizys, and R.W. Davies. Line shape and the water vapor continuum. *Atmospheric research*, 23(3-4):229–241, 1989.

S.A. Clough, M.W. Shephard, E.J. Mlawer, J.S. Delamere, M.J. Iacono, K. Cady-Pereira, S. Boukabara, and P.D. Brown. Atmospheric radiative transfer modeling: a summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91(2):233–244, 2005.

W.D. Collins and M.G. Mlynczak. Prospects for measurement of far infrared tropospheric spectra: Implications for climate modeling. *AGUFM*, 2001:GC32A–0210, 2001.

C.V. Cox, J.E. Murray, J.P. Taylor, P.D. Green, J.C. Pickering, J.E. Harries, and A.E. Last. Clear-sky far-infrared measurements observed with TAFTS during the EAQUATE campaign, September 2004. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133(S3):273–283, 2007.

C.V. Cox, J.E. Harries, J.P. Taylor, P.D. Green, A.J. Baran, J.C. Pickering, A.E. Last, and J.E. Murray. Measurement and simulation of mid-and far-infrared spectra in the presence of cirrus. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 136(648):718–739, 2010.

D.P. Dee, S.M. Uppala, A.J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M.A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S.B. Healy, H. Hersbach, E.V. Holm, L. Isaksen, P. Kallberg, M. K´ohler, M. Matricardi, A.P. McNally, B.M. Monge-Sanz, J.J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thepaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656): 553–597, 2011. doi: 10.1002/qj.828. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.828`.

J.S. Delamere, S.A. Clough, V.H. Payne, E.J. Mlawer, D.D. Turner, and R.R. Gamache. A far-infrared radiative closure study in the Arctic: Application to water vapor. *Journal of Geophysical Research: Atmospheres*, 115(D17), 2010. doi: 10.1029/2009JD012968. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012968`.

B.M. Dinelli, E. Castelli, B. Carli, S. Del Bianco, M. Gai, L. Santurri, B.P. Moyna, M. Oldfield, R. Siddans, D. Gerber, W.J. Reburn, B.J. Kerridge, and C. Keim. Measurement of the tropical UTLS composition in presence of clouds using millimetre-wave heterodyne spectroscopy. *Atmospheric Chemistry and Physics*, 9(4):1191–1207, 2009.

A. Dudhia, V.L. Jay, and C.D. Rodgers. Microwindow selection for high-spectral-resolution sounders. *Applied Optics*, 41(18):3665–3673, 2002.

K.F. Evans. The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. *Journal of the Atmospheric Sciences*, 55(3):429–446, 1998.

D.R. Feldman, W.D. Collins, R. Pincus, X. Huang, and X. Chen. Far-infrared surface emissivity and climate. *Proceedings of the National Academy of Sciences of the United States of America*, 111(46):16297–16302, 2014. doi: doi:10.1073/pnas.1413640111. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4246346/`.

R.R. Gamache, C. Roller, E. Lopes, I.E. Gordon, L.S. Rothman, O.L. Polyansky, N.F. Zobov, A.A. Kyuberis, J. Tennyson, S.N. Yurchenko, A.G. Csaszar, T. Furtenbacher, X. Huang, D.W.Schwenke, T.J. Lee, B.J. Drouin, S.A. Tashkun, V. I. Perevalov, and R.V. Kochanov. Total internal partition sums for 166 isotopologues of 51 molecules important in planetary atmospheres: Application to HITRAN2016 and beyond. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 203:70–87, 2017.

I.E. Gordon, L.S. Rothman, R.R. Gamache, D. Jacquemart, C. Boone, P.F. Bernath, M.W. Shephard, J.S. Delamere, and S.A. Clough. Current updates of the water-vapor line list in HITRAN: A new "Diet" for air-broadened half-widths. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 108(3):389–402, 2007.

I.E. Gordon, L.S. Rothman, C. Hill, R.V. Kochanov, Y. Tan, P.F. Bernath, M. Birk, V. Boudon, A. Campargue, K.V. Chance, B.J. Drouin, J.-M. Flaud, R.R. Gamache, J.T. Hodges, D. Jacquemart, V.I. Perevalov, A. Perrin, K.P. Shine, M.-A.H. Smith, J. Tennyson, G.C. Toon, H. Tran, V.G. Tyuterev, A. Barbe, A.G. Csaszar, V.M. Devi, T. Furtenbacher, J.J. Harrison, J.-M. Hartmann, A. Jolly, T.J. Johnson, T. Karman, I. Kleiner, A.A. Kyuberis, J. Loos, O.M. Lyulin, S.T. Massie, S.N. Mikhailenko, N. Moazzen-Ahmadi, H.S.P. Ḿüller, O.V. Naumenko, A.V. Nikitin, O.L. Polyansky, M. Rey, M. Rotger, S.W. Sharpe, K. Sung, E. Starikova, S.A. Tashkun, J. Vander Auwera, G. Wagner, J. Wilzewski, P. Wcislo, S. Yu, and E.J. Zak. The HITRAN2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 203:3 – 69, 2017. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2017.06.038. URL `http://www.sciencedirect.com/science/article/pii/S0022407317301073`. HITRAN2016 Special Issue.

J. Harries, B. Carli, R. Rizzi, C. Serio, M. Mlynczak, L. Palchetti, T. Maestri, H. Brindley, and G. Masiello. The far-infrared earth. *Reviews of Geophysics*, 46(4), 2008.

X. Huang, X. Chen, D.K. Zhou, , and X. Liu. An Observationally Based Global Band-by-Band Surface Emissivity Dataset for Climate and Weather Simulations. *Journal of the Atmospheric Sciences*, 73(9):3541–3555, 2016. doi: 10.1175/JAS-D-15-0355.1. URL `https://doi.org/10.1175/JAS-D-15-0355.1`.

Y. Huang and V. Ramaswamy. Evolution and trend of the outgoing longwave radiation spectrum. *Journal of climate*, 22(17):4637–4651, 2009.

J. Humlíček. Optimized computation of the voigt and complex probability functions. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 27(4):437–444, 1982.

T. Inoue. On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bi-spectral measurements in the 10μm window region. *Journal of the Meteorological Society of Japan. Ser. II*, 63(1):88–99, 1985.

N. Jacquinet-Husson, R. Armante, N.A. Scott, A. Chedin, L. Crepeau, C. Boutammine, A. Bouhdaoui, C. Crevoisier, V. Capelle, C. Boonne, N. Poulet-Crovisier, A. Barbe,

D. Chris Benner, V. Boudon, L.R. Brown, J. Buldyreva, A. Campargue, L.H. Coudert, V.M. Devi, M.J. Down, B.J. Drouin, A. Fayt, C. Fittschen, J.-M. Flaud, R.R. Gamache, J.J. Harrison, C. Hill, O. Hodnebrog, S.-M. Hu, D. Jacquemart, A. Jolly, E. Jimenez, N.N. Lavrentieva, A.-W. Liu, L. Lodi, O.M. Lyulin, S.T. Massie, S. Mikhailenko, H.S.P. Müller, O.V. Naumenko, A. Nikitin, C.J. Nielsen, J. Orphal, V.I. Perevalov, A. Perrin, E. Polovtseva, A. Predoi-Cross, M. Rotger, A.A. Ruth, S.S. Yu, K. Sung, S.A. Tashkun, J. Tennyson, Vl.G. Tyuterev, J. Vander Auwera, B.A. Voronin, and A. Makie. The 2015 edition of the GEISA spectroscopic database. *Journal of Molecular Spectroscopy*, 327: 31 – 72, 2016. ISSN 0022-2852. doi: https://doi.org/10.1016/j.jms.2016.06.007. URL `http://www.sciencedirect.com/science/article/pii/S0022285216301011`. New Visions of Spectroscopic Databases, Volume II.

D. Johnson, W. Traub, K. Chance, and K.W. Jucks. Detection of HBr and upper limit for HOBr: Bromine partitioning in the stratosphere. *Geophysical Research Letters - GEOPHYS RES LETT*, 22:1373–1376, 06 1995. doi: 10.1029/95GL01349.

J.T. Kiehl and K.E. Trenberth. Earth's Annual Global Mean Energy Budget. *Bulletin of the American Meteorological Society*, 78(2):197–208, 02 1997. ISSN 0003-0007. doi: 10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2.

D. Lubin, B. Chen, D.H. Bromwich, R.C.J. Somerville, W.H. Lee, and K.M. Hines. The impact of Antarctic cloud radiative properties on a GCM climate simulation. *Journal of climate*, 11 (3):447–462, 1998.

M. Matricardi. The generation of RTTOV regression coefficients for IASI and AIRS using a new profile training set and a new line-by-line database. *Technical Memorandum, European Centre for Medium-Range Weather Forecasts*, 2008.

R.A. McClatchey, W.S. Benedict, S.A. Clough, D.E. Burch, and R.F. Calfee. AFCRL atmospheric absorption line parameters compilation. *Air Force Cambridge Research Labs HANSCOM AFB MA*, 1973.

M.B. McElroy, S.C. Wofsy, and N.D. Sze. Photochemical sources for atmospheric H2S. *Atmospheric Environment (1967)*, 14(2):159 – 163, 1980. ISSN 0004-6981. doi: https://doi.org/10.1016/0004-6981(80)90274-7. URL `http://www.sciencedirect.com/science/article/pii/0004698180902747`.

D. B. Millet, M. Baasandorj, D. K. Farmer, J. A. Thornton, K. Baumann, P. Brophy, S. Chaliyakunnel, J. A. de Gouw, M. Graus, L. Hu, A. Koss, B. H. Lee, F. D. Lopez-Hilfiker, J. A. Neuman, F. Paulot, J. Peischl, I. B. Pollack, T. B. Ryerson, C. Warneke, B. J.

Williams, and J. Xu. A large and ubiquitous source of atmospheric formic acid. *Atmospheric Chemistry & Physics*, 15(11):6283–6304, Jun 2015. doi: 10.5194/acp-15-6283-2015.

M.I. Mishchenko, L.D. Travis, and A.A. Lacis. *Scattering, absorption, and emission of light by small particles*. Cambridge university press, 2002.

E.J. Mlawer, V.H. Payne, J.L. Moncet, J.S. Delamere, M.J. Alvarado, and D.C. Tobin. Development and recent evaluation of the MT_CKD model of continuum absorption. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370 (1968):2520–2556, 2012.

E.J. Mlawer, D.D. Turner, S.N. Paine, L. Palchetti, G. Bianchini, V.H. Payne, K.E. Cady-Pereira, R.L. Pernak, M.J. Alvarado, D. Gombos, J.S. Delamere, M.G. Mlynczak, and J.C. Mast. Analysis of water vapor absorption in the far-infrared and submillimeter regions using surface radiometric measurements from extremely dry locations. *Journal of Geophysical Research: Atmospheres*, 124(14):8134–8160, 2019.

R.H. Norton and R. Beer. New apodizing functions for Fourier spectrometry. *JOSA*, 66(3): 259–264, 1976.

L. Palchetti, G. Bianchini, G. Di Natale, and M. Del Guasta. Far infrared radiative properties of water vapor and clouds in Antarctica. *B. Am. Meteorol. Soc., 96, 1505–1518*, 96, 2015.

F. Parol, J.C. Buriez, G. Brogniez, and Y. Fouquart. Information content of AVHRR channels 4 and 5 with respect to the effective radius of cirrus cloud particles. *Journal of Applied Meteorology*, 30(7):973–984, 1991.

T. Reiner and F. Arnold. Stratospheric SO3: Upper limits inferred from ion composition measurements - implications for H2SO4 and aerosol formation. *Geophysical Research Letters*, 24(14):1751–1754, 1997. doi: 10.1029/97GL01758. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97GL01758`.

J.J. Remedios, R.J. Leigh, A.M. Waterfall, D.P. Moore, H. Sembhi, I. Parkes, J. Greenhough, M.P. Chipperfield, and D. Hauglustaine. MIPAS reference atmospheres and comparisons to V4.61/V4.62 MIPAS level 2 geophysical data sets. *Atmospheric Chemistry and Physics Discussions*, 7:9973–10017, 2007. doi: 10.5194/acpd-7-9973-2007. URL `https://www.atmos-chem-phys-discuss.net/7/9973/2007/`.

M. Ridolfi, B. Carli, M. Carlotti, T. Von Clarmann, B.M. Dinelli, A. Dudhia, J.M. Flaud, M. Höpfner, P.E. Morris, P. Raspollini, G. Stiller, and R.J. Wells. Optimized forward model and retrieval scheme for MIPAS near-real-time dataprocessing. *Applied Optics*, 39(8):1323–1340, Mar 2000. doi: 10.1364/AO.39.001323. URL `http://ao.osa.org/abstract.cfm?URI=ao-39-8-1323`.

M. Ridolfi, S. Del Bianco, A. Di Roma, E. Castelli, C. Belotti, P. Dandini, G. Di Natale, B.M. Dinelli, L.C. Labonnote, and L. Palchetti and. FORUM Earth Explorer 9: Characteristics of Level 2 Products and Synergies with IASI-NG. *Remote Sensing*, 12(9):1496, 2020.

C.D. Rodgers. *Inverse methods for atmospheric sounding: theory and practice*, volume 2. World scientific, 2000.

L.S. Rothman, I.E. Gordon, A. Barbe, D. Chris Benner, P.F. Bernath, M. Birk, V. Boudon, L.R. Brown, A. Campargue, J.-P. Champion, K. Chance, L.H. Coudert, V. Dana, V.M. Devi, S. Fally, J.-M. Flaud, R.R. Gamache, A. Goldman, D. Jacquemart, I. Kleiner, N. Lacome, W.J. Lafferty, J.-Y. Mandin, S.T. Massie, S.N. Mikhailenko, C.E. Miller, N. Moazzen-Ahmadi, O.V. Naumenko, A.V. Nikitin, J. Orphal, V.I. Perevalov, A. Perrin, A. Predoi-Cross, C.P. Rinsland, M. Rotger, M. Simeckova, M.A.H. Smith, K. Sung, S.A. Tashkun, J. Tennyson, R.A. Toth, A.C. Vandaele, and J. Vander-Auwera. The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(9):533 – 572, 2009. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2009.02.013. URL `http://www.sciencedirect.com/science/article/pii/S0022407309000727`. HITRAN.

L.S. Rothman, I.E. Gordon, Y. Babikov, A. Barbe, D. Chris Benner, P.F. Bernath, M. Birk, L. Bizzocchi, V. Boudon, L.R. Brown, A. Campargue, K. Chance, E.A. Cohen, L.H. Coudert, V.M. Devi, B.J. Drouin, A. Fayt, J.-M. Flaud, R.R. Gamache, J.J. Harrison, J.-M. Hartmann, C. Hill, J.T. Hodges, D. Jacquemart, A. Jolly, J. Lamouroux, R.J. Le Roy, G. Li, D.A. Long, O.M. Lyulin, C.J. Mackie, S.T. Massie, S. Mikhailenko, H.S.P. Müller, O.V. Naumenko, A.V. Nikitin, J. Orphal, V. Perevalov, A. Perrin, E.R. Polovtseva, C. Richard, M.A.H. Smith, E. Starikova, K. Sung, S. Tashkun, J. Tennyson, G.C. Toon, Vl.G. Tyuterev, and G. Wagner. The HITRAN2012 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 130:4 – 50, 2013. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2013.07.002. URL `http://www.sciencedirect.com/science/article/pii/S0022407313002859`. HITRAN2012 special issue.

R. Saunders, J. Hocking, E. Turner, P. Rayer, D. Rundle, P. Brunel, J. Vidot, P. Roquet, M. Matricardi, A. Geer, N. Bormann, and C. Lupu. An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11(7), 2018.

C. Serio, G. Masiello, F. Esposito, P. Di Girolamo, T. Di Iorio, L. Palchetti, G. Bianchini, G. Muscari, G. Pavese, R. Rizzi, B. Carli, and V. Cuomo. Retrieval of foreign-broadened water vapor continuum coefficients from emitted spectral radiance in the H2O rotational band from 240 to 590 cm-1. *Optics Express*, 16(20):15816–15833, 2008.

R. Spurr. *LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems*, pages 229–

275. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-48546-9. doi: 10.1007/978-3-540-48546-9_7. URL `https://doi.org/10.1007/978-3-540-48546-9_7`.

K. Stamnes, S.C. Tsay, W. Wiscombe, and K. Jayaweera. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Applied Optics*, 27(12):2502–2509, Jun 1988. doi: 10.1364/AO.27.002502. URL `http://ao.osa.org/abstract.cfm?URI=ao-27-12-2502`.

W.T. Sturges, T.J. Wallington, M.D. Hurley, K.P. Shine, K. Sihra, A. Engel, D.E. Oram, S.A. Penkett, R. Mulvaney, and C.A.M. Brenninkmeijer. A potent greenhouse gas identified in the atmosphere: SF5CF3. *Science*, 289(5479):611–613, 2000. ISSN 0036-8075. doi: 10.1126/science.289.5479.611. URL `https://science.sciencemag.org/content/289/5479/611`.

D.C. Tobin, F.A. Best, P.D. Brown, S.A. Clough, R.G. Dedecker, R.G. Ellingson, R.K. Garcia, H.B. Howell, R.O. Knuteson, E.J. Mlawer, H.E. Revercomb, J.F. Short, P.F.W. van Delst, and V.P. Walden. Downwelling spectral radiance observations at the SHEBA ice station: Water vapor continuum measurements from 17 to 26μm. *Journal of Geophysical Research: Atmospheres*, 104(D2):2081–2092, 1999.

D.D. Turner, E.J. Mlawer, G. Bianchini, M.P. Cadeddu, S. Crewell, J.S. Delamere, R.O. Knuteson, G. Maschwitz, M. Mlynczak, S. Paine, L. Palchetti, and D.C. Tobin. Ground-based high spectral resolution observations of the entire terrestrial spectrum under extremely dry conditions. *Geophysical research letters*, 39(10), 2012.

J.H. Van Vleck and D.L. Huber. Absorption, emission, and linebreadths: A semihistorical perspective. *Reviews of Modern Physics*, 49(4):939, 1977.

Acknowledgements

I wish to thank my supervisor **Bianca Maria Dinelli** from ISAC-CNR (Bologna - Italy) and my co-supervisor **Tiziano Maestri** from the University of Bologna for their constant support and constructive criticism during the development of the present thesis work.

I am very grateful to a number of people who has been very helpful in allowing this work to be completed:

Luca Palchetti from INO-CNR (Florence - Italy), Principal Investigator of the FORUM mission for supplying the REFIR-PAD measurements above Teresina (Brazil)

Hilke Oetjen from ESA-ESTEC (European Sace Agency) for helpful discussions

Marco Ridolfi from INO-CNR (Florence - Italy) that has helped me in many different parts of the work performed during the PhD Course

Jonathan Tennyson from UCL (University College of London, London, UK) for supplying the updated spectroscopic parameters of the water vapour transitions in the FIR range

Helen Brindley and **Jonathan Murray** from Imperial College (London, UK) for welcoming me at the Imperial College and giving me chance to work on TAFTS data

Elisa Castelli and **Enzo Papandrea** for the willingness, technical support and encouragement

Samuele Del Bianco from IFAC-CNR (Florence - Italy) for supplying me the KLIMA RTM code

Luca Sgheri from IAC-CNR (Florence - Italy) for supplying me the retrieval code used for the analysis of TAFTS spectra

All my colleagues at CNR-ISAC for coffees and always interesting discussion

Alba I really care you even if you fall asleep at 10 pm...