Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Corso di Laurea Magistrale in Fisica del Sistema Terra

Antarctic cloud spectral emission from ground-based measurements, a focus on Far InfraRed signatures

Relatore:
Prof. Rolando Rizzi

Correlatore:
Prof. Tiziano Maestri

Presentata da:
Carlo Arosio

Sessione II
Anno Accademico 2013/2014

Sommario

Il lontano infrarosso (FIR) è una regione spettrale fortemente caratterizzata dalla presenza della banda rotazionale pura del vapore acqueo. L'importanza del FIR, in relazione al bilancio energetico terrestre, emerge con forza in presenza di nubi alte che modulano il flusso uscente di radiazione in onda lunga. Il FIR, per difficoltà tecnologiche, è a tutt'oggi ancora poco esplorato e il progetto PRANA (Proprietà Radiative del vapore Acqueo e delle Nubi in Antartide) è la prima campagna di misure in cui si osservano spettri di emissione fino a 100 cm^{-1} su un arco temporale di oltre due anni. Lo strumento principe è REFIR-PAD, uno spettrometro a trasformata di Fourier, per mezzo del quale sono studiate anche le proprietà delle nubi sul Plateau Antartico.

In questo lavoro, è stato analizzato il dataset di misure raccolte da PAD per tutto il 2013, insieme alle misure di altri strumenti installati in loco, quali quick-looks Lidar e radiosondaggi. Per prima cosa, sono state fissate una serie di soglie per la selezione di spettri di buona qualità, considerando le radianze e le loro oscillazioni in alcuni intervalli spettrali chiave. In questo modo, sono stati ottenuti migliaia di spettri misurati nelle più disparate condizioni meteorologiche. Questi spettri sono stati descritti in maniera sintetica mediando dapprima le radianze misurate in alcuni intervalli spettrali (come le 'dirty-windows'), trasformando queste in temperature di brillanza e considerando infine le differenze tra ciascuna coppia di esse. Un algoritmo di Feature Selection supervisionato è stato implementato, con lo scopo di selezionare quali, tra le features appena definite, fosse effettivamente informativa riguardo la presenza, la fase e la tipologia di nube osservata. A questo scopo, sfruttando i quick-looks Lidar, sono stati costruiti training e test datasets per ogni mese. La classificazione del database è stata compiuta a livello bimestrale allenando una Support Vector Machine sui training datasets.

Sulla base dei risultati di questa classificazione e con l'aiuto delle misure Lidar, 29 casi studio di nubi di ghiaccio non precipitanti sono stati accuratamente selezionati su tutto l'anno. Un singolo spettro per ciascun caso, o al più una media su pochi spettri, è stato processato da un algoritmo di retrieval (RT-RET), che sfrutta alcuni canali nella finestra atmosferica, in modo da estrarre le proprietà della nube in esame. Raggi efficaci e spessori ottici ottenuti sono analizzati, per confrontare i risultati con la letteratura disponibile e per individuare possibili trend stagionali. Infine, i profili atmosferici forniti in output dal retrieval sono stati utilizzati come input per simulazioni, assumendo due differenti forme di cristalli, con lo scopo di indagare la nostra capacità nel riprodurre le radianze nel lontano infrarosso. Differenze sostanziali per le micro-windows nel FIR sono state osservate: dal confronto tra i vari casi, emerge una grande variabilità di pattern spettrali per le differenze tra simulazioni e spettri osservati; è stato tentato uno sforzo per legare questa variabilità alle proprietà microfisiche delle nubi.

Abstract

The Far InfraRed (FIR) is a region of the spectrum strongly influenced by the presence of the pure rotational water vapor absorption band. The role of the FIR, in the Earth's energy balance frame, becomes particularly relevant in the presence of high clouds, that affect the outgoing long-wave energy flux. The FIR, due to technological difficulties, is under-explored still today and the PRANA project has been the first extensive field campaign measuring atmospheric spectra down to 100 cm^{-1} for more than two years. The principal deployed instrument is REFIR-PAD, a Fourier transform spectrometer, used by us to study Antarctic cloud properties.

In the present work, a dataset covering the whole 2013 has been analyzed, with the help of other in situ measurements such as quick-looks Lidar and radiosonde profiles. Firstly, a selection of good quality spectra is performed, using, as thresholds, radiance values, with estimated errors, in few chosen spectral regions. As a result, thousands of good quality spectra are obtained, measured in the presence of the most various meteorological conditions. These spectra are described in a synthetic way averaging radiances in selected intervals (such as the 'dirty-windows'), converting them into brightness temperatures and finally considering the differences between each pair of them. A supervised feature selection algorithm is implemented with the purpose to select the features really informative about the presence, the type of cloud and the phase of the particulate. With this aim, for each month, a training and a test sets of spectra are collected, by means of Lidar quick-looks. The supervised classification step of the overall monthly datasets is performed using a Support Vector Machine.

On the base of this classification and with the help of Lidar observations, 29 non-precipitating ice cloud case studies are accurately selected over the year. A single spectrum, or at most an average over two or three spectra, is processed by means of the retrieval algorithm RT-RET, exploiting some main IR window channels, in order to extract cloud properties. Retrieved effective radii and optical depths are analyzed, to compare them with literature studies and to evaluate possible seasonal trends. In the end, retrieval output atmospheric profiles are used as inputs for simulations, assuming two different crystal habits, with the aim to examine our ability to reproduce spectral radiances in the FIR region. Substantial mis-estimations are found for FIR micro-windows: a high variability is observed in the spectral pattern of simulation deviations from measured spectra and an effort to link these deviations to cloud parameters has been performed.

Contents

Introduction **1**

1 The Far InfraRed spectral region and its role in atmospheric physics **3**
1.1 Spectral properties of the Earth in the FIR and Radiative transfer summary 4
1.2 Climatic role of the FIR: energy balance 9
1.3 Remote sensing of water vapor profiles and ice cloud spectral signatures . . . 12
1.4 Instrument development 16
1.5 REFIR project and EU feasibility study 17
1.6 The ECOWAR campaign pilot study 19

2 Antarctic cloud properties **23**
2.1 From radiometric measurements 24
2.2 From in situ measurements 28

3 The PRANA project: REFIR-PAD and other instruments **35**
3.1 REFIR-PAD instrument details 36
3.1.1 Martin-Puplett configuration 37
3.1.2 Mach-Zehnder configuration 38
3.1.3 REFIR-PAD configuration 39
3.2 Error Analysis 42
3.3 Lidar 47
3.4 Vertical profile data 51

4 From pre-processing to spectra classification **53**
4.1 A look at the available REFIR-PAD dataset and at the observed spectra . . 53
4.2 REFIR-PAD Data Pre-processing 57
4.3 Data grouping 62
4.3.1 Manual selection of 'examples' 62
4.3.2 Un-Supervised Feature Extraction 64
4.3.3 Supervised Feature Selection and Support Vector Machine 68

5 Cloud properties retrieval and simulation **85**
5.1 Case studies selection and description 85
5.2 Retrieval methodology and results 92
5.2.1 Clear-sky gaseous optical depth computation 93
5.2.2 Spectral particle radiative properties 95
5.2.3 Model approach and phase function calculation 97
5.2.4 Retrieval methodology 99
5.2.5 Retrieval results 101
5.3 Simulations: how well cloud spectra are reproduced? 107
5.3.1 Simulation accuracy 112
5.3.2 Simulations in the FIR 114

Conclusions **127**

Abbreviations **129**

Bibliography **130**

Acknowledgements **137**

List of Figures

1.1.1 Line optical depths for water vapor, CO_2 and O_3 in the IR and FIR regions. 5
1.1.2 Line absorption optical depths for other atmospheric gaseous components in the FIR region. 5
1.1.3 Spectral brightness temperature difference and radiance percentage difference for the F (full scattering) and A (only absorption) type of simulations for downwelling radiance at the ground 9
1.2.1 Spectral heating rates as function of altitudes, for clear sky conditions. From Harries et al.(Harries et al., 2008) 10
1.2.2 Spectral layer energy balance for a reference cirrus and net flux profiles, from (Maestri and Rizzi, 2003) 13
1.3.1 Spectral variations of the Real and the Imaginary part of the ice refractive index. From Warren (1984) 14
1.3.2 The sensitivity of BTD (250.0 - 559.5 cm^{-1}) and BTD (250.0 and 410.2 cm^{-1}) to optical thickness, from Yang et al. (2003) 15
1.3.3 Sensitivity of the BTD 8.52 - 11.0 μm and BTD 250.0 - 559.5 cm^{-1} to effective particle size. From Yang et al. (2003) 15
1.6.1 Mean differences between simulations and observations in selected micro-windows, from Maestri et al. (2014) 21
1.6.2 Mean differences between simulations and observations for selected FIR micro-windows vs FOV, from Maestri et al. (2014) 22

2.1.1 Mahesh et al. retrieved effective radii and optical depths vs time 26
2.2.1 Relative frequency of occurrence and contribution of the different ice crystal types, from Walden et al. (2003) 29
2.2.2 Mean effective radii from Lawson et al. (2006) 31
2.2.3 Examples of Cloud Particle Imager observations classified into habits categories. From Lawson et al. (2006). 32
2.2.4 Lawson et al. ice crystal classification results 33
2.2.5 Mean equivalent radii from Lawson et al. (2006) 33

3.0.1 Antarctic map with Concordia Station position 36
3.0.2 Physical Shelter where Lidar system and REFIR-PAD instrument are located and on which the VAISALA station is set 37
3.1.1 Optical configuration of a polarizing interferometer in the Martin-Puplett configuration 38
3.1.2 Optical configuration of a Mach-Zehnder interferometer type 39
3.1.3 New proposed configuration of double input-output Mach-Zehnder interferometer with double signal and tilt compensation. 40
3.1.4 Compact design of REFIR-PAD 41
3.2.1 Example of uncertainties on uncalibrated spectra and NESR on calibrated spectra from Bianchini and Palchetti (2008) 45
3.2.2 Radiometric performance of REFIR-PAD measurements, NESR, CAL and STD for a clear and a cloudy case 46
3.3.1 Example of quicklook Lidar, with ice and water clouds 48
3.3.2 Example of quicklook Lidar, with non-precipitating ice clouds 49
3.3.3 Example of Lidar profile 50
3.4.1 Example of radiosonde and ECMWF profiles 52

4.1.1 An example of measured spectrum in clear conditions 54
4.1.2 An example of measured spectrum in the presence of an ice cloud 55
4.1.3 The same example of clear spectrum converted in BT 56
4.1.4 The same example of cloudy spectrum converted in BT 56
4.2.0 Example of parameter used to select good spectra 60
4.2.1 Case of spectrum considered as good but deformed by an anomalous slope. 61
4.3.1 PCA result for January observations 65
4.3.2 Unsupervised Spectral Regression algorithm result for the overall January dataset 66
4.3.3 Unsupervised Spectral Regression algorithm result for January cloudy cases 67
4.3.4 Unsupervised Spectral Regression algorithm result for June cloudy cases . . 68
4.3.5 Projection example linked to Fisher discriminant analysis 70
4.3.6 Example of scatter plot for training points selected for May and June 72
4.3.7 Example of scatter plot for training points selected for May and June 72
4.3.8 Two linear classifiers and the associated margin lines for a 2-class classification problem, from Theodoris and Koutroumbas (2010). 74
4.3.9 Example of SVM clear-cloudy classification: scatter plot of the overall measured set, with training and test examples superimposed. 76
4.3.10 Example of scatter plot showing the final result of the clear-cloudy classification task 76
4.3.11 Lidar quick-look in the presence of water cloud and smog episode 78

4.3.12 Spectral absorption coefficient for water and bulk ice 79
4.3.13 Example of SVM ice-water clouds classification: scatter plot of the overall January dataset, with distinction of training and test examples 81
4.3.14 Example of SVM precipitating/non-precipitating ice cloud classification . . 82

5.1.1 Example of some cloud events chosen as case studies 91
5.2.1 Example of measured clear spectrum and the relative simulation 94
5.2.2 Example of weighting functions for selected windows in a zenith observation geometry 94
5.2.3 Differences between simulations performed using $\sigma = 0$ and $\sigma = 5$ 96
5.2.4 Two examples of particle size distributions for plates and columns 98
5.2.5 Spectral differences between two simulations performed using a different number of zenith angles 99
5.2.6 Example of retrieved OD and respective AOD in the 800-1000 cm^{-1} region, for different values of r_{eff}. From Maestri and Holtz (2009) 101
5.2.7 Retrieved effective radii vs time. 104
5.2.8 Retrieved optical depths with uncertainties vs time. 105
5.3.1 Sensitivity test to IWC variations 107
5.3.2 Spectral sensitivity to cloud IWC variations 108
5.3.3 Three simulations fixing the cloud IWC profile with different radii 109
5.3.4 Spectral sensitivity to particle effective radius variations 109
5.3.5 Spectral pattern of the scattering coefficient 110
5.3.6 BTDs identified by Yang et al.Yang et al. (2003) as informative about cloud properties 111
5.3.7 Mean differences between simulated and measured radiances in the IR main window 113
5.3.8 Mean differences between simulated and measured radiance slopes in the IR main window 113
5.3.9 Three examples of simulated spectra and deviations from observed ones . . 115
5.3.10 Mean FIR absolute simulation deviations from observed radiances over selected dirty windows 116
5.3.11 Spectral differences between simulated and measured spectrum assuming plate habit, for the 27/12/13 case. 117
5.3.12 Averaged spectral differences between simulated and measured spectra . . . 118
5.3.13 Mean differences between simulated and measured spectra, on five spectral intervals. 118
5.3.14 Differences between simulated and measured spectra, averaged on intervals between 410 and 500 cm^{-1} for plate and column habit, sorted by radius . . 120

5.3.15 Differences between simulated and measured spectra, averaged on intervals between 410 and 500 cm^{-1} for plate and column habits, sorted by optical depth 120
5.3.16 Mean deviations on 410 - 500 cm^{-1} region of simulated spectra from measurements, relative to cloud with $OD > 0.7$, for plate and column habits, sorted by radius 121
5.3.17 Mean deviations on 410-500 cm^{-1} region of simulated spectra from measurements, relative to cloud with $OD < 0.7$, for plate and column habits, sorted by radius 122
5.3.18 Differences in the FIR region between simulated and observed spectra for cloud with $OD > 0.7$ and selected increasing radii. 122
5.3.19 Deviations of column crystal simulated radiances from plate ones, averaged over five selected intervals and sorted by increasing OD. 123
5.3.20 Differences between plate and column simulations averaged around 450 cm^{-1} vs optical depth 124
5.3.21 Measured and simulated spectra using both new and old 125
5.3.22 Spectral differences between simulations using new and old atmospheric profiles and deviations from measured spectrum. 125

List of Tables

3.1 REFIR-PAD instrument main characteristics in the amplitude-division operating mode used in the campaign 42
3.2 Lidar main characteristics 47

4.1 List of the selected spectral intervals used to features definitions 63
4.2 List of the most important features for clear-cloudy case discrimination . . . 73
4.3 List of the most important features for ice-water clouds discrimination 79
4.4 List of the most important features for precipitating-non precipitating ice clouds discrimination 83

5.1 Description of the 29 non-precipitating ice cloud case studies 90
5.2 Retrieval results for the 29 selected cloud cases, performed assuming columns and plates as ice habits 103
5.3 Correlation between retrieved parameters and other cloud properties. 105
5.4 Literature results in terms of retrieved effective radii together with RT-RET obtained values. 106

Introduction

High clouds are an important element of the Earth climatic system, due to their strong modulation of the total outgoing long-wave radiation (Liou, 2002). The magnitude and sign of their radiative effect is strongly affected by microphysical properties, difficult to study and simulate. A consistent fraction of the Earth's outgoing energy flux occurs in the so called Far InfraRed (FIR) spectral region, a low photon-energy part of the spectrum between the microwave region and the CO_2 ν_2 absorption band. The presence of an ice cloud, due to its low temperature, increases the importance of this spectral region (Rizzi and Mannozzi, 2000). Hence, the study of the FIR is strictly related to climatic and energy transfer considerations; however, the high sensitivity of the observed radiance to water vapor distribution and to cloud microphysical properties, make this spectral interval an extremely interesting region for many remote sensing applications (Harries et al., 2008).

Despite the general interest, until the end of the last century the FIR has represented a frontier in the atmospheric sensing: technological difficulties avoided observations at these wavelengths and new solutions were studied in order to achieve the desired radiometric range, resolution and accuracy. The REFIR project is a milestone in this framework and has provided, as final product, a Fourier Transform Spectrometer meeting required accuracy, lifetime and simplicity requested for a satellite mission. Fading this perspective, some measurement campaigns were organized, as (Maestri et al., 2014), and the PRANA project on the Antarctic Plateau is the first extensive field campaign collecting atmospheric spectra in the FIR. Main purposes are the improvement of retrieval methodologies and simulation models for this spectral region and the study of water vapor and cloud signatures at these wavelengths. Microphysical properties and climatology of under-explored Antarctic clouds have been studied.

The present work regards the analysis of a full year of downwelling atmospheric emission spectra, measured from the ground by the REFIR-PAD instrument deployed at the Concordia scientific station. A pre-processing step for the selection of quality criteria for this database is implemented and an analysis of spectral radiances is elaborated, to derive a classification of spectra regarding the presence, the phase and the type of clouds. A retrieval methodology from Maestri and Holtz (2009) is applied on a number of study cases, representative of the most important non-precipitating cloud events over the year. Simulations are performed, on the basis of the retrieved profiles, and differences from measured radiances are obtained, in order to evaluate our ability in the reproduction of radiances in FIR micro-windows. Cloud properties are evaluated in relation to precedent studies and linked to the accuracy of simulated radiances in the FIR.

Chapter 1

The Far InfraRed spectral region and its role in atmospheric physics

Conventionally, the Far InfraRed (FIR) spectral region starts from a wavelength of about 15 μm (corresponding to the CO_2 band centered at 667 cm^{-1}), as a rough boundary, and lies at longer wavelengths until 100 cm^{-1}, before the millimeter microwave region; in fact, does not exist a unique and formal definition of this spectral interval. What does matter is that FIR is of great importance for planetary energy balance considerations and in a number of key climate processes, for example, water vapor and cloud feedbacks; but also it is of great practical value in applications concerning with remote sounding (Harries et al., 2008). The region of the FIR has not been exploited until the end of the last century because it seemed to be at a disadvantage in terms of low photon energy with respect to the Mid-Infrared, while not benefiting, for technical reasons, from the multiplexing or heterodyne techniques usable in the mm region. In fact, on the one hand, Mid-InfraRed (MIR) typically used photo-detectors can be modified with sensitivity to longer wavelength but with the implication of complex and expensive lower temperature cooling systems; on the other hand, microwave systems can be designed for higher frequencies, but with difficulty to built oscillators at FIR higher frequencies. So, new technical solutions have been studied in order to allow reliable observations in this spectral region.

1.1 Spectral properties of the Earth in the FIR and Radiative transfer summary

From the spectroscopic point of view, the FIR is characterized by the transition between vibrational energy states and pure rotational ones: therefore, it is a cool, low-energy part of the spectrum. The key feature concerns the fact that molecules, such as H_2O and CO_2, absorb and emit infrared radiation at discrete frequencies, through vibrational or rotational transitions; these are regulated by specific quantum mechanical selection rules that determine the wavelengths of the observed absorption lines. Hence, the measured atmospheric spectrum is a highly complex merge of individual lines and bands from all of the absorbing molecules, depending also on molecular density and atmospheric temperature and pressure profiles.

The strength of absorption can be measured by a parameter known as absorption coefficient $k_\nu\,[m^2/kg]$, which is related to the fractional transmittance of a path of length l and absorber density ρ, by

$$\tau_\nu = \exp\left[-\int_0^L k_\nu(l)\rho(l)dl\right];$$

the integral is known as optical path. This quantity can be computed for each spectroscopic line; in particular, we define the line strength S as the integral of the respective absorption coefficient over wavelength and, for each gaseous specie, an average vertically integrated density q can be obtained from climatological measurements or measured profiles. Using HITRAN 2004 spectroscopic dataset (Rothman et al., 2009), the total strength for an average amount of each gas atmospheric component can be computed for each line. Fig.1.1.1 and Fig.1.1.2 report such optical depth defined as the product $S \cdot q$. Each value is plotted at the resonance frequency of the relative transition line in a logarithmic scale, allowing a comparison between absorption contributions of various gas species.

Fig.1.1.1 refers to the spectral range covered by the FTS measurements considered in this work. The most important spectroscopic features in this region are the water vapor and CO_2 absorption bands. In particular, the line structure below about 500 cm^{-1} is mainly due to the water vapor that presents a broad absorption band comprising many hundreds of strong individual pure rotational lines and extending from the microwave region to and beyond FIR wavenumber upper boundary. The absorption intensity peaks, linked to the strongest rotational transitions, are around 100-200cm^{-1}. As a consequences of this spectral figure, from space the Earth surface in completely obscured in the FIR, except in the coldest and driest conditions near the poles. Water vapour presents another important absorption band: the ν_2 vibrational band centered around 1250 cm^{-1}.

The feature centered at 667 cm^{-1} is the CO_2 ν_2 vibrational band, characterized by a close sequence of very strong absorption lines, visible in the figures as red dots. The O_3, represented with green dots, shows several spectral figures such as the band centered at 1045 cm^{-1}, usually well detectable in measured spectra, and another band toward the mm region.

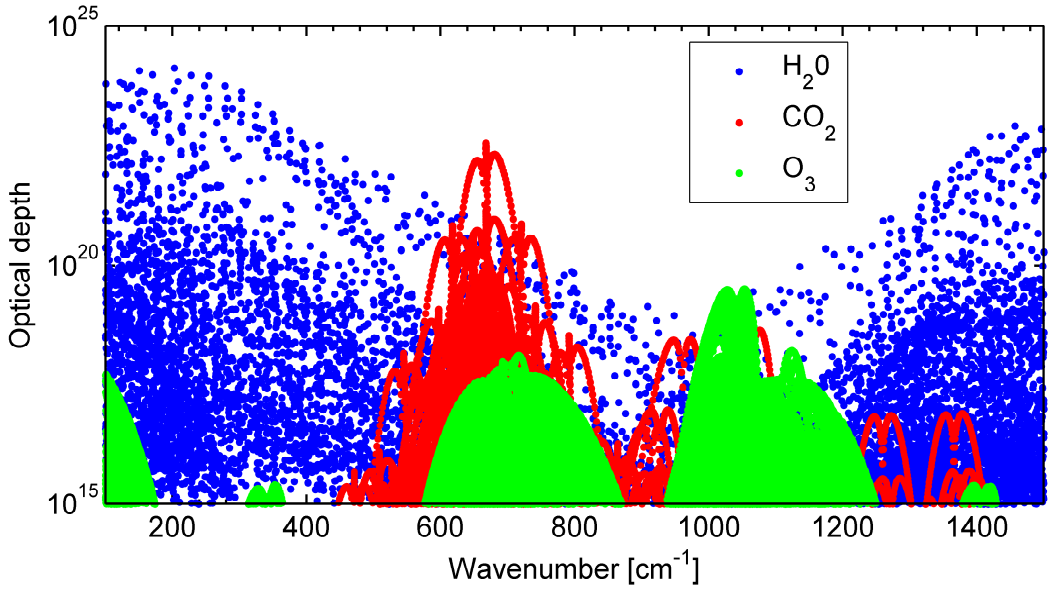


Figure 1.1.1: *Line optical depths for water vapor, CO_2 and O_3 in the IR and FIR regions.*

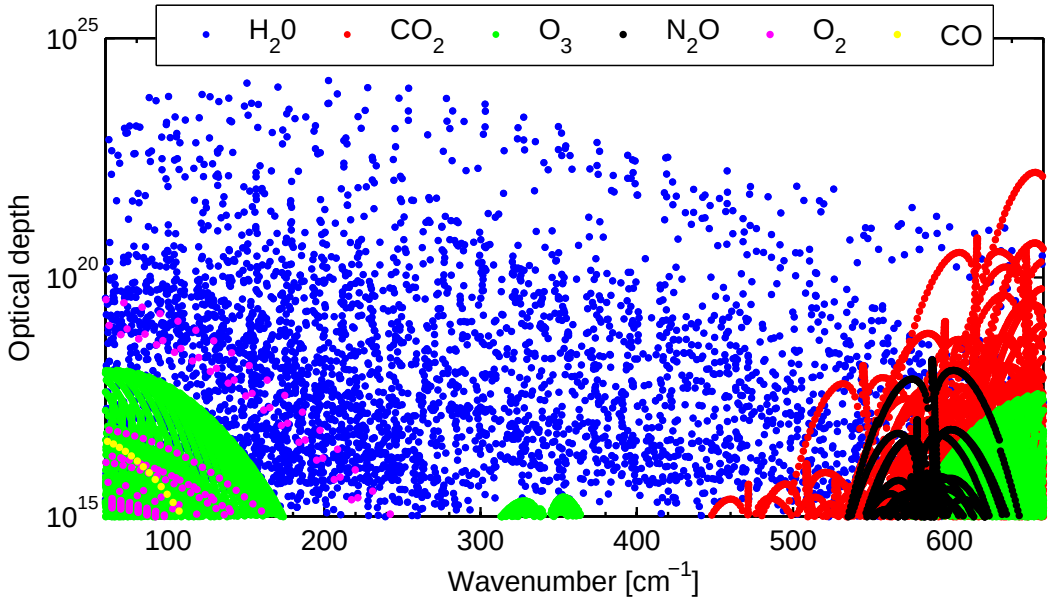


Figure 1.1.2: *Line absorption optical depths for other atmospheric gaseous components in the FIR region.*

Also other gases, such as O_2, N_2O, HNO_3 and CO have rotational or low-vibrational transitions at this wavelengths. Fig.1.1.2 shows optical depths for these molecular species with a focus on the FIR region: as can be seen, these other gases have negligible effects in terms of radiative energy transfer compared with water vapor. However, the existence also of these spectral features, provide the possibility of using FIR region for the retrieval of atmospheric composition.

One aspect not well understood theoretically, concerns the precise shape and properties of the absorption in the "far wings" of water vapor lines, which is stronger than theoretical predictions.This field is actually object of research and different approaches have been proposed: actually, the most successful one defines the water vapour continuum as all absorption due to water vapor not attributable to a Lorentz line within 25 cm^{-1} of each line center (Clough et al., 1989). This model describes the contribution from each spectral line as a sub-Lorentzian line shape and an additional super-Lorentzian absorption in the intermediate line wings.

A part from strong absorptive regions, in some spectral intervals, region called "atmospheric windows", the atmosphere is indeed highly transparent. An important window region can be noted in Fig.1.1.1 between 800 and 1250 cm^{-1}, referred to as "main atmospheric window". In particular, for this IR window region, the brightness temperature seen at TOA is close to the assumed surface brightness temperature. Moreover, looking in the FIR, between 400 and 600 cm^{-1} water vapor transitions are weaker and less dense, so that some semi-transparent windows occurs; these are indeed named as FIR "dirty windows".

The radiative transfer of the atmosphere is controlled by the spectra of atmospheric molecules and is described by means of the Radiative Transfer Equation. This equation describes how radiation spectrally propagates along a path through an atmosphere where scattering, absorption and thermal emission occur due to interactions with the molecules and cloud particles along the path. In the Infrared Region (IR), for clear condition, the two fundamental processes are absorption and emission: considering a beam of radiation, its intensity I_ν will be attenuated by molecular absorption and strengthened by thermal emission (Liou, 2002). The general equation for an emitting and absorbing medium can be written in terms of differential changes in the intensity in the form:

$$-\frac{1}{k_\nu \rho_a}\frac{dI_\nu}{ds} = I_\mu - J_\nu \tag{1.1.1}$$

where k_ν denotes the absorption coefficient, ρ_a is the density of the absorbing gas, s the slant path and J_ν the source function. In this treatment, the atmosphere is considered plane-parallel and in local thermodynamic equilibrium.The last assumption allows to consider the Planck intensity for the source function, thank to Kirchhoff's law, while the first one lead to vertical only variations of atmospheric parameters. As a consequence the basic equation

that governs the thermal IR radiation in the height coordinate can be written as:

$$-\mu \frac{dI_\nu(z,\mu)}{k_\nu \rho_a dz} = I_\nu(z,\mu) - B_\nu(z) \tag{1.1.2}$$

Defining also the optical depth as:

$$\chi = \int_z^{TOP} k_\nu(z')\rho_a(z')dz'$$

the (1.1.2) can be written in χ coordinate:

$$-\mu \frac{dI_\nu(\chi,\mu)}{d\chi} = I_\nu(\chi,\mu) - B_\nu(\chi)$$

The formal solution for the previous differential equation for downward intensity, considering as boundary conditions isotropic emission at the surface and zero downwelling contribution at TOA, is:

$$I_\nu(\chi,-\mu) = \int_0^\chi B_\nu(\chi') e^{\left(-\frac{(\chi-\chi')}{\mu}\right)} \frac{d\chi'}{\mu} \tag{1.1.3}$$

Moreover, introducing the monochromatic transmittance as:

$$\tau_\nu(\chi/\mu) = \exp\left(-\frac{\chi}{\mu}\right)$$

the formal solution for the intensity can then be expressed in the form:

$$I_\nu(\chi,-\mu) = \int_0^\chi B_\nu(\chi') \frac{d}{d\chi'} T_\nu \left[\frac{(\chi-\chi')}{\mu}\right] d\chi' \tag{1.1.4}$$

Now, this equation must be modified in the presence of cloud, because also scattering processes inside the cloud have to be considered. Let the scattering coefficient of the cloud be β_s, the absorption coefficient for cloud particles plus water vapor within the cloud be β_a and the source function associated with scattering be J_ν. So, the radiative transfer equation can then be written as:

$$\mu \frac{dI_\nu}{dz} = -\beta_a(I_\nu - B_\nu) - \beta_s(I_\nu - J_\nu) = -\beta_e(I_\nu - S_\nu) \tag{1.1.5}$$
$$\text{with} \quad \beta_e = \beta_a + \beta_s \quad \text{the extinction coefficient}$$

The source function is an average of the two separate source functions, weighted by their respective absorption and scattering coefficients. Using the definition of the single-scattering

albedo $\widetilde{\omega} = \beta_s/\beta_e$, the source function may be written as:

$$S_\nu = (1 - \widetilde{\omega})B_\nu + \widetilde{\omega}J_\nu$$

The source function for scattering is associated with multiple scattering processes and, in the thermal infrared, it suffices to take the azimuth-independent component:

$$J_\nu = \frac{1}{2}\int_{-1}^{1} P(\mu, \mu')I_\nu(\chi, \mu')d\mu'$$

If the cloud as a whole behaves as a blackbody, it would be just like the Earth surface: radiation below and above the cloud would not be able to penetrate it. Except for thin cirrus, clouds composed of water droplets or cloud containing large ice crystals behave like blackbodies with little variation in the window spectral region. On the contrary, infrared radiative transfer through thin cirrus in the window region must account for scattering processes in order to allow interpretation of the observed spectrum. In particular if the scattering role in the window region between 800 and 1000 cm^{-1} is negligible, in the FIR region both absorption and scattering play an important role. For example, Di Giuseppe and Rizzi (Giuseppe and Rizzi, 1999) illustrated the importance of scattering in radiative transfer processes in the thermal part of the spectrum, performing two set of transfer calculations: the first using cloud optical properties including scattering coefficients, and the second considering only absorption and emission processes. They considered only cirrus clouds, described by means of spherical particles of various effective radii. It is shown that scattering increases the optical depth of the cloud and reduces the OLR at all wavelengths, but with the largest values of the difference in the region around 400 cm^{-1} where OLR is reduced of 5%. In fact, the region around 400 $cm^{-1}$is at the margin of both he CO_2 vibro-rotational band and of the H_2O rotational band and is characterized by relatively high values of single scattering albedo.

Additional considerations are shown for ground-based FIR measurements during the ECOWAR campaign (Maestri et al., 2014). In order to evaluate the relative importance of scattering and absorption, a set of radiative transfer simulations for three different effective particle diameters were performed: in the first case scattering is fully accounted for, while in the second case is considered only gaseous and crystals emission/absorption processes within the cloud. The difference between these two simulations represent an estimate of the importance of the scattering processes. Results reported in Fig.1.1.3 show the importance of the scattering in the FIR interval, with differences that generally increase as the effective diameter of the PSD decreases.

In addition, differences in the region between 800 and 1000cm^{-1} are weakly dependent on the assumed PSD, which means that in this spectral interval the features of the radiance field are driven by absorption processes. On the contrary, FIR differences varies greatly with effective radius.

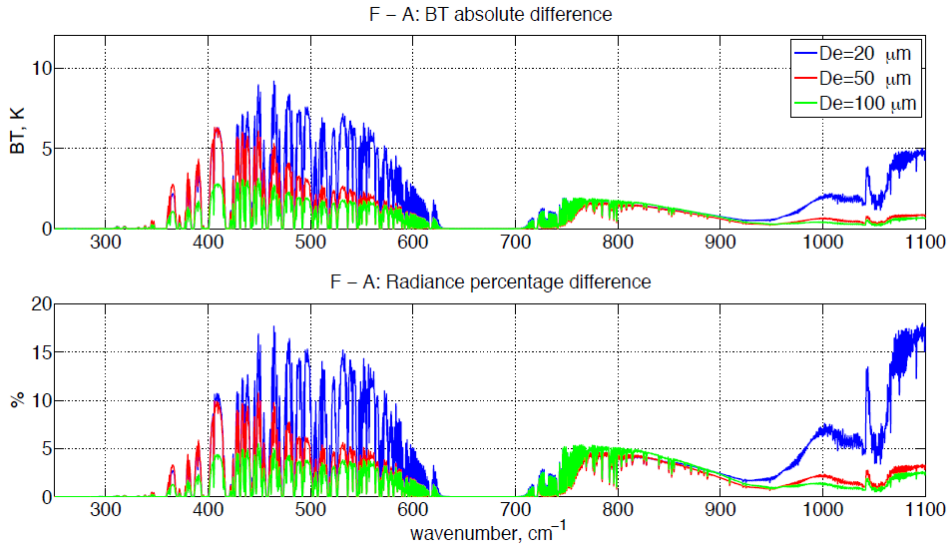


Figure 1.1.3: *Spectral brightness temperature difference and radiance percentage difference for the F (full scattering) and A (only absorption) type of simulations for downwelling radiance at the ground*

1.2 Climatic role of the FIR: energy balance

If we observe the brightness temperature (BT) of the Earth with an instrument in space, we would measure something like 255K; because the peak of the Plank function for a blackbody with this temperature occurs at about 500cm^{-1}, we find that up to 35% of the energy escaping to space is at wavenumbers below 500cm^{-1} and even the 45% below 667cm^{-1}. As a consequence of the spectral structure shown by CO_2 and H_2O, for much of the FIR, the atmosphere becomes partially transparent only in the upper troposphere and stratosphere; so, much of the energy that is emitted and which cools the Earth to space comes not from the surface but from the upper troposphere because the lower one is totally opaque over most of the planet. In the end, we can say that water vapor acts in the FIR as an important greenhouse gas: it can absorb energy from the warm surface and re-emit it to space at a lower temperature.The energy that the Earth emits to space is strongly modulated at different wavenumbers and different heights by water vapor. Therefore, water vapor exhibit a strong feedback, in which any greenhouse warming due to increasing levels of CO_2 is amplified if more water vapor enters the atmosphere as the result (Harries et al., 2008).

For energy transfer considerations, to deal with the concept of heating rates is very useful. The heating rate of a volume of the atmosphere, defined as $dQ/dt = \dot{Q}$, is a measure of the rate of gain of energy per unit volume, per unit volume, per unit time. The spectral heating

rate, in units [J m^{-3} s^{-1} $(cm^{-1})^{-1}$] is equal to the vertical gradient of the net radiative flux. For an atmosphere in hydrostatic equilibrium:

$$\dot{Q}_\nu(z) = \rho(z) c_p \frac{dT(z)}{dt} = \frac{dF_{\nu,N}(z)}{dz} \tag{1.2.1}$$

where c_p is the specific heat and $F_N = F_{down} - F_{up}$: F_{down} and F_{up} are the spectral upward and downward flux densities or spectral irradiances.
In particular, it is interesting to have a look at spectral heating rates as a function of altitude, reported in Fig.1.2.1.

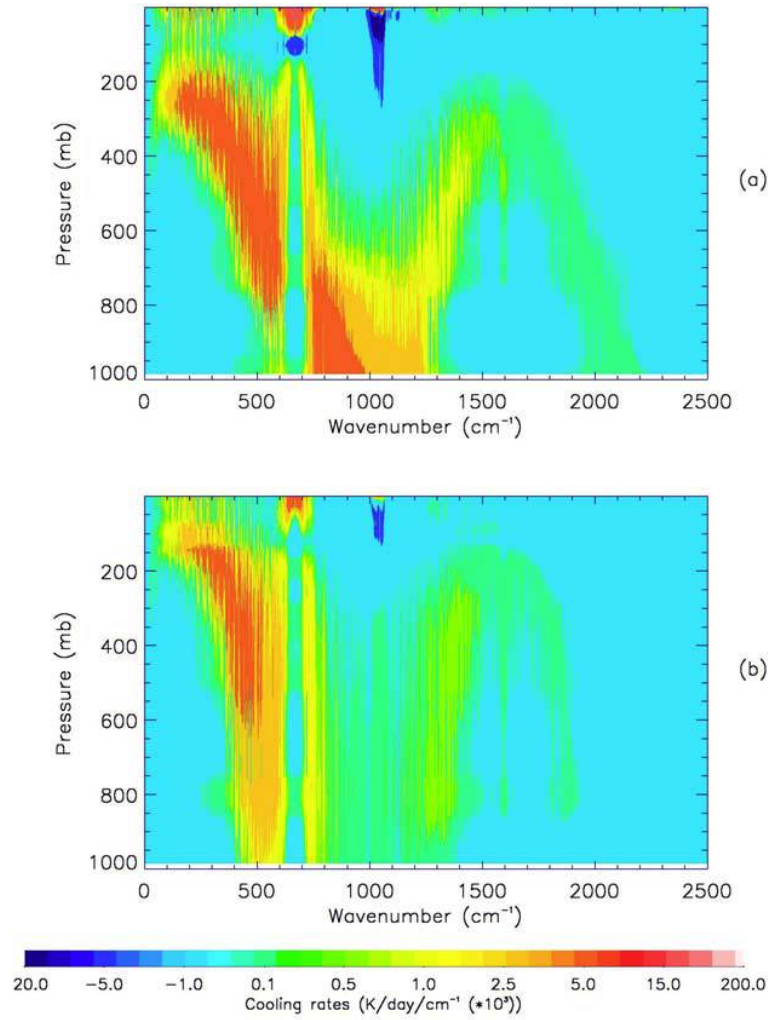


Figure 1.2.1: *Spectral heating rates as function of altitudes, for clear sky conditions. From Harries et al.(Harries et al., 2008)*

The wavenumber dependence of transmissivity of the pure rotational band implies that the height of maximum cooling to space oscillates from middle- to upper- troposphere and has the peak around 100-200 cm^{-1} where the absorption is strongest. At any given wavenumber the cooling rate profile can be calculated from the vertical gradient of net fluxes. This quantity is determined by the vertical gradient of the atmospheric transmission to space, that in the FIR, in absence of clouds, is controlled by the vertical distribution of water vapor in the atmosphere. The level of peak cooling occurs where the transmissivity gradient is a maximum. Above this level the transmissivity to space tends to one, since water vapor amount decreases exponentially, while below it tends to zero. Other interesting features regard the the strong absorption and local heating in the lower stratosphere by the ν_3 band of O_3 and the extreme absorption in the ν_2 CO_2 band: a strong cooling is experienced at stratospheric height around 670 cm^{-1} (Harries et al., 2008).

Concerning with Earth's energy balance, the presence of cold clouds increases considerably the role of the FIR within the total emission to space. Rizzi and Mannozzi (Rizzi and Mannozzi, 2000), calculated spectral radiance emitted to space taking into consideration six standard atmospheres and various top heights and particulate densities. The radiance is integrated in the range from 100 to 600 cm^{-1} and 50 to 2700 cm^{-1} and the focus is on the ratio of these two quantities. In clear sky conditions, this ratio attains values ranging from 0.38 for the tropical atmosphere to 0.48 for the subarctic winter profile. But, in presence of clouds the ratio increases with the opacity of the cloud and its height, exceeding in some cases the value of 0.50. In this sense clouds play a fundamental role in the regulation of the Earth's energy balance.

The fundamental importance of the ice phase is simply that to replace a warm earth surface with a very cold ice cloud has a larger energy balance impact at the TOA than to replace by a low, warm, liquid-phase cloud, especially in the tropics. Bulk radiative studies of cirrus clouds show that they may radiatively cool or heat the upper atmosphere in the thermal infrared wavelengths depending upon height and geometrical and microphysical features. The effect of cloud is to make the range below 600 cm^{-1} energetically more important than in the clear-sky case, because at typical cloud top temperatures, the cold cloud emission is strongly dominated by emission at FIR wavelength (Rizzi and Mannozzi, 2000).

Rizzi and Maestri (Maestri and Rizzi, 2003) showed that flux divergence, integrated over the whole cloud depth for a tropical cirrus, reveals two well-defined spectral regions: the FIR, where net emission of radiation occurs, and the atmospheric window region and NIR, where net absorption dominates; there are two large contributions of opposite sign that influence the heat balance and the net diabatic effect on the cloud. The proportion of absorbed and emitted energy by the layer is strongly affected by cloud ice water content (IWC) but, the crossover wavnumber between cloud net absorption and net emission is well defined quite independently from cloud transmittance, see Fig.1.2.2. They also pointed out that the presence of a cloud changes diabatic heating of all layers below the clouds and state

that the atmospheric concentration of water vapor has an important impact in the cloud's energy balance, since the more the lower atmospheric levels are transparent, the more energy is absorbed by the cloud.

Rizzi and Maestri (Rizzi and Maestri, 2003) in order to study and quantify the FIR contribution of clouds to the radiative balance introduced the parameter known as the cloud radiative forcing. It can be computed by differencing the clear-sky and total-sky radiative fluxes, a positive CRF indicates that the clouds cause a warming of the overall Earth-atmosphere system and vice versa. CRF is a fundamental quantity needed to constrain the effects of clouds within short- and long-term climate models and should be determined as accurately as possible.

In conclusion, the greenhouse effects of both water and ice clouds have been recognized to be very important in the MIR window and in the FIR regions. In Maestri (2000) calculations of the spectral greenhouse parameter g_ν and G_ν are repeated for various values of optical depth: as the transmittance through the ice cloud became smaller, so the radiance escaping to space decreases and G_ν gets larger. In absolute terms the greenhouse trapping is found very important in the FIR.

1.3 Remote sensing of water vapor profiles and ice cloud spectral signatures

Water vapor is on the one hand the most important greenhouse gas and on the other, the most difficult atmospheric component to monitor on a global scale, due to its vertical (and horizontal) and time variability operating on various scales. We have seen the importance of the rotational band of water vapor in the planetary energy balance; the sensitivity of the FIR spectrum at the top of the atmosphere to changes in atmospheric water vapor is well addressed by some studies, as Rizzi et al. (2002). It is estimated that the sensitivity of the rotational band to water vapor perturbations, as measured by the matrix of the derivative of spectral radiance with respect to water vapor amount (Jacobian), may reach values 6-7 times greater than that exhibited using the MIR $\nu 2$ H_2O band. In addition, a comparison between the retrieval performance achieved with a realistic sounder in the MIR and one in the FIR shows that the last one has a better performance than the MIR sounder in the tropopause/lower stratosphere region (Harries et al., 2008).

The FIR spectral region has another important application linked to remote sensing of cloud properties that strictly concerns this work; in fact, signal in the FIR dirty windows can be exploited to derive information about clouds, even if further studies are needed to resolve ambiguities, linked to the complexity of radiative transfer in this region, and to assess how much information on an unknown cloud can be extracted from FIR spectral measurements.

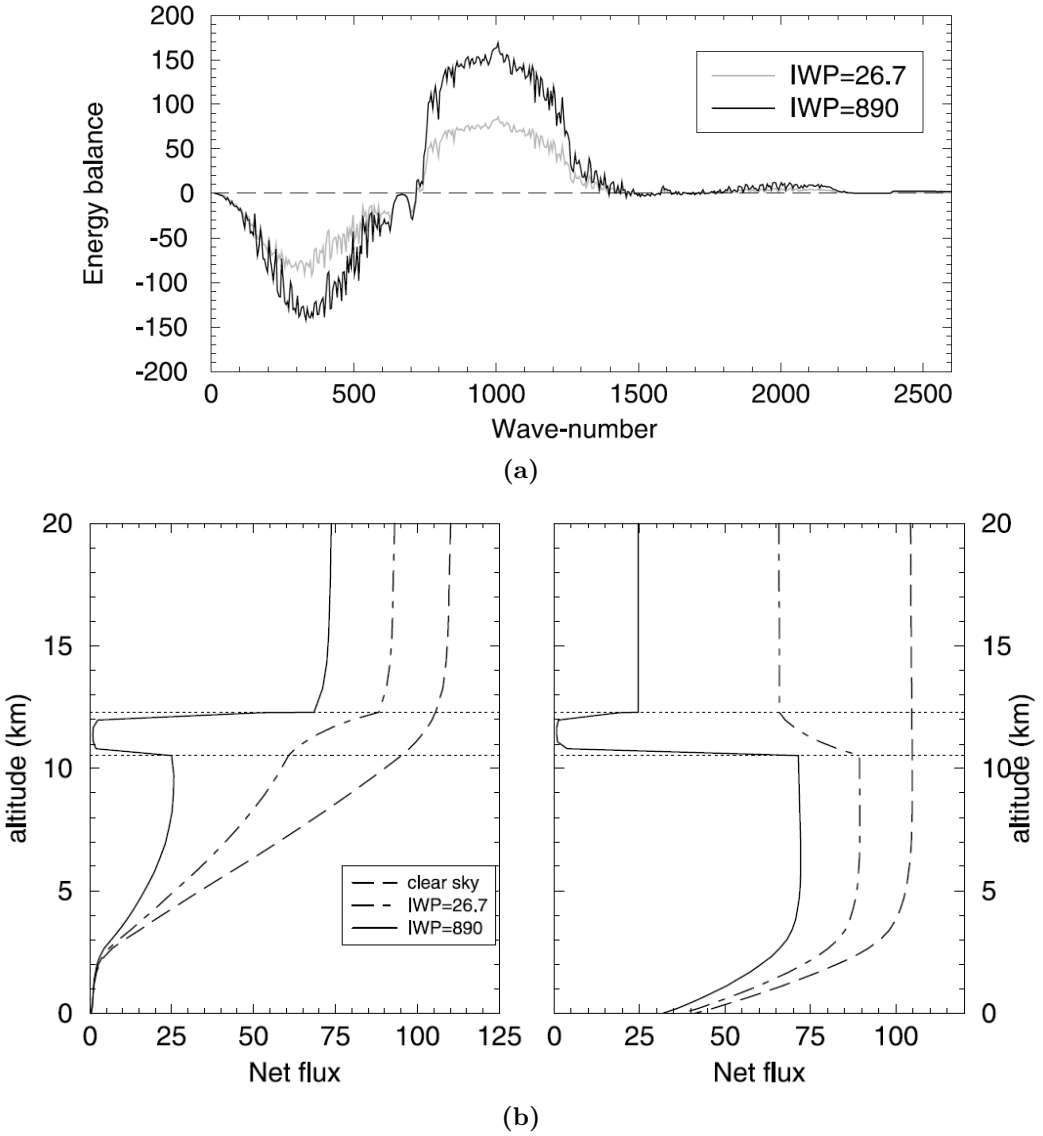


Figure 1.2.2: *In (a) spectral layer energy balance* $[mW/m^2cm^{-1}]$ *for a reference cirrus; in (b) are plotted vertical profiles of the net flux* $[W/m^2]$ *again for the reference cirrus with two ice IWP and for clear sky; on the left FIR band, on the right WIN band. From (Maestri and Rizzi, 2003)*

The study of cloud spectral signatures in the far infrared is based on the important role of the scattering at these wavelengths, as mentioned above. For example, this region can be used for the identification of ice clouds using remote sensing since the imaginary part of the ice refractive index has a local minimum around 450 cm^{-1}, while the real part shows values comparable to the main window ones (Giuseppe and Rizzi, 1999), as reported in Fig.1.3.1.

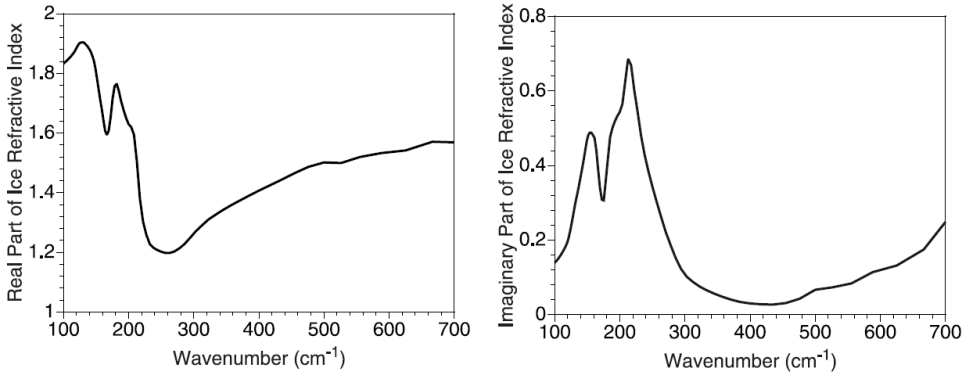


Figure 1.3.1: *Spectral variations of the Real and the Imaginary part of the ice refractive index. From Warren (1984)*

Due to the reduced absorption, the radiance signal originates deeper within the cloud and therfore additional information may be extracted. Related to this issue, Naud et al. (2001) state that the FIR bands around 410 and 550 cm^{-1} offer great advantage for particle size identification in thick clouds and these signals can characterize large particles better than 11-12 μm band.

Moreover, Yang et al. (2003) are able to show that FIR spectral signatures of ice are useful in the retrieval of ice cloud properties. They calculate individual particle scattering properties for a size range from 1 to 10000 μm and bulk scattering properties for 30 PSD. The radiative properties of ice clouds are input to a radiative transfer model together with clear-sky optical thickness and the sensitivities of far-IR spectra to ice cloud optical thickness and effective particle size are investigated. The results show that some brightness temperature differences (BTD) are effectively informative about cloud properties: for example, they show that the brightness temperature near 400 cm^{-1} is sensitive to ice crystal size while, as shown in Fig.1.3.2 the BTD between 250 and 560 cm^{-1} is sensitive to optical thickness for optically thin clouds but shows little variations with the effective particle size when $\tau < 2$. On the contrary,especially for optically thick cloud the BTD between 250 and 410 cm^{-1}, reported in Fig.1.3.2 too, is sensitive to the effective particle size. As can be seen in Fig.1.3.3 for ice clouds having a small optical thickness, this parameter may be inferred with a better accuracy using far-IR data compared to retrievals based on the mid-IR spectrum.

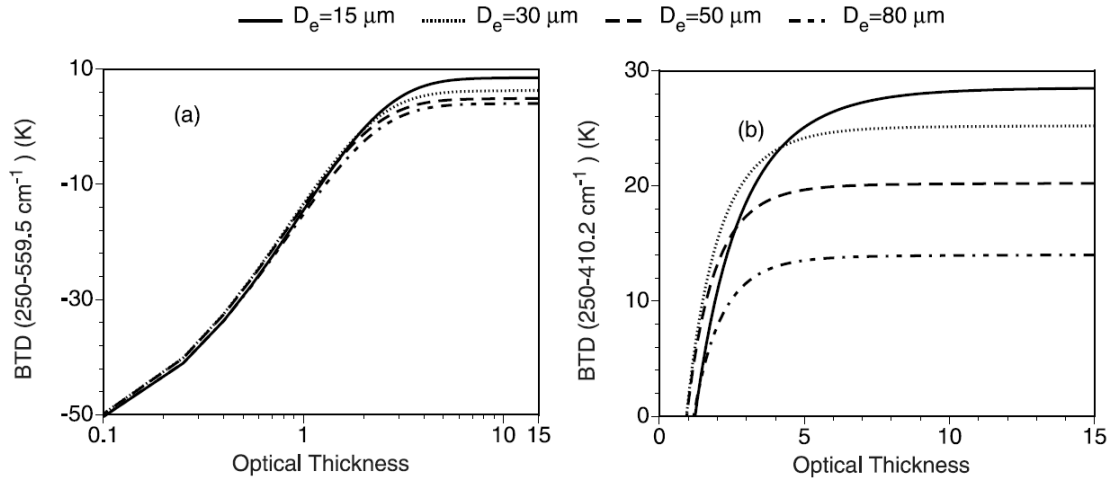


Figure 1.3.2: *The sensitivity of BTD (250.0 - 559.5 cm^{-1}) and BTD (250.0 and 410.2 cm^{-1}) to optical thickness for four effective particle sizes. From Yang et al. (2003).*

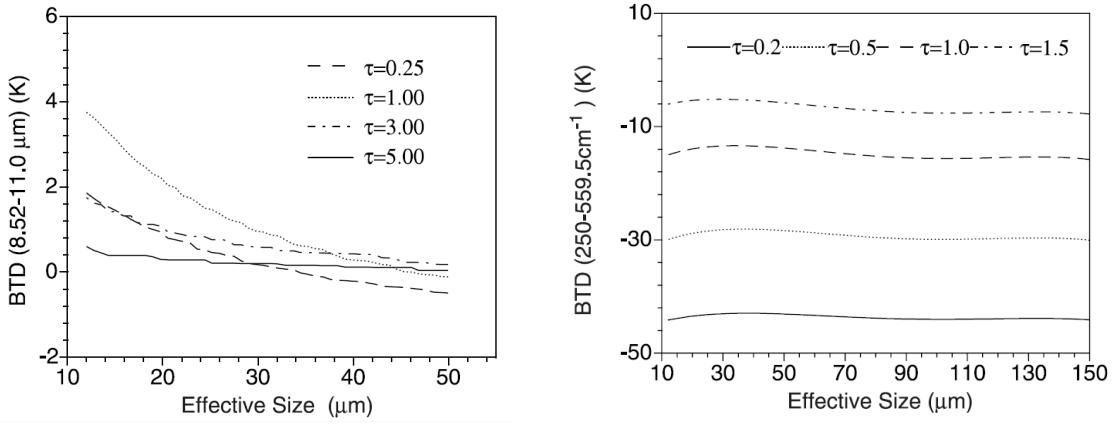


Figure 1.3.3: *In (a) the sensitivity of the BTD between 8.52 and 11.0 µm to effective particle size for four optical thickness, in (b) the same for the two channel at 250.0 and 559.5 cm^{-1}. From Yang et al. (2003)*

Space borne measurements of the thermal IR atmospheric spectra are generally used to retrieve the cirrus cloud properties such as height, phase, optical thickness and particle size, since ice particles have stronger absorption at wavelengths between 11 and 13 μm than wavelengths between 8 and 10 μm and absorption depends on the ice crystal size. The limitations regarding this method is that the BTD 8.5-11 μm is sensitive to small particles over a range of optical thickness; hence, the properties of optically thin or optically thick clouds are difficult to infer.

An important study about FIR simulation quality, is presented by Cox et al. (2010): they perform measurements of mid-infrared and FIR spectra in the presence of cirrus and state that cloudy radiance simulations are not able to consistently reproduce the observed spectral radiances across the entire infrared region and are particularly poor in the 330-600cm^{-1}. Possible causes are inadequate sampling of the cloud structure and of atmospheric parameters.

A general issue related to this difficulty is how well we are able to simulate radiance among different spectral regions, from SW to IR. A number of different studies addresses this problems and discrepancies are found between state of the art simulations and measurements in cloudy conditions.
Ham et al.(Harries et al., 2008) highlight that, comparing satellite radiances measured by MODIS to simulations, results that radiances in SW bands are quite accurate while for MIR window bands show important discrepancies. These differences are linked, according to the authors, to lack of accuracy in the estimate of cloud-top and cloud-base heights.
Bozzo et al. (Bozzo et al.) perform a similar comparison, using airborne data and a retrieval methodology exploiting IR signal in order to retrieve optical depth and effective radius. Results show excellent agreement in the LW band but a sensible underestimation of SW radiances. The authors point out that this problem concerns the consistency of the database of single-scattering properties of ice crystals over the wavelength domain from SW to FIR.

1.4 Instrument development

Measurements in the FIR region have the disadvantage that the signal is intrinsically weak because the Planck function falls to low values at long wavelengths; on the other hand, the atmospheric emission is observed in the Rayleigh-Jeans region of the blackbody distribution where the dependence on temperature is quite linear, so that any errors in our knowledge of the atmospheric temperature cause less error in composition retrieval than in other spectral regions. The typically weak signal in the FIR may be efficiently detected either using cryogenic detectors or heterodyne techniques. A number of field campaigns have been based on an instrument whose basic design is due to the ABB Bomem MR100 series interferometer. In its basic version it is an FTS with liquid nitrogen cooled sandwich MCT/InSb detector

to cover the 500-5000cm^{-1} spectral range (Harries et al., 2008). A short description of three of the principal instrument developed for FIR observations are presented below.

1. TAFTS was developed specifically for aircraft platforms, with the aim of studying the in situ radiative properties of the upper troposphere. TAFTS is a differential, dual-input, polarizing FTS of the Martin-Pupplet kind. This type of interferometer is able to perform simultaneous differential measurements of spectrally resolved radiance from two input ports; it covers a wide spectral range: 80-660cm^{-1}.

2. Radiation Explorer in the Far InfraRed (REFIR) was proposed with the aim of a satellite mission. In the initial design, the FTS was a polarizing interferometer with a novel optical scheme with double-port configuration and a resolution of 0.5cm^{-1}. The configuration is described in more detail in Carli et al. (1999) and in Ch.3.
   Two prototypes have been used to study different solutions: REFIR-Breadboard (BB) and REFIR-Prototype for Applications and Development (PAD). The main design requirement of wide-band coverage with uncooled operation can be met by using efficient wide-band beam splitters (BS) and high-performance room temperature detectors. If the input/output polarizer is removed, the interferometer configuration becomes the Mach-Zehnder scheme (Jenkis and White, 1981). The other optical characteristics, such as the double-port configuration and the full tilt compensation of the moving mirror are maintained. The REFIR-BB instrument was tested in June 2004 in a ground-based field campaign in south Italy. This first measurements were compared to simulated ones for clear sky using LbLRTM: the good results obtained demonstrate that the REFIR concept of using uncooled components is feasible.

3. The FIRST instrument was developed in USA; it is a Michelson FTS designed to demonstrate the ability to measure the spectrum between 1000 and 100 cm^{-1} globally, on a daily basis at high resolution.

1.5 REFIR project and EU feasibility study

REFIR instrument is the final product of a long research began at the end of '90 in the framework of a feasibility study funded by the European Union, with a contract started on January 2007 and lasted on April 2000. The main scientific objectives of the REFIR project are (REFIR Executive Summary, 2000):

- the measurement of the outgoing spectral FIR radiation at the top of the atmosphere in the context of the Earth's radiation budget, monitoring the FIR region not covered by any current or planned mission;

- the improvement of our knowledge of the principal atmospheric constituents that modulate, to a large extent, the FIR emission to space:
  - mid- and upper-tropospheric water vapor, exploiting its strong spectral signature in the FIR;
  - mid- and upper-tropospheric clouds, such as cirrus.

The core instrument for the REFIR purpose, is a Fourier-Transform Spectrometer (FTS) able of high-resolution spectral measurements of the radiative flux from the Earth, with a high Signal-to-Noise Ratio (SNR) in broad-band operations. In particular, a number of requirements about instrumental characteristics are formulated (Carli et al., 1999):

- a small and compact instrument capable of continuous operations between 100 and 1000 cm^{-1} with a spectral resolution of 0.5 cm^{-1};
- dual input and dual output in the instrument, in order to have a good accuracy of the radiometric calibration, with the acquisition of dual-sided interferograms: the two output ports give complementary interferograms, which allow the increase of the SNR by averaging (it is required a $SNR > 100$);
- detectors with moderate cooling and simple and reliable mechanisms, suitable for space mission;
- a tilt-compensated optical configuration for interferometric alignment, to meet the lifetime requirements (at least 3 years);
- less stringent requirements for shear compensation.

The first requirement leads to take in consideration a polarizing interferometer; in fact, a constant efficiency is offered by this kind of instrument, such as the Martin-Puplett type, which uses a polarizer as a beam splitter (BS) (Martin and E.Pupplet, 1969). This type of interferometer offers the important advantage of dual input and dual output which is a valuable feature for making accurate radiometric measurements. Possible drawbacks concern its exploitation of only a fraction of the source signal, measuring one polarized component of the beam, and the incompatibility with fully tilt-compensated optics.

On the other hand, to maintain the quality of the optical alignment during the interferometric scan, one must use, as requested for space mission, an optical layout in which tilt and shear of the returning beam are unaffected by inaccuracies in the movements. This is one of the most important requirement and looking at other proposed instrumental configurations for space-borne missions (ATMOS, MIPAS, TES, FIRAS for example), none of them meets all the desired requirements. Multiple configurations can be studied and arranged (Carli

et al., 1999), but the only clear solution for tilt compensation in both axes is that of a dual-rooftop mirror, by means of which full tilt compensation without losing the feature of dual input-output is obtained. In the end, a configuration able to meet all the three important requirements (full tilt compensation, dual input and output ports and detection of total signal) was proposed and is described in section 3.2.

In addition to this FTS a set of other instruments are proposed for a space mission: a embedded imager operating in a IR window for scene/cloud identification, a total energy radiometer for referring the FIR fractional component to the total spectrum and a multi-channel imager. Data from this instrumental set would have been used for a number of fields, related to:

- energy balance study, exploiting the wide spectral coverage to monitor climate changes;
- simulation of the REFIR TOA spectrum, to test out ability in the reproduction of FIR radiances;
- determination of cloud properties;
- retrieval of temperature and water vapor profiles.

1.6 The ECOWAR campaign pilot study

The Earth COoling by WAter vapouR emission (ECOWAR) campaign was set up to start filling the paucity of FIR measurements and took place on the Alps, at the Testa Grigia station in March 2007. The main scientific objective was to improve the description of water vapor continuum in the FIR region; in cloudy conditions the experimental setup allowed to study the FIR properties of cirrus clouds. This campaign represents a pilot study for the PRANA project and provides the very first set of far-infrared spectral downwelling radiance measurements, in dry atmospheric conditions.

Looking at the deployed instrumentation, ECOWAR project exploits observations of two FIR Fourier transform spectrometers, able to measure the downwelling atmospheric emitted radiance in the spectral range from 100 to 1100 cm^{-1} with a spectral resolution of 0.5 cm^{-1}. REFIR-PAD instrument was operated in Testa Grigia at an altitude of 3500m a.s.l, while the I-BEST interferometer was deployed in Cervinia, at 2000m a.s.l. where a second base camp was located; this two sites are 5 km apart.
Ancillary information about atmospheric state and cloud geometry are obtained with the help of:

1. a Ground-Based Millimeter-wave Spectrometer (GBMS) for observations of precipitable water, located at the Testa Grigia station;

2. the University of BASILicata Raman Lidar system (BASIL), located at Cervinia, able to perform high-resolution accurate measurements of atmospheric temperature and water vapor and used to derive the cloud geometry and optical depth;

3. a Vaisala radiosonde system located in Cervinia.

In (Maestri et al., 2014) the analysis procedure followed in occasion of a set of measurements in the presence of a cirrus cloud is discussed. This single event is observed through 8 measurement sequences, called 'FOV', covering the time from 12:30 to 13:50 UT on 9 March 2007, when a cirrus cloud of optical depth less than 1 and base at 4.7 km a.s.l. was reported by BASIL. The methodology used in the analysis of the eight FOVs can be summarized in five main steps:

1. determination of the temperature and gas concentration profiles and computation of the gaseous optical depths;

2. determination of cloud geometrical boundaries and profile;

3. retrieval of cloud optical depth and PSD effective dimension from REFIR-PAD spectral radiances, exploiting information in the IR main window;

4. forward simulation of the REFIR-PAD downwelling radiances from 250 to 1100 cm^{-1} using the retrieved cloud parameters and evaluation of residuals between simulations and data;

5. final sensitivity studies.

The atmospheric state is determined using two radiosondes launched from Cervinia during the morning; they show that temperature and water vapor mixing ratio profiles are changing due to the cloud evolution. The selected FOVs fall during a gap in GBMS PWV measurements. CO_2 and other gases concentration profiles are taken from the mid-latitude standard atmospheric model; hence, gaseous optical depths are computed with LbLRTM (Clough et al., 1989) using the high-resolution HITRAN 2004 database.

In order to constrain the simulation scheme, a priori information concerning cloud ice vertical distribution were supplied by the Lidar system; however, the climatological profile is preferred to BASIL one, because the instrument was located 5 km away from REFIR and the cloud structure fairly variable. The cloud geometrical thickness is instead derived from the Lidar system.

Cloud optical depth is retrieved from REFIR-PAD radiances using a limited number of channels in the range 820-960 cm^{-1}, by means of the RT-RET retrieval methodology, described in section 5.2.4. A mixture of habits typical of Mid-latitude cirrus clouds is assumed as a priori information (Bozzo et al.). Regarding RT-RET retrieval uncertainties, the REFIR-PAD measurement errors and the uncertainties in the atmospheric state are the predominant

source when the cloud is optically thin, while the assumptions on particle shape and PSD are the largest source when the cloud is optically thicker.

Cloud parameters derived from RT-RET are used to simulate the observed cloudy scenes over the whole spectrum: in this way, both the accuracy of the retrieval methodology and the consistency of the cloud optical properties can be evaluated. Residuals in selected micro-windows are compared with the 3-σ uncertainty of the mean associated to the REFIR-PAD data in each interval. In Fig.1.6.1, residuals in the FIR are reported: they show a large spectral variability and important differences, while, as expected, in the main IR window the differences simulation-observation (not shown) are smaller than the 1-σ uncertainties.

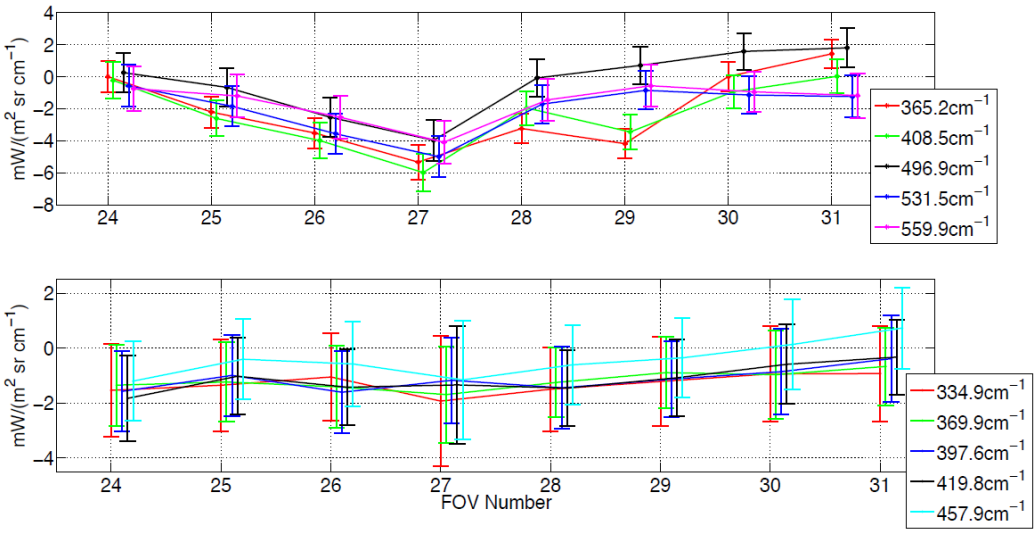


Figure 1.6.1: *Mean differences between simulations and observations and* $3-\sigma$ *uncertainties of the mean for: selected FIR micro-windows (upper panel) and for highly absorbing water vapor regions (lower panel). From Maestri et al. (2014).*

For an assumed temperature profile, the residuals in the FIR are linked to the guessed cloud features and water vapor's profile; the higher the water vapor content, the higher is the downwelling radiance at the ground. The cloud signal is mostly detected in the so called 'dirty windows' where the signal is not saturated by water vapor; however, also in these regions both cloud and water vapor contribute to the measured downwelling radiance. We can say that the difference between REFIR-PAD and simulated data, in the FIR region, is larger when channels of lower wavenumber and cloud optical depth less then 0.5 are considered. With increasing cloud optical depth, the differences are smaller but still larger than the total uncertainties. Looking at the same differences for selected highly absorptive water vapor regions, reported again in Fig.1.6.1, they reflect the accuracy of the assumed temperature and water profile near the ground and show smaller values.

Finally, the sensitivities of the results to particle habit changes and water vapor profile modifications are also studied. Assuming pristine hexagonal columns leads to higher simulated radiances, with smaller differences with measurements; the average residuals are reported together with the result of the control case (ML mixture) in Fig.1.6.2. FIR windows at high wavenumbers show larger sensitivity to assumed habit, with differences increasing with optical depth.
On the other hand, variations of water vapor profile are more important when the sensitivity to particle properties are smaller. The uncertainties related to the water vapor and temperature profiles result to be of the same order as the sensitivity to the a priori assumption on particle habits; but, also varying particle habit and water vapor profile, the differences found in FIR micro-windows are not generally explained.
The main result is that simulation in the FIR range do not show differences whose magnitude should be due to the use of unrealistic ice particle properties.

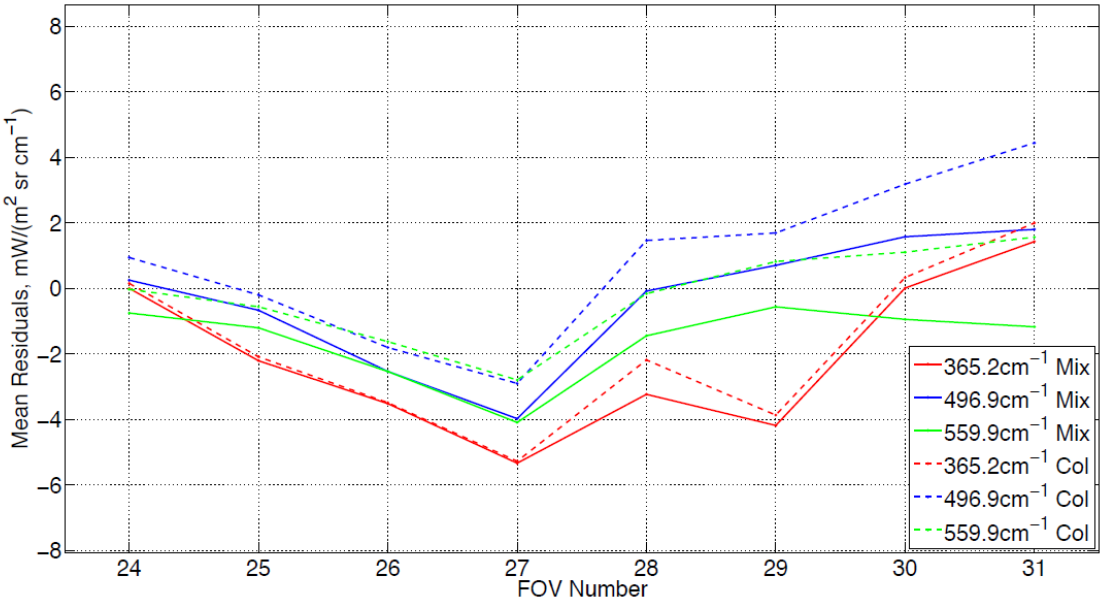


Figure 1.6.2: *Mean differences between simulations and observations for selected FIR micro-windows vs FOV. Results are for the mixture of particle habits (full line) and for pristine solid columns (dashed lines). From Maestri et al. (2014).*

Chapter 2

Antarctic cloud properties

The isolation of Antarctica and the extreme nature of its climate mean that the study of clouds in this area is less advanced than in many other regions of the world. The surface synoptic network is on the whole more widely spaced over Antarctica, with large areas of the continent without any in situ measurement and rare visual observations. Satellite retrieval of cloud presence is also difficult over a snowy surface, especially during dark months, when infrared passive measurements cannot easily discriminate between cloud top and surface emission, mainly due to the frequent low-level temperature inversion. As a consequence, seasonal and area coverage of clouds are very difficult to track (Lachlan-Cope, 2010).

Measuring cloud microphysical properties, such as particle size, particle phase and crystal habit, is particularly difficult in this extreme environment. Hence, less is known about the microphysics of Antarctic clouds than of mid-latitudes ones.
Some measurements have been made in the South Pole region using both remote sensing techniques and occasionally in situ methods. Several groups have used LIDAR to investigate the cloud properties using ground-based instruments and airborne ones, as Morley et al. (1989). During a flight from Antarctic coast to the Plateau in January 1986, cloud properties changed radically: the dense mid-layer water clouds found over the Ross Ice Shelf were replaced by higher ice clouds, predominantly cirrus (Morley et al., 1989). Other groups used radiation measurements to infer microphysical properties, which include ground-based interferometers. Another approach is to measure ice crystal properties directly, either by in situ collection of cloud particle replicas from the clouds that reach down to the surface or from in situ aircraft measurements. There are no aircraft in situ measurements of size on Antarctic Plateau; most information on particle size and habit comes from radiometric measurements, both surface-based interferometer and radiometric sondes, and from ground-based crystal sampling campaigns.

2.1 From radiometric measurements

One of the first studies of cloud properties with remote sensing techniques is developed by Stone (1993), using data collected between 1959 and 1963 at South Pole, during several flights of a radiometersonde, able to measure upward and downward IR irradiances; ancillary meteorological observations are available. The database is quite limited, as only 8 cases of winter clouds are considered. These clouds are generally optically thin with integrated optical depths in the IR equal to 1 or less and bottom heights coinciding with the top of the surface temperature inversion. The principal radiative property for each cloud layer is the effective longwave emissivity, computed using cloud fluxes directly measured during the balloon ascent; a mean emissivity is found equal to 0.6.
The author compares these derived radiative properties with theoretically computed values for model clouds assuming various particle sizes and ice water contents. These calculations are performed using Mie theory for cloud particles represented as ice spheres with surface areas equal to those of the ice-crystals. From this comparison, effective particle radii between 4 and 16 μm and IWCs of 0.3-6 mg/m^3 are found.

Another important study (Mahesh et al., 2001) is conducted at the South Pole Station (SPS) during 1992. Twice-daily longwave atmospheric emission spectra are measured from the surface by means of a FTS with 1 cm^{-1} resolution. The aim is to develop a remote sensing technique to derive cloud properties from High Spectral Resolution (HSR) radiance measurements in the IR region. A year-long dataset of cloud properties representative of the Antarctic Plateau is presented and seasonal variations of these parameters are highlighted.
Knowledge of the cloud-base temperature and the vertical distributions of ozone and temperature are required for the retrieval methodology and are available at SPS from radiosondes and ozonosondes.All spectra are analyzed in this work as if the clouds consisted of ice crystals only, although some of the summer clouds probably contained super cooled liquid water droplets.
The retrieval methodology combines three pieces of information:

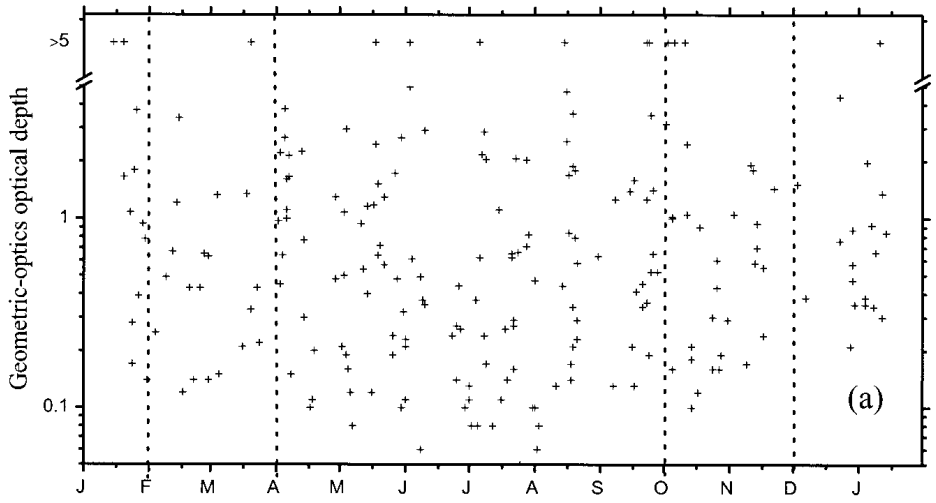
- cloud's emissivity at 903 cm^{-1},
- cloud's emissivity at 988 cm^{-1},
- cloud's transmissivity in the 9.6 μm ozone band.

The variation in the absorption efficiency with particle size is necessary to determine the particle effective radius: the two microwindows at 903 and 988 cm^{-1} are selected because they guarantee both unique solutions and an extreme sensitivity to a broad range of particle dimensions (1-25 μm). Extinction efficiencies, single-scattering albedos and phase function moments are computed using Mie theory for ice spheres. DISORT code is indeed adopted

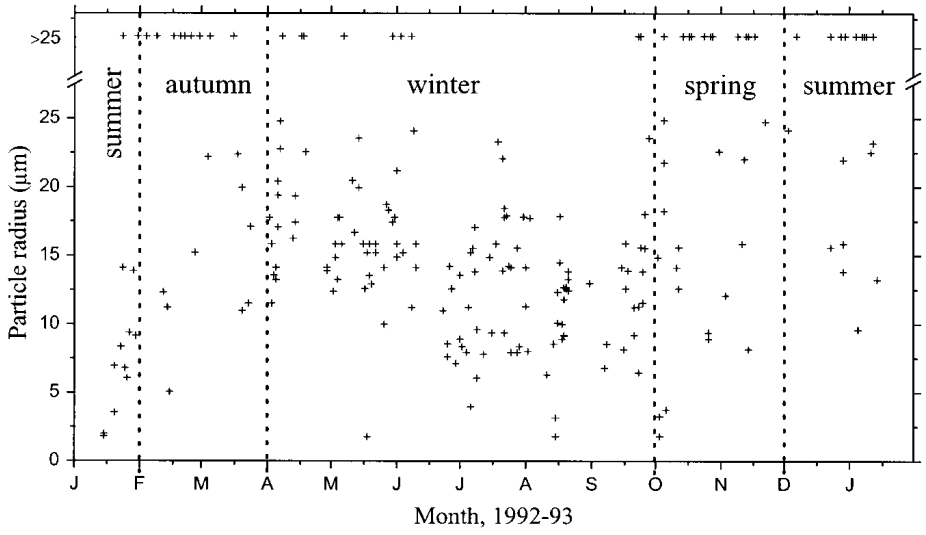
to calculate cloud emissivities, simulating thermal emission radiances from clouds and using observed cloud base temperature. On the one hand, Mie and DISORT simulations are performed for a number of different effective radii and optical depths values; on the other, from each measured radiance spectrum, the above listed parameters are computed. Hence, a comparison between model computations and observations can be performed to obtain best fit r_{eff} and τ. The combination of particle radius and geometric optical depth is chosen so that the differences between measured and computed values of the three parameters, weighted over the reciprocal of the relative uncertainty magnitude, are minimized.
Effective radius of the particle distribution, obtained from spectral measurements, has a median value of 15.2 μm. Since the method is not sensitive to effective radii greater than about 25 μm and approximately 20% of the clouds have retrieved radius > 25 μm, the obtained value could be an underestimation of the real one.

Looking at the cloud properties seasonal cycle, Fig.2.1.1 shows retrieved values of effective radius vs time, for winter and non-winter months separately: $r_{\text{eff}} > 25\mu m$ are rarely seen in winter, while $r_{\text{eff}} < 10\mu m$ are rarely seen during the rest of the year. Comparison of the median retrieved effective radius with ice crystal dimensions collected in situ is also made and described in the next section.
Cloud optical depth τ_g, considered in the geometric-optics limit, is not uniquely determined for extremely thick clouds ($\tau_g > 5$); for these cases, only a lower limit is indicated. From an overall point of view, fewer than 10% of the clouds have $\tau_g > 5$ and about two-thirds have $\tau_g < 1$. But, separating observations for the winter and non-winter months, it is evident that optical depths greater than 5 are twice as common in non-winter months as in winter months. A predominance of thin clouds is on average observed with respect to clouds on the Antarctic coast (Ricchiazzi et al., 1995). Three tendencies are found: larger crystals in summer than in winter, larger optical depths in summer than in winter and larger optical depths at the coast than in the interior; these results conform to expectation, because the atmosphere contains more water vapor at the coast than in the interior and more water vapor in summer than in winter.

Lubin and Harper (1996) estimated effective cloud particle radius using the Advanced Very High Resolution Radiometer (AVHRR) channels at 11 and 12 μm; this study was conducted in overlap with Mahesh et al. (2001) ground-based measurements, so that independent information about clouds can be derived.
About 300 satellite datasets are exploited to calculate radiative properties of ice clouds. Ebert and Curry (1995) ice cloud parameterization is assumed; Mie theory is then used to calculate optical properties for an ice crystal PSD n(L), where L is the crystal length, characterized by an effective radius that corresponds to spheres of equal surface area. In this way, cloud emissivities can be modeled and expressed as functions of r_{eff} and Ice Water Path



(a)



(b)

Figure 2.1.1: *Upper panel (a): retrieved cloud optical depth in the geometric-optics limit at the SPS during 1992-1993; lower panel (b): retrieved effective particle radii r_{eff} vs time, with winter and non-winter months distinguished. From Mahesh et al. (2001).*

[g cm^{-2}], using the formula:

$$\epsilon_\nu = 1 - \exp -\beta k_\nu IWP$$

where β is the diffusivity factor fixed and k_ν the mass absorption coefficient [cm^2 g]. Brightness temperature in the two selected AVHRR channels are then computed, given a surface temperature value from the South Pole Weather Office and an estimate for the effective cloud temperature T_c. In the retrieval methodology, a trial value for the cloud temperature and a pair of cloud emissivities are given; then the values of T_c, r_{eff} and IWP that best represent AVHRR observation are those that minimize the squared difference between satellite-measured brightness temperature and the one given by the model at TOA.

Considering the lack of sensitivity in the retrieval method for large ice particles, the authors obtain a mean effective radius value of 12.3 μm in summer and 5.6 μm in winter; these values are smaller than the ones found by Mahesh et al. (2001) over the same period. They suggest that this discrepancy may be due to smaller crystals at cloud top than at cloud bottom. However, a clear seasonal cycle is found, since ice summer crystals are larger than winter ones. In addition, cloud top temperatures about 180-220K and optical depths about 1 are reported.

Another remote sensing study using satellite instrument is reported by Grenier et al. (2009). In this work satellite observation of CloudSat radar reflectivity and of CALIPSO Lidar backscattering are presented, for July 2007 over the Antarctic continent; the satellite observations are used to heuristically classify polar thin ice clouds into two groups on the base of their crystal dimensions. In fact, while Lidar is sensitive also to small cloud particles, its signal saturates in the presence of thick clouds contrary to the CloudSat radar, whose wavelength is 3mm. These different capabilities allow the discrimination of cloud composed of particles with $r_{\text{eff}} < 28 - 30\mu m$ from clouds with bigger ice particles. Hence, two main categories are identified, separating precipitating from non-precipitating clouds. On the Antarctic Plateau high and thin clouds are mainly observed within 25% of the analyzed profiles around 6-9 km a.s.l, while low clouds (under 5-6 km) are detected only 10% of the observations.
The authors suggest that absolute estimates of r_{eff} from CloudSat-CALIPSO are possible, but no comprehensive investigations of this property over Antarctica have been published to date.

2.2 From in situ measurements

In situ measurements of cloud ice crystals on the Antarctic Plateau do not exist until today, but some campaigns have been organized at the SPS to sample precipitating ice crystals, collected and classified on the base of size and habit. These crystals result to be bigger than clouds ones, generally between 2 and 1000 μm.
The presence of atmospheric ice crystals near the ground is a quite ubiquitous phenomenon with three different main origins:

- Blowing snow: snow particles lifted from the surface when the wind speed is greater than about 7 m/s; these particles tend to be rounded, due to several collisions. This phenomenon undergoes a seasonal cycle in frequency of occurrence, particle size and vertical extent.
- Diamond dust: small ice crystals forming in the persistent temperature inversion layer, particularly strong during winter; in this layer, vertical mixing causes the boundary layer air to become supersaturated with respect to ice. Frequency of occurrence of this phenomenon reach 63% during winter months (Walden et al., 2003).
- Snow grains: these are larger ice crystals falling from clouds.

First studies that focused on analysis of atmospheric ice crystals at the SPS, were performed during the '70s, especially covering summer months. Hogan (1975) and Ohtake (1978) show that during antarctic summer months, most observed habits are columns, bullets and bullet clusters, named 'rosette shaped', but also hexagonal plates and pyramid shaped (small diamond dust); the former author notes that the 650-600 hPa layer is supersaturated with respect to ice during the entire observation period and that cirrus cloud cover is observed at all times when precipitation occurs. On the contrary, Kikuchi and Hogan (1979) during a campaign of observation between January and February 1974, observe the 'Diamond Dust' phenomenon also in clear sky conditions, due to homogeneous nucleation under $-40°C$.
Ohtake and Yogi (1979) study precipitating ice crystals during winter months June and August 1979. They suggest that bullet clusters are the most common crystal shapes also for winter, with dimensions up to 1 mm, hence representing the major contribution in terms of mass. This habit is observed predominantly in association with high cirrus clouds and cirrostratus around 1-3km a.g. and temperature at 500 hPa about $-40/-55°C$. Moreover, the authors observed combinations of plates, bullets and columns in correspondence with lower clouds, with base at about 300-1000 m, linked to humid air advection in middle troposphere and temperature of $-35\,to-45°C$. The fall of little ice crystals (with dimension up to 100 μm) is reported too, i.e. diamond dust in the form of hexagonal plates, columns or polyhedral shapes often with clear sky or low clouds.

One of the most important studies on the collection and analysis of atmospheric ice crystals is performed by Walden et al. (2003) during a campaign at the SPS between June and September 1992. A gridded microscope slide is used to collect ice crystals: 84 images are studied, detecting 20000 ice crystals that represent a significant sampling of wintertime ice crystals because they are collected under all types of meteorological conditions.
Authors divide shapes into three main categories linked to atmospheric formation process: 'Diamond dust' (especially columns and hexagonal plates), 'Snow grains' (as bullet clusters) and 'Blowing snow' (as grains/little ice spheres). Fig..2.2.1 reports relative frequencies of occurrence of the different ice crystal types together with the relative contribution to total surface area and total volume.

Crystal type	Relative frequency (%)	Contribution to total surface area (%)	Contribution to the total volume (%)
Diamond dust			
Solid columns	10	15	7
Plates	9	12	9
Pyramids	<1	<1	<1
Hollow columns	<1	<1	<1
Blowing snow			
Blowing snow	**72**	**46**	31
Residual blowing snow	8	6	5
Snow grains			
Sector-plate clusters	<1	8	5
Bullet clusters	<1	11	**39**
Snow-grain columns	<1	1	3

Figure 2.2.1: *Relative frequency of occurrence of the different ice crystal types from the collected photomicrographs, as well as the relative contribution to the total surface area and total volume. Values in bold type are the maximum values. From Walden et al. (2003).*

For each collected crystal, its dimensions are used to calculate the respective surface area and volume. The size distributions for each crystal type are generated by first representing the non-spherical ice particles by a collection of 'equivalent' spheres. Spheres with the same volume-to-area ratio (V/A) as the crystals are used by the authors, because of the importance of V/A in calculating radiative fluxes and heatings rates in atmospheric radiation models. Grenfell and Warren (2010) have shown that equal V/A spheres provide a good approximation for the extinction efficiency, the single-scattering albedo and the asymmetry

parameter over a wide range of wavelengths from the ultraviolet through the IR. Hence, once calculated volume V and area A for each crystal, the non-spherical particle is represented by means of a collection of independent 'equivalent' spheres, each of which has a radius

$$r_{\mathrm{VA}} = 3\frac{V}{A}$$

and so the same volume-to-area ratio as the original crystal. In order to conserve the ice mass a collection of independent spheres is indeed used, collection that has the same total volume and the same total area of the original crystal. The number of these identical spheres is computed as:

$$n_{\mathrm{s}} = \frac{3V}{4\pi r_{VA}^3}$$

Finally, for each crystal type, is obtained a dimensional distribution in terms of equivalent spheres and the effective radius for this PSD is calculated as the area-weighted mean radius from Hansen and Travis (1974) formulation:

$$r_{eff} = \frac{\int_{r_{min}}^{r_{max}} r^3 n(r) dr}{\int_{r_{min}}^{r_{max}} r^2 n(r) dr}$$

in fact, this single value for the PSD proves to be very significant for radiative transfer calculations. Fig.2.2.2 reports the effective radius for each crystal type and also for the three major categories of crystals. Snow grains are larger than both diamond dust and blowing snow, with an effective radius of 24 μm.

Finally, the authors suggest that, although the diamond dust formation in layer near the surface can suggest the possibility of an indirect study of cirrus cloud inner processes, diamond dust measurements may not be representative of high cloud particles, since cirrus cloud experience greater turbulence and faster crystal growth. From mid-latitude measurements, it appears that cirrus crystals could be similar to the snow grains component, especially for higher temperature than these.

The above procedure, to derive r_{eff} from a crystal collection, is validated by Mahesh et al. (2001) that, in addition to inferring effective radius from ground-based IR remote sensing, also has collected atmospheric ice crystals for 100 days from June to October 1992. Dimensions of 14000 crystals are measured and equivalent spheres are determined for each non-spherical particle using three different specifications: equal area, equal volume and equal V/A. On average, the particle effective radius retrieved by remote sensing results to be in agreement with the one obtained for the PSD computed using the last specification: a median value about 14.1 μm is found. On the contrary, equal-volume and equal-area radii are unrealistically large. Accordingly with the authors, this agreement is quite surprising, because one would expect crystals falling out of the cloud to be larger than the average for

Crystal type	Effective radius, r_{eff} (μm)
Diamond dust	12.2
Solid columns	10.1
Plates	15.1
Pyramids	9.4
Hollow columns	6.4
Blowing snow	11.0
Blowing snow	10.8
Residual blowing snow	11.9
Snow grains	23.6
Sector-plate clusters	11.9
Bullet clusters	25.2
Snow-grain columns	67.0

Figure 2.2.2: *Effective radius for PSDs of the individual crystal types and the three major categories. From Walden et al. (2003).*

the cloud. However, the radii from the photographs are not correlated with the remotely sensed radii, due to the difference of 2-6 h between the time the crystals were collected and the time of spectral measurements.
In addition, the authors consider also the measurements of crystals dimensions from Kikuchi and Hogan (1979) and computed the average effective radius, in terms of equal-V/A spheres: they obtain r_{eff}=17.5 μm for summer months.

Another important campaign of in situ collection of precipitating ice crystals is presented by Lawson et al. (2006). In this study, 900000 precipitating ice crystals digital images are recorded at the SPS, between 1 and 8 February 2001. A cloud particle imager (CPI) is used and an automatic crystal habit classification algorithm is developed, considering crystals with dimension bigger than 50 μm only. Three classes are identified:

- 'Rosette Shaped' such as bullet clusters and aggregations of plates; these are polycrystals typically formed from rapid freezing of a supercooled water drop, hence linked to mixed phase clouds.

- 'Diamond Dust' typically observed under thin, high clouds that are penetrated by the sun's rays and produce optical effects; columns, thick plates and hexagonal plates are classified under this label.

- 'Irregular' shaped crystals are both non-pristine precipitating particles and blowing snow grains; the contamination from blowing snow make impossible to obtain a quantitative assessment about the fraction of non-pristine precipitating crystals.

In Fig.2.2.3 examples of different crystal shapes are reported .

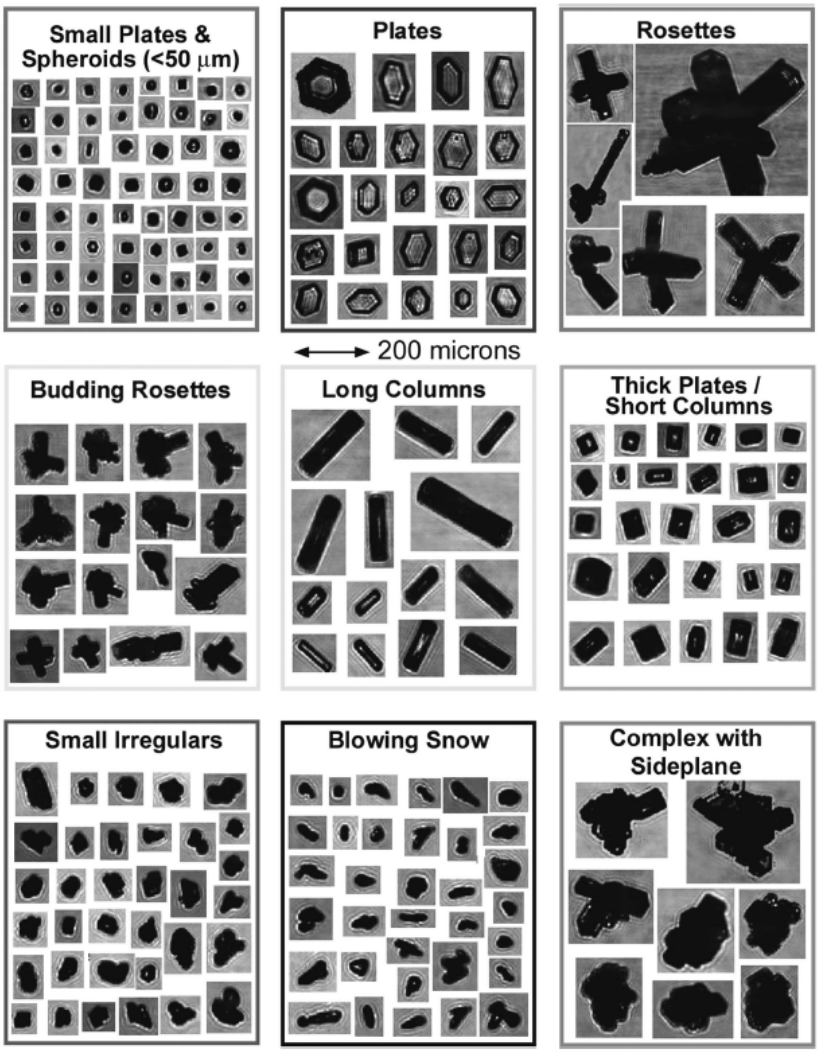


Figure 2.2.3: *Examples of Cloud Particle Imager observations classified into habits categories. From Lawson et al. (2006).*

Crystals classified as Diamond Dust result to have dimensions generally less than 150 μm and are observed in the presence of faint cirrus clouds but also in clear conditions. On the contrary, Rosette shaped crystals show dimensions up to 250 μm ; although rosettes

are the less frequent in number, are the most important for overall precipitation mass, confirming observation from Walden et al. (2003) and Ohtake and Yogi (1979). Results of this classification are reported in Fig.2.2.4, which shows histogram of ice crystal habits as percentage weighted by concentration, area and mass.

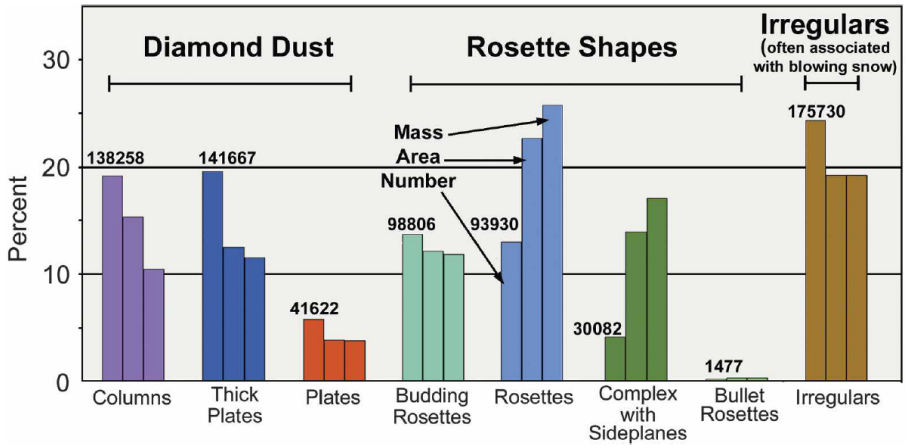


Figure 2.2.4: *Histogram of ice crystal habits observed at SPS during 1-8 Feb 2001. The percentage weighted by concentration, area and mass is shown for each habit category*

Following the methodology used by Walden et al. (2003), authors compute PSDs for each habit class as a function of the radius of equivalent spheres, adopting the V/A prescription; but, differently from Walden et al. (2003) and Mahesh et al. (2001), they do not compute also effective radii for the obtained PSDs but deals with r_{VA} only. Hence, these results should not be directly compared, as authors do, with the ones from the other two articles. However, results expressed in terms of equivalent sphere radius are reported in Fig.2.2.5.

	Equivalent Radii (μm)		
	Summer		
Diamond dust	17.0		
Columns	10.5		
Plates	19.2		
Bullet clusters	25.3	27.2	27.9
Blowing snow	17.0		

Figure 2.2.5: *Mean equivalent radii for various ice crystal habits; the three values for bullet clusters respectively refer to budding rosettes, rosette shapes and complex rosette crystals. Form Lawson et al. (2006).*

Chapter 3

The PRANA project: REFIR-PAD and other instruments

All the measurements examined in this work belong to the framework of the 'Proprietà Radiative del vapore acqueo e delle Nubi in Antartide' (PRANA) research project; PRANA activities started in December 2011 with the XXVII scientific expedition and are supported by the Italian 'Programma Nazionale di Ricerche in Antartide' (P.N.R.A.). Measurements still continues and the field station is located at the Italian-French Concordia scientific Station, Dome-C, on the Antarctic Plateau ($74°30'S, 123°00'E, 3.280m\,a.s.l.$), located in the map in Fig.3.0.1

The main scientific objective of the PRANA project is the study of radiative properties of water vapor and clouds in the thermal spectral region with particular attention to the unexplored FIR region, monitoring in this way downwelling longwave radiation emitted by the atmosphere in different sky conditions. In detail, specific aims attains to:

- improvement of the spectroscopic knowledge of the pure rotational water vapor band in the FIR;
- study of the spectral features of clouds and PSC at these wavelengths;
- improvement of radiative transfer forward model and retrieval methodology with regard to FIR spectral region;
- validation of atmospheric retrieved parameters by means of data provided by other on-site or satellite-based sensors.

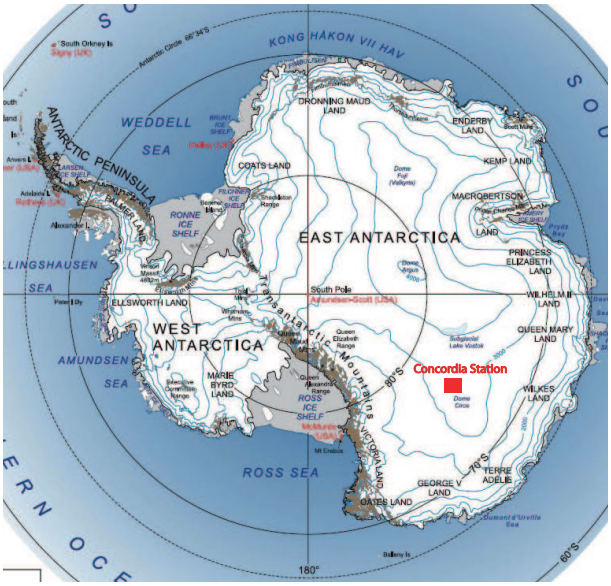


Figure 3.0.1: *Antarctic map with Concordia Station position* ($74° 30' S, 123° s00' E$)

In this contest, the spectral measurements of the downwelling radiance is performed by REFIR-PAD instrument (REFIR-Prototype for Applications and Development), which is a prototype developed at IFAC for field applications. Cloud characterization are performed by a Lidar system and ICE-CAMERA, a scan camera that acquires a high resolution image of ice crystals fallen on a glass every hour, in this way a continuous monitoring and classification of precipitating ice crystals can be performed (but no article about this research field has been published since now). All instruments are installed in the Physics shelter devoted to atmospheric physics, located 500 m from the main base, in the so called "Clean Air Area"; moreover, a VAISALA weather station is installed on the roof of this shelter, look at Fig.3.0.2, providing continuous observations of local meteorological conditions.

3.1 REFIR-PAD instrument details

The REFIR-PAD instrument was developed at IFAC-CNR as a demonstration of the feasibility of the REFIR project; this FTS is suitable to fly on a stratospheric balloon and comes after a prototype named REFIR-BB, used to test the base concept of the instrument (Palchetti et al., 1999).
REFIR-PAD, as used during PRANA project, provides spectrally-resolved zenith-sounding radiance measurements in the 100-1500 cm^{-1} range with a 0.4 cm^{-1} resolution, covering in

Figure 3.0.2: *Physical Shelter where Lidar system and REFIR-PAD instrument are located and on which the VAISALA station is set*

this way the most part of the Earth's long-wave emission and including both the far-infrared and the better known mid-infrared region.

REFIR-PAD has been designed to operate with uncooled detectors and optics; moreover, the optics and the electronics modules are both integrated in a compact package. The instrument is capable of working both in a polarizing mode, using photo-lithographic wire-grid beam splitters (a double Martin-Puplett scheme) and in an amplitude mode, in which the input polarizers are removed and two Germanium coated Mylar beam splitters are used (a Mach-Zehnder optical scheme).

3.1.1 Martin-Puplett configuration

The optical layout shown in Fig.3.1.1 refers to a polarizing interferometer in its classical configuration: it is composed of an input polarizer P_1 that transmits the vertical plane of the polarization of the analyzed source I_1 and reflects the horizontal plane of polarization of a reference source I_2 (two input ports). Polarizing BS P_2 has its principal axis oriented at 45° with respect to the direction of polarization of the two sources: in this way each source is split into two polarized components of equal amplitudes which enter the 2 arms of the interferometer. A rooftop mirror in each arm of the interferometer reflects the beam back from that arm and introduces a folding of the wave front about the dihedral edge of the mirror. This folding is about a vertical axis (because the mirror is vertical) and introduces a rotation of the plane of polarization by 90°. So, when the polarized components return to the beam splitter, the one that was originally transmitted is now reflected and the other is now transmitted; in this way the full input signal is transmitted to the output side and

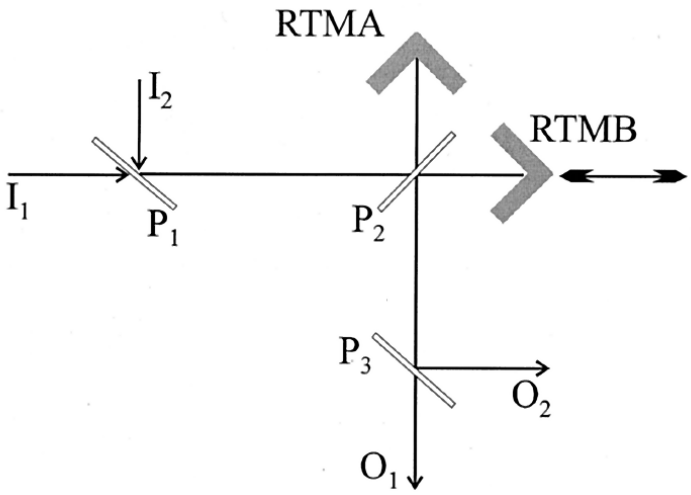


Figure 3.1.1: *Optical configuration of a polarizing interferometer in the Martin-Puplett configuration*

the components that travel different paths do not interfere because they have perpendicular polarization. It is necessary a third polarizer P_3 that creates two output (transmitted and reflected O_1 and O_2) in which the components that traveled different paths now have the same polarization and can interfere; in fact its principal axis is at 45° with respect to that of the BS, (Martin, 1982).

3.1.2 Mach-Zehnder configuration

As shown if Fig.3.1.2, this configuration is characterized by the use of two beam splitters and two mirrors to divide (into sample beam SB and reference beam RB) and recombine them finally. One important and distinctive feature concerns the fact that the two paths are widely separated and are traversed only once. Recalling that a beam splitter is substantially a piece of glass with a dielectric or metal coating on the surface, a reflection of the beam on it induces a phase shift of π, whereas a transmitted photon picks up no phase shift. The total path lengths the light travels in the upper and lower paths respectively, are indicated as l_1 and l_2, while the light that passes through the glass of the beam splitter picks up an extra phase shift, called $2\pi t/\lambda$, where t is linked to its optical thickness and refractive index (Zetie et al., 2000). In the absence of a path difference in the two arm path length ($l_1 = l_2 = l$), both the sample beam SB and the reference beam RB will arrive in phase at Output A, yielding constructive interference. In fact, both SB and RB will have undergone a phase

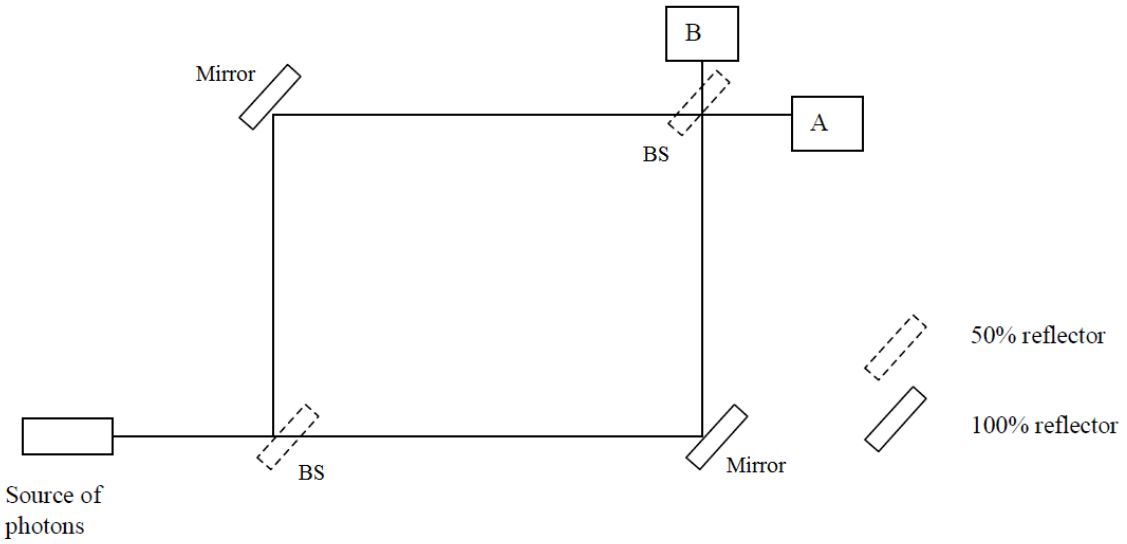


Figure 3.1.2: *Optical configuration of a Mach-Zehnder interferometer type*

shift of:

$$2\pi + 2\pi\left(\frac{l+t}{\lambda}\right)$$

due to two front-surface reflections each of which gives a 180° phase shift and one transmission through a beam splitter.

At detector B, in the same situation, the sample beam and reference beam will arrive with a phase difference of 180°, yielding complete destructive interference. The RB arriving at detector B will have undergone one front-surface reflection and two transmissions while the SB arriving at detector B will have undergone two front-surface reflections and one rear-surface reflection. So, the phase difference of the two paths on their way to detector B are:

$$2\pi + 2\pi\left(\frac{l+2t}{\lambda}\right) - \pi + 2\pi\left(\frac{l+2t}{\lambda}\right) = \pi$$

Therefore only detector 1 receives light. If a path difference is introduced moving the mirror ($l_1 \neq l_2$), the intensities of the beams entering the two detectors will change, allowing the formation of two complementary interferograms.

3.1.3 REFIR-PAD configuration

In this campaign the instrument is used in the second configuration. The beam-splitter were specifically studied for REFIR-PAD, in order to meet the above requirements and the gain in

theoretical efficiency given by the $0.85\mu m$ Germanium coated onto the $2\mu m$ Mylar substrate with respect to Mylar alone is substantial.

The instrument is built on two floors that have equal optical components. The beam goes from the lower floor to the upper floor by means of the two rooftops mirrors, which are part of a solid unit based on double-parallelogram design with flexural pivots that provides the interferometric sweep, with alignment compensated for tilt in both directions and for shear in one direction. The beam that enters the BS on either side will exit on a different floor but from the same side. An asymmetry and a small misalignment are introduced on purpose between the two paths to ensure that the residual coherence does not lead to measurable interference. The optical layout also permits the compensation of eventual deviations of RTMA and RTMB from perfect 90° rooftops: BS_2 is made from two separated units. The relative orientations of these two units can be used for the correction. Fig.3.1.3.

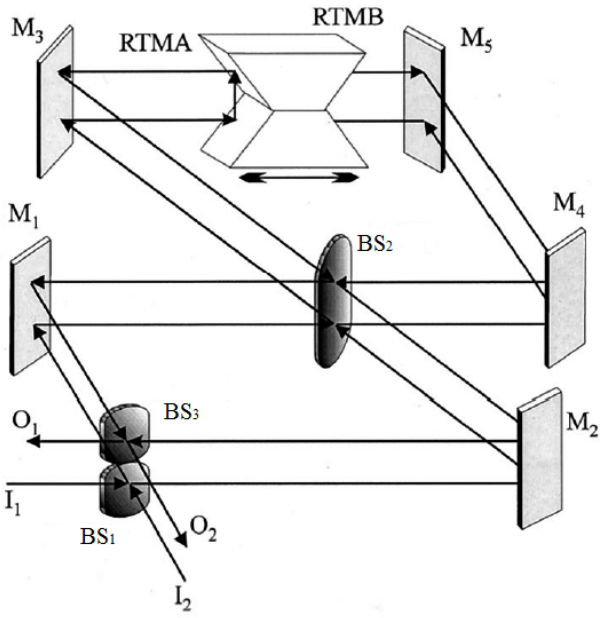


Figure 3.1.3: *New proposed configuration of double input-output Mach-Zehnder interferometer with double signal and tilt compensation.*

One of the input port is coupled to a reference black body (RBB), while the other, through a scanning mirror, can be switched between two calibration sources at different temperatures, i.e. a hot and a cold blackbodies (HBB and CBB), and three different possible atmospheric views, for our case a zenith one, Fig.3.1.4.

The detectors are placed on a plate thermally stabilized at around $30°C$ and two Winston cone provide the coupling between the output paraboloids and the detectors. Both the

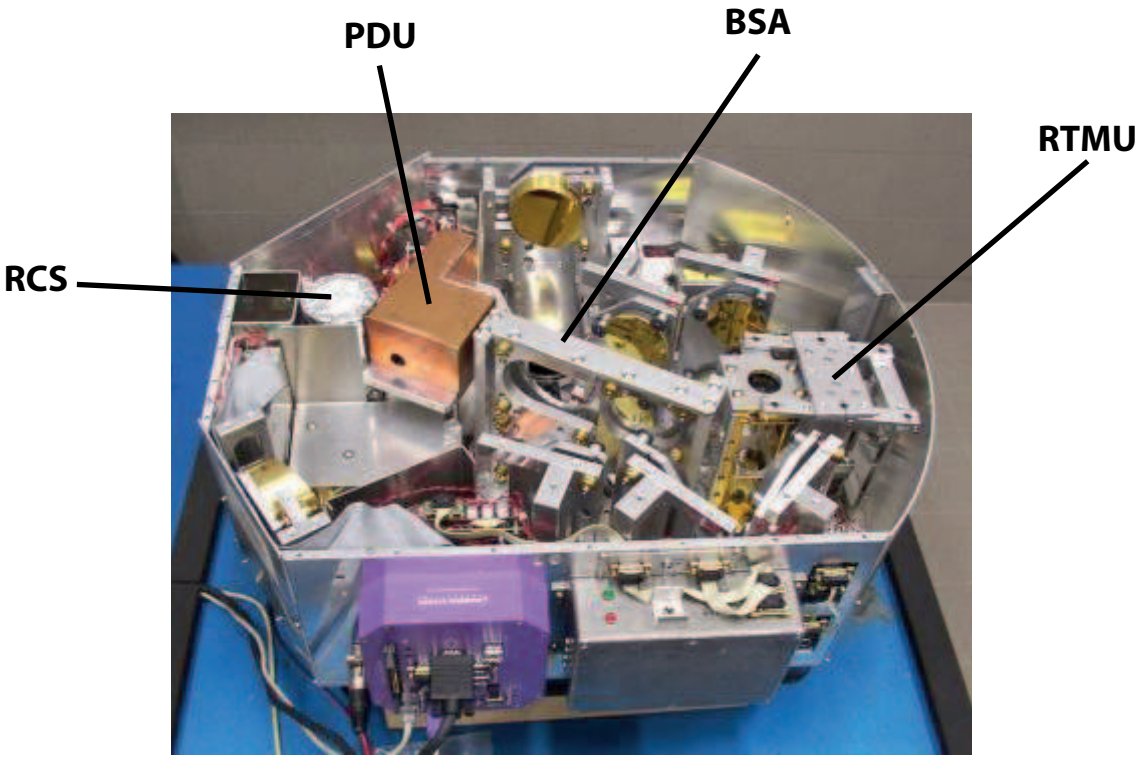


Figure 3.1.4: *Compact design of REFIR-PAD, are shown the roof-top mirror unit (RTMU), the beam splitter assembly (BSA), the infrared detector unit (PDU) and the input optics with reference calibration system (RCS)*

detectors are deuterated L-alanine-doped triglycene sulfate (DLATGS) pyroelectrics, modified removing the original CsI window and enclosed in a sealed case with a polypropylene window, in order to extend the response to low wavenumbers down to 100 cm^{-1} (Bianchini et al., 2006).

REFIR-PAD is a fast scanning spectroradiometer with signals acquired in the time-domain and resampled in post-processing at equal intervals in optical path difference. The interferometer scanning mirror positions is monitored by means of a reference laser interferometer using a stabilized solid state semiconductor laser source at wavelength of 780 nm. Tab.3.1 reports the main features of REFIR-PAD as used during the measurement campaign in Antarctica.

The resampling of the interferogram signal in the optical path difference domain, after the acquisition in time domain, requires that no phase delay exists between the infrared interferogram and the laser which provides the reference interference fringes. This is not simply achievable because the pyroelectric detectors show strongly frequency-dependent response

Instrument type	Mach-Zehnder type non-polarizing FTS
Beam splitterr	Broadband Ge-coated Mylar
Operating spectral bandwidth	100-1400 cm^{-1}
Operating spectral resolution	0.4 cm^{-1} (double-sided)
Operating throughput	0.01 $cm^2 sr$
Field of view	133 mrad
Detector type	Room temperature pyroelectric (DLATGS)
Acquisition time	30 s
Acquisition frequency	20 kHz
Weight	55 kg
Power consumption	$\sim$ 50 W

Table 3.1: *REFIR-PAD instrument main characteristics in the amplitude-division operating mode used in the campaign*

and a consequence dephasing. The infrared detector preamplifiers flatten the response of the system and partly address the problem, but a numerical correction is needed. Interferogram not perfect sampling, in particular as regard the zero path difference point, causes a shift of the entire spectrum, because the mis-evaluated mean radiance value. The noise characteristics of REFIR-PAD depend not only on the pyroelectric detector noise figure, but also on the scanning speed that controls the acquisition time.

3.2 Error Analysis

The REFIR-PAD data analysis work-flow, here identified as level 1a, is peculiar of FTS instruments and has the purpose of obtaining the uncalibrated spectra and the housekeeping data from the instrument raw data (Bianchini and Palchetti, 2008).

As regard detector system, a linear response is assumed through the full dynamic range experienced during REFIR-PAD measurements. The non-linearity estimate error is less than 0.2%, which produces a negligible effect on the calibrated radiance compared to radiometric uncertainties due to detector noise and the accuracy with which reference blackbody temperatures are known. The electronics contribution is equal for the two channels and about 20 dB below detectors noise through the operating frequency range of the REFIR-PAD instrument.

The interferogram sampling method used in REFIR-PAD is at equal-time intervals, so the first part of the level 1 data analysis is devoted to the filtering and resampling on equally spaced optical path difference intervals. The interferograms from the two infrared detectors and the reference interferometer first are band pass filtered to remove baseline effects and high frequency noise. A procedure of identification and correction of the transient phenomena

is applied to the infrared signals; also a compensation of the acquisition system response is needed, since a time delay between reference laser and infrared signals leads to significant sampling errors. Finally, the infrared signal is resampled in the OPD domain on the zero crossing points of the reference interferometer signal. If a constant resolution is needed, the interferograms are cut in order to obtain the desired resolution.

The second part performs the Fourier transform of the resampled interferograms and the phase correction, in order to obtain the uncalibrated spectra. The acquisition point of the interferogram that is nearest to the Zero Point Difference is found and a complex FFT is performed using that point as ZPD. The same interferogram with a greatly reduced resolution is used to obtain the frequency dependent phase, with which the full resolution spectrum is phase corrected.

The calibration of the housekeeping data is performed in the third part of the level 1a procedure. Auxiliary information used for diagnostic and instrument calibration is acquired through 20 housekeeping channel, sampled at a lower rate than the detector. Housekeeping signals can be divided in three categories: environmental monitoring, instrument diagnostics and blackbody temperature measurements. Some parameters are used to provide information about possible systematic errors and others as diagnostic tools in case of instrument failure.

The level 1a data, i.e. uncalibrated spectra and calibrated housekeeping data, are then processed in order to calibrate spectra in terms of radiance, thus obtaining the level 1b products. Frequency calibration is performed using reference laser (Bianchini et al. 2010). Radiometric calibration is however more complex and is obtained calculating the instrumental response using on-board blackbody sources. Generally the response of the two input are different, depending on the optical layout characteristics, so we consider two calibration function F1(σ) and F2(σ), i.e a frequency dependent complex proportionality factor. In ideal conditions, the instrument output is proportional to the difference of the two input weighted with the calibration function. For each the uncalibrated spectrum S(σ), defined as the phase-corrected Fourier transform of the equal-spaced interferogram, is related to the calibrated radiance spectrum L(σ) through the following relationship:

$$S(\sigma) = F1(\sigma)L(\sigma) - F1(\sigma)B_r(\sigma)$$

Where $B_r(\sigma)$ is the RBB radiance. F1(σ) and F2(σ) can be obtained from a two point radiometric calibration procedure performed by means of the measurement of the radiance of the two calibration blackbodies HBB and CBB. Finally the calibrated radiance L(σ) is calculated from the uncalibrated spectrum S(σ) and the $B_r(\sigma)$ expression, by means of the following formula:

$$L(\sigma) = \Re\left\{\frac{S(\sigma)}{F1(\sigma)} + \frac{F1(\sigma)}{F2(\sigma)}B_r(\sigma)\right\}$$

The absolute radiometric uncertainty on the calibrated spectra depends on both the measu-

rement precision (linked to S(σ)) and the accuracy with which the calibration procedure is performed (linked to $B_r(\sigma)$). The measurement precision is calculated in term of the Noise Equivalent Spectral Radiance (NESR), which is dominated by the detector noise and produces independent fluctuations for each spectral element, whereas the Calibration Accuracy (CAL) is dominated by the measurement accuracy of the reference blackbody temperatures (systematic error component), which produces correlated radiometric error among different spectral elements (Bianchini and Palchetti, 2008).

The NESR is calculated from the 1-σ uncertainty on the uncalibrated spectrum as a function of the frequency. While acquiring multiple spectra, if enough stability in the reference and in the measurement source can be maintained, the uncertainty for each spectral element in an uncalibrated spectrum is obtained from the standard deviation of multiple measurements, smoothed to eliminate high frequency dependence. The final value of NESR on each output channel can be obtained from the uncalibrated error through error propagation. Considering a general case in which the measured spectrum S is the average of N acquisitions and calibration measurements S_h and S_c (HBB and CBB radiances respectively) that come from an average of n single spectra (i.e. calibration sequences), then:

$$NESR = \sqrt{\frac{1}{N} + \frac{2}{n}\left(\frac{S}{S_h - S_c}\right)^2}\frac{\Delta S}{F1} \tag{3.2.1}$$

where S, Sh and Sc are the averages respectively of the measured radiance, the HBB radiance and the CBB radiance, while ΔS is the 1-σ uncertainty on the uncalibrated measurement. The term $\frac{S}{(S_h - S_c)}$ contains a spectral dependence and when the measured radiance is near to the reference blackbody emission, the number of measurements N predominates. In the other case, a reduction of uncertainty is better obtained through a greater number n of calibration sequences. The condition N=4 and n=2 gives the best compromise with the minimum spectral variability on the calibrated uncertainty. For zenith-looking ground-based measurements, the uncalibrated error is lower due to the higher integration time that was allowed by slower atmospheric variability observed by the instrument from a fixed position. An example of the result of the uncertainty estimation procedure is shown in Fig.3.2.1

The calibration accuracy is obtained through the error propagation of the temperature 1σ error measured on the reference blackbodies. If we consider temperature error of each blackbody as independent, we obtain the calibration error CAL from the equation:

$$CAL = \sqrt{\Delta B_r^2 + \left(\frac{S}{S_h - S_c}\right)^2 (\Delta B_h^2 + \Delta B_c^2)} \tag{3.2.2}$$

The final spectrum is obtained from the weighted average of both output channels for a single acquisition sequence including 4 zenith observations and 2 calibrations. Each observation has a duration of 64s, so the total acquisition time is about 10 minutes. The estimated

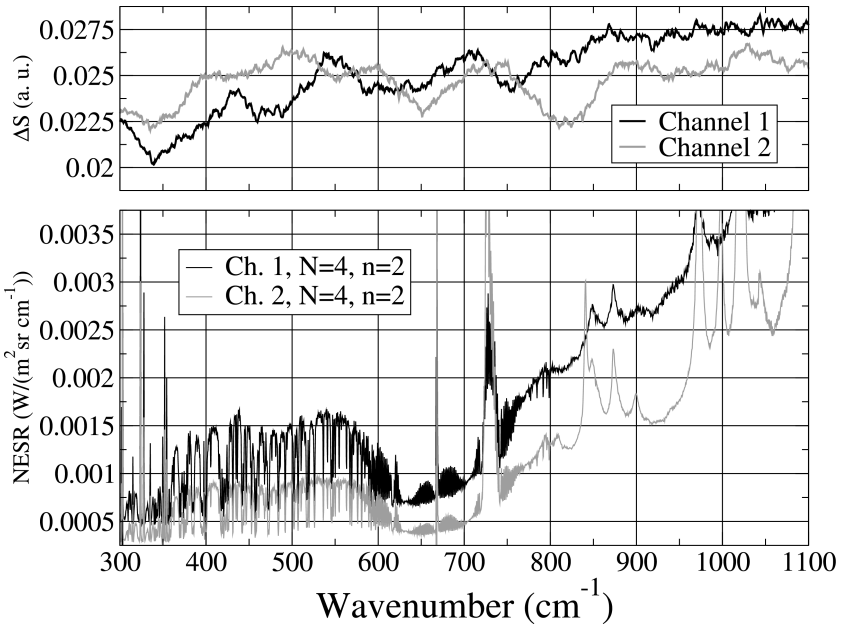
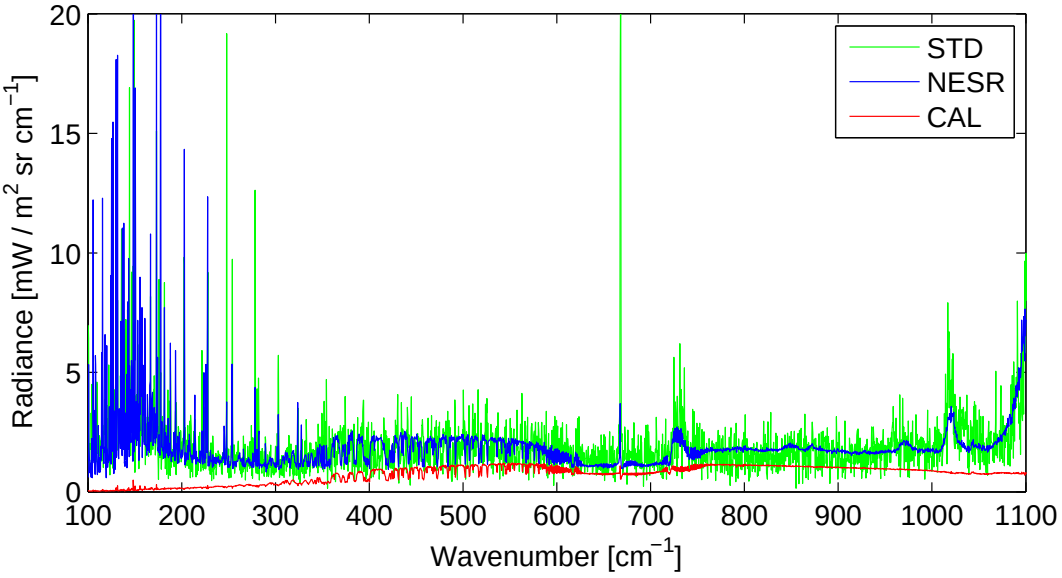


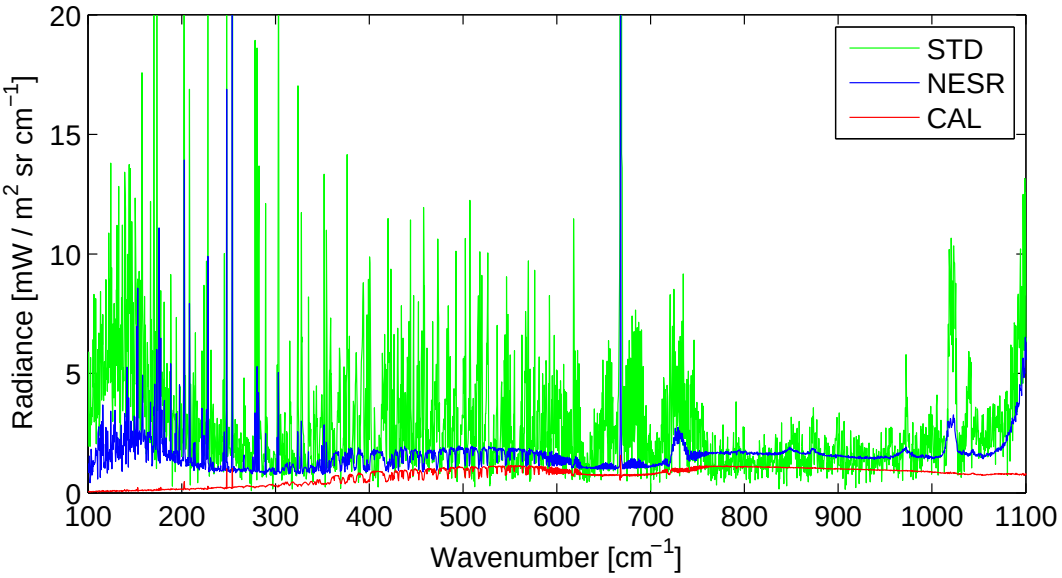
Figure 3.2.1: *Example of uncertainties on uncalibrated spectra (upper panel) obtained as standard deviation of multiple measurements, and NESR on calibrated spectra from error propagation (bottom panel) in a case of 64s acquisitions. From Bianchini and Palchetti (2008).*

uncertainty in the measured spectra can also be validated, through the comparison between the estimated NESR and the statistical standard deviation of the observed spectra (STD, i.e. the simple standard deviation for a given set of measurements), in the presence of a stable atmospheric scene. This comparison can be made, in particular, between the NESR calculated for N=4 and n=2 and the STD calculated from 6 measurements each obtained by averaging the spectra of a single sequence (4 zenith observation and 2 calibration). A good agreement results a part for window region, influenced from residual atmospheric variability. An example is shonw in Fig.3.2.2.

A further product of level 1 data is the instrumental line shape (ILS), important for the retrieval of atmospheric parameters. It is found that the instrumental self-apodization effect, due to the beam aperture, is negligible and that the ILS is a pure Sinc function corresponding to the nominal resolution 0.4 cm^{-1}. The results show that instrumental performances are well up to the scientific requirements that were posed from the REFIR preliminary study.



(a)



(b)

Figure 3.2.2: *Radiometric performance of REFIR-PAD measurements. Noise equivalent spectral radiance (NESR), calibration accuracy (CAL) and standard deviation of the average spectrum (STD) are shown for a clear sky case (upper panel) and a cloudy scene (low panel)*

3.3 Lidar

The second instrument deployed at Concordia Station during PRANA campaign is a tropospheric linear depolarization Lidar system, important for our purpose due to the fundamental information provided about cloud geometry. The Lidar is mounted on an aluminum frame inside the above mentioned Physical Shelter. Double optical windows flushed with pre-heated external air make the operation of the Lidar possible in all weather conditions, even on $-80°C$ cold polar nights. This instrument is able to provide tropospheric profiles of aerosol and clouds every 5 min, 24 hours/24, by means of a fully automated system that collect together with Lidar observations also housekeeping and local meteorological data. Principal instrumental characteristics are reported in Tab. 3.2

Laser type	Quantel Brio, 532 nm, linearly polarized
Telescope	Refractive with 10 cm diameter
Data acquisition	Licel GmbH, analogical with 2 channels
Vertical resolution	7.5 m
Altitude range	30 m - 12000 m above ground

Table 3.2: *Lidar main characteristics*

For the thesis work only quick-looks daily images of Lidar observations were initially available, covering the entire 2013. An example in Fig. 3.3.1, refers to 25/02/2013.

This figure shows, on the top and in a logarithmic scale, the backscattering signal due to molecular and cloud or aerosol particle scattering. The signal is very strong in the first hundred meters a.g. and, for the considered day, one can observe the presence of a precipitating cloud during the night and in the morning, while in the first afternoon (13-15 GMT) appears a cloud stratus at about 1000m. The presence of the cloud is easily detected by means of the lower panel, that shows the depolarization ratio of the signal and gives information also about the phase of the cloud. In fact, spheric cloud droplets cause little change in the polarization of the laser beam, while, non-spherical ice crystals, due to their asymmetry, leads to a higher depolarization signal. In this example the stratus cloud is a water or mixed one, because the depolarization values are well below the ones shown for the night cloud, clearly made of ice crystals. Theoretical and experimental studies have been made during past decades about the depolarization Lidar techniques in order to remotely sense the phase and, possibly, the habit of cloud particles.
In addition to quicklooks, for selected days, daily Lidar matrices were provided, to establish with better accuracy the base and top height of the cloud. An example is given in Fig.3.3.3: the profile refers to the non-precipitating cloud visible in the quicklook 3.3.2. Lidar profiles are quite noisy due to both low signal and horizontal striping (due to a weld joint partially

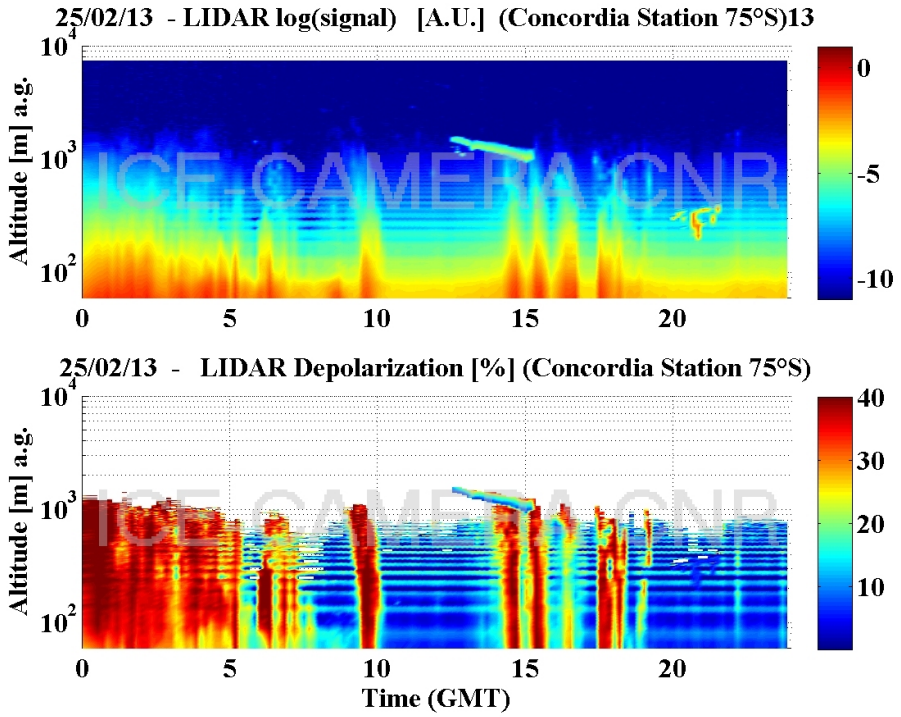


Figure 3.3.1: *Example of quicklook Lidar, with ice and water clouds*

detached): hence, the profile is filtered out in order to smooth oscillations that covers in part the low cloud signal.

Theoretical and experimental study have been made during past decades about the depolarization Lidar techniques, used to remote sense the phase and, possibly, the habit of cloud particles (Sassen, 1974). We are interested with the representation of radiation scattered into the exact backscattering direction from linearly polarized laser light, thus described by a Stokes vector proportional to {1,1,0,0}. So, we define a scattering plane with reference to the laser beam, which is 100% linearly polarized parallel to this plane. The simplifying assumptions that must be made regard the random orientation of the non-spherical particles in 3D space and the reasonable diversity of shape and symmetry displayed by particles. Under these conditions, the relation between incident and scattered beams are expressed by the equation:

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & 0 & 0 \\ P_{21} & P_{22} & 0 & 0 \\ 0 & 0 & P_{33} & P_{34} \\ 0 & 0 & P_{43} & P_{44} \end{pmatrix} \begin{pmatrix} I_0 \\ Q_0 \\ 0 \\ 0 \end{pmatrix} \tag{3.3.1}$$

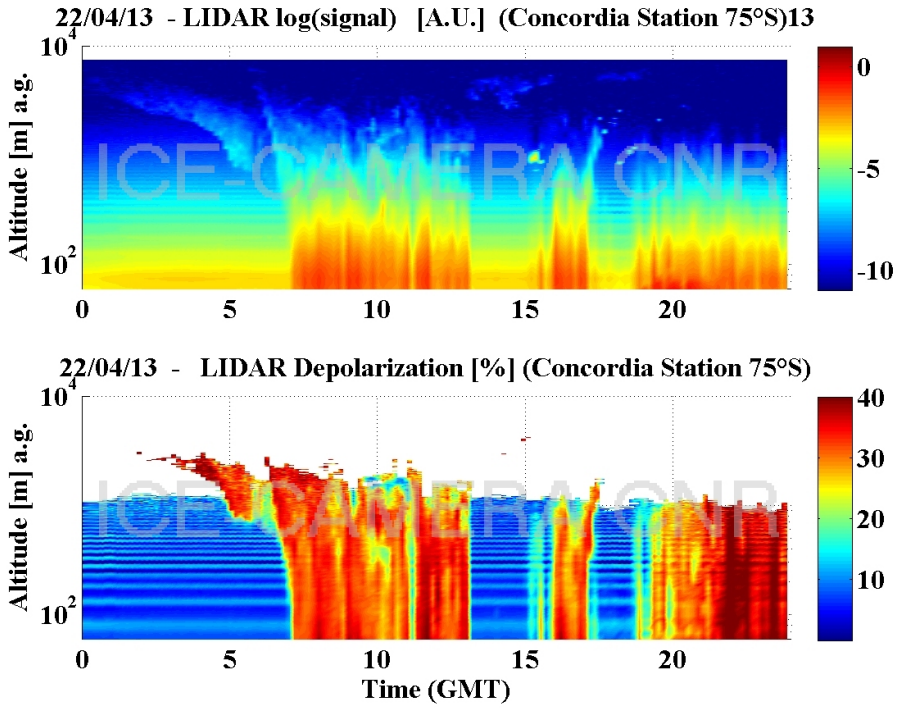


Figure 3.3.2: *Example of quicklook Lidar, with non-precipitating ice clouds*

but, the scattering phase matrix becomes diagonal for exact backscattering because the elements P_{12} and P_{43} are zero when the scattering angle is 180° and the linear depolarization ratio δ, defined as the ratio of the perpendicular-to-parallel polarization components of backscattered light, is given by:

$$\delta = \frac{(F_{11}(180°) - F_{22}(180°))}{(F_{11}(180°) + F_{22}(180°))}$$

Because for the spheres $F_{11}(180°) = F_{22}(180°)$ we have the finding of $\delta = 0$ for single scattering by cloud droplets. On the contrary, an ice crystal may produce a non-depolarizing specular reflection when the crystal is fortuitously aligned perpendicular to the laser beam direction, but it is considerably more likely that internally refracted ray paths will be chiefly responsible for backscattering. These processes result in the reorientation of the incident polarization vector at every interface, leading to depolarization when the backscattered ray is transported into the initial plane of polarization. When precipitating ice particles were probed from out of a laboratory window, the Lidar backscatter depolarization technique was shown to enable the separation of various ice particle types. Snowflakes, composed

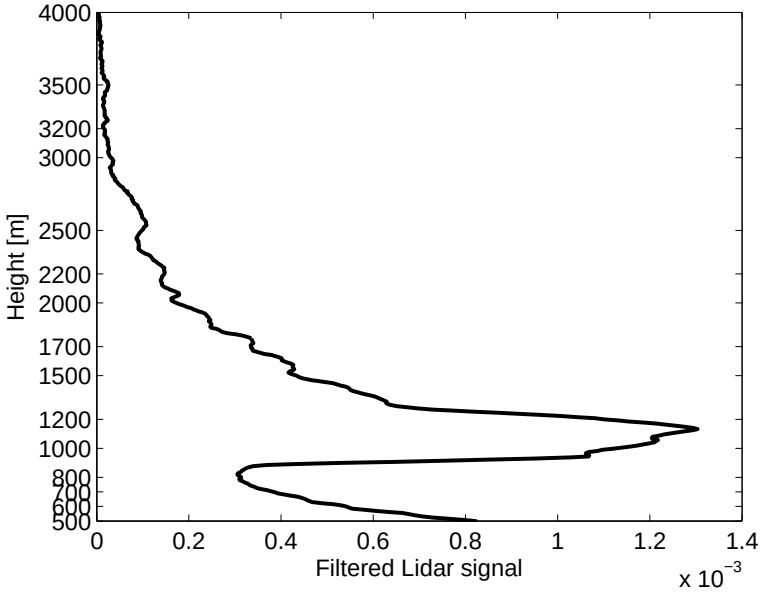


Figure 3.3.3: *Example of Lidar profile, referred to the ice cloud of 04/22 quicklook at 05:00 a.m.*

essentially of randomly oriented dendritic ice crystals, tend to produce the same $\delta \approx 0.5$ as randomly oriented laboratory crystals of generally mixed habits. However, as frozen cloud droplets begin accumulating on the ice crystal faces (i.e. riming, roughness is important at these wavelengths) the increase in surface complexity leads to a depolarization increase. Moreover, the droplet accretion process results in the formation of low-density graupeln particles: these opaque particles generate $\delta \approx 0.65$. The action of the final microphysical process, melting, is revealed to produce a strong increase in depolarization for snowflakes, due to changes in surface complexity again.

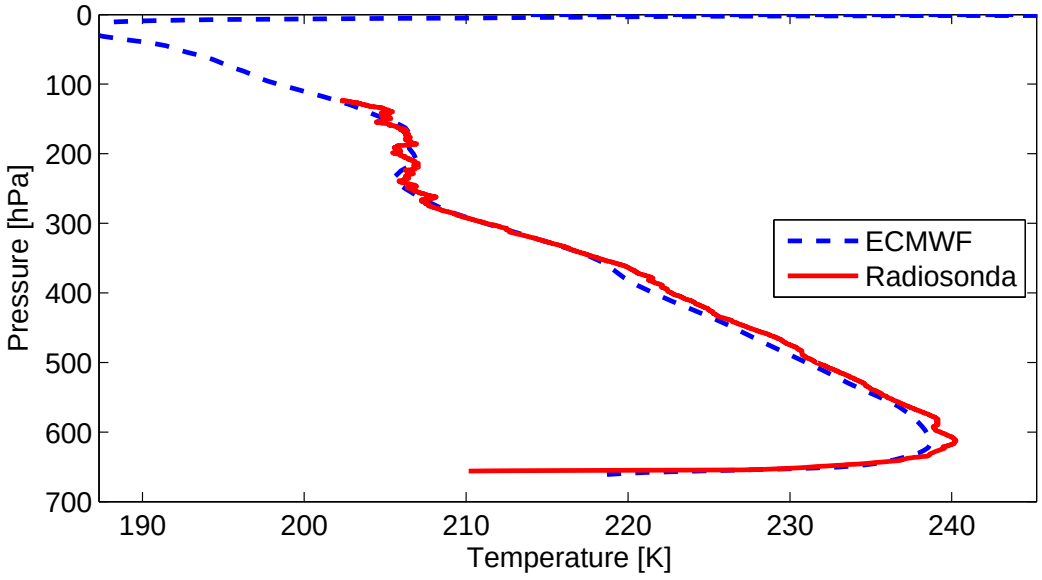
Moreover, if a phase distinction is possible, can also the habit type be classified from those measurements? Sassen (1977) analyzes the possibility of habit discrimination from the study of single crystal backscattering signal obtaining positive results. Simple ice crystals were classified according to whether the dominant direction of growth has occurred along the a axis in the plane of the basal face (plate crystal) or along the c axis in the direction of the prism face (column crystal). Ice crystal that display neither high nor low c/a axis ratios are designated as thick plate ones. All these crystal habits show hexagonal crystallographic symmetry. Temperature of the environment is the dominant factor in determining crystal habit, although the amount of water vapor above ice saturation available during growth also has an influence. It may seem likely that the amount of depolarization produced during scat-

tering should depend on crystal habit and as a consequence on cloud temperature. However, (Sassen, 2000), the δ value generating by viewing large assemblies of randomly oriented ice crystals of any of the hexagonal habits is characteristically equal to ≈ 0.5 in the absence of a water droplet cloud. Thus, when numerous hexagonal crystals are simultaneously sampled, ice crystal habit cannot be identified on the basis of linear depolarization ratio: the average δ value observed from sampling many individual hexagonal crystals which are randomly arrayed should similarly be 0.5, regardless of habit. However, there are features of scattering behavior sensitive to crystal shape which can be used to differentiate ice crystal habit when the particles are considered singly.

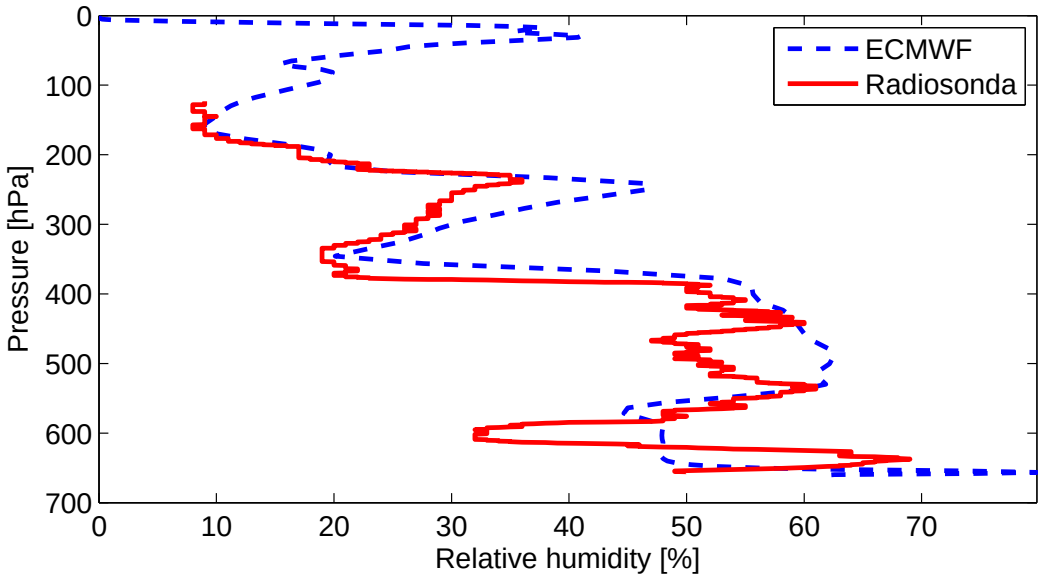
3.4 Vertical profile data

At Concordia Base, a radiosonde system is also installed and during 2013 one radiosonde profile/day is available, launched at 12 p.m. Looking at the entire dataset, not all the radiosonde data are complete: in fact, measurements are often available only for few hundred meters in troposphere or, for example, relative humidity is missing. Because the need of accurate and constantly available atmospheric profiles for the retrieval of cloud parameters, a complete dataset of ECMWF analysis for the DomeC site is also available. This dataset contains 4 profiles/day: at 00, 06, 12 and 18 UTC. A preliminary work was done in order to assess the reliability of this model products: we have compared radiosonde data with ECMWF profiles and concluded that in the majority of the cases the two sources are in agreement, with different greater in relative humidity field than in temperature one; an example is in Fig. 3.4.1.

In this way, in order to create a good atmospheric profile for simulation requirements, if the radiosonde data is available, these are considered up to 8km and then a climatological profile is merged; on the contrary if radiosonde observations are missing, the time-interpolated ECMWF profiles are used. Hence, for each identified cloud case an atmospheric profile is derived.



(a)



(b)

Figure 3.4.1: *Example of radiosonde and ECMWF profiles, referred to 27/06 at 12 UTC, (a) Temperature profile and (b) Relative Humidity profile*

Chapter 4

From pre-processing to spectra classification

4.1 A look at the available REFIR-PAD dataset and at the observed spectra

The available dataset of REFIR-PAD measurement in the framework of PRANA project, covers 2 years from the end of 2011, with the installation of the instrument, until the end of 2013, even if the instrument is still working and will continue under other projects. Spectra analyzed in this work cover the whole 2013; in this dataset, both single channel spectra and weighted averages of both channels are available and, in order to smooth noise, we have used the latter. In fact, the second channel is noisier than the first one, due to problems in the phase calibration or a deformed sinc function, difficult to correct.

The instrument is operating 24 hours a day, alternating cycles of 6 hours of measurements, during which groups of 27 sequences of observations are typically collected, with 3 hours of analysis. Each sequence of observation lasts about 14 minutes, during which 4 zenith measurenments are taken, each lasting 80s, while 4 calibration scans are performed using the two blackbodies (HBB and CBB). So, for each sequence, the average of 8 spectra (four for each channel) is collected and used for the successive analysis. The unapodized spectral resolution at which REFIR-PAD operates is equal to 0.4 cm^{-1} while the nominal spectral range is between 100 and 1500 cm^{-1}.

Instrument measurements are organized in .dat files characterized by a five-column structure, respectively containing wavenumbers (cm^{-1}), measured radiances ($W/m^2\,sr\,cm^{-1}$) and spectral uncertainty estimates with different meaning: NESR, CAL and STD, as described in section 3.2.

Examples of measured spectra for clear sky conditions and in the presence of a non-precipitating ice cloud, are reported respectively in Fig.4.1.1 and Fig.4.1.2:

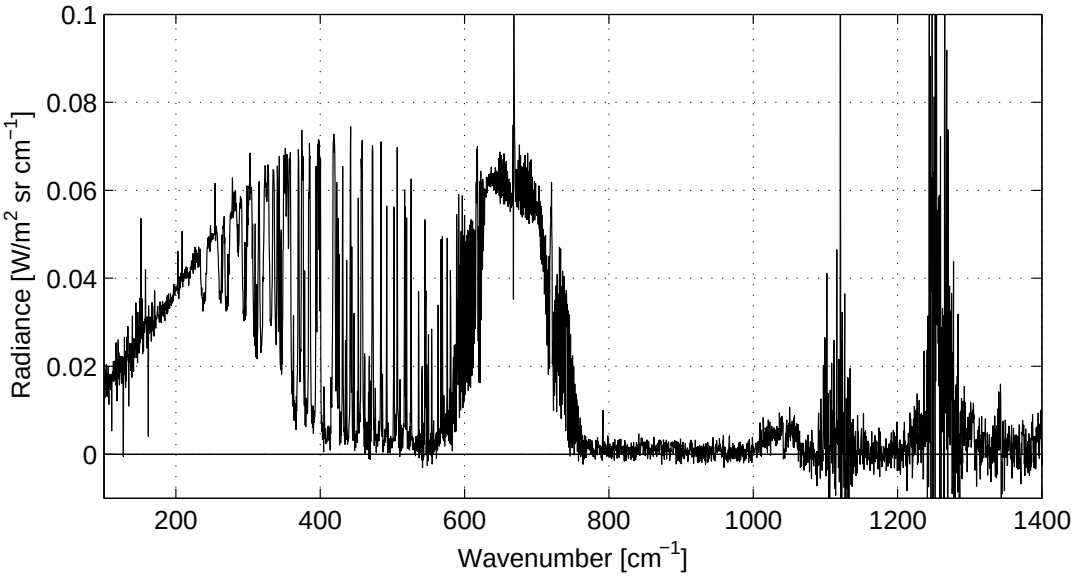


Figure 4.1.1: *An example of measured spectrum in clear conditions on 06/06/13 at 12:12 pm*

In these two figures, the most important spectral features that appear, as described in Ch.1, are the strong signature related to the CO_2 ν_2 band centered at 667 cm^{-1}; at the center of this band a net peak of radiance appears: this signal is due to extremely strong absorption in the center of the band that make the instrument measure the emission of air inside the instrument (at a temperature well above external environment) which causes intuitable calibration problems. Related to this, a little shift over the entire range of frequencies can be observed in some spectra; this shift causes single absorption peak to be distorted into two peaks one upward and one downward.

In clear conditions, downwelling radiance in the IR window is very low, with the presence of many negative values in single channels, perfectly compatible with an interferometric type of measurement. Also in the FIR region a number of micro-windows open in very dry air conditions, as the ones observed especially during the Antarctic winter; in the first selected case dirty windows are clearly detectable also at wavenumber as low as 220 cm^{-1}, but in extreme cases even down to 190 cm^{-1}. Some spikes are visible at very low wavenumbers due to both extremely strong absorption H_2O lines and detector noise. On the other side, around 1050 cm^{-1} the O_3 band can be noted, but also two regions characterized by very high

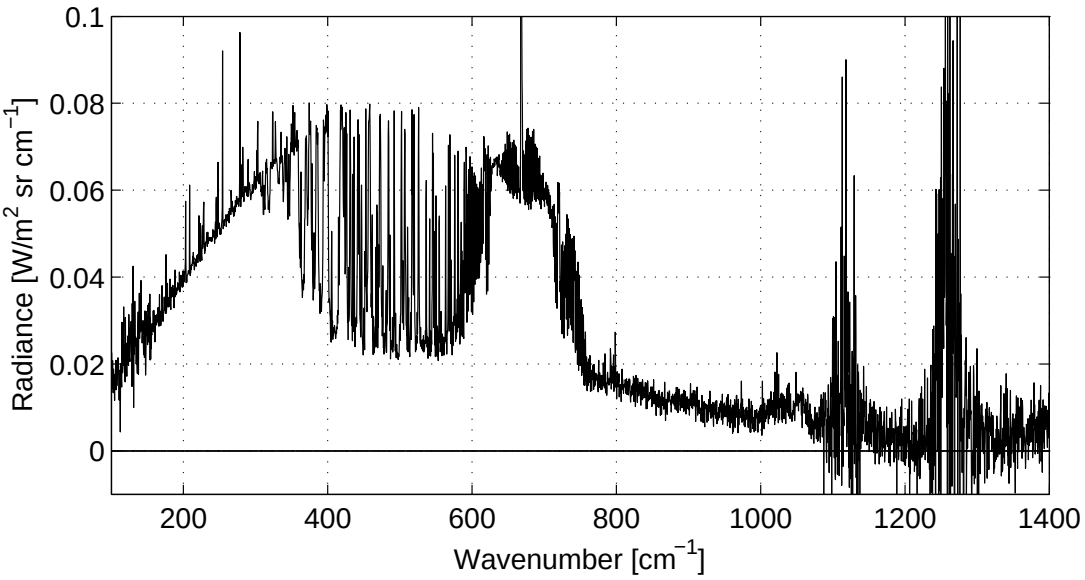


Figure 4.1.2: *An example of measured spectrum in the presence of an ice cloud, on 23/04/13 at 06:39 am*

noise leap out: the first centered at 1110 cm^{-1} and the second at 1250 cm^{-1} are linked to the absorption features of the beam splitters.

The spectrum in Fig4.1.2 shows that the radiance, measured in IR window region and in FIR micro-windows, definitely changes in the presence of cloud, with a clear slope between 800 and 1000 cm^{-1}, where absorption/emission processes are of primary importance, which depend on cloud optical depth; looking instead at FIR micro-windows, signal in this regions depends on several parameters, such as particle size and habit, due to the relative relevance of scattering processes at these wavenumbers.

In Fig.4.1.3 and Fig.4.1.4, the same clear and cloudy spectra are reported, but radiance measurements are converted into brightness temperatures (BT), avoiding negative values. These plots are interesting because they give an idea of the level the radiance arrives from: in fact, it is evident that in highly absorptive regions the measured radiance corresponds to the same brightness temperature at about 240K, with higher peaks in the strongest absorption lines; on the contrary, in the cloudy case, the BT in the main window is about 200K, temperature that can be assumed as the cloud one if it has a sufficiently high optical depth in order to be considered a blackbody. Moreover, the BT around 800-1000 cm^{-1} results higher than in the FIR windows, evidence that in the latter regions the cloud is more transparent. The absorption coefficient has indeed a minimum for wavenumbers around 400 cm^{-1} and a maximum around 800 cm^{-1}, see Fig.4.3.12.

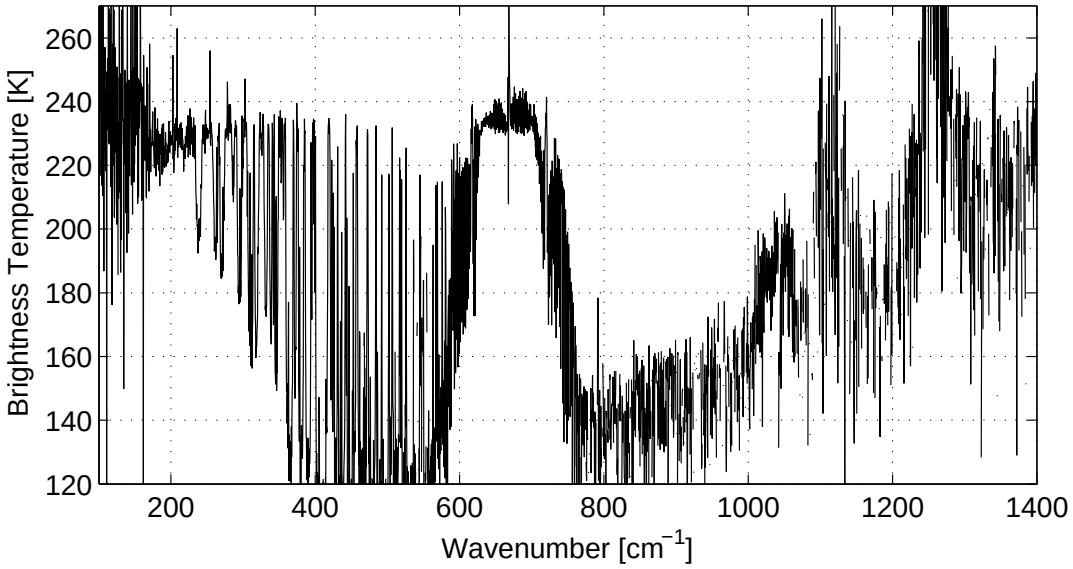


Figure 4.1.3: *Same spectrum of Fig.4.1.1 for clear conditions in terms of brightness temperature.*

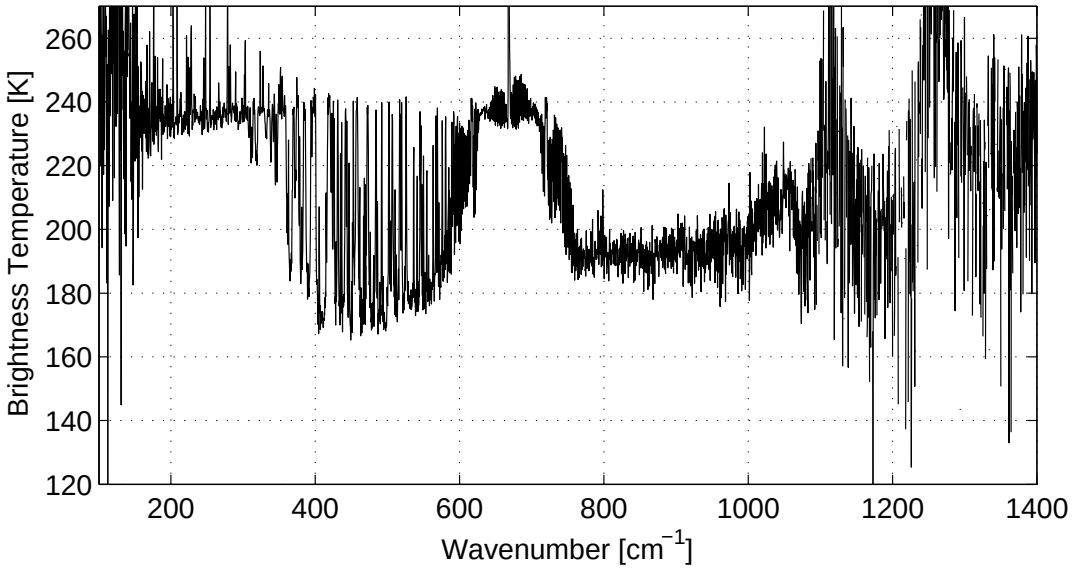


Figure 4.1.4: *Same spectrum of Fig.4.1.2 for cloudy conditions in terms of brightness temperature.*

4.2 REFIR-PAD Data Pre-processing

First of all, .dat files are downloaded and copied in a specific directory month by month, for computational requirements; so, a first pre-processing is executed (file *PreprocessingA.m*), with which a list of all available files, concerning each month, (*Elenco'mm'Completo.txt*) and a 3D matrix (file *PAD-DomeC-'mm'-C13.mat*) are created. This matrix is composed by n (number of available spectra) 2D matrices each of which has a number of rows equal to the amount of wavenumber at which REFIR-PAD measures and 5 columns, maintaining the previously described original structure. A marginal note concerns headings related to .dat files that are maintained in the first version of .mat files and then removed with also measurements transformed in ($mW/m^2\ sr\ cm^{-1}$) for retrieval analysis requirements. Another observation related to the dataset concerns the fact that for December 2013, the REFIR-PAD resolution, fixed during the rest of the year and equal to 0.4 cm^{-1}, changes in some sequences to 0.5 cm^{-1}; in addition, also the range of measurement results extended up to 1500 cm^{-1}. These spectra are related to instrumental tests and are not considered in our analysis. Finally, during the period from mid-November until mid-January, because the presence of people in loco, working on the instrument, spectra have to be chosen with particular attention.

Once data matrices are built for every single month, the first step involves the criteria to mark the spectra that should be rejected. In particular, 6 "indicators" are chosen in order to evaluate the goodness of a spectrum.

1. Mean radiance in the window between 895 and 900 $cm^{-1} > 0$.
   Some of the spectra that show a negative mean radiance in this wide interval presents an anomalous slope beginning in the IR window and extending towards increasing wavenumbers; hence are rejected. This problem is probably linked to instrumental mis-alignments or difficulties in the sampling of the Zero Point Difference of the interferogram.

2. Mean radiance in the spectral interval between 1195 and 1205 cm^{-1} + error > 0.
   In this case the selection criterion is more tolerant, because these wavenumbers are not of particular interest in our study and the variability showed in this window is very important. For these two motivations, only spectra for which a mean negative radiance in the selected interval is obtained after adding associated error are rejected; spectral uncertainty is calculated as the maximum value between STD error, linked to the scene variability over the measurement time range, and the root mean square of the calibration error (CAL) and the NESR.

3. Radiance variance in the window around 900 $cm^{-1} < 3$ std of the variance distribution.
   Looking firstly at the graph of this parameter for all the 2013, the presence of some

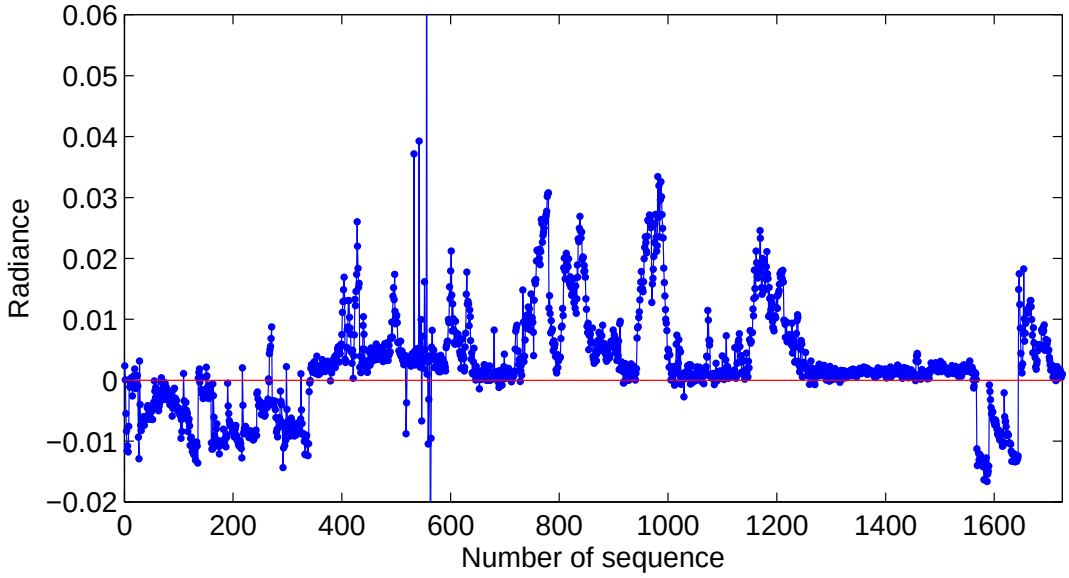
abrupt changes in the spectral noise can be noted and a general trend towards increasing spectral noise is evident over the year, until December when the instrument was re-calibrated. As a consequence, 2013 is divided into 5 periods (January-March, March-June, June-October, October-November and December); for each period, the mean radiance and its mean variance is calculated, in the considered spectral interval. In addition, as one can expect, the mean variance for each spectrum plotted vs time assumes a Gaussian distribution with some outliers. Hence, considering each month singularly, a robust variance (std_g) for this hypothetic Gaussian distribution is calculated and spectra that show a mean variance around 900 cm^{-1} that exceed 3 std_g are rejected.

4. Radiance in the spectral interval around $675\,cm^{-1} < 0.1\,W/m^2\,sr\,cm^{-1}$.
   This parameter results linked to instrumental conditions of work: few spectra report radiances in the CO_2 ν_2 band too high to be plausible, so a threshold of $0.1\,W/m^2\,sr\,cm^{-1}$ for winter months and $0.08\,W/m^2\,sr\,cm^{-1}$ for others is fixed. These spectra are related to non-sky observations or are extremely noisy.

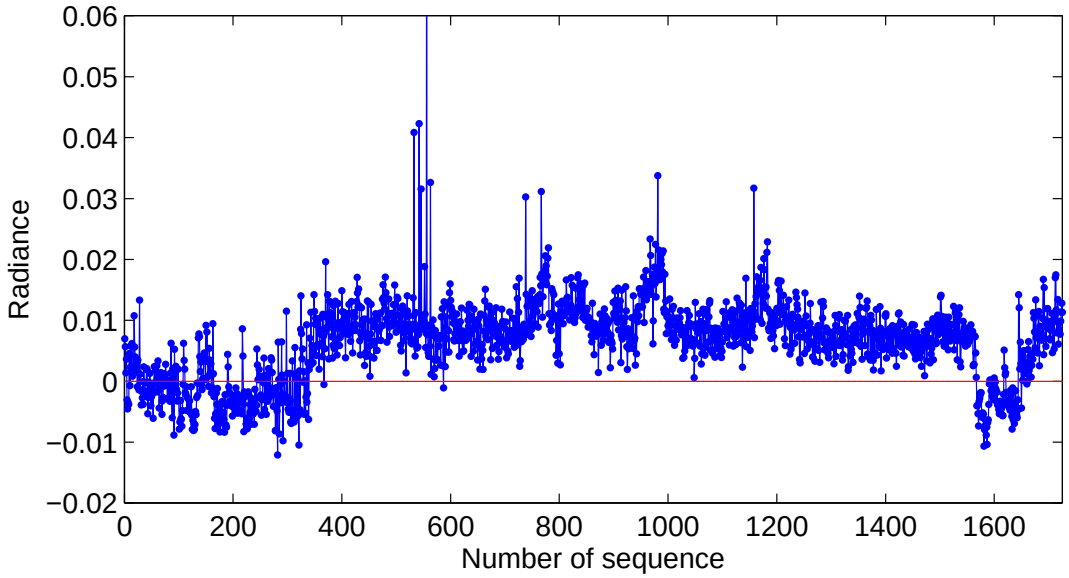
5. The difference (jump) between the mean radiance around 675 cm^{-1} for the considered spectrum and its value for the previous one, both belonging to the same set of measurements, must be less than a fixed threshold.
   In some cases important jumps in the measured mean radiance are reported from one spectrum to the other, belonging to the same measurement set; these variations seem to be linked to some problems in the stability of shelter inside temperature and are removed. The relative threshold is linked again to the observed monthly distribution of these jumps, that is likely to be considered a Gaussian; in this way, spectra that show jumps that exceed 3 standard deviation from the mean of their distribution are rejected

6. Mean radiance around 530 $cm^{-1} > 0$.
   This last criterion is similar to the first one; spectra that show negative mean value in this dirty window are frequently affected by problems in the FIR region as a whole.

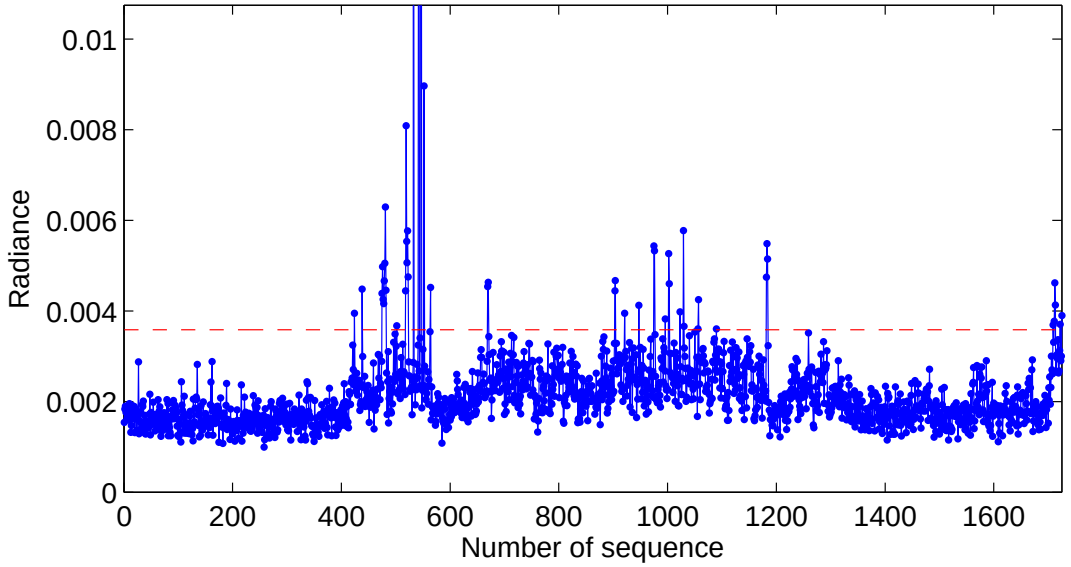
An example of these criteria is reported in Fig.4.2.0 that shows the distribution of the items 1, 2, 3 and 5 listed above, for August 2013. As can be noted, during this month some instrumental problems occurred; in fact, a lot of spectra show negative values for averaged radiances in IR window, as can be seen in the first and in the second figures. Looking at the panel (a) periods characterized by clear conditions, linked to low radiance values, and cloudy episodes, related to the peaks in the graphic, can be visibly detected. Moreover, the distribution of jumps between consecutive observations shows outliers well above the fixed threshold, related to high instability in the shelter temperature. In these figures, the threshold for rejection is plotted by a horizontal line.


0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
−0.02
Radiance
0
200
400
600
800
1000
1200
1400
1600
Number of sequence

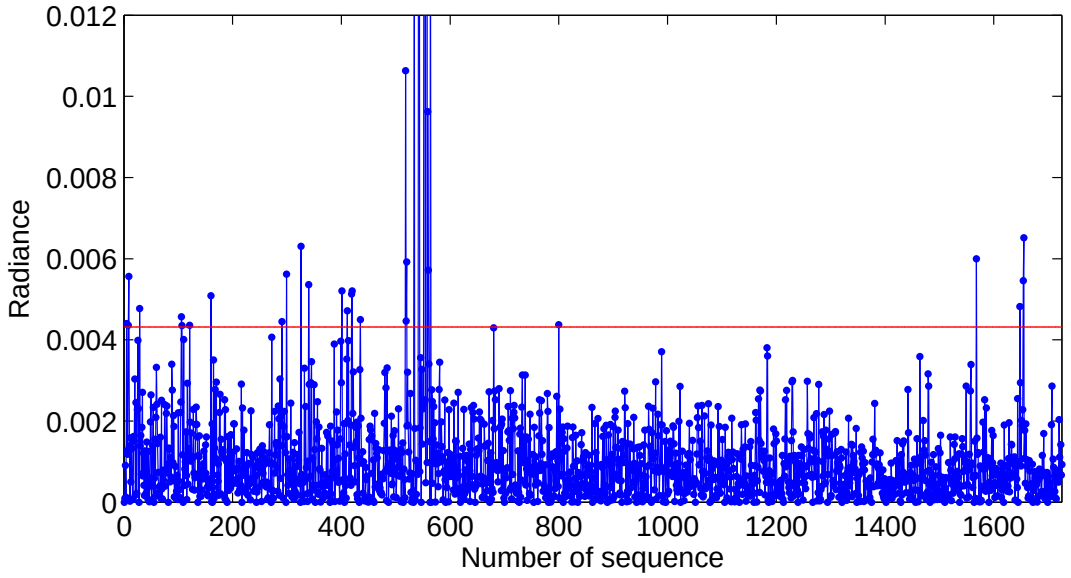
(a)


0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
−0.02
Radiance
0
200
400
600
800
1000
1200
1400
1600
Number of sequence

(b)



(c)



(d)

Figure 4.2.0: *Examples of parameters used to select good spectra for August 2013. In panel (a) the average radiance measured in the interval around 900 cm^{-1} (item 1) with the fixed threshold; in panel (b) the average radiance summed with error around 1200 cm^{-1} (item 2). In (c) the variance of the radiance around 900 cm^{-1} (item 2) and in (d) radiance jumps in the CO_2 band between consecutive spectra (item 5).*

Despite these selection criteria, few spectra characterized by deficiencies are still detected; in particular, spectra which show anomalous slopes or distortions, especially for wavenumber in the IR window, can be found and these cases of uncertain validity are rejected. An example is displayed in Fig.4.2.1: this is a cloudy spectrum measured on 29/08/13 that shows a strong slope beginning from CO_2 band, with a clear distortion and extensively negative radiances over 1000cm^{-1}; this spectrum represents an observation that passes the 6 fixed criteria for selection.

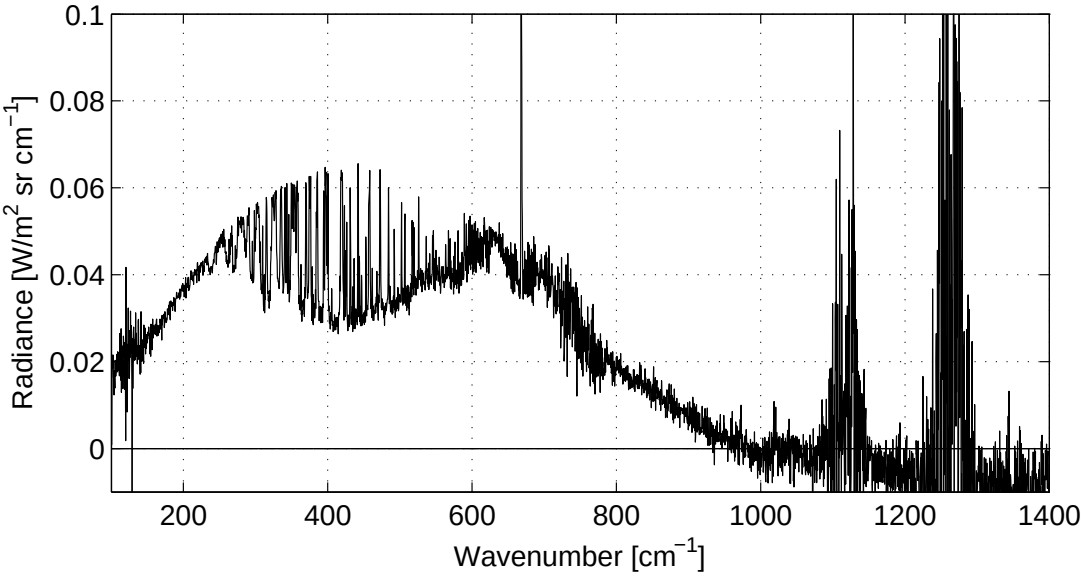


Figure 4.2.1: *Case of spectrum considered as good but deformed by an anomalous slope.*

The script *PreprocessingB.m* generates new 3D matrices for each month (*file PAD-DomeC-'mm'-13.mat*) that contain only good observations.

A preliminary discrimination among 'good' cases into clear or cloudy classes can be done, using data in the interval 895-905 cm^{-1}. The selection uses the mean measured radiance and the SNR, with errors computed as in step 2 of section 4.2. The mean measured radiance is converted to BT and spectra are classified as cloudy when BT is larger than a threshold, fixed at 165K for summer or 170K for winter, and, at the same time, the SNR is larger then a second threshold. These values are empirically chosen from the distributions of BTs and SNRs for each month.

4.3 Data grouping

The first purpose is grouping of cases that belong to the same class, i.e. are similar to each other. The first classification is between clear and cloudy cases, while later cloudy cases will be separated according to phase (ice or water cloud) and to type (precipitating or non-precipitating ones). Any selection method is based on the choice of a set of features that are considered important for the classification task. The features we have a priori selected are based on measured data on specific channels. Two selection approaches have been adopted:

1. Un-supervised feature extraction, using Principal Component Analysis (PCA) and Spectral Regression (SR);

2. Supervised feature extraction, using concepts linked to Linear Discriminant Analysis (LDA) and a Support Vector Machine.

A shortlist of 26 channels is selected and defined in Table.4.1, whose widths ranges from 2 to 10 cm^{-1} in the spectral range between 180 and 1200 cm^{-1}. These channels are assumed as information synthesis of the data (considering all 3500 wavenumbers for analysis is inconceivable). Most of these spectral intervals coincide with atmospheric windows (or micro-windows).

The mean radiances measured in these 'super-channels' are calculated and transformed into brightness temperatures; this process is done keeping out from each interval all negative values: in fact, as discussed in section 4.1, an interferometric observation is allowed to present negative values of radiance, but in the process of converting it into BT, these values are not admissible. In other words, only positive values are considered in order to avoid some spectra to present mean negative radiance values over the selected intervals. This problem is frequently met for the window around 1200 cm^{-1}, due to the fact that only negative values of radiance + error are rejected, which still leave negative radiances in the dataset.

4.3.1 Manual selection of 'examples'

A preliminary step consists in the selection of a group of spectra among all monthly available observations, that form the so called training dataset, i.e. 'example' spectra that are used to "train" the classifier in the supervised approach or to analyze grouping patterns from the un-supervised methods; in addition, another set has to be prepared, set that is used to test the classification ability (test dataset). Those spectra for training and test are defined in the *calcolaEsempi.m* file, which has a switch structure according as one is discriminating the presence, the phase or the type of cloud. Chosen observations are listed in vectors in the form of numbers, each one corresponding to the position in the list of all the available spectra for each month (*Elenco'mm'Completi.txt*). Contextually, in the same file, the relative class

N	**wavenumber interval** $[cm^{-1}]$	**width** $[cm^{-1}]$
1	190-193	3
2	216-219	3
3	235-240	5
4	259-262	3
5	269-271	2
6	293-297	4
7	317-320	3
8	330-333	3
9	337-339	2
10	363-367	4
11	379-382	3
12	388-392	4
13	407-413	4
14	437-440	3
15	448-451	3
16	462-466	4
17	476-482	6
18	495-500	5
19	521-524	3
20	529-534	5
21	558-562	4
22	572-575	3
23	675-685	10
24	830-836	6
25	895-905	10
26	1195-1205	10

Table 4.1: *List of the selected spectral intervals: radiance averaged over them is used to synthesize information content of the measurements.*

labels for each example are also present: equal to 1 or -1 depending on whether the example is cloudy or clear, in the first step. Their choice is obtained manually, using available daily quick-looks of Lidar profiles, trying to select example in the most scattered and random way, with, at the same time, the aim to preserve the widest record as possible, in order to consider a fair variety of observed conditions. Because spectra numeration in this file is referred to the complete lists, those numbers remain fixed also varying the thresholds for initial spectra selection. On the other hand, having to consider hereafter good spectra only, it's necessary, at the beginning of the analysis, to find for each chosen spectrum its relative index in the list of the "good" spectra and also verify its effective presence in this list.

As regards this first classification (clear-cloudy) step, generally 10-20 clear and just as many cloudy spectra are chosen, in order to form the training dataset for each month; cloudy ones are already classified by phase and type so that they can be used in the next steps. Similarly, spectra are chosen also for test dataset. In February, March and November, their number is substantially reduced, due to the lack either of many quicklooks Lidar or of REFIR-PAD spectra, linked to instrumental problems; as a consequence, the number of spectra effectively available for a complete analysis will be very limited for these months. On the contrary, for other months the number of spectra is around 1500 units.

4.3.2 Un-Supervised Feature Extraction

Dimensionality reduction is a key problem in the wide field of data mining and machine learning and has the aim of mapping high dimensional data onto a space of lower dimensionality, preserving at the same time the maximum information content. This problem can be dealt with two methodologies: the feature extraction that operates a linear or a non-linear mapping of the data space, transforming the original features (i.e. variables that describe data) into a set of new low dimensional features able to well explain observations; the other is the feature selection approach that tries to find a subset of the original variables that alone can well describe data, in this case the physical meaning of the features can be maintained. In the first group, we find a lot of methods that exploit eigen-decomposition techniques: one of the most popular dimensionality reduction algorithm is the Principal Component Analysis (PCA) that performs dimensionality reduction by projecting the original n-dimensional data onto the $d(<< n)$-dimensional linear subspace spanned by the leading eigenvectors of the data's covariance matrix. Its goal is to find a set of mutually orthogonal basis functions that capture the directions of maximum variance in the dataset; if data are embedded in a linear subspace, PCA is guaranteed to discover the dimensionality of the subspace and produces a compact representation (Cai et al., 2007a), but could not be relevant for classification purposes, as in our case. As an example, in Fig.4.3.1 the result of the Principal Component Analysis performed on the measured dataset of January is shown: the data matrix, each row of which is a measured spectra, is passed to a Matlab routine that computes its representation in the principal component space.

The figure is a scatter plot of the overall dataset represented by means of the first two principal components, which retain the maximum variance. In red and blue examples of clear and cloudy spectra are highlighted : it can be seen that this method does not lead to bad results, even if a sharp boundary between clear and cloudy cases seem difficult to draw. Hence, for classification purpose this method is probably not the best one: in fact, PCA leads to the most 'spanned' representation of data not to the best separation of them.

Spectral methods have recently emerged as alternative and powerful tools for dimensionality reduction and machine learning. These methods use information contained in the eigen-

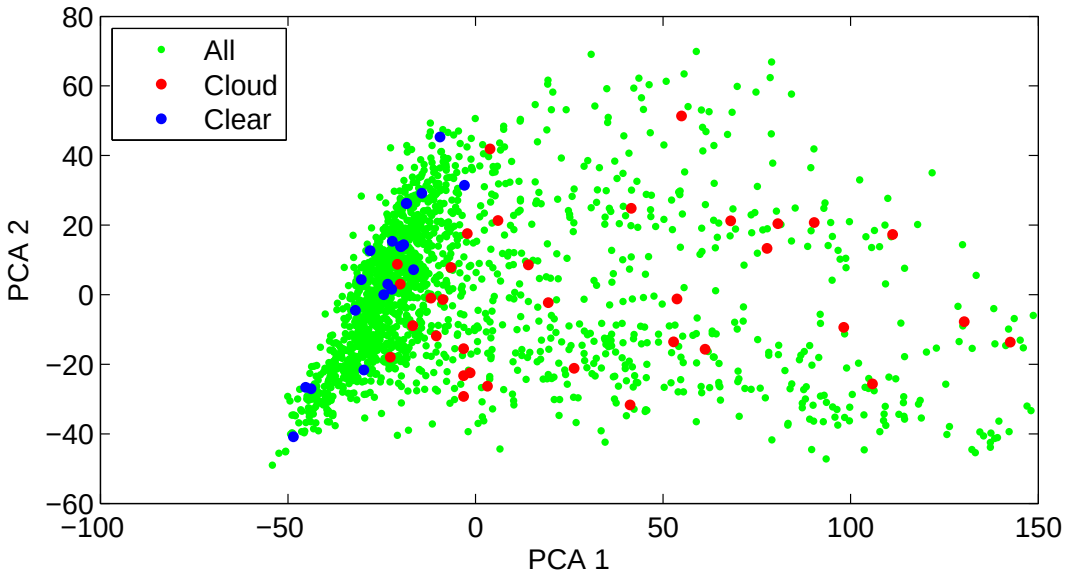


Figure 4.3.1: *PCA result for January observations*

vector of a data affinity matrix (i.e. item-item similarity matrix) to reveal low dimensional structure in high dimensional data.

A novel dimensionality reduction method called Spectral Regression (SR) casts the problem of learning an embedding function into a regression framework, which avoid eigen-decomposition of dense matrix requested in other methods. In fact, a lot of dimensionality reduction methods, also for non-linear data space, have been proposed but most of them have difficulties in provide a functional mapping between the high and the low dimensional space that are valid on and off the training set. Moreover, as said above, eigen-decomposition is frequently requested and it is almost unfeasible to apply these approaches to large datasets. The SR algorithm is fundamentally based on regression and spectral graph analysis; firstly an affinity graph is built over points, to discover the intrinsic discriminant structure in the data. This graph is used to learn response for unlabeled points. Once the responses are obtained, the ordinary regression is then applied for learning the embedding function.

Given m samples $\mathbf{x_i}$ with i=1,...,m $\in R_n$, dimensionality reduction aims at finding $\boldsymbol{z}_i \in R_d$ (with $d << n$) such as $\mathbf{z_i}$ can represent $\mathbf{x_i}$. Many algorithms of dimensionality reduction are based on graphs: a graph is build among m data points, called vertexes of the graph. Let $\mathbf{W}$ be a symmetric $m \times m$ matrix whose elements are the weights of the edge joining points i and j. A graph G and the W affinity matrix can characterize some geometric or statistical properties of a data set. The purpose of the graph embedding is to represent each vertex of a graph with a low dimensional vector that preserves similarities between vertex

pairs, where similarity is measured by edge weights. For details and an analytic treatment of Spectral Regression theory and algorithm refer to Cai et al. (2007a) and Cai et al. (2007b).

In our case, in order to implement this algorithm, a Matlab code written by Deng Cai (2012) is used. This code implements a Spectral Regression algorithm in an unsupervised setting, that embed the input high dimensional data (expressed in terms of the BT calculated in the intervals defined in Tab.4.1) into a d-dimensional subspace, where d is a parameter to be specified. In particular, in Fig.4.3.2 the result of this methodology applied to January observations, choosing d=3, is reported.

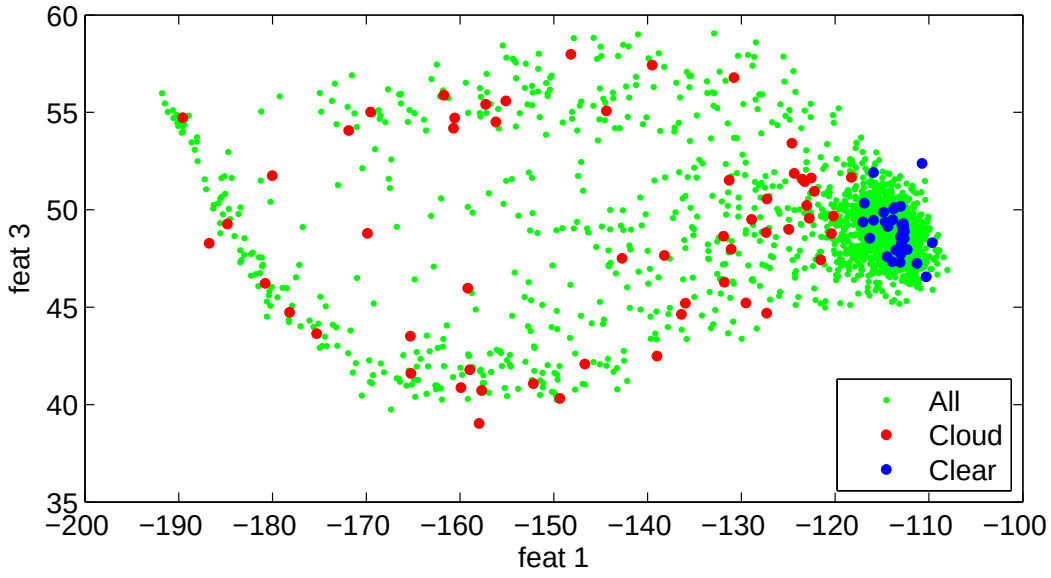


Figure 4.3.2: *Projection of data points onto the subspace individuated by two features calculated by means of the Unsupervised Spectral Regression algorithm; clear and cloudy examples are superimposed, January dataset.*

It represents a scatter plot feature 1 vs. feature 2, where this last two quantities are two among the three new features by means of which observations are described. In this way, data similarities in the original high dimensional space are conserved and data points are mapped in the 3D subspace defined by the new features. The result is pretty good: plotting on the overall scatter plot clear and cloudy examples, results that a dense group of points lay where the clear cases fall, while cloudy cases are scattered but fairly well separated from the clear ones. Based on this result and knowing these classified points, a classification procedure can be developed using, for example, a k-nearest-neighbor method.

This method can be performed on cloudy spectra only, in order to see if some embedded structure in the dataset can be identified. Fig.4.3.3 reports the result obtained considering

only spectra classified as cloudy one, by means of the above discussed preliminary thresholds (using BT and SNR, section 2.1), for the January dataset:

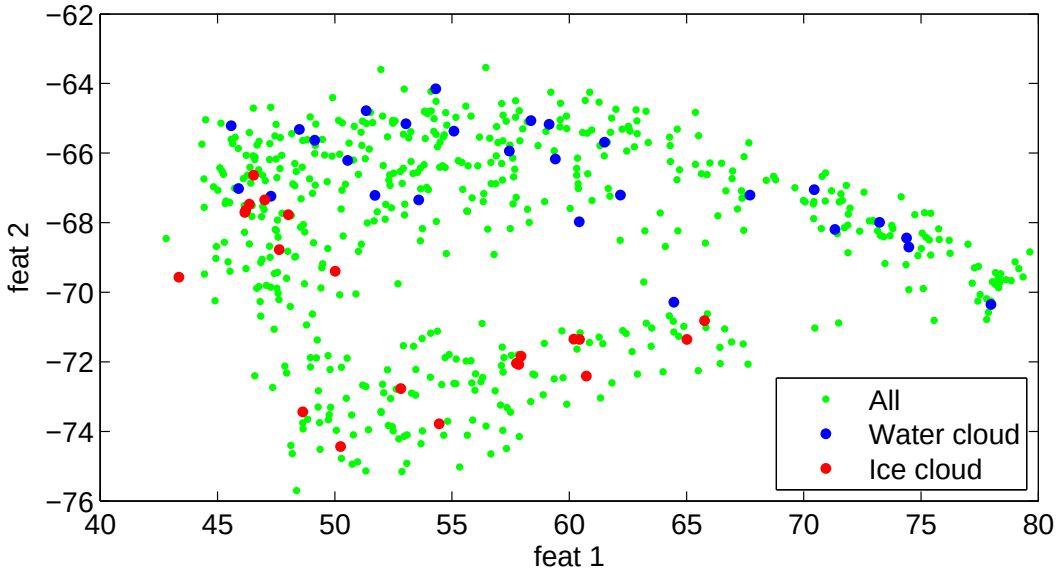


Figure 4.3.3: *Projection of data point onto the subspace individuated by two features calculated by means of the Unsupervised Spectral Regression algorithm, water and ice cloud examples are superimposed, January dataset.*

Also in this case, it seems that two 'groups' emerge from the overall distribution and when ice and water cloud examples are superimposed it seems that this structure can be related, even partially, to a cloud phase distinction. But, as above, the new features loose any physical meaning, as a consequence of the performed mapping.

In the end, the Spectral Regression algorithm is applied to June, when no water cloud can be detected looking at quick-looks Lidar: the result is plotted in Fig.4.3.4: an embedded structure is detected but result difficult to relate this with some cloud property, having at disposal only quicklooks Lidar. In this figure, examples referring to spectra measured in the presence of precipitating and non-precipitating ice clouds are superimposed. These two groups show overlapping regions and the structure may be related to other cloud characteristics such as effective radius or optical depth.

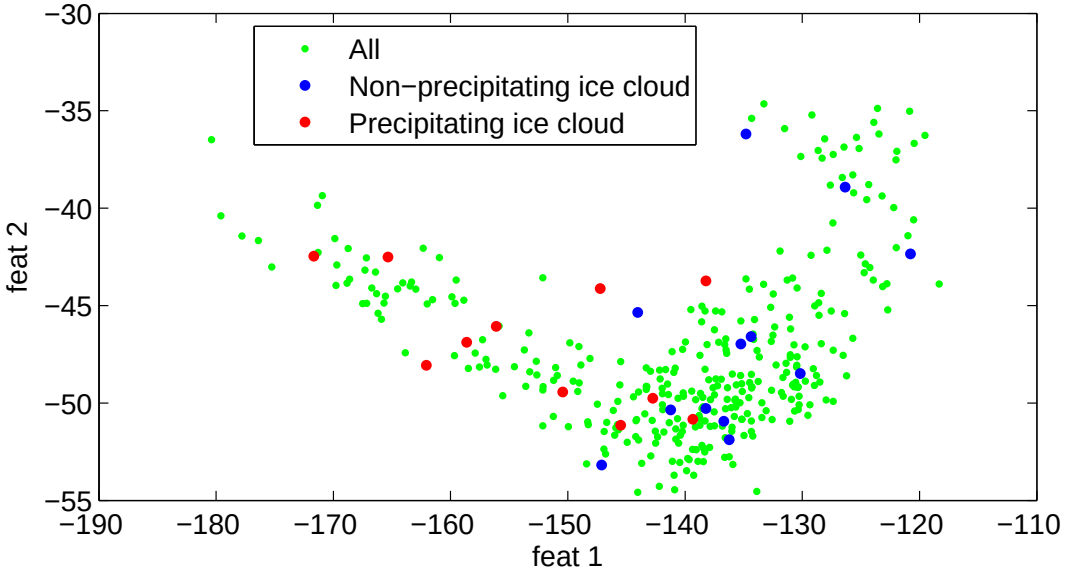


Figure 4.3.4: *Projection of data point onto the subspace individuated by two features calculated by means of the Unsupervised Spectral Regression algorithm; precipitating and non-precipitating ice cloud examples are superimposed; June dataset.*

4.3.3 Supervised Feature Selection and Support Vector Machine

The second approach is linked to the feature selection techniques with supervised learning: a feature selection problem is firstly dealt with, in order to obtain those spectral features that seem to be more useful for the observation classification; secondly, a Support Vector Machine is run in order to build the very classification stage. In fact, despite that also with the unsupervised methods some reasonable results are obtained, the supervised approach has been chosen in the end, in order to guarantee a well-defined physical meaning to features chosen for data representation and classification.

For this analysis, the differences between each pair of BTs computed in the selected interval reported in Table.4.1 are calculated; in this way, about 300 features are found: each measure is described and synthesized by these BT difference (BTD) values. Differences are considered in order to obtain a major number of features describing spectra and, above all, to free the considerations from global fluctuations of spectra, linked both to instrumental instability and temporal variability. Now, the question concerns how to find among these defined features which are truly important for describing the observations and for classification purposes.

The first step of this supervised learning has as purpose to separate clear from cloudy spectra; the same procedure will be repeated, with little modifications, also for the following steps. The aim is to choose, from all the obtained BT differences, those that are effectively meaningful for the desired classification. The code is in the *FeatSelection.m* script. As a consequence of the lack of training examples in some months, in order to separate cloudy from clear spectra, two contiguous months are each time considered, so that there are a sufficient number of training and test spectra also for those months when Lidar is less available.

At the beginning of the *FeatSelection.m* file, a matrix (**TB**) is built: it includes all data belonging to the two months under consideration; each row represents a PAD observation, described by means of the above mentioned features. Moreover, a second matrix (**X**) is defined, containing the training dataset, and a third one (**X2**) that includes the test dataset; the two vectors **Y** and **Y2** enumerate the classes (1 or -1) to which each example belongs and are in association to the last two matrices. Once matrices are structured, begins the feature selection algorithm that has the aim to find the few features resulting truly significant for the measures classification.

The ideas on which the feature selection algorithm is based on are linked to Linear Discriminant Analysis, that treats a linear classification problem as a dimensionality reduction one. Considering the case of two classes and supposing to have a D-dimensional input vector **x**, one can project it down to one dimension using

$$y = \boldsymbol{w}^T \boldsymbol{x}$$

where **w** is the weight vector. If a threshold on y is placed and y classified as C_1 class member when $y \geq -\omega_0$ and vice versa, we obtain a standard linear classifier; for our purpose, it's interesting to note that the projection onto one dimension can lead to a considerable loss of information, and classes that are well separated in the original D-dimensional space may become overlapped. However, by adjusting the components of the weight vector **w**, a projection that maximizes the class separation can be found.

Considering a two class problem in which N_1 poinst are of class C_1 and N_2 points are of class C_2, the simplest measure of the separation of the classes, when projected onto the direction given by **w**, is the separation of the projected class means. So, it's natural to choose **w** so as to maximize

$$\mu_2 - \mu_1 = \boldsymbol{w}^T(\boldsymbol{m}_2 - \boldsymbol{m}_1)$$

where $\mathbf{m_1}$ e $\mathbf{m_2}$ are the mean vectors of the two classes while μ_1 e μ_2 are the mean of the projected points. Even constraining **w** to have unit length, there is still a problem with this approach: as illustrated in Fig.4.3.5 we can have two classes that are well separated in the original space but when projected onto the line joining their means, in order to maximize mean distance, a considerable overlap is shown. This difficulty arises from the strongly non-diagonal covariances of the class distributions.

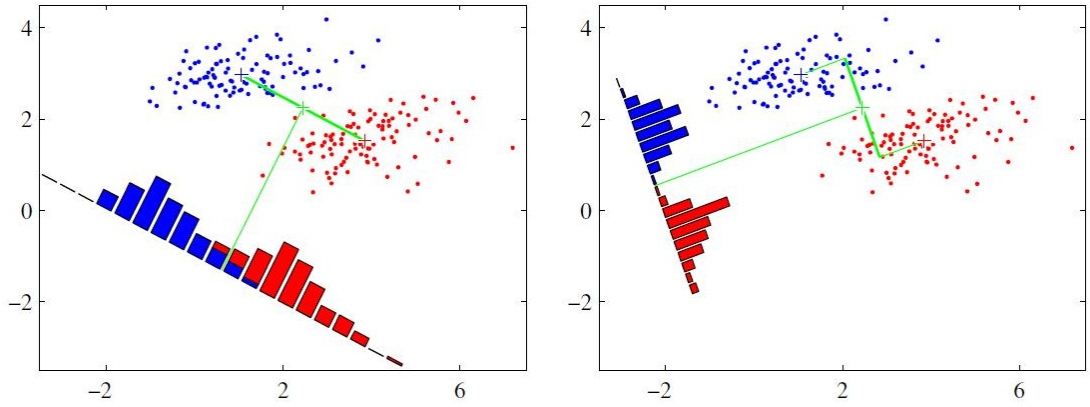


Figure 4.3.5: *On the left panel samples from the two classes are projected onto the line joining the class means with considerable overlapping; on the right panel is shown the corresponding projection based on Fisher approach. Form Bishop Bishop (2010)*

The idea proposed by Fisher (Bishop, 2010) is to maximize a function that will give a large separation between the projected class means while also giving a small variance within each class, so that the class overlap is minimized. The so-called within-class or intra-class variance of the transformed data from class C_k is given by:

$$s_k^2 = \sum_{n \in C_k} (y_n - \mu_k)^2;$$

$$d_{12} = \frac{(\mu_{clear} - \mu_{cloud})^2}{(s_{clear} - s_{cloud})^2}; \tag{4.3.1}$$

where $y_n = \boldsymbol{w}^T \boldsymbol{x}_n$ and the denominator of the expression for d_{12} is a measure of the scattering of these points in the two belonging classes, the so-called intra-class distance.

Based on these considerations, the aim of the implemented feature selection algorithm is to find a set of features that, used together, not only lead to a minimum classification error for training points, but also represent data at best, i.e features characterized by least correlation, in order to avoid redundant information, and, satisfying the Fisher discrimination idea.

The implemented feature selection algorithm, involves a sequential procedure. Initially, each features is singularly considered and a k-fold cross validation is performed, using a Matlab routine: training observations (contained in the **X** matrix) are partitioned in 10 sets (folds), 9 are considered as training examples, while the last fold is cast as test set. In this way, 10 different datasets are obtained each of which samples and uses examples contained in the **X** matrix in a different way. Considering each of these 10 artificial dataset

singularly, the code classifies each measure cast as test example, using the k-nearest-neighbor technique (with k=5), i.e. the test example is classified as cloudy or clear looking at the most recurrent labels among which of the nearest 5 measures to it. The classification error can be determined as the fraction of test examples mis-classified (knowing the true class of this cast test set) and averaging over all the 10 generated set. In this way, for each feature an error estimate, if considered alone, is obtained: at the first iteration, the feature that results to have the minimum classification error is taken or, in case of joint winners, the one that maximizes the normalized distance between the two class examples (inter-class distance) is chosen, following the Fisher formulation (Eq.4.3.1).

For the choice of the second feature, the classification error is again considered, but now evaluated on each pair of features, the first of which is the one already chosen. In addition, because with the increasing of the index of the iteration a lot of features result to lead to the same classification error, other two parameters are considered: the correlation between the new feature and the one already selected and the inter-class distance again as defined in Eq.4.3.1. With this aim, rank are considered: in fact, from correlation and separation number values the algorithm pass to consider their position in their sorted list for all equally good features; in this way the algorithm choose the feature that has the best rank in both of this lists.

When other features are added, the correlation is calculated for each feature with respect to all the chosen ones; moreover, a weight is inserted (around 1.5) in order to consider of greater importance a feature that can best distinguish the members of the two groups. In fact, the aim is the classification, so the features chosen must be able to separate well the two groups; the little correlation is also important but less critical.

This sequential process is interrupted when either the classification error remains stable for the last two iterations, or it increases with respect to the previous iteration value. The final set of features (BT differences) identified are the most relevant for the classification task and are a small number, usually 3 or 4 members.
Once the features are selected, scatter plots, with a BT difference versus another one, can be realized, plotting all the training observations at the same time and visualizing the two group distributions. Two example are reported in Fig.4.3.6 and Fig.4.3.7 showing the scatter plots of training examples as function of two selected BT differences, for the two months May and June. It can be seen that the means of the two classes are well separated with a little overlapping.

This process is characterized by a basic randomness: in fact, once the training examples are fixed, the partition of the training dataset for the cross-validation process has a random character. As a consequence, repeating the algorithm, different features are generally chosen at each run. But, taking 10 runs for the same pairs of month and considering about 20 different pairs during the year, it can be qualitatively shown that some features are recurrent (see Table4.2): the BT difference between the channel around 650cm^{-1} (CO_2 absorption

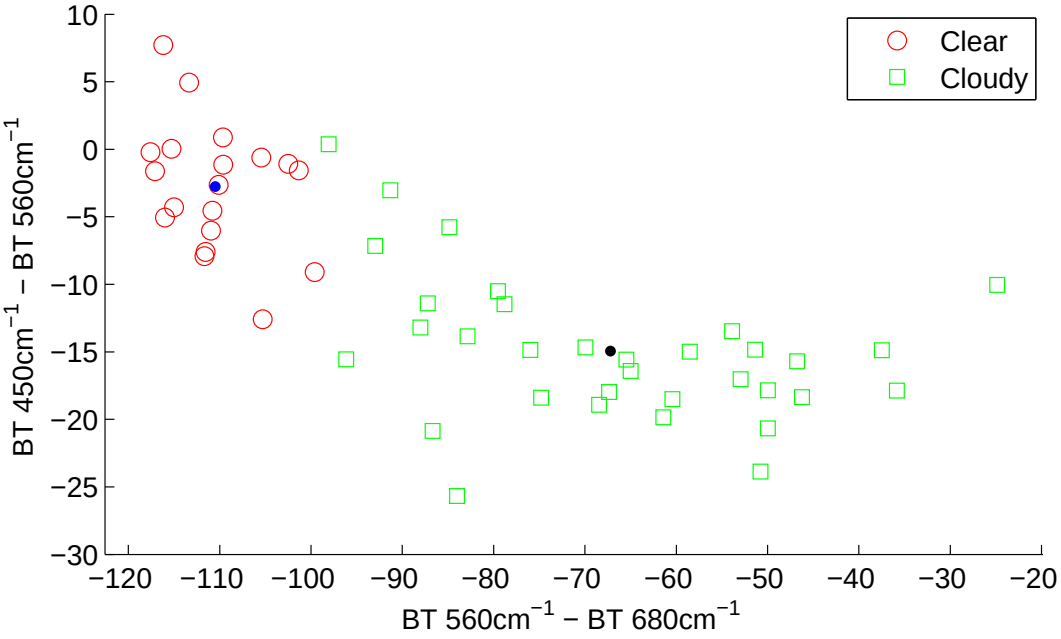


Figure 4.3.6: *Example of scatter plot for training points selected for May and June*

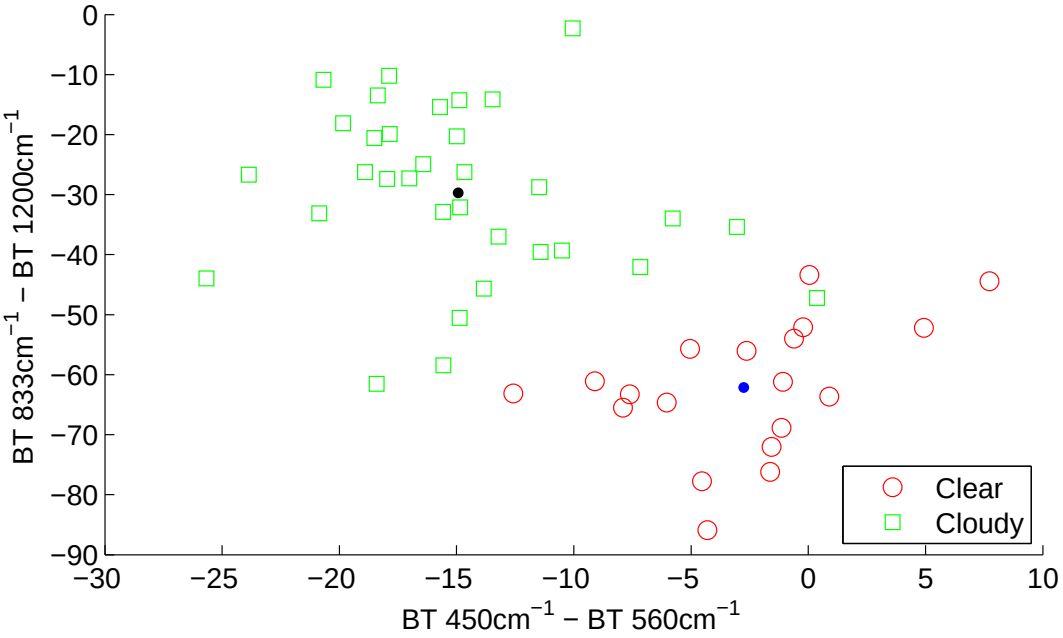


Figure 4.3.7: *Example of scatter plot for training points selected for May and June*

feat $N°$	**occurrences**	**BT1**	**BT2**	**type**
14	51	560	680	FIR-CO_2
61	34	450	560	FIR-FIR
6	30	680	833	CO_2-WIN
2	27	833	1200	WIN-WIN
9	27	573	833	FIR-WIN
34	23	498	560	FIR-FIR
36	21	498	523	FIR-FIR
1	20	900	1200	WIN-WIN

Table 4.2: *List of the most important features selected after running the feature selection algorithm 10 times for 17 different couples of months, in order to discriminate among clear and cloudy cases.*

band) and the one around 560 cm^{-1} (FIR) results to be very significant, since it occurs almost for all the considered month combinations. As can be expected, the relevant information for this classification regards the brightness temperature in window regions, very low for clear scene higher vice versa; in fact, in third position there is another similar difference: the one between BT in the CO_2 band and around 833 cm^{-1}. However, in the following list the BT difference between two selected intervals in the IR main window and the one between two intervals in the (near) FIR region (450 - 560 cm^{-1} micro-windows) appear very significant too. This features are related to the fact that in clear conditions the measured radiance in the window region, especially between 800 and 1000 cm^{-1} show a plateau with very low values, while in the presence of a cloud the radiance field assumes some slope, giving a peculiar effect on the above mentioned BT differences. The above mentioned randomness, added to subjectivity about Support Vector Machine threshold described below, has the consequence that the classification varies, not too much fortunately, for every iteration.

As a corollary, it can be said that this discrimination process can be done also considering only the FIR spectral region, even if the signal measured in dirty windows is more complex.

It must be said that this approach is quite simplified; in fact, radiance uncertainties are not taken into consideration: a complete treatment of the data analysis that accounts also for instrumental errors would be more complex.

Once the most important features for the classification are found, in order to be formally correct, the observations used as training or test datasets are removed from the complete set. The new is passed to the Support Vector Machine (SVM) for the final classification, which is done using the training dataset in terms of the most important features already selected only. A linear kernel is used in this case and the code for SVM is an optimized one, written by Michael Mavroforakis (2003).

At the heart of SVM learning method there is the margin concept: SVM is a classifier that aims at separating the two class elements, not only looking at the classification error, but also at the maximum distance between elements belonging to the two classes. Considering a classifier as:

$$\boldsymbol{w}^T\boldsymbol{x} + w_0 = 0;$$

the margin is defined as the region between the two hyperplanes:

$$\boldsymbol{w}^T\boldsymbol{x} + w_0 = 1; \qquad \boldsymbol{w}^T\boldsymbol{x} + w_0 = -1$$

It can be demonstrated that the distance between each point on these two hyperplanes and the classifier is equal to $\frac{1}{\|\boldsymbol{w}\|}$. From an heuristic point of view, the situation is the one reported in Fig.4.3.8: the elements of the two classes are to be separated by means of a plane; two linear classifiers are shown in the figure with respective margins and both of them leads to the same classification error but their quality is different.

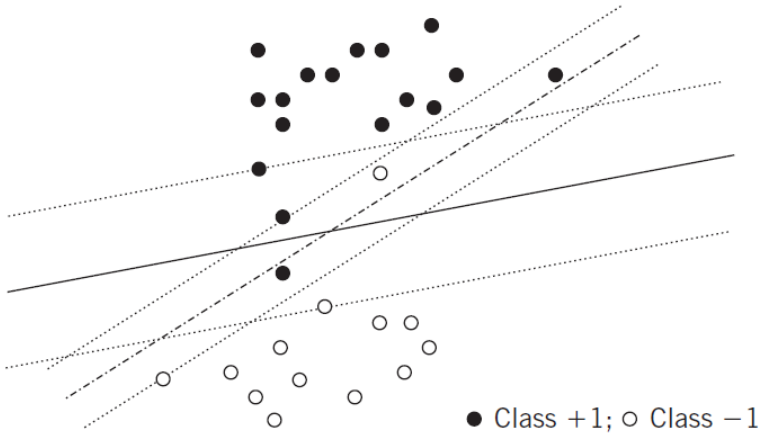


Figure 4.3.8: *Two linear classifiers and the associated margin lines for a 2-class classification problem, from Theodoris and Koutroumbas (2010).*

The best plane is indeed the one identified with a solid line in Fig.4.3.8, because the distance from the regions where the points belonging to the two class lie is maximum and it is the one that probably guarantee a good generalization ability, that is the ability to operate satisfactorily with data outside the training dataset.

With this background, in the SVM frame, the cost function corresponding to the classification error does account not only for those points on the wrong side of the classifier, but

also for any point that lies inside the margin, even if it is on the correct side of the classifier. Only points that lie outside the margin and on the correct side of the classifier make no contribution to the error-counting cost. So, minimizing the cost function means finding that particular separating hyperplane that has the maximum distance between the two class elements. This characteristic guarantees to the SVM good performance in the generalization ability (Theodoris and Koutroumbas, 2010).

From the mathematical point of view, given a set of training points $\mathbf{x_i}$ with $i = 1, \ldots, N$, with respective class labels $y_\mathrm{i} \in \{-1, 1\}$, in a linear 2-class classification task, the SVM computes a hyperplane so that the following cost function is minimized:

$$\boldsymbol{J}(\boldsymbol{w}, w_0, \boldsymbol{\delta}) = \frac{1}{2} \parallel \boldsymbol{w} \parallel^2 + C \sum_{i=1}^{N} \delta_i$$

with as constrains:

$$\begin{aligned} \boldsymbol{w}^T \boldsymbol{x}_i + w_0 &\geq 1 - \delta_i \qquad \text{if } \boldsymbol{x}_i \in \omega_1 \\ \boldsymbol{w}^T \boldsymbol{x}_i + w_0 &\leq -1 + \delta_i \qquad \text{if } \boldsymbol{x}_i \in \omega_2 \\ \delta_i &\geq 0 \end{aligned} \tag{4.3.2}$$

where C is a constant that weights the classification errors δ_i that are :

- $\delta_i = 0$ for point laying on the right side of the classifier and outside the margin;
- $0 < \delta_i < 1$ for point correctly classified but laying inside the margin;
- $0 < \delta_i < 1$ for point mis-classified.

So, the aim expressed with this cost function is to make the margin as large as possible, in fact minimizing $\parallel \boldsymbol{w} \parallel$ corresponds to maximize the margin width, but at the same time keeping the number of points with $\delta_i > 0$ as small as possible.
It is found that the solution is given by:

$$\boldsymbol{w} = \sum_{i=1}^{N} \lambda_i y_i \boldsymbol{x}_i$$

with λ_i are the Lagrange multipliers of the optimization process: they are different from zero only for points that lie on the wrong side of the margin or inside it; these last points are the very observations that determine the separator and are called support vectors (Theodoris and Koutroumbas, 2009). Moreover, the classification error is calculated for another group of spectra, chosen as a test for the classifiers, regulated accordingly to the constant C value in order to minimize this error.

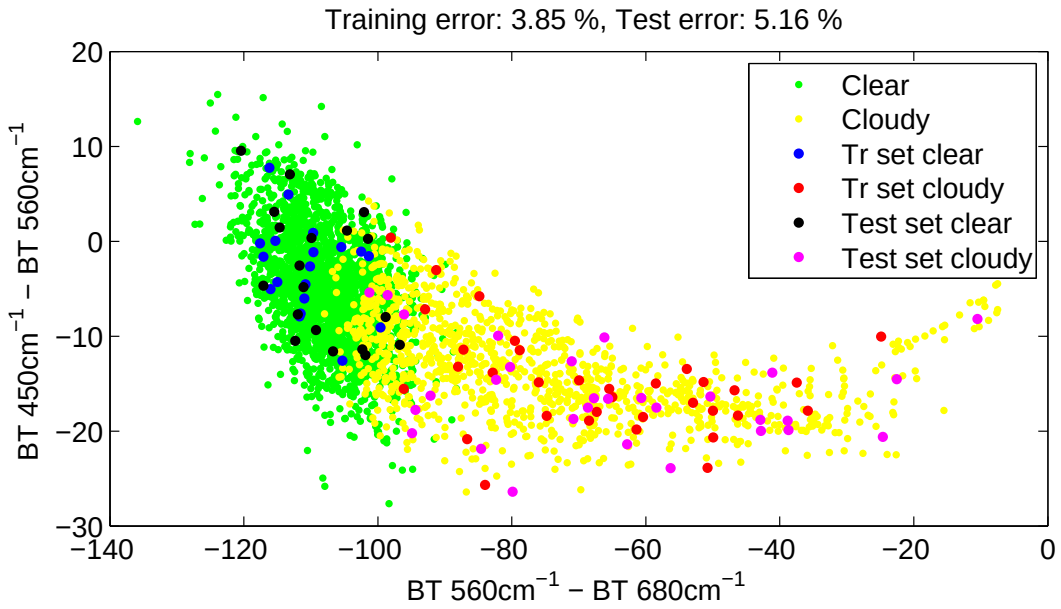


Figure 4.3.9: *Example of SVM clear-cloudy classification: scatter plot of the overall measured set, with training and test examples superimposed.*

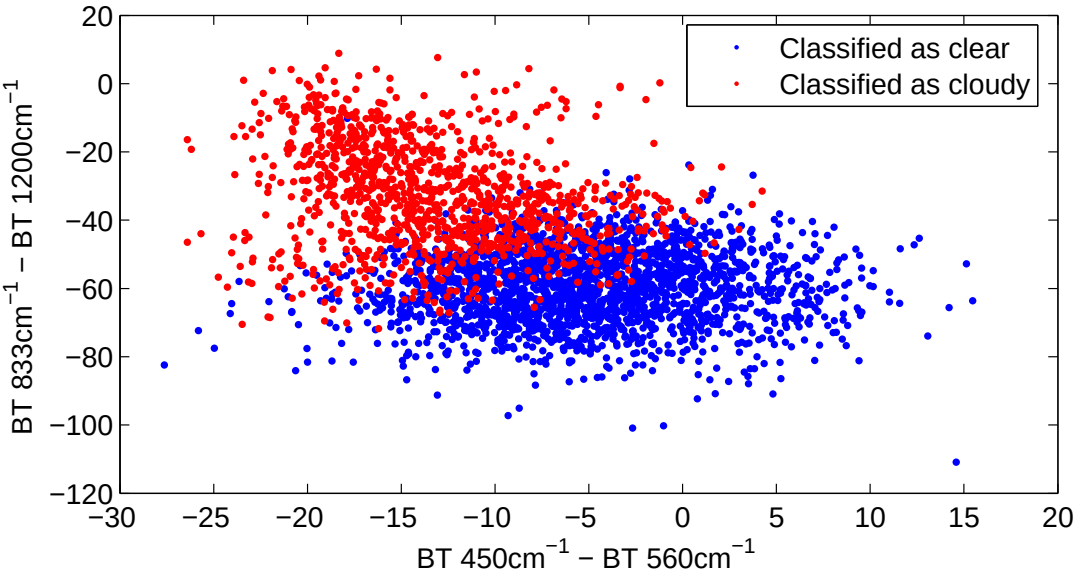


Figure 4.3.10: *Example of scatter plot showing the final result of the clear-cloudy classification, comparable with the Fig.4.3.7 (reporting only training examples).*

An example of the result of SVM classification is reported in Fig.4.3.9, where also training and test points are plotted; the test error is reported together with the training error in the title of the figure. In addition, in Fig.4.3.10 the final classification of the overall observation set for April and May is shown: the scatter plot is comparable with the one in Fig.4.3.7; the absence of a net boundary between the two classes is evident.

In this way, it's possible to obtain a final classification vector that contains all the observations of the two months under consideration. For a final further and useful check, it is possible to generate a list of spectra classified as clear and as cloudy after this procedure, so that the output can be compared with quick-looks Lidar again. In addition, having the need to give a list of high confidence clear spectra for the entire 2013, once fixed the SVM classifier and chosen the C constant, only cases that lie outside and on the right side of the margin are considered. The comparison of this final shortlist with Lidar data appears to be very good; nevertheless, few observations that are taken in the presence of weak diamond dust episodes or faint high clouds are rarely classified as clear.
On the contrary, if all features are passed to SVM the result turns out to be of less quality and more confused.

At the end of the analysis, a comparison is performed between clear spectra classified by SVM and those classified after the fixed thresholds concerning radiance around 900cm^{-1} and it is verified that no spectrum considered potentially cloudy after that analysis is now classified as clear.

Once clear and cloudy cases are discriminated, it is necessary to separate water from ice clouds. During summer months cloud layers, also above 1000m a.g., characterized by a low depolarization ratio, are found. In fact, water clouds can sometimes be detected with Lidar at Concordia Station: these clouds are composed of supercooled water droplets that can survive until $-40°C$, when homogeneous nucleation occurs. These water clouds often produce ice crystal precipitations, clearly visible from Lidar again. In addition, a low depolarization signal, linked to the passage over the instrument of the plume produced by the Dome C power station, can be detected in particular during summer month. In fact, aerosol with a base component of black carbon produces a depolarization signal similar to that of water droplets and wind direction, observed by the in loco weather station, can be used as a confirmation hint.
As an example, in Fig.4.3.11 the quick-look for 11 December 2013 is reported: a short smog episode is visible during the night, characterized by high backscatter signal and low depolarization, while during the morning and early afternoon, a water or mixed-phase stratus cloud appears above 1000m a.g., with low depolarization values. From the lower panel (Lidar depolarization) the precipitation of ice crystals from this mixed-phase stratus is well detectable.

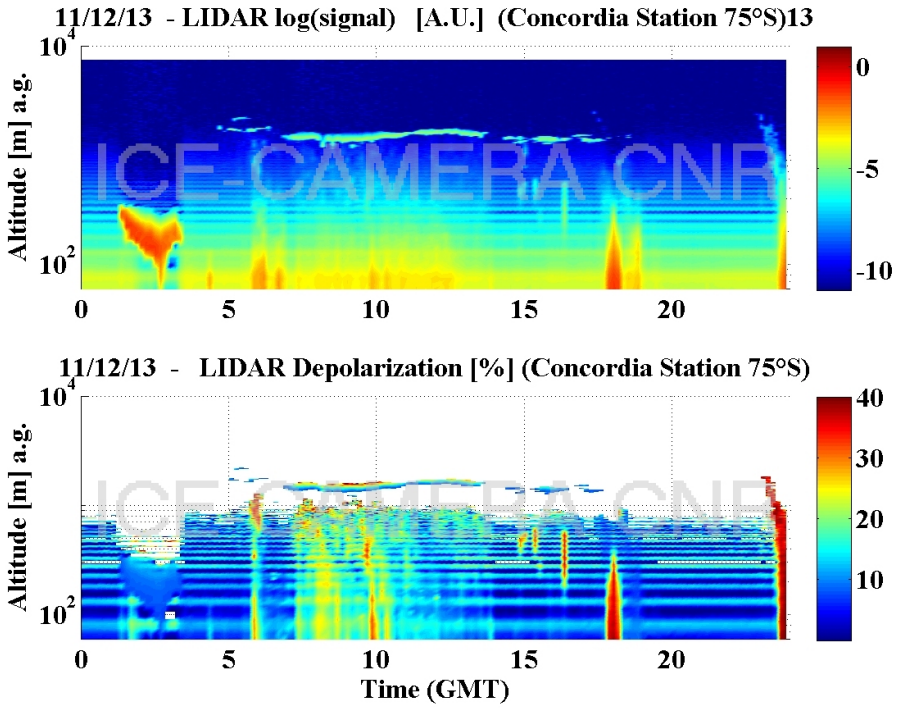


Figure 4.3.11: *Lidar quick-look for 11/12/13 with smog (during the night) and water cloud episodes*

The Matlab script *FeatSelectionfase.m* is used to separate water from ice clouds, on a monthly basis; due to the low number of spectra really available during February, March and November, the only two months that see an important presence of liquid phase clouds are January and December; so, for these two months only, an analysis, similar to the one previously discussed, is implemented.

20 iterations of the algorithm are performed to extract the most significant features. Results are shown in Tab.4.3: the most recurrent BT differences regard two FIR channels, beginning with the one between the interval around 380 cm^{-1} and 498 cm^{-1}. This fact may be related to the strong difference between the absorption coefficient for water and bulk ice in the FIR spectral region (Fig.4.3.12). This different trend may modulate the spectral signal from cloud of different thermodynamic phase. Other important features include FIR-WIN or CO_2 -WIN BT differences.

Once the most useful features for classification are detected, generally 3 or 4 again but in some runs their number can reach 7-8 units, the SVM is processed. Since the separation between the two groups is more confused, the margin of the SVM is built using a non-linear kernel, based on Radial Basis Functions (RBF).

feat $N°$	**occurrences**	**BT1**	**BT2**	**type**
114	18	380	498	FIR-FIR
42	14	480	560	FIR-FIR
104	14	390	440	FIR-FIR
110	10	380	573	FIR-FIR
62	9	450	530	FIR-FIR
90	9	410	450	FIR-FIR
87	8	410	498	FIR-FIR
211	8	270	1200	FIR-WIN
4	7	680	1200	CO_2-WIN
37	7	480	1200	FIR-WIN
91	7	410	440	FIR-FIR

Table 4.3: *List of the most important features selected after running the feature selection algorithm 20 times for December and January only, to discriminate ice from water clouds.*

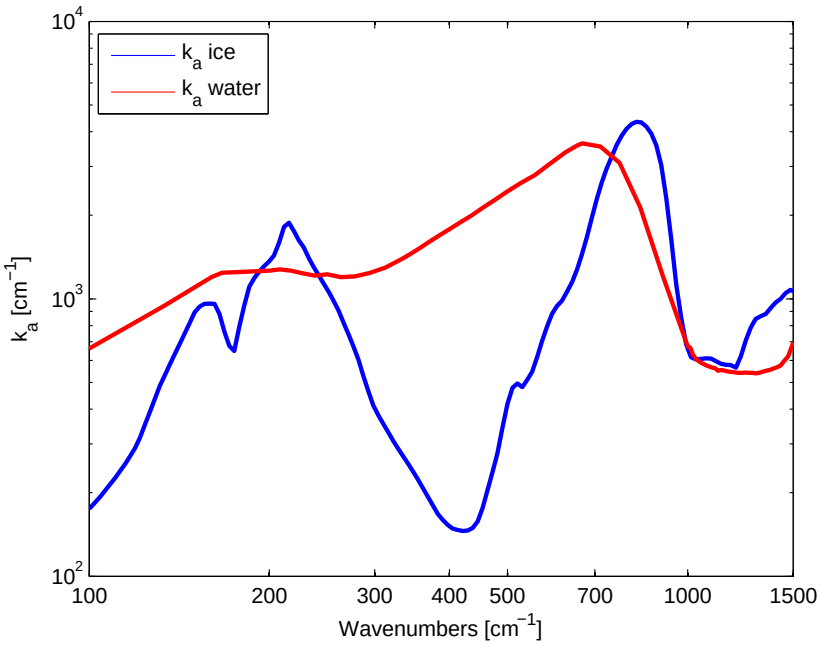


Figure 4.3.12: *Spectral absorption coefficient for water and bulk ice, respectively from Zolotarev et al. (1969) and revised Warren (1984) data.*

From an historical point of view, Radial Basis Function were introduced in the frame of exact function interpolation: given a set of input vectors $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ along with corresponding target values $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$, the goal is to find a smooth function $f(\boldsymbol{x})$ that fits every target value exactly. This is achieved by expressing $f(\boldsymbol{x})$ as a linear combination of specific functions $h(\parallel \boldsymbol{x} - \boldsymbol{x}_n \parallel)$ (RBF indeed), one centered on every data point $\mathbf{x_n}$ and whose value depends only on the distance from it:

$$f(\boldsymbol{x}) = \sum_{n=1}^{N} \omega_n h(\parallel \boldsymbol{x} - \boldsymbol{x}_n \parallel)$$

In our contest, RBF are used as kernel functions to perform a non-linear classification task. Radial Basis Function are defined as:

$$h(\boldsymbol{x}) = \exp(-\frac{(\parallel \boldsymbol{x} - \boldsymbol{c}_i \parallel)^2}{\sigma^2})$$

where $\mathbf{c_i}$ is the mean point that is the center of the RBF. The resolution of a non-linear classification problem is generally done by mapping vectors from the original space into a high dimensional one, where the class results linearly separable; but, using this special functions, the problems is reduced to a kernel evaluation in the original space (Theodoris and Koutroumbas, 2009). In addition to the C parameter, linked to the margin width, in this case another free parameter σ is introduced, which correspond to the radius of the RBF. Intuitively, a small value for σ implies that the RBF drops rapidly toward zero around each $\boldsymbol{x}_i \in \boldsymbol{X}_i$, which leads to an increase in the number of support vectors and, as a consequence, a low training error with overfitting. The SVM won't be able to predict the test points due to the incapability of covering the space. On the contrary, if σ is too large, the results will be poor for both training and test datasets. In fact, all RBF remains almost constant in the area where the data points lie and the summation in the classification rule shows slow variation for the various x's, which leads to reduced discrimination capability. In conclusion, low values of σ gives a strictly local classification too tuned on training example and unable to generalize, while large value of RBF radii leads to flat classifiers, undesirable for classification tasks. So, both C and σ have to be tuned in order to minimize empirically test error and have a good classification (Theodoris and Koutroumbas, 2010).

An example of SVM classification result is shown in Fig.4.3.14, for the January month.

In the end, it's possible to build the list of spectra classified as water or ice clouds, so that a final control can be performed, looking at the Lidar quick-looks again. In general the classification gives good results, but difficulties are noted, linked to clouds that have small optical depth: in these cases features selected as meaningful are not helpful in the discrimination, since spectra of both cloud types tend to converge to a clear spectrum, losing their specific characteristics.

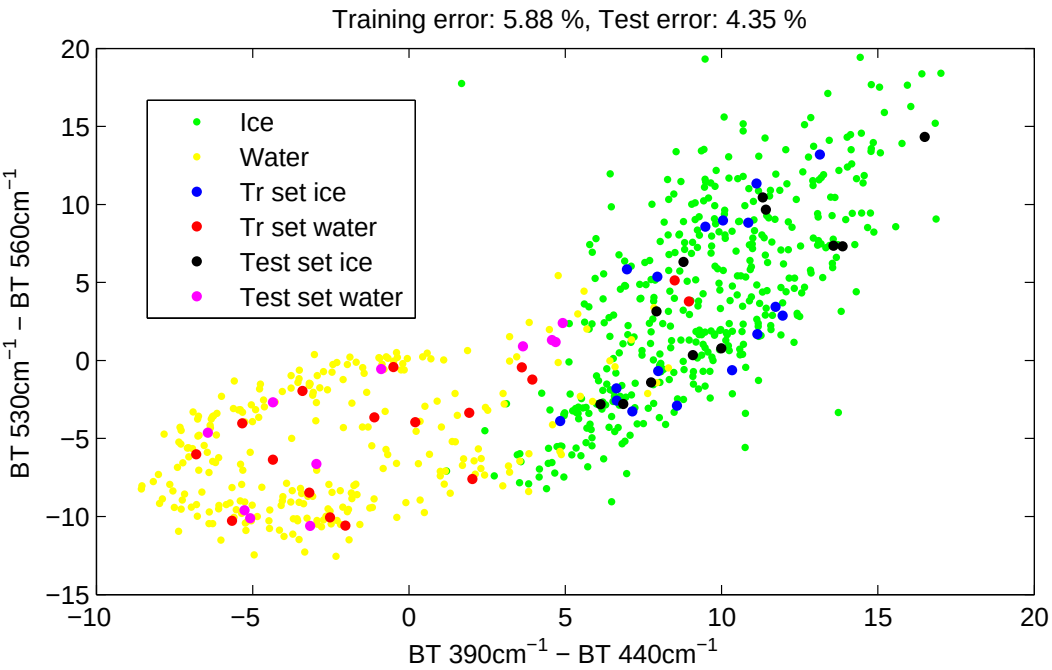


Figure 4.3.13: *Example of SVM ice-water clouds classification: scatter plot of the overall January dataset, with distinction of training and test examples*

For months during which important water clouds are not present, and for January and December after the phase separation, the last part of the classification procedure is performed to separate spectra relative to precipitating ice clouds and to non-precipitating/high ice clouds. The same procedure is applied: features that are able to separate these two groups, if possible, are searched, with a supervised learning approach. As at the beginning, months are considered in pairs, in order to have a valuable number and variety of conditions also for those months for which available spectra are fewer. This classification is more problematic than the previous ones; in fact, any feature, either alone nor with others, isn't able to obtain a really satisfying separation of observations. As a consequence, in the majority of the cases the classification error stays around 10% for the training dataset and is higher for the test dataset: the risk is the overfitting, that leads a classifier adherent to training example but not so useful for other observations.

Selected features generally result in a limited number, rarely over 4 units; in addition, the random component of the process on the one hand and, on the other, the similarity in the performances of many features imply that the selected group varies at each iteration. However, in the majority of the cases, the chosen features tend to correspond to BT differences between two FIR micro-windows or, at most, between a window in the FIR region and the

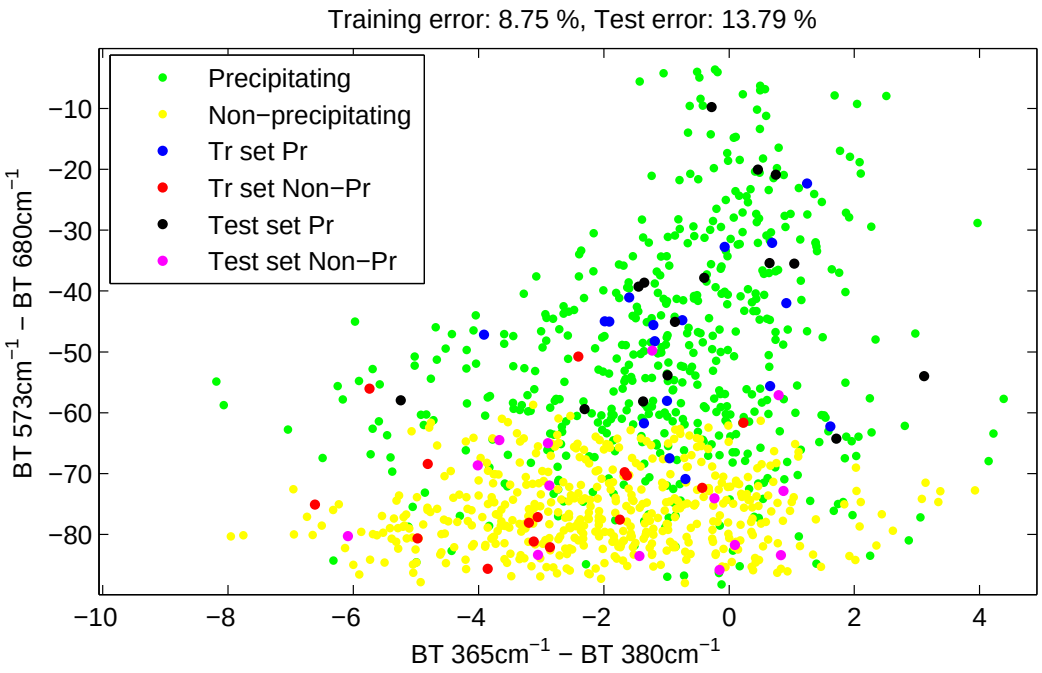


Figure 4.3.14: *Example of SVM classification precipitating/non-precipitating ice cloud: scatter plot of the overall May and June dataset, with distinction of training and test examples*

interval chosen around 900 cm^{-1} or again the slope in the main window region. The 573 cm^{-1} is a window frequently selected and is located in the region of maximum sensibility to cloud signal (it is chosen at least in a third of the total number of runs). In the end, this time it is difficult to highlight possible significant features or link the resulted BT differences to particular cloud properties, due the complexity of radiative phenomena that occurs in this spectral region. Results from 10 iterations for 17 pairs of months are shown in Tab.4.4.

As regards the development of the SVM, the kernel is based on Radial Basis Function again, in order to determine the discriminant and the margin; it's important to modify by hand the constant C and the width of RBF to minimize error on test dataset and have a reasonable classification. Once a list of spectra classified by means of this routine is obtained, it is possible to check the results with Lidar quick-looks again. Difficulties are evident in the clear identification of these two groups: if the most important cases are generally well classified, problems are more important for faint precipitating clouds or non-precipitating clouds with high optical depth.

feat $N°$	**occurrences**	**BT1**	**BT2**	**type**
20	23	530	573	FIR-FIR
26	23	523	573	FIR-FIR
136	23	365	380	FIR-FIR
229	17	270	330	FIR-FIR
276	17	238	260	FIR-FIR
28	12	523	530	FIR-FIR
33	12	498	573	FIR-FIR
36	12	498	523	FIR-FIR
3	10	833	900	WIN-WIN
64	10	450	498	FIR-FIR
231	10	270	295	FIR-FIR
255	10	238	900	FIR-WIN

Table 4.4: *List of the most important features selected after running the feature selection algorithm 20 times for 17 different couples of months, to discriminate precipitating from non-precipitating ice clouds.*

Moreover, the classification seems to be particularly poor for October and November, I think due to the elevate noise that characterizes spectra during this period. For these two months in particular, but in general for the entire year, it's necessary, on the base of Lidar measurements, to perform an accurate check and to correct the mis-classifications and evident errors that occur. In fact, spectral signatures depend on a lot of parameters, starting with particle size and habit, but also optical depth, water vapor profile, aerosol concentration: so, it's difficult to think that on the base of this type of approach, it is possible to infer a satisfying classification in the two final groups.

Chapter 5

Cloud properties retrieval and simulation

5.1 Case studies selection and description

The classification process, has provided a limited number of cloudy case studies: on these cases a detailed study will be performed, by means of a retrieval methodology to extract cloud properties, followed by radiance simulations, to test our ability in the reproduction of the measured spectra.

The selection of case studies begins with the inspection of the obtained monthly lists of non-precipitating ice clouds: in fact, in absence of Lidar derived IR optical depth profiles, of fundamental importance for the simulation of precipitating clouds, we have focused our attention on non-precipitating ice clouds only, excluding also water or mixing phase clouds. The inspection is driven by the consultation of the quick-looks Lidar dataset, in order to detect the non-precipitating ice cloud most interesting events: in fact, the lists of spectra, a part from mis-classified events, also contain episodes that could be very difficult to understand, such as very thin clouds, diamnond dust episodes and so on.

Only about thirty non-precipitating ice cloud episodes are detected for the entire 2013, for each of which both quick-looks Lidar and good quality REFIR measurements are available. For each case, Lidar profiles are plotted and analyzed in order to evaluate the structure, temporal variability and geometry of the cloud: in fact, important parameters that must be supplied in the retrieval scheme are the top and the base heights of the cloud. As a result, among those spectra pertaining to a selected cloud case, only one of them is chosen: this spectrum is evaluated as the best one, looking at quality and noise of measurements, signal intensity and sharpness of the cloud structure. Eventually, for noisy observations, two or three spectra are averaged, but only if the cloud structure results sufficiently stable.

In the Table5.1 below, the 29 case studies finally selected are listed and described.

N	Date Hour Type	Description	Top - Base [m]
1	02/01/13 12:04 S	Short lasting non-precipitating cloud, characterized by a variable structure, low thickness and a very high base at the beginning, lowering with time. Only 3 REFIR good spectra are available with signal quite weak and stable.	7730 - 6830
2	03/01/13 06:05 S	Non-precipitating cloud between precipitating moments; its structure is highly variable, so a single spectra is considered while the cloud base is decreasing in height. Lidar and REFIR signal well-reveals the geometry and the presence of the cloud.	6030 - 5030
3	19/01/13 10:49 S	Short lasting episode of cloud characterized by precipitating ice crystals not reaching the ground; the cloud structure shows variability and some Lidar profiles present multiple layers. A single spectrum with well defined cloud profile and evident differences from the simulated clear one is chosen.	6230 - 4130
4	19/04/13 08:43 A	High, thin cloud not followed by precipitating phase; Lidar backscattering signal is very weak, due to elevated base height, while REFIR spectra show a well-defined and very stable cloud signal. It has a structure not so variable, so an average of two spectra is considered.	7030 - 5930
5	22/04/13 05:07 A	Cloud followed by a precipitating phase, with increasing optical depth and decreasing base height, sign of slowly falling ice crystals. Lidar depolarization signal results very strong. Cloud spectra considerably differ from simulated clear ones and are quite similar among them; so a couple of spectra are averaged.	5230 - 4130
6	22/04/13 14:14 S	One of the highest cloud clearly detected over the year, the signal is weak but defined, the structure variable with layer appearing and disappearing over 20 min. The depolarization signal is quite undetectable. The only spectrum evidently different from the clear one is chosen.	8230 - 6030

7	23/04/13 06:39 S	Episode of non-precipitating cloud between precipitating moments; the cloud structure is highly variable, so not to be averaged. Lidar and REFIR signals are well defined, due to an optical depth quite high.	6030 - 5230
8	24/04/13 05:55 S	High non-precipitating thin cloud with base height slowly lowering. Lidar signal results very weak as REFIR one, especially in IR window, while FIR micro-windows are more saturated.	6830 - 6230
9	25/04/13 08:12 S	This is the most important case of non-precipitating cloud not followed by a precipitating phase. The spectrum is chosen in a moment of high optical depth, high depolarization signal and cloud structure quite variable. So, the spectral signal is well-defined and a single observation is chosen.	5730 - 4330
10	09/05/13 07:04 S	Long lasting pre-precipitating cloud phase: in the first period results thin and difficult to detect a part from the considered period around 7 o'clock. Few spectra can be chosen, and the variable structure prevents from averaging.	6230 - 4430
11	09/05/13 12:11 S	In the second period, the thickness increases and the cloud-base decreases in height, is followed by a precipitation event; depolarization signal is quite low. Again a single spectrum is chosen due to variability.	5730 - 4630
12	20/05/13 06:05 S	Important and typical case of ice cloud with decreasing base, followed by a short lasting precipitating phase. Up to 1500 m the Lidar signal is well-defined: the top is here fixed, even if the cloud presence can be inferred from depolarization signal up to 2000m. The cloud has a relevant geometrical depth and a variable structure.	4730 - 3730
13	25/05/13 13:04 S	Short episode of non-precipitating cloud, characterized by a strong Lidar and spectral signal; it seems to have an important optical depth and is characterized by strong depolarization signal. The structure is quite variable and a single spectrum is chosen before the rapid cloud base decrease that occurs at the end of the sequence.	4930 - 4030

14	29/05/13 02:29 S	Non-precipitating thin cloud with Lidar signal quite weak and REFIR spectra that don't show significant difference from clear one; the signal is very weak and some spectra are not of good quality: a borderline case.	6430 - 5430
15	10/06/13 06:49 S	Non-precipitating cloud lowering and followed by a precipitation event, in a typical configuration. Lidar signal is quite weak at the beginning but rapidly increasing, while REFIR signal well-reveals the cloud presence. Depolarization signal results quite strong.	4430 - 3730
16	23/06/13 06:58 S	Short lasting non-precipitating cloud followed by very little precipitations, Lidar signal is defined but not so much REFIR spectra. The structure is highly variable, so a single spectrum with defined Lidar cloud profile is chosen.	6230 - 5230
17	24/06/13 22:56 S	Non-precipitating cloud phase, with rapidly variable structure, low optical thickness and depolarization ratio. In particular, a spectrum with an interesting strong slope in the FIR region is chosen.	5630 - 4630
18	17/08/13 06:07 S	Long lasting case of non-precipitating cloud followed by weak ice crystal falling. Looking at the REFIR measurements, the cloud signal appears quite weak and also Lidar shows a cloud difficult to track. An average can be done on some spectra, but a single spectrum with a quite-defined signal is chosen.	5930 - 4930
19	18/08/13 10:31 S	Typical case of ice cloud followed by a precipitating moment, with base decreasing in height. The cloud structure is quite stable but a single spectrum is considered, for which the Lidar profile is particularly defined. REFIR spectra show great saturation in FIR region, in particular towards the end of the period.	5530 - 4730
20	03/09/13 09:24 S	Long lasting non-precipitating cloud layer that gives a quite defined spectral signal, especially for the chosen last observation of the sequence; on the contrary, Lidar shows a weak signal, with cloud that is sometimes at the same magnitude of instrumental noise, fact maybe linked to a temporary reduced power of the instrument.	5730 - 4730

21	16/09/13 13:56 S	Pre-precipitating cloud, with variable structure characterized in some moments by two layers and a base heights in sharp decrease. Looking at the intense signal from REFIR and the well-defined profile from Lidar, a single spectrum is selected.	6430 - 4730
22	27/09/13 21:13 S	High cloud clearly detectable between the presence of a very low mixed phase cloud; the signal is well-defined both from REFIR and Lidar. The cloud shows a strong depolarization signal; its profile remains fairly constant, despite the bottom tends to lower: a single spectra is chosen.	5930 - 4730
23	06/10/13 18:31 S	Pre-precipitating high cloud with lowering bottom, for which are available few spectra due to the concurrent presence of a low cloud. Spectral signal is very intense, with strong attenuation of FIR micro-windows; Lidar profile shows a straight peak under 1500m with a weaker signal above and a depolarization ratio extremely intense.	5430 - 3930
24	15/10/13 04:28 A	Isolated cloud, characterized by falling ice crystals not reaching the ground. REFIR signal is weak, especially at the end of the series. Lidar signal results quite well-defined, with cloud structure fainting at the end and cloud base lowering; however an average over two spectra is done in order to reduce noise and avoid too many negative window values.	4730 - 3830
25	18/10/13 07:31 S	A short lasting non-precipitating cloud between precipitating moments, with highly variable structure. Lidar signal is defined only for spectra around 7 a.m. o'clock, one of which is chosen, while REFIR dirty windows are quite closed.	5330 - 4130
26	30/10/13 09:41 A	Low faint ice cloud with Lidar backscatter signal that between 8 and 11 a.m. indicates a base not under 500m and that is sometimes confused due to strong molecular scattering in the first hundreds meters. The structure is variable, but an average is forced in order to reduce the high noise of these spectra, which are not so different from clear ones.	5430 - 4030

27	26/12/13 17:09 S	Partially non-precipitating cloud characterized by a bottom in fast lowering during the afternoon, with defined signal from both REFIR and Lidar. Highly variable structure and spectra characterized by particularly low noise prevent from averaging.	6230 - 4330
28	26/12/13 21:06 S	During the evening of the same day, a thin cloud appears: it results particularly difficult to track looking at the weak Lidar signal, a part from the last spectrum, that has been chosen. Also REFIR measurements show little differences with simulated clear spectra.	5130 - 4730
29	27/12/13 16:28 S	Long lasting pre-precipitating cloud, with typical optical depth that tend to increase and the base height to lower; spectral signal well-reveal the cloud presence while Lidar profiles are more confused.	5830 - 4730

Table 5.1: *Description of the 29 case studies relative to non-precipitating ice cloud events; in the first column a progressive chronological numeration, in the second the date and type (S=single, A=averaged spectrum), in the third a brief description and in the last top and base heights.*

The above listed cases, present a quite high variability in terms of cloud structure and duration: four examples are reported in Fig.5.1.1. The case (a) is a high cloud, not followed by a precipitating phase, not so frequently seen by Lidar because of its short vertical range. Panels (b) and (c) show two clouds characterized by a pre-precipitating phase during which the spectrum is chosen; cloud base height gradually decreases until falling ice crystals reach the ground. These are quite frequent situations that occur during all the seasons. Finally, panel (d) shows a cloud that seems characterized by precipitating ice-crystals in mid-troposphere, not reaching the ground.
Selected cases generally fall within one of these typology, with bottom height that spans between 500m a.g. up to 3500m a.g. for the higher detected cloud on 02/01/13. The typical geometrical depth is around 1000m, with a pair of very thin cloud cases, such as the 26/12/13 with a thickness of 400m, and rare thick clouds, such as the 06/10/13 around 1500m).

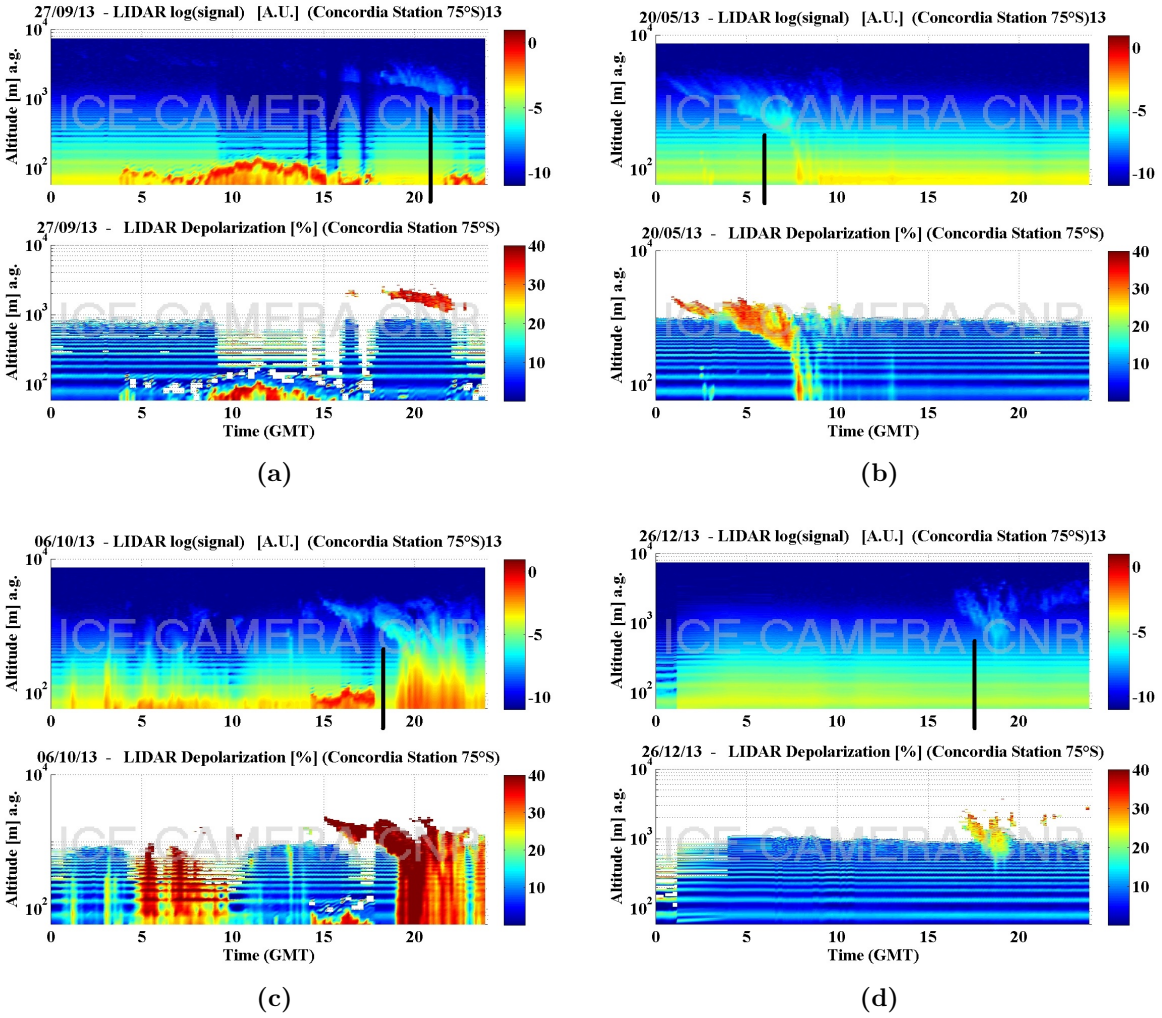


Figure 5.1.1: *Example of some cloud events chosen as case studies: (a) high non-precipitating ice cloud, (b) and (c) pre-precipitating clouds, (d) falling ice crystals in mid troposphere. The black lines indicates the time of the selected spectrum.*

5.2 Retrieval methodology and results

Various methods have been developed to retrieve cloud properties from radiance measured by means of passive sensors, exploiting information, in most of the cases, provided by short wave channels in the near IR region; but, this shortwave retrieval methodologies are applicable only during daytime, due to the use of scattered solar radiation. Recent advances in satellite instrument technology, clear and cloudy radiative transfer models, new parameterization of ice cloud and aerosol optical properties and computer performances offer the opportunity to develop sophisticated cloud microphysical retrievals using passive HSR-IR data.

The retrieval methodology adopted in this study, named RT-RET (Maestri and Holtz, 2009) uses HSR measurements in the IR window, which are sensitive to spectral variations of cloud OD, linked to wavenumber dependence of the absorption coefficient of ice and water; REFIR-PAD measurements have the advantage of providing micro-window channels with minimal gas absorption. The algorithm retrieves the spectral OD from PAD measurements by solving the multiple-scattering radiative transfer equation and, unlike previous techniques, allows to obtain the particle size by fitting the retrieved spectral absorption OD (AOD) in IR window channels with a precomputed database. This new technique handles complex cloud structures, since allows the introduction of layers at different temperatures. This method requires, in addition to HSR measurements, a description of vertical atmospheric state, surface skin temperature (and emissivity) and cloud geometrical boundaries.
In summary, the algorithm is divided into 3 steps:

- Clear-sky gaseous OD computations: using temperature and gaseous profiles, the HSR gaseous ODs are calculated and then convolved to match the interferometer's spectral resolution;

- Cloudy retrieval – first iteration: a First Guess is used for the cloud Scattering Properties (FGSP) for a first iteration of the cloudy simulation. Simulated radiances are fitted to the measured radiances in selected micro-window channels between 800 and 1000 cm^{-1}. The retrieved AOD is compared with the AODs stored in a precomputed database.

- Identification of cloud optical and microphysical properties - final iteration: The optical properties retrieved in the first iteration are used as the second-guess scattering properties (SGSPs): RT-RET is initialized for the second iteration, that is necessary to obtain a spectral OD independent of the FGSP.

The following section deals with the first point, while the next three address retrieval methodology and other issues related to the last two items.

5.2.1 Clear-sky gaseous optical depth computation

The first step concerns the calculation of gaseous optical depths for a clear atmosphere; the software used to this scope is LbLRTM: this is a line-by-line radiative transfer model extensively validated against atmospheric radiance spectra from the ultraviolet to the sub-millimeter region (Clough et al., 2005). LbLRTM model demonstrates speed, reliability and errors associated with computational procedures five times less than those associated with the line parameters. The spectroscopic gas parameters are taken from HITRAN 2008 database (Rothman et al., 2009), with continuum absorption for different gas types accounted for.

As mentioned at the end of Ch.3, for each case study a temperature and a humidity profile are derived from radiosonde and ECMWF datasets. Depending on the availability of in situ weather balloon measurements, daily radiosonde (RS) profiles or ECMWF reanalysis are selected and a temperature and a water vapor profiles are time-interpolated to the measurement time. In situ RS profiles are used up to 3 km below the maximum height reached by the balloon, above which the ECMWF profile is merged with RS one and finally only reanalysis are used up to the top level. If radiosondes are partially or not available, only ECMWF profiles are exploited and interpolated. For other molecular species, the profile is taken from climatological mean concentrations (Anderson et al., 1986).

The final profile is interpolated vertically from the surface to 66 km at a fixed grid, characterized by height levels that are quite close near the surface while increasing going upward. In this way, the first hundred meters in troposphere, where strong inversions take place especially during winter, are accurately described.

Clear-sky simulations can be preliminarily performed and compared to measured clear spectra, in order to evaluate the ability of the used radiative transfer model to reproduce measurements in absence of cloud. An example is shown in Fig.5.2.1 (26/12/13). Obviously, the measured spectrum shows noisy features and many negative values in the window region, but the overall spectrum seems to be well reproduced: systematic differences are evident only below 200 cm^{-1}, linked to noise and mis-accounting for strong water vapor lines in this region, but also above 1100 cm^{-1} where beam splitter absorption produces two particularly noisy features. In addition, both the CO_2 band around 667 cm^{-1} and peaks measured in correspondence of FIR water vapor lines result quite bad reproduced: in fact, the first few meters of geometrical path inside the instrument, where the temperature is higher than outside and from which signal in these strong absorptive regions originates, are not taken into account in the profile.

It is interesting to evaluate clear-sky weighting functions for selected micro-windows, in order to understand where the signal detected at different wavenumbers arrives from. In Fig.5.2.2 weighting functions in clear sky conditions are shown, averaged over some FIR microwindows and IR windows, as reported in the legend. As one can expect, all of them are peaked in the lower troposphere, but, for the most saturated dirty windows, the peak

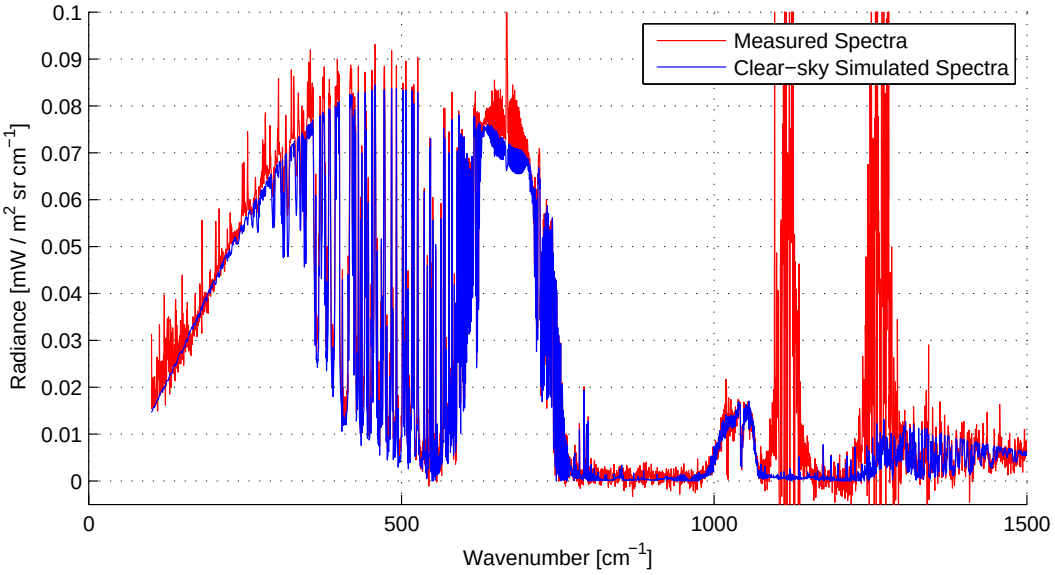


Figure 5.2.1: *Example of measured clear spectrum and the relative simulation for the 26/12/13.*

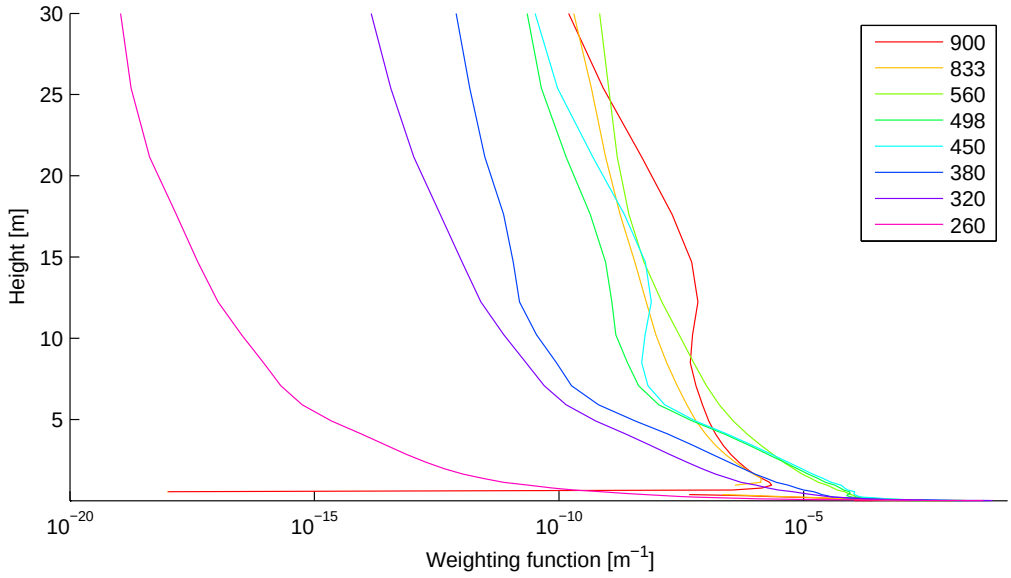


Figure 5.2.2: *Example of weighting functions for selected windows in a zenith observation geometry, for the case of 24/04/13.*

is very close to the ground level, whereas for very transparent windows the curve is slightly less peaky and the decrease with height slower. As a consequence, for dirty windows, the signal measured from the instrument in the presence of a cloud, is linked both to water vapor emission and cloud emission and scattering, while for IR windows (see line for 900 cm^{-1}) the signal arrives principally from the cloud.

5.2.2 Spectral particle radiative properties

As mentioned above, RT-RET model uses a precomputed Particle Size Distribution (PSD) scattering property database (SPD), in order to derive cloud properties. This database contains spectrally resolved extinction and absorption coefficients (K_e and K_a), single-scattering albedos (ω), and asymmetry parameters (g) for multiple PSDs normalized. The database was created using non-spherical single particles and single-scattering properties computed by Yang et al. (2013).

When comparing retrievals from sensor measurements using different methods such as solar wavelength techniques or infrared ones, but also when cloudy simulation are performed in a wide spectral range, the need for spectral consistency in cloud optical properties becomes of fundamental importance. Numerous articles and data libraries were presented about single-scattering properties in relatively limited domains, such as Yang et al. (2003) and Yang et al. (2005). These works contained several inconsistencies in the solar and thermal IR spectral regions, due to differences in particle shape definitions and in computational methodologies. For example, an empirical approach known as the composite method was employed to merge the extinction and absorption efficiencies in the region of overlapping of the two used scattering computation models, with not so satisfying results.

Yang et al. (2013) has developed a spectrally consistent data library containing the scattering, absorption and polarization properties of a set of 11 randomly oriented ice crystal habits at wavelengths from 0.2 to 100 μm; these habits were selected on the base of in situ measurements and, among them, column-type and plate-type are used in this work. The maximum diameter for each habit range from 2 to 10000 μm, spanned by 189 discrete values. Information relating to the volume and projected area of each habit as well as the asymmetry parameter, single-scattering albedo, extinction and absorption cross sections and the six independent elements of the phase matrix are provided.
This library adds to previous work in four ways:

- the used scattering models have been improved;
- the calculations employ the most recent compilation of the real and the imaginary indexes of refraction for ice;
- the aspect ratios used in the calculations are consistent for the whole spectral range;

- the effect of surface roughness is considered with the introduction of the parameter σ

As regards the last item, three values are fixed for this parameter: $\sigma = 0$ for smooth particles, $\sigma = 0.03$ for moderate surface roughness and $\sigma = 0.5$ for severe one. We have performed two simulations choosing column like habit, firstly considering smooth particles and then severe rough ones, in order to understand the radiance sensitivity to this parameter over the FIR region, where scattering process are quite relevant. In Fig.5.2.3 the spectral differences between this two now described simulations are shown, for a cloud with particles of $r_{\text{eff}} = 16\mu m$.

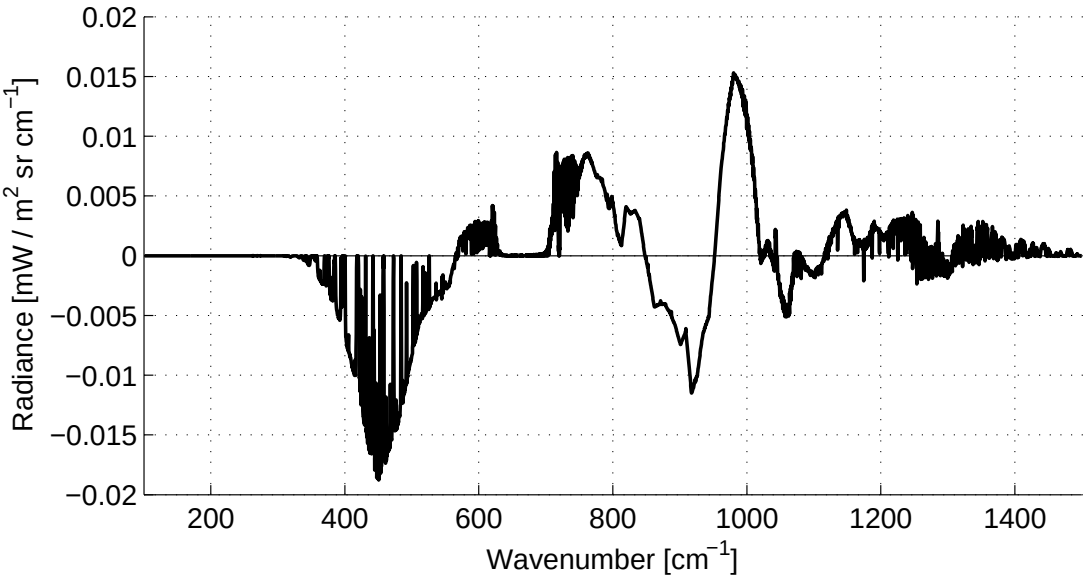


Figure 5.2.3: *Differences between simulations performed for 13/09/03 case, in the presence of ice cloud layer characterized by columnar crystals with $r_{eff} = 16\mu m$, considered the first time as smooth particles and the second time as severely rough ones.*

The differences are very small, about 0.1% of the measured radiance and so totally negligible with respect to, for example, instrumental uncertainties, that, for the same wavenumbers, result over $1\, mW/m^2\, cm^{-1}\, sr$. As a consequence, for all the following retrievals and simulations the roughness parameter σ is set equal to zero.

In the retrieval framework, the cloud properties are computed using modified gamma-type PSDs. The gamma distribution is defined as:

$$n(D) = N_0 D^{\mu} e^{-\lambda D}$$

where, D is the particle's maximum dimension, the λ parameter is linked to the slope of the PSD, μ identifies the variance of the distribution and is fixed at 7, while N_0 is the intercept and determines the total number of particles in the PSD, constraining the IWC. The PSD bulk optical properties are uniquely identified by three parameters: the crystal habit, λ, and μ.

A parameter that synthesize PSD information is the effective radius, whose classical definition was formulated for the first time by Hansen and Travis (1974) as:

$$r_{eff} = \frac{\int_{r_{min}}^{r_{max}} r^3 n(r) dr}{\int_{r_{min}}^{r_{max}} r^2 n(r) dr} \tag{5.2.1}$$

This definition holds for a distribution of spheres, where r is indeed the radius of the sphere. A more general formulation for effective dimension of a PSD composed by non-spherical particles was introduced by Foot (1988) and is used in the retrieval frame:

$$D_{eff} = \frac{3}{2} \frac{IWC}{\rho_i \int_{D_{min}}^{D_{max}} P(D) n(D) dD} = \frac{3}{2} \frac{\int_{D_{min}}^{D_{max}} V(D) n(D) dD}{\int_{D_{min}}^{D_{max}} P(D) n(D) dD} \tag{5.2.2}$$

where P(D) and V(D) are respectively, the projected area and the volume of a particle with maximum dimension D, while ρ_i is the bulk ice density, assumed constant.
Fig.5.2.4 reports several gamma distributions, for two habits. Despite the selected effective dimensions (r_{eff} is roughtly equal to half D_{eff}) are equal for plates and columns, the PSDs are different: in fact, the V(D)/P(D) ratio in the 5.2.2 depends on habit, hence the same D_{eff} is obtained from different PSDs for diverse habits.

5.2.3 Model approach and phase function calculation

Both RT-RET and RTX models, respectively used for the retrieval and the forward simulations, are based on a doubling and adding method (Liou, 2002), adopted to solve the multiple-scattering RT equation. As explained in Evans and Stephens (1991), a plane-parallel, vertical inhomogeneous scattering atmosphere is taken into consideration, with the presence of randomly oriented particles. RTX calculates the monochromatic polarized radiation emerging from such an atmosphere, with the angular dependence of the radiation represented by means of a Fourier series in azimuth angles and discretization of zenith angles. The doubling and adding approach reveals numerical stability for a large interval of optical depth and computes the radiative properties of the medium rather than the radiance field itself: in this way, the radiance exiting the atmosphere may be found for different boundary conditions after the complete solution is computed. For our purpose, having a ground-based instrument, the radiance field is requested only at the surface level.

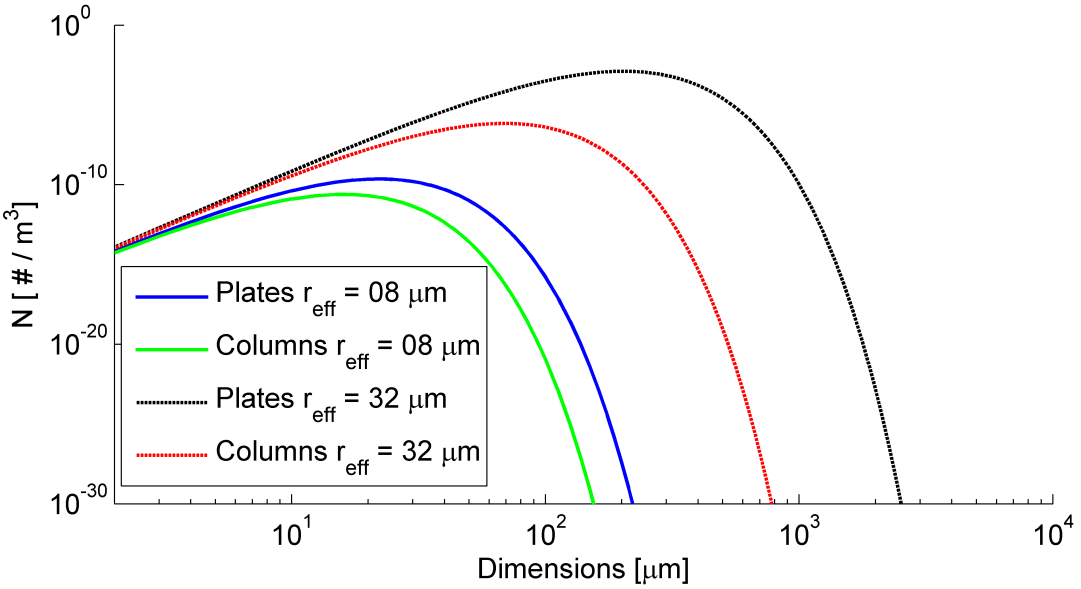


Figure 5.2.4: *Two examples of particle size distributions for plates and columns, with effective radius reported in legend*

This kind of approach, requires the description of the phase function relative to ice particles by means of a truncated series of Legendre polynomials; the largest coefficients are the most important terms for the reconstruction of the phase function and their number increases with wavenumber and PSD effective dimension. Since scattering processes play an important role in the FIR region, an accurate description of crystal phase function is required. This aim is achieved if a sufficient number (nc) of Legendre coefficients is used in the computational process; in fact, as shown in Maestri et al. (2014), this parameter is related to the number of zenith angles for each hemisphere (nz) at which the radiance integrations are performed in the model scheme; using a Gauss-Legendre quadrature scheme, the following relation holds: $nc = 4 \cdot nz - 3$. Even though in the Infrared spectral region the phase function forward lobes results much broader with respect to those at shorter wavelengths, the representation of the phase function is still challenging: Maestri et al. (2014) show that a typical choice of $nz = 16$ is insufficient for an accurate description of ice crystals scattering properties, especially for large r_{eff}.

As an example, in Fig5.2.5 the difference is shown between two simulations performed for a cloud layer composed of plate crystals with $r_{\text{eff}} = 28\mu m$. For both retrieval and simulation calculations, in the first case $nz = 16$ and in the second one $nz = 60$. The differences are quite relevant especially for dirty windows in the FIR region, where cloud signal and scattering processes have great relevance. In particular, around 500 cm^{-1} differences have a

magnitude of $1\,mW/m^2\,sr\,cm^{-1}$, corresponding to a tenth of the typically measured radiance. As a consequence, in the present work, all the retrievals and simulations are performed with $nz = 60$ corresponding to 237 Legendre coefficients.

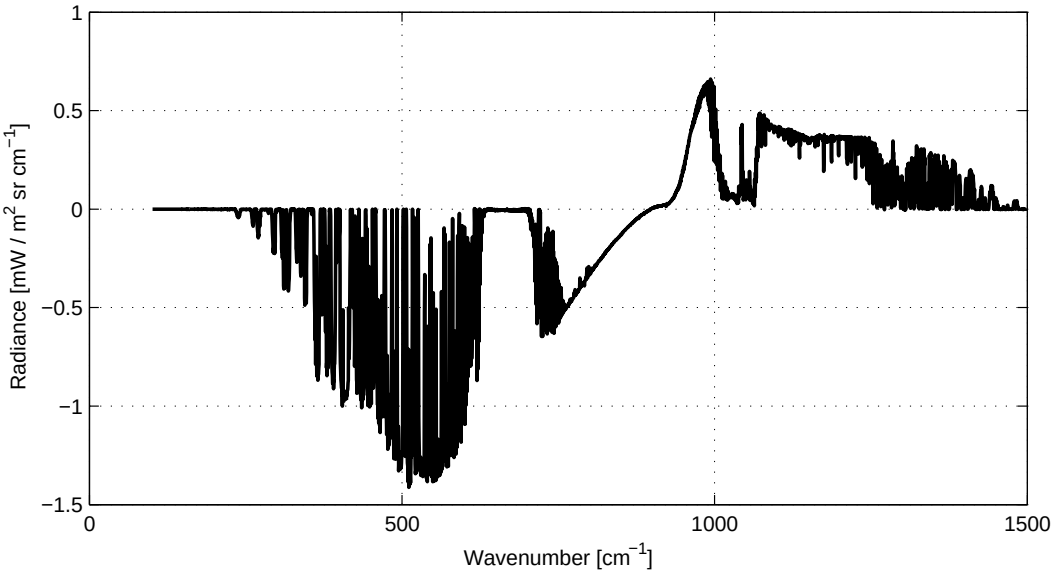


Figure 5.2.5: *Spectral differences between two simulations performed using a number of zenith angles for radiance integration, respectively equal to 16 and 60; specifically the figure refers to the 26/12/13 cloud case.*

Since RT-RET performs the retrieval procedure in the spectral interval between 800 and 1000 cm^{-1}, where the cloud absorption coefficient is large and scattering is a second-order process, the simplified Henyey–Greenstein phase function is used as a close approximation of the scattering processes.

5.2.4 Retrieval methodology

It's interesting to describe in some depth the implemented retrieval algorithm (Maestri and Holtz, 2009).
Once gaseous ODs are calculated, the layer-to-space transmittances are then computed and finally convolved with the instrument spectral response function (SRF), this represents the only instrument-dependent operation in the retrieval process.
The layer ODs are required as RT-RET input, together with the cloudy spectral measurements, surface temperature, a FGSP and cloud geometry. In addition, the algorithm assumes that the cloud extinction vertical distribution is described by a climatological profile, typi-

cal for cirrus clouds (Veglio and Maestri, 2011). In situ measurements are usually used to determine the FGSP including the habit type of the crystals in the PSD, but, since this information is not available for PRANA, two habits are chosen: hexagonal columns and plates. This choice is based on the relationship between temperature and crystal habit: simplifying the considerations in Bailey and Hallett (2009), column-type crystals prevail at temperature less than $-50°C$ while plate-type ones are more common at higher temperatures. In addition, MODIS team assumes column crystals in their cloud properties retrieval algorithm (King et al., 1997).

A number (currently 16) of REFIR-PAD single channels are selected and for each of them RT-RET determines the optimal value of OD required to fit the simulated radiances with the cloudy-sky measurements. The retrieved OD from the first channel is then used for the next wavenumber to optimize the convergence. The spectral AOD is computed by using the following formula:

$$AOD = OD_{tot} - OD_{tot} \cdot \tilde{\omega}.$$

where the OD_{tot} is the total OD retrieved by the algorithm and $\tilde{\omega}$ is the assumed single-scattering albedo.

The relevant fact, is that FGSPs used to model the scattering properties have a minimal impact on the 800-1000cm^{-1} slope of the first retrieved AOD. In fact, as can be seen in Fig.5.2.6, the retrieved slope of the AOD in the 800–1000 cm^{-1} region is minimally impacted by the assumed first-guess optical properties; on the contrary, the total OD is strongly dependent on them.

Comparing the AOD slope with those in the theoretical SPD refines the scattering properties: a root mean square fitting methodology is applied to identify the best fit of the retrieved AOD slope to the slope of the precomputed AOD. The optical properties ($\tilde{\omega}$ and g) retrieved from this comparison are then used as the SGSP to initialize the second iteration of the retrieval.

As shown by Maestri and Holtz (2009), results from an IR cloud property retrieval procedure, such as RT-RET, are affected by uncertainties originating from multiple sources including: forward model precision, the ice-crystal database's uncertainties, measurement noise, lack of knowledge about atmospheric and surface properties and assumptions made in the algorithm. The magnitude of the errors associated with the algorithms and the database results very small with respect to that produced by the assumption in the retrieval code and by the lack of knowledge of the atmospheric state. In order to investigate the sensitivity of RT-RET to uncertainness in the input parameters, the retrieval can be run with varying input parameters and the resulting modified retrievals can then be compared to the baseline case to quantify the effects on the retrieved OD and r_{eff}. A cloud temperature perturbation is instead obtained by moving up or down the cloud boundaries.
Results demonstrate that uncertainties in the cloud geometrical properties (affecting mean

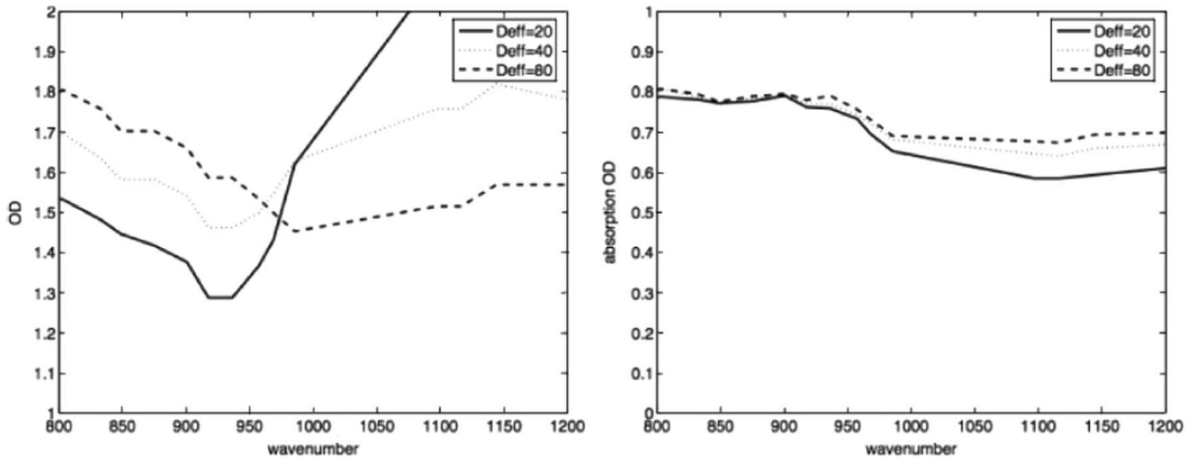


Figure 5.2.6: *Example of retrieved OD (left panel) and respective AOD (right panel) in the 800-1000 cm^{-1} region, for different values of r_{eff}. From Maestri and Holtz (2009)*

cloud temperature) dominate the uncertainties in the retrieved OD. The surface properties (temperature and emissivity) are negligible on OD results for an up-looking ground geometry. On the contrary, the situation changes when the effect of water vapor is varied, since water vapor primarily affects layers between the surface and the cloud.

In Maestri and Holtz (2009), sensitivity tests are also applied to AERI and AIRS r_{eff} retrievals: the effects are generally very small. Moreover, assuming realistic modifications of various parameters, the largest source of error is related to uncertainties concerning the cloud height, which produces at worst deviations from the baseline retrieval of 15%.

5.2.5 Retrieval results

In the next Table.5.2 are reported results obtained from RT-RET retrievals of cloud properties for the 29 case study selected and for column and plate crystal habits.

Date - hour	habit	$r_{eff}[\mu m]$	$r_{eff}[\mu m]$	IWC $[g/m^3]$	IR OD
02/01 - 12:04	c	120	120	0.0317	0.29 ± 0.11
	p	120	120	0.0196	0.18 ± 0.07
03/01 - 06:05	c	46	50	0.0141	0.59 ± 0.06
	p	34	28	0.0106	0.60 ± 0.06
19/01 - 10:49	c	8	8	0.0006	0.18 ± 0.06
	p	8	6	0.0006	0.18 ± 0.06
19/04 - 08:43	c	32	34	0.0116	0.72 ± 0.14
	p	26	26	0.0098	0.75 ± 0.14
22/04 - 05:07	c	16	16	0.0059	0.67 ± 0.09
	p	16	16	0.0063	0.71 ± 0.09
22/04 - 14:14	c	22	22	0.0085	1.26 ± 0.37
	p	16	18	0.0065	1.28 ± 0.37
23/04 - 06:39	c	12	12	0.0128	0.62 ± 0.11
	p	12	12	0.0136	0.65 ± 0.11
24/04 - 05:55	c	16	18	0.0061	0.33 ± 0.23
	p	16	18	0.0065	0.35 ± 0.24
25/04 - 08:12	c	38	40	0.0149	0.77 ± 0.15
	p	38	34	0.0155	0.81 ± 0.16
09/05 - 07:04	c	10	10	0.0010	0.24 ± 0.11
	p	10	10	0.0010	0.25 ± 0.12
09/05 - 12:11	c	120	120	0.0589	0.69 ± 0.16
	p	120	120	0.0311	0.42 ± 0.10
20/05 - 06:05	c	120	120	0.0315	0.35 ± 0.13
	p	120	120	0.0228	0.26 ± 0.09
25/05 - 13:04	c	10	10	0.0038	0.37 ± 0.08
	p	10	10	0.0039	0.39 ± 0.08
29/05 - 02:29	c	120	120	0.0163	0.23 ± 0.17
	p	24	24	0.0033	0.22 ± 0.17
10/06 - 06:49	c	10	10	0.0109	0.92 ± 0.10
	p	10	10	0.0114	0.94 ± 0.10
23/06 - 06:58	c	42	10	0.0163	0.54 ± 0.12
	p	32	10	0.0128	0.56 ± 0.12
24/06 - 22:56	c	32	34	0.0220	1.09 ± 0.11
	p	26	24	0.0186	1.14 ± 0.11
17/08 - 06:07	c	28	28	0.0049	0.34 ± 0.17
	p	26	26	0.0048	0.35 ± 0.18

18/08 - 10:31	c	16	14	0.0113	1.08 ± 0.24
	p	12	12	0.0089	1.08 ± 0.25
03/09 - 09:24	c	16	14	0.0091	0.87 ± 0.23
	p	16	14	0.0097	0.92 ± 0.24
16/09 - 13:56	c	16	16	0.0033	0.65 ± 0.22
	p	16	16	0.0036	0.69 ± 0.23
27/09 - 21:13	c	32	34	0.0111	0.94 ± 0.17
	p	24	24	0.0087	0.97 ± 0.17
06/10 - 18:31	c	10	10	0.0049	1.20 ± 0.16
	p	10	8	0.0052	1.23 ± 0.16
15/10 - 04:28	c	24	24	0.0053	0.25 ± 0.02
	p	20	20	0.0046	0.26 ± 0.08
18/10 - 07:31	c	12	12	0.0088	1.11 ± 0.19
	p	12	12	0.0093	1.15 ± 0.20
30/10 - 09:41	c	120	120	0.0259	0.55 ± 0.28
	p	120	120	0.0172	0.37 ± 0.18
26/12 - 17:09	c	34	36	0.0130	1.04 ± 0.06
	p	28	26	0.0112	1.08 ± 0.07
26/12 - 21:06	c	16	18	0.0026	0.15 ± 0.02
	p	16	16	0.00283	0.16 ± 0.03
27/12 - 16:28	c	60	60	0.0281	0.76 ± 0.05
	p	42	32	0.0198	0.76 ± 0.06

Table 5.2: *Retrieval results for the 29 selected cloud cases, performed assuming columns and plates as ice habits*

The third and the fourth columns report retrieved effective radii, intended as half effective dimensions for non-spherical particles (5.2.2), after the first and the second algorithm iteration: the one to be considered is the first value, while the second can be used as an indication about the convergence of the algorithm; in fact, the stability of these two values indicates that the algorithm has found a good PSD able to match with measured radiances and vice versa. Retrieval convergence seem guaranteed in the totality of the cases, a part from the last one that oscillates a bit for plates. In addition, radii values are generally consistent between the two habits; however, in the presence of particularly thin cloud, as the 29/05/13 morning case, the two retrieved values can be fairly different, due to the difficulty in the handling of these clouds.

Looking at the variability of these results, we found values from $8\mu m$ up to $120\mu m$, the latter being the maximum effective dimension considered in the PSD database and quite unexpected for non-precipitating clouds.

In the fifth column are reported retrieved IWC and in the last one IR optical depths: respective values are generally included between 10^{-3} and $3 \cdot 10^{-2}\, g/m^3$ for IWC and between 10^{-1} and 1.5 for OD: as expected these are small values, appropriate for non-precipitating ice clouds in a very dry and cold environment, as found by Mahesh et al. (2001) and Stone (1993). Retrievals errors are reported for IR ODs and indicate the presence of very good retrieved values with a relative uncertainty under 10% linked to less noisy spectra; on the other hand, some cases show very high relative errors, even over 50% such as 09/05 and 24/04 cases, both of which are very thin clouds. Generally, relative errors increase with decreasing optical depth.

In order to look for a possible seasonal trend, retrieved effective radii and optical depths vs time are shown in Fig.5.2.7 and Fig.5.2.8. No significant trend can be found: in fact, selected cases refer to cloud types very different one from the other and this inhomogeneity among case studies prevent from the possibility of extrapolating seasonal considerations.

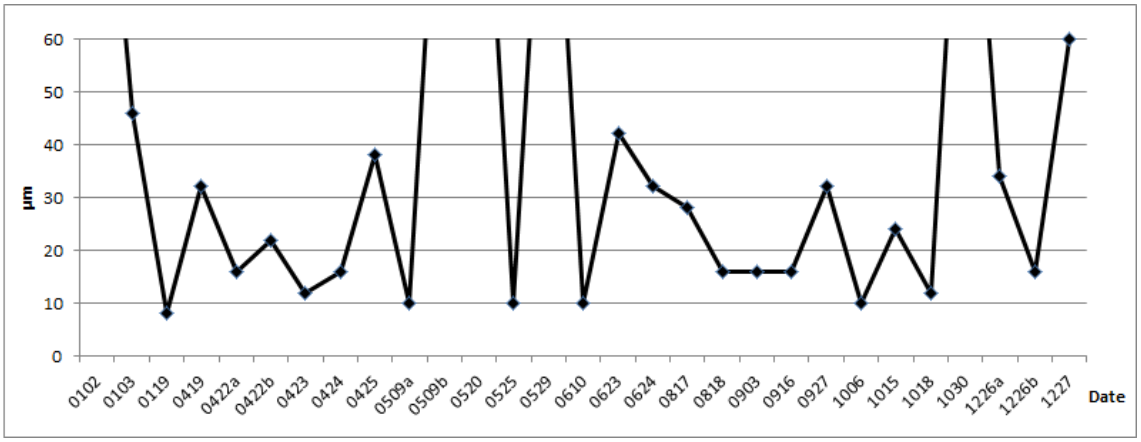


Figure 5.2.7: *Retrieved effective radii vs time.*

Correlations between various retrieved parameters and other cloud properties such as cloud base temperature and cloud bottom height, are reported in the Table.5.3. No strong correlation is evident between retrieved quantities (radius, OD and IWC) and other parameters (Base height, Base temperature or date); the only meaningful weak signal is linked to the expected positive correlation with date: summer cloud tend to be optically thicker than winter ones. By the way, looking at other correlations emerge the obvious negative one between base height and base temperature and the little seasonal signal for what concerns base temperature: the last negative value indicate higher cloud temperatures during summer months.

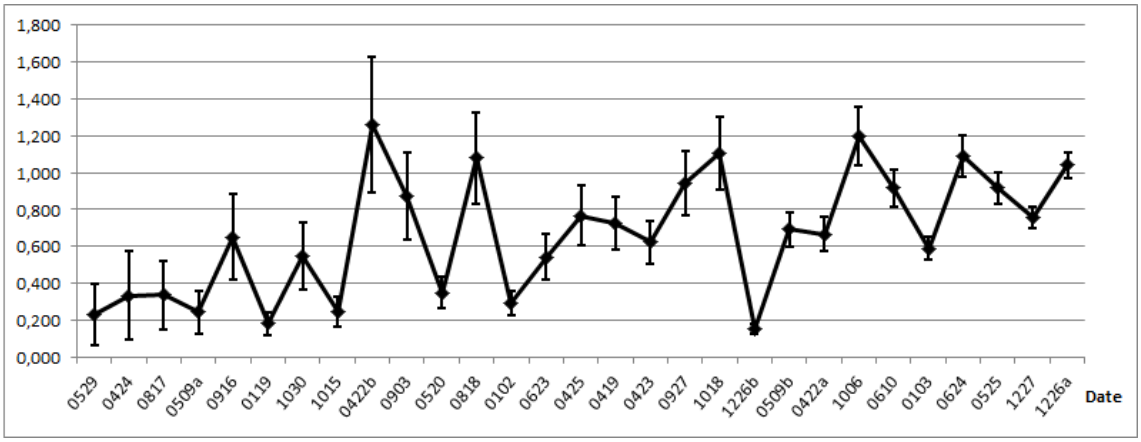


Figure 5.2.8: *Retrieved optical depths with uncertainties vs time.*

	Radius	**OD**	**IWC**	Base height	Base T
Radius					
OD	-0,307				
IWC	0,803	0,027			
Base heigtht	**0,235**	**-0,165**	**0,201**		
Base T	**-0,336**	**0,125**	**-0,275**	-0,702	
Date	**-0,080**	**0,218**	**-0,057**	-0,088	-0,387

Table 5.3: *Correlation between retrieved parameters and other cloud properties.*

Another interesting analysis can be done averaging retrieved values with the aim to compare them with ones reported in the literature. As mentioned in Ch.2 about Antarctic clouds studies, one of the most significant works is the measurement campaign described in Mahesh et al. (2001), covering one entire year of ground-based observations. In this work, not only remote sensed effective radii are reported but also calculated effective dimensions for sampled ice crystal PSDs. Also Walden et al. (2003) and Lawson et al. (2006) report two antarctic campaigns of ice crystal collection, which were classified on habit base; effective radii for each type PSD were calculated, following the procedure described in section 2.2.

Table.5.4 reports some results obtained by these authors together with our retrieved results. Mean effective values reported in the literature span from 5 to 27 μm depending on crystal habit and season. Caution must be used comparing Lawson et al. (2006) results with the others, because they refer to mean values of equivalent sphere radii, while Walden et al. (2003) refer to effective radii for PSD composed of equivalent spheres (see section 2.2). **?** retrieve small particle sizes using AVHRR observations.
RT-RET values, averaged over all the cases, are larger but, if we exclude from the calculation

those cases characterized by effective radii of 120 μm, the obtained values are more compatible. In fact, a part from the case of 05/20, the other effective radii = 120 μm are attributed to very thin clouds, for which the retrieval of properties may be particularly challenging. So, excluding those values from this calculation could be a reasonable idea.

Moreover, the median effective radius value of 15.2μm computed by Mahesh for remote sensed clouds over the entire year, results consistent with the median value obtained from RT-RET retrievals of 16μm (last row of the table), without considering extreme radii as above; in addition, this median value is independent of crystal shape.

Authors	Values	Notes
Lubin and Harper (AVHRR)	12.3μm	Mean effective radius for summer months, with 50μm as maximum retrievable dimension
	5.6μm	Mean effective radius for winter months, with 50μm as maximum retrievable dimension
Stone (radiometer-sonde)	4/16 μm	Mean effective radius for 8 summer clouds
Mahesh et al. (ground-based FTS)	15.2μm	Median remote sensed value over the entire year, with 25μm as maximum retrievable dimension
Walden et al. (ground-collection)	10.1μm	Mean effective radius for column crystal PSD, winter months
	15.1μm	Mean effective radius for plate crystals PSD, winter months
	25.2μm	Mean effective radius for bullet clusters PSD, winter months
Lawson et al. (ground-collection)	10.5μm	Mean equivalent spheres radius for column crystals PSD, summer months
	19.2μm	Mean equivalent spheres radius for plate crystals PSD, summer months
	27μm	Mean equivalent spheres radius for bullet rosettes crystals PSD, summer months
RT-RET	40μm	Mean value for columns, all cases
	23μm	Mean value for columns, leaving out $r_{eff} = 120\mu m$
	34μm	Mean value for plates, all cases
	20μm	Mean value for plates, leaving out $r_{eff} = 120\mu m$
	16μm	Median value for both habits, leaving out $r_{eff} = 120\mu m$

Table 5.4: *Literature results in terms of retrieved effective radii, computed effective or equivalent radii for collected ice crystal PSDs (adopting equal V/A spheres formulation), together with RT-RET obtained values.*

5.3 Simulations: how well cloud spectra are reproduced?

The retrieval procedure, among other files, outputs the atmospheric profile, where the cloud, geometrically bounded by means of a priori knowledge, is vertically described by retrieved properties, i.e. IWC and r_{eff}. These profiles are now used for simulation purpose on the entire REFIR spectral range, in order to test both the accuracy of the retrieval itself in the IR window region and the ability in the reproduction of measured spectra in the FIR region.

As a preliminary consideration, it could be interesting give a couple of examples about sensitivity of simulated radiances to r_{eff} and IWC variations. With this aim, the 03/09/13 case is considered: this cloud is characterized by a retrieved $r_{eff} = 16\mu m$ and, starting from the baseline profile, the cloud IWC is varied of the 30%. Results are reported in Fig.5.3.1.

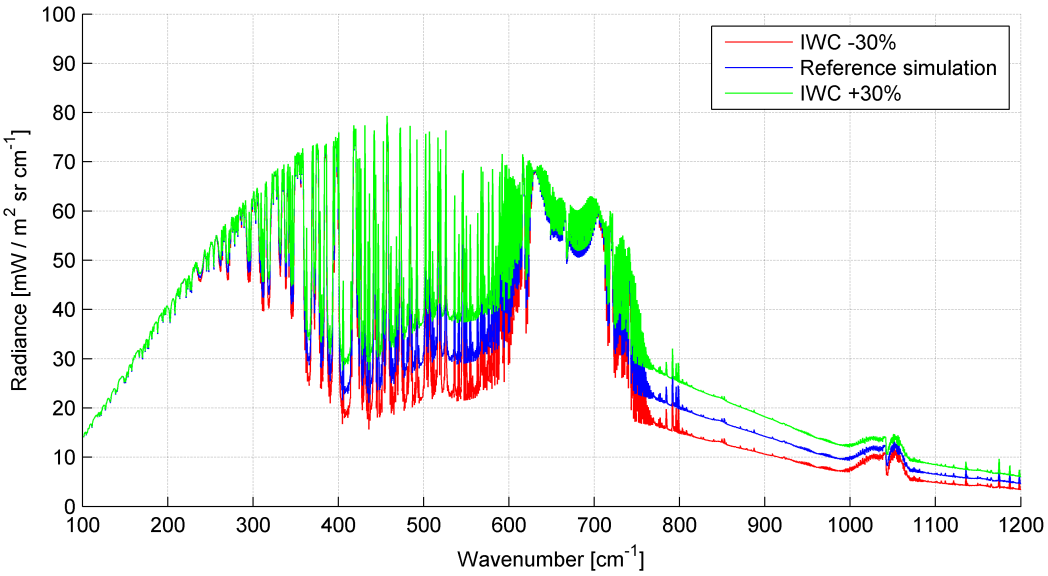


Figure 5.3.1: *Sensitivity test to IWC variations for the 03/09/13 cloud case: simulations are performed for the baseline profile and considering a 30% increase and an equal decrease of the cloud IWC*

With the increasing of cloud IWC, radiance simulated in window region increases, due to the stronger cloud emission and scattering contributions. But, as can be seen in the Fig.5.3.2, if selected micro-windows are considered and the differences is computed between

the baseline and the cloud with -30% of IWC, it strongly emerges that higher differences are found for windows around 500 cm^{-1}: in this region the cloud signal is better detected, and spectral radiances show very high sensitivity this parameter.

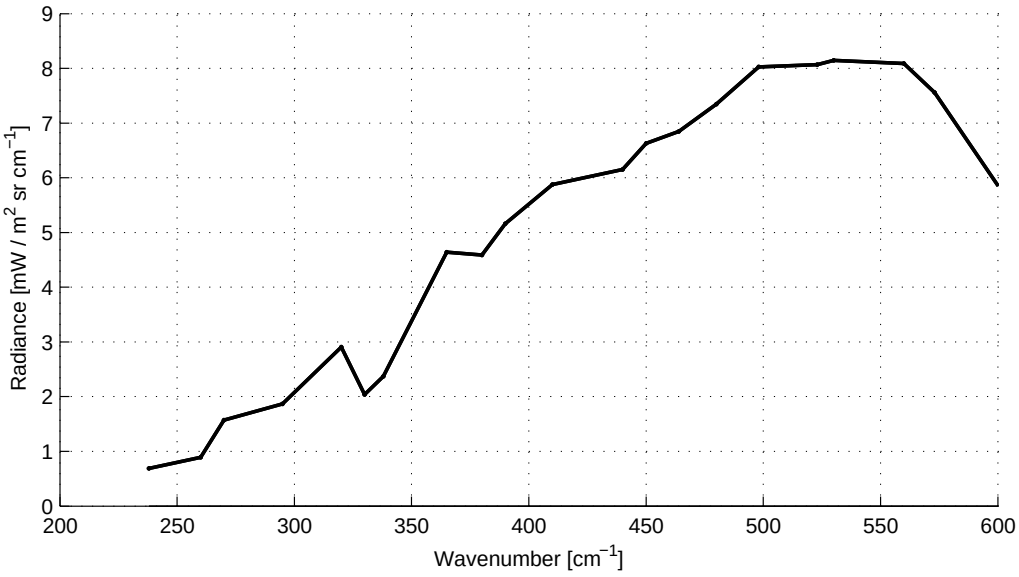


Figure 5.3.2: *Sensitivity to cloud IWC variations: spectral differences between the blue and the red line in Fig.5.3.1.*

Another sensitivity test is performed for effective size variations: simulations are run for radius equal to the retrieved value, to $4\mu m$ and to $60\mu m$. Results are plotted together in Fig.5.3.3.

As can be seen, spectral radiance shows very important sensitivity to this parameter: the most evident change concerns the slope differences in the IR window region and in the near FIR. It is interesting the behavior below 600 cm^{-1}: on the one hand, the sensitivity shown in more transparent dirty windows around 500 cm^{-1} resembles the one in the main window, with spectrum that, keeping IWC fixed, gets closer to a blackbody one; on the other hand, looking at the far FIR, changes become smaller, especially for region around 410 cm^{-1}. The Fig.5.3.4 shows the differences between the simulations for this three selected radius: the peak is again found around 550cm^{-1} that seems the region with the best sensitivity also to the effective radius.

In order to understand better this different spectral sensitivity, it's interesting to plot also the scattering coefficient, in order to evaluate the region of maximum scattering contribution, again relative to the same cloud case. The Fig.5.3.5 shows the spectral scattering coefficients for plate and column crystals. On the one hand, we verify that in the IR main window and

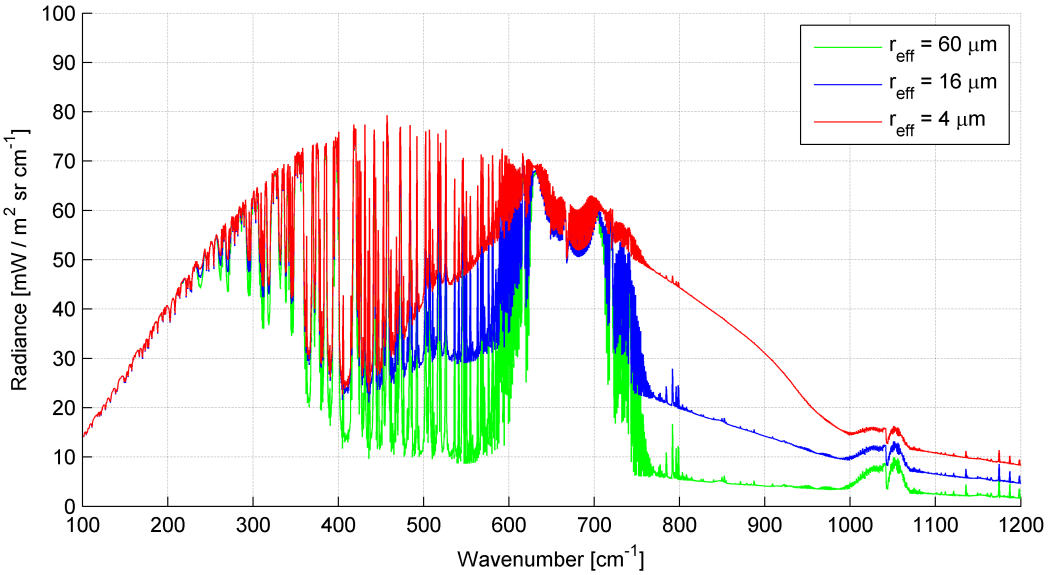


Figure 5.3.3: *Three simulations for the 03/09/13 cloud case with the same IWC profile but r_{eff} set to 4, 16 and 60 μm. Simulations are performed for column habit.*

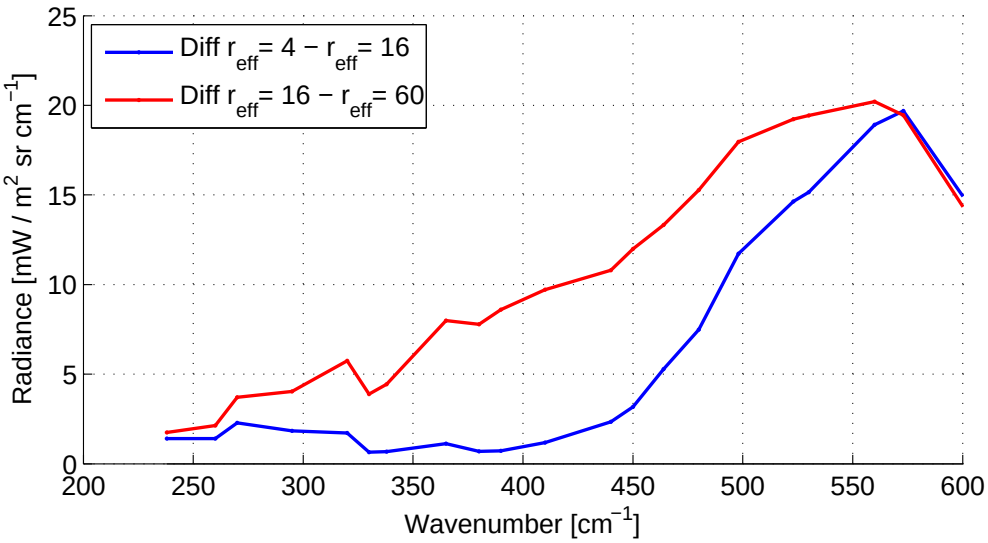


Figure 5.3.4: *Sensitivity to particle effective radius variations: spectral differences between the red and the blue lines in Fig.5.3.3 and between the blue and the green lines.*

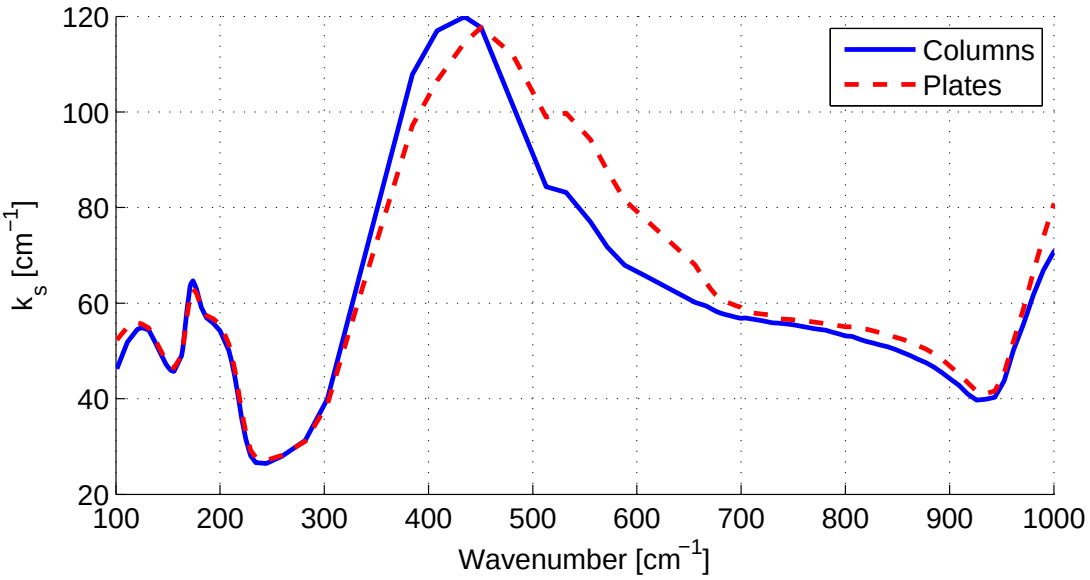


Figure 5.3.5: *Spectral pattern of the scattering coefficient, calculated for 03/09/13 cloud case assuming plate and column crystals.*

between 200 and 300 cm^{-1} scattering processes have a minor role. In particular, this latter spectral region is less affected by changes in cloud properties, since water vapor increasing absorption strength make this range sensitive to water vapor vertical distribution, especially in the low troposphere. On the other hand, the maximum k_s is reached between 400 and 450 cm^{-1}, with a little horizontal shift between the two habit types. Hence, the contribution of scattering processes is accountable for the different sensitivity to the effective radius seen in Fig.5.3.3 around 400 cm^{-1} and around 560cm^{-1}.

Yang et al. (2003), as discussed in section 1.3, indicate the BTD 250-410 cm^{-1} as strongly correlated to effective radius, specifying that this correlation holds only for $OD > 2$. In our case this consideration cannot be done, due to low optical depth; in addition, the pattern found in Fig.5.3.3 is more reasonably linked to the increasing OD (keeping IWC fixed) than to a real sensitivity to effective radius. In fact, the Fig.5.3.6 shows that neither the BTD 250-410 cm^{-1} sensitivity to particle size is detected (panel (b)), due to low ODs, nor the proposed sensitivity of the region around 550 cm^{-1} to particle size is straightforward (BDT 250-530 cm^{-1} panel (c)).
On the contrary, the sensitivity of 250-560 cm^{-1} BTD to optical depth is very clear (top panel): in fact, as said by Yang et al. (2003), this strong relation particularly holds for not too thick clouds.

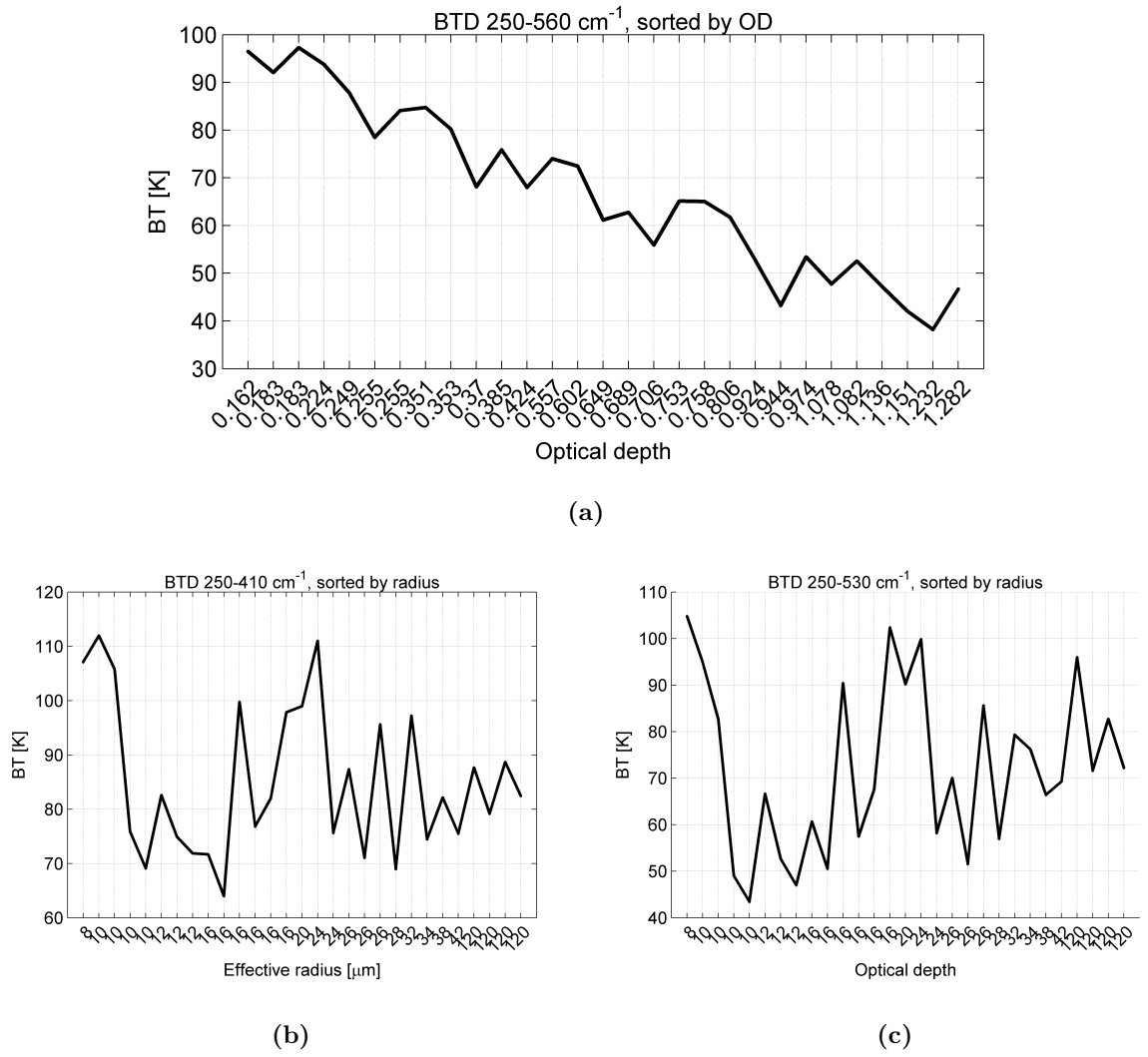


Figure 5.3.6: *BTD identified by Yang et al.Yang et al. (2003) as informative about cloud properties (first two panels): in (a) BTD 250-560 cm^{-1} sorted by optical depth and in (b) BTD 250-410 cm^{-1} sorted by radius. In (c) BTD 250-530 cm^{-1} sorted by radius.*

5.3.1 Simulation accuracy

Once simulations are performed for the overall set of case studies and for both column and plate crystal habits, the first consideration can be done about the accuracy in the reproduction of measured spectra in the main IR window.

About this point, the first accuracy measure can be the evaluation of the mean differences between simulated and measured radiances in the channels used by the retrieval scheme only. In order to have a wider view about simulation quality in the main IR window other two simple parameters are computed:

- the difference between simulated and observed radiances averaged between 800 and 1000 cm^{-1};
- the difference between the slope of simulated and measured spectra over the same spectral region.

Fig.5.3.7 shows the first two parameters, linked to the presence of a systematic shift of simulated radiance with respect to the measured ones. Differences are generally very small for RT-RET channels, with the tendency to overestimate the measured spectral radiance. On the contrary, in some cases, such as 05/20 and 10/06, there are evident mis-estimation of the mean radiance in IR window as a whole. In fact, mean differences that reach $1\,mW/m^2\,cm^{-1}\,sr$ correspond to a tenth of the measured radiance and are linked to the fact that the number of channels used in the retrieval algorithm could sometimes be insufficient to have a good evaluation of this region, due to observation noise.

However, the systematic shifts for all IR the window region are, for the worst case, less then half of the mean instrumental error.

Fig.5.3.8 instead, shows the third parameter linked to the different simulated and measured window slope. Slope differences are generally not larger than $1-2\,mW/\,m^2\,cm^{-1}\,sr$ a part from few cases that show some peaks. These differences can be linked to some difficulty of the retrieval method to give, in few cases, a correct effective radii. In fact, when these values are sorted by increasing effective radius, negative differences are generally found for big size while positive ones for smaller particles.
In conclusion, substantial differences are not found in any cloud case, hence, also for what concerns this parameter, retrieval methodology shows a good accuracy in the window region. However, in a limited number of events, retrieval may not output the very correct radii value to match IR window observed slope.

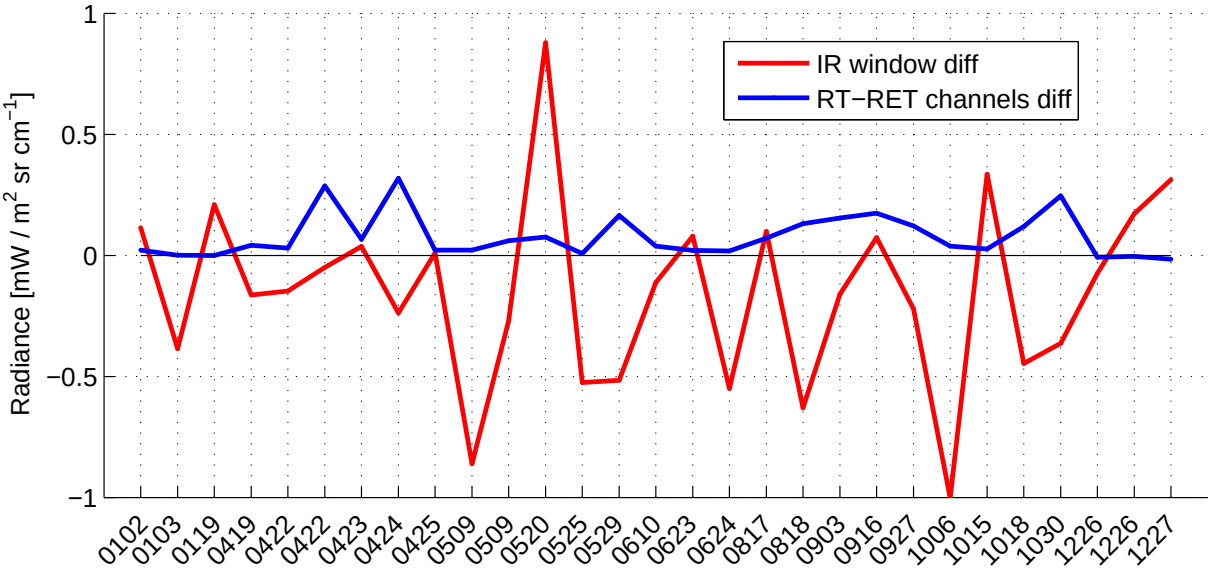


Figure 5.3.7: *Differences for all the cases between simulated (using column) and measured radiances averaged between 800 and 1000 cm^{-1} together with same differences calculated only over RT-RET exploited channels.*

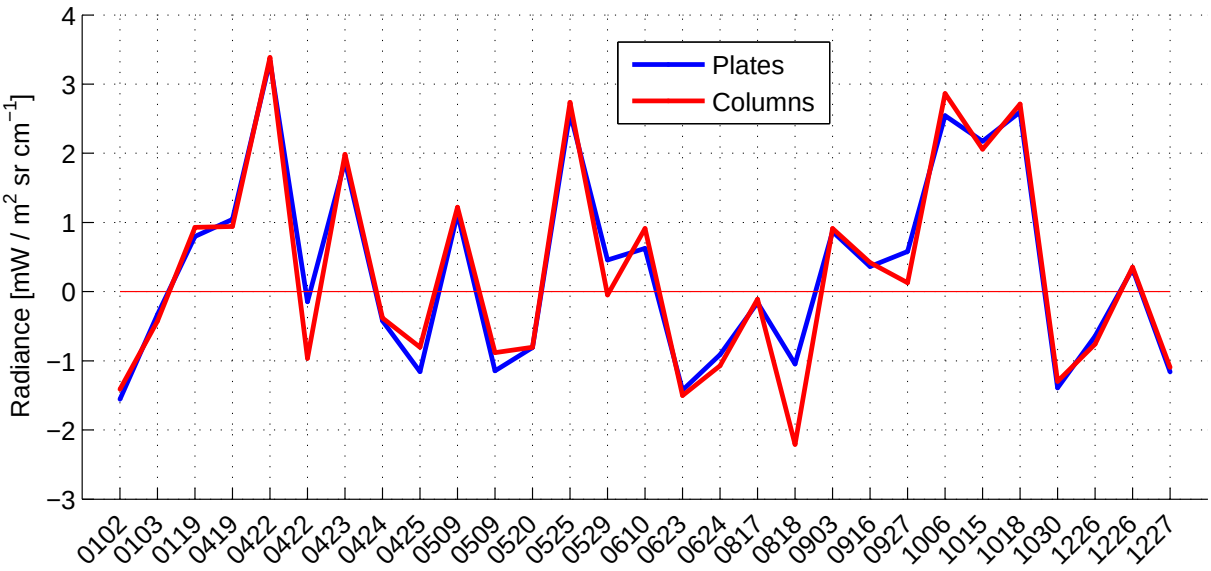


Figure 5.3.8: *Differences for all the cases between simulated (both plate and column) and measured radiance slope in the 800-1000 cm^{-1} region.*

5.3.2 Simulations in the FIR

Once the quality of the retrieval/simulation methodology has been proven, the next step concerns the evaluation of our ability to reproduce FIR radiances and the estimation of eventually detected correlations between simulated-observed spectral differences and retrieved parameters.

In Fig.5.3.9 results are reported for three simulations, using the column habit. It's interesting not only to plot the simulated and the measured spectra (right panel), but also to select micro-windows both in the FIR region, so that differences averaged over some intervals can be drawn. With this aim, the spectral intervals defined during features selection phase are again considered and in the Fig.5.3.9 (left side) simulated - measured differences averaged on these "windows" are shown, only for FIR wavenumbers. In addition, they are compared to respective REFIR-PAD mean $1-\sigma$ uncertainties, superimposed as red dots, in order to have an idea of the relative magnitude of these differences.

Looking at the figures, the green line refers to spectrally resolved differences, while blue error bars represent the variance of calculated differences inside each selected interval. In some dirty windows these differences are well above total instrumental error and reach 20% or more of the relative measured radiance, but the sign of these deviations are opposite for the first two cases. The first high thin cloud simulation shows a net over-estimation in the whole FIR region (considered between 240 and 570 cm^{-1}). The second case instead, referred to a high cloud between precipitating moments, show very large under-estimation, especially from 300 to 450 cm^{-1}. The last case instead reveals a good performance with differences almost always inside REFIR-PAD mean errors and quite small spectral variances (short error bars). A singular detail regards wavenumbers around 400 cm^{-1}: in the first two cases, this region show a peak in the deviation plot, while for the last one it result the best simulated interval. The simulated radiance in the CO_2 band results every time underestimated, due to the lack of description of the very warm first three meters of the optical path inside the instrument and the shelter.

Hence, great variability among cloud cases, affecting spectral simulation deviations from measurements, is evident, even from these three cases only. Looking now at all the simulated cases, few ones show good simulations in the FIR region: in most cases FIR simulated radiances reveal important differences, if compared with REFIR-PAD errors, but spectral difference patterns are quite varied, as it has been already shown.

In Fig.5.3.10 are reported the mean differences in the FIR region, calculated by averaging absolute differences between simulated and measured radiances over the spectral interval among defined, from 238 until 573 cm^{-1}. In the plot, mean PAD errors are superimposed again, in order to have an idea of the overall quality of simulations in the FIR as a whole. Both column and plate habits are considered and reported together in the figure. The trend shown by the two curves is quite similar but some cases are evident for which columns lead to very different performance with respect to plate crystals.

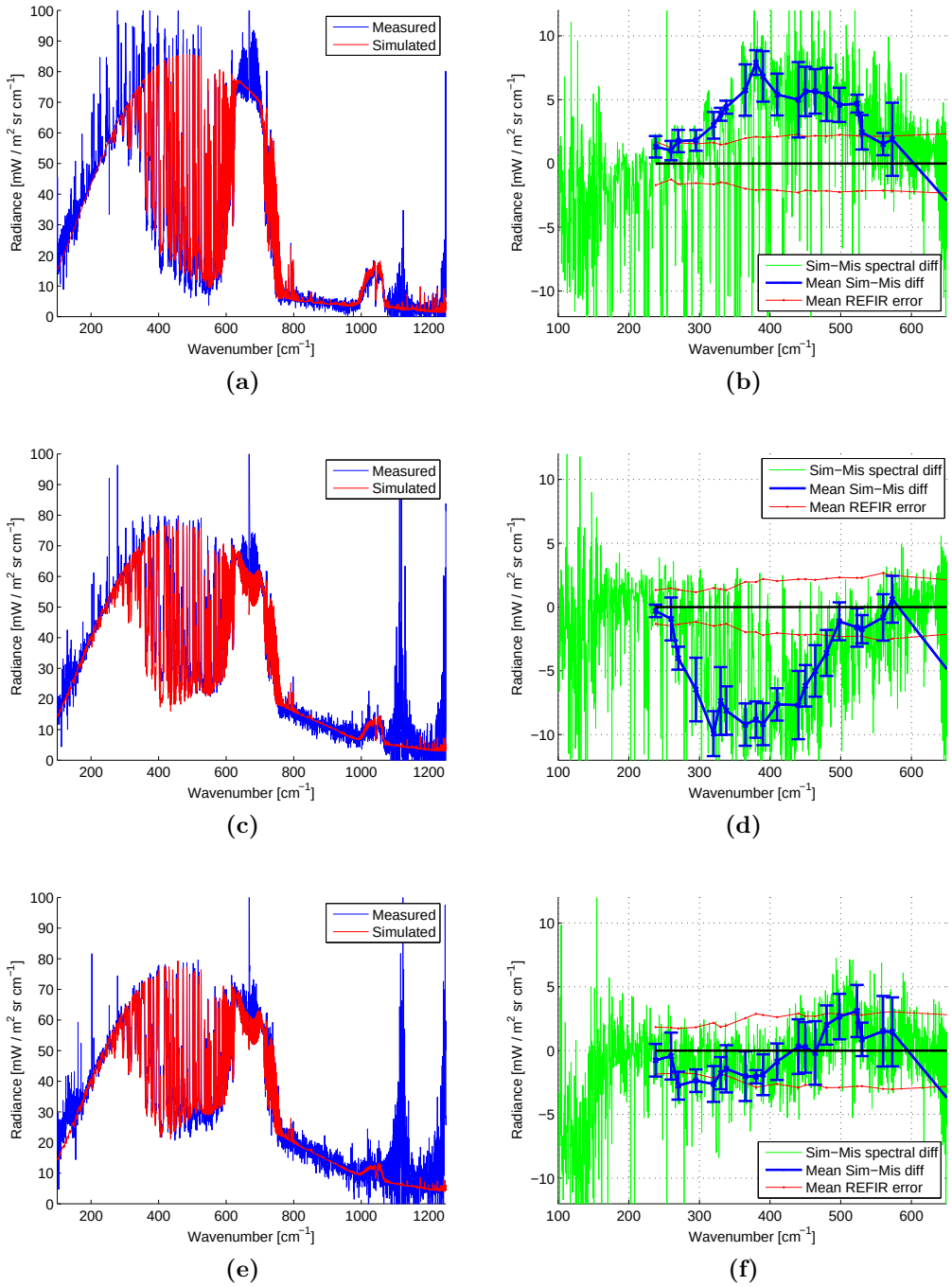


Figure 5.3.9: *Examples of three chosen case studies (02/01, 23/04 and 03/09): on the left side the simulated spectrum and the retrieved one are shown together, on the right side differences in selected windows between the two spectra are reported for the FIR region, with REFIR-PAD averaged errors superimposed.*

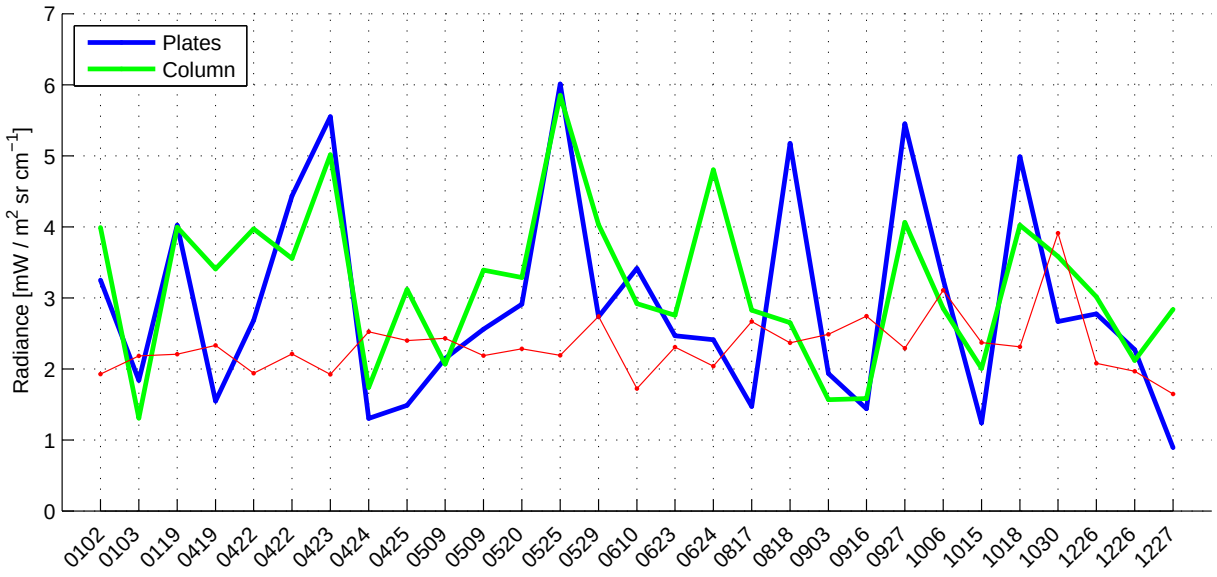


Figure 5.3.10: *Mean FIR absolute simulation deviations from observed radiances calculated by averaging over selected 'dirty windows', for all the considered cases, assuming both columns and plates as crystal habits. REFIR-PAD mean uncertainties are superimposed.*

Despite various attempts, no meaningful correlations have been found between the magnitude of these spectral deviations and specific cloud types. Nevertheless, comparing mean FIR absolute differences from observed spectra with cloud-base height, it can be seen that high clouds simulations generally show smaller deviations than low level clouds. This signal is very weak and perhaps tied to selected case studies. Looking at Fig.5.3.10, only six simulation for column habit show mean FIR differences from measurements below REFIR-PAD errors: among them, only one has a bottom below 1100m a.g., while 40% of selected cases have such a low base. Looking at plate simulations, the situation appears slightly better with ten cases showing deviations less then PAD errors: again only two of them are characterized by a cloud base height below 1100m a.g.

From this mean perspective, plate crystals seem to lead to better simulations than column ones, with generally smaller deviations from observed spectra. In particular, the best simulation results are for the 27/12/13 pre-precipitating high cloud, using plate habit. The Fig.5.3.11 shows spectrally resolved and selected intervals differences with the reference REFIR-PAD $1-\sigma$ error boundaries.

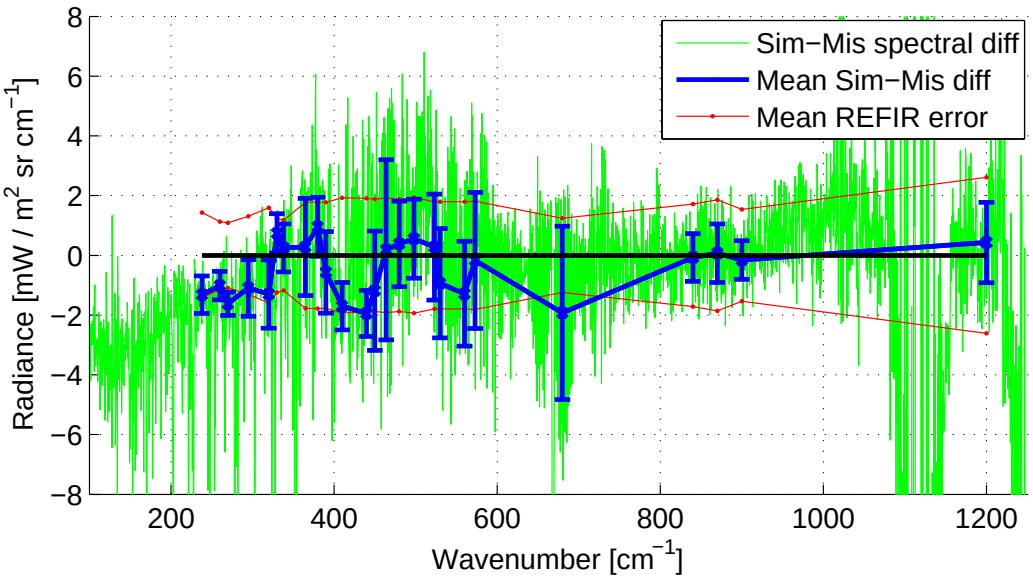


Figure 5.3.11: *Spectral differences between simulated and measured spectrum assuming plate habit, for the 27/12/13 case.*

Instead of looking only at mean differences over the FIR region, let's also consider in more details the spectral structure of these deviations. To begin with, an average over all the simulation deviations from measurements in the above selected intervals is reported in Fig.5.3.12. The average spectral pattern for both columns and plates is characterized by a mean under-estimation in the region below 400 cm^{-1} and a smaller over-estimation above 450 cm^{-1}; this spectral feature appears, however, quite blurred looking at the 1-σ error around the mean, which is smaller at lower wavenumber and higher around 400 cm^{-1}. This is an expected signal, after the considerations about the importance of scattering processes in this region. In addition, the two curves appears slightly different: a systematic vertical shift is evident, especially above 350cm^{-1}, shift that is no longer present going to the IR window and to less transparent dirty windows. Looking at simulated spectra, assuming plate crystal habit leads to radiance systematically reduced with respect to the ones obtained assuming column habit.

The last step for this spectral analysis concerns the consideration of a limited set of wavenumber intervals, in order to look at the deviations trend over all the cases. Fig.5.3.13 reports simulations-observations differences for five chosen intervals from 260 to 900 cm^{-1} Differences around 900 cm^{-1} are verified to be much lower than in the FIR region, as expected. From 260 cm^{-1}, differences seem to increase until 450 cm^{-1}, then the curve referred to the 560

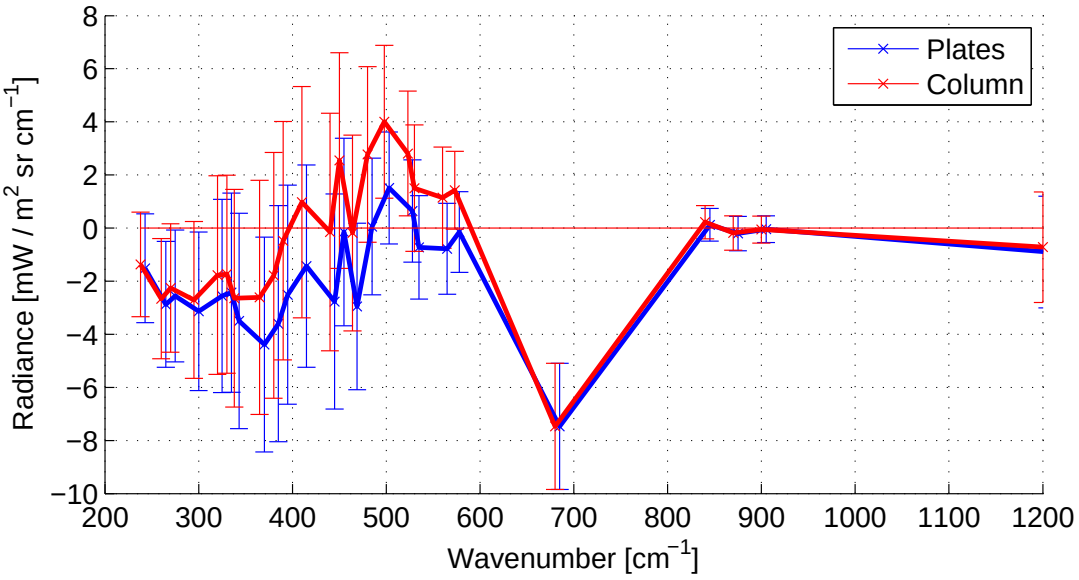


Figure 5.3.12: *Spectral differences between simulated and measured spectra, averaged over all cases: both results for plate and columns simulations are reported with a small horizontal shift provide a better representation.*

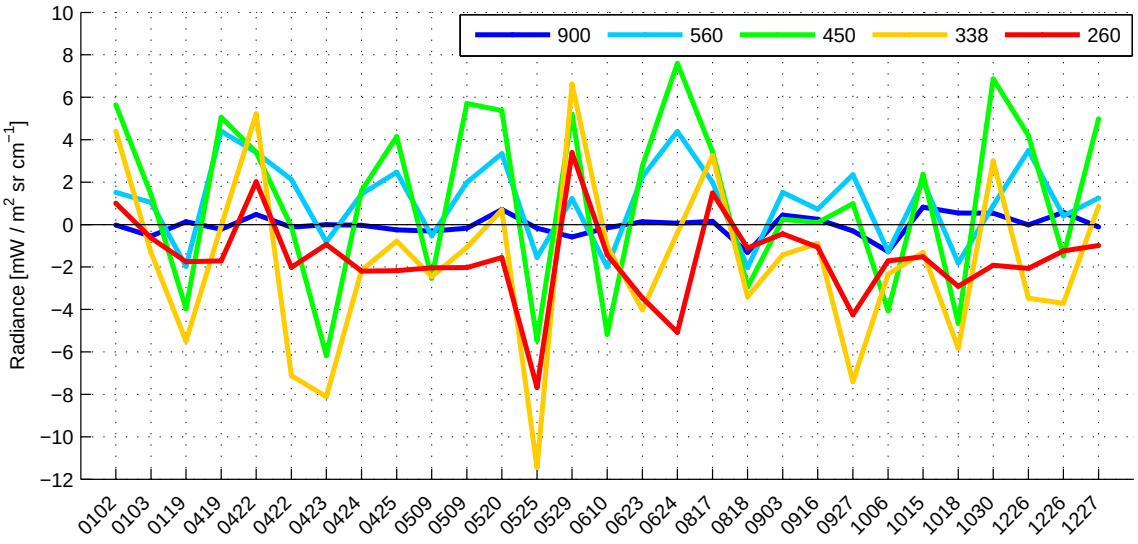


Figure 5.3.13: *Mean differences between simulated (with column) and measured spectra, averaged on five spectral intervals.*

cm^{-1} shows less peaks. In addition, the trend described in Fig.5.3.12, is now confirmed and the high variability previously represented by means of error bars is now plainly shown: in fact, the red curve, linked to far FIR deviations, shows negative values in most of the cases, the green curve around 450 cm^{-1} presents both sign differences, while the cyan one, linked to wavenumbers towards the CO_2 band, generally assumes positive values.

Based on considerations reported at the beginning of this section about the spectral relevance of scattering processes and the best channels for the detection of cloud signal, we can select the region between 410 and 500 cm^{-1} as the most interesting for our considerations, region that show also the highest variability from case to case and from plate to column simulations.

Simulation deviations in this spectral range can be linked to other cloud characteristics, such as retrieved effective radius, OD, cloud temperature and base height, in order explain the variability seen in this figures.

Fig.5.3.14 is a plot similar to Fig.5.3.13 but differences are now averaged over the new range and values obtained for plate and column habits are sorted by respective effective radius. The great variability seen before results now a bit organized; in fact, a trend appears, even if quite confused: deviations from observations tend to change sign with increasing radius. In particular, simulations for clouds characterized by very little effective radii generally show a net under-estimation; on the contrary, clouds with high retrieved radii show over-estimations, especially for column simulations. A detail to be noted is the presence of repeated radius values that should be averaged to smooth the variability. This plot highlights that FIR simulations are better for mid-sized PSDs.

Fig.5.3.15 shows the same quantities of the previous one, but now respectively sorted by IR optical depth. Here an incoherent pattern seems to emerge, a part from the little negative trend shown by the curve for plate habit, where simulations for high optical depths clouds seem to show negative biases from observations. This pattern is completely confused for column simulations.

A further step is to separate cloud cases into two groups as regards optical depth, using a threshold fixed at the mean retrieved value ($OD = 0.7$). Fig.5.3.16 and Fig.5.3.17 are plots same as Fig.5.3.14 but for each OD group; differences on 410-500 cm^{-1} range are sorted by increasing radius (this time the one retrieved with column simulations, so that both curves refer in each point to the same case). Especially looking at the high optical depth cases, Fig.5.3.16, the half-see bias trend in Fig.5.3.12 here appears more clearly; in addition the averaged curve of the results for each radius for column simulations is superimposed (thin black line) and describes a quasi monotonic relationship between bias and particles radii.

For low optical depth clouds, Fig.5.3.17 this behavior is less straightforward, in particular for plate crystals for which the trend is flatter.

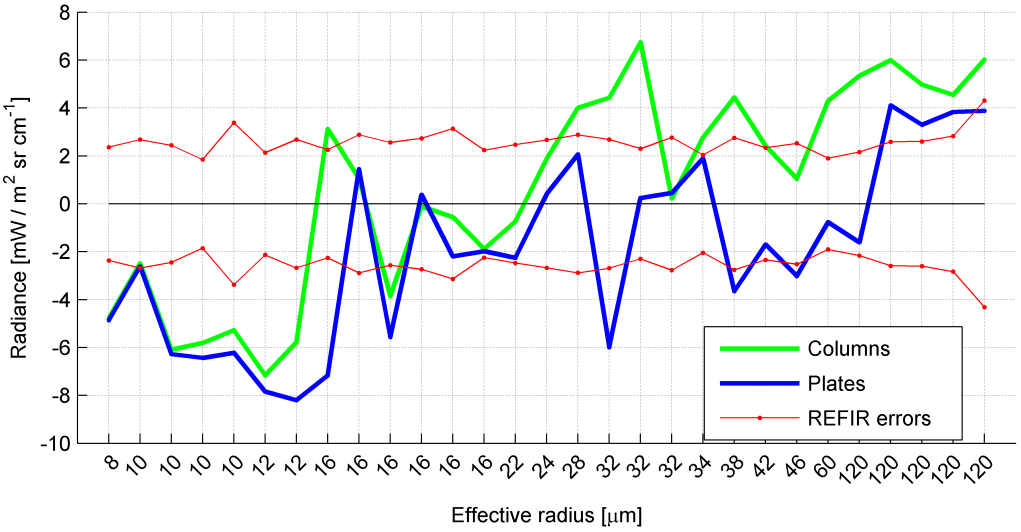


Figure 5.3.14: *Differences between simulated and measured spectra, averaged on intervals between 410 and 500 cm^{-1} for plate and column habit, sorted by radius*

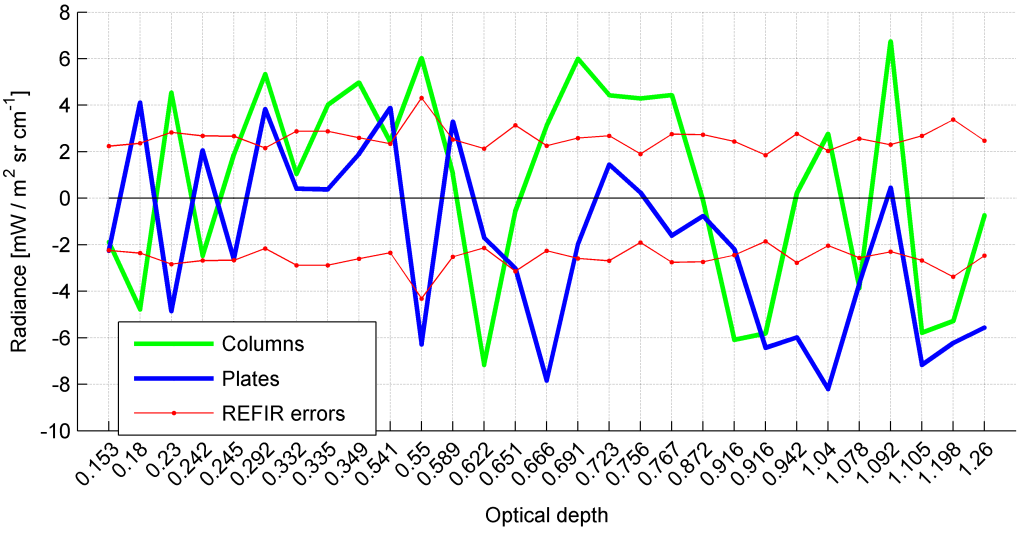


Figure 5.3.15: *Differences between simulated and measured spectra, averaged on intervals between 410 and 500 cm^{-1} for plate and column habits, sorted by optical depth*

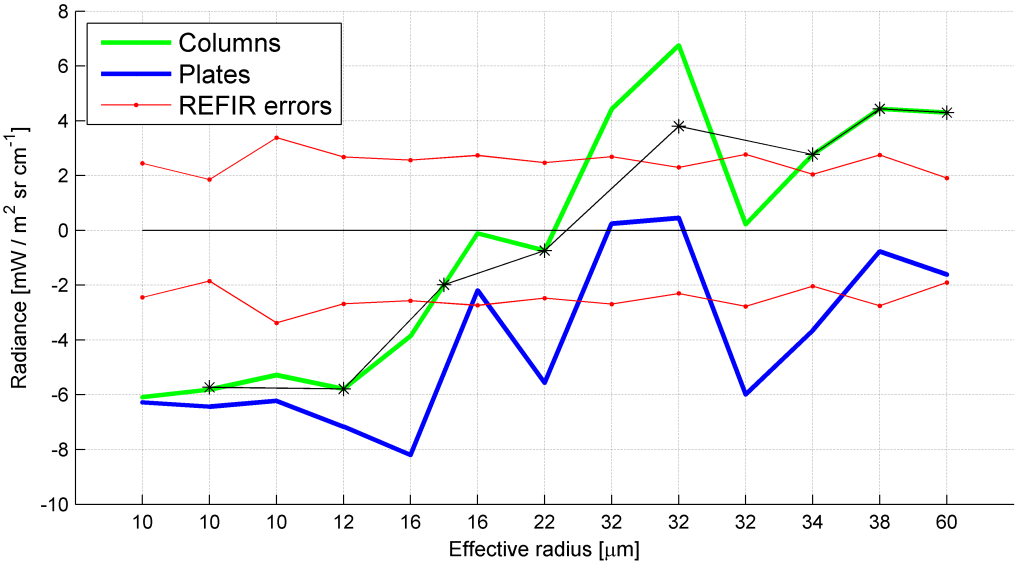


Figure 5.3.16: *Mean deviations on 410 - 500 cm^{-1} region of simulated spectra from measurements, relative to cloud with $OD > 0.7$, for plate and column habits, sorted by radius*

Fig.5.3.16 and Fig.5.3.17 also show a different behaviors of plate vs column simulations: the two curves are very near when radii are small, whereas differences increase as the size becomes bigger. As a consequence, clouds characterized by small particles are generally badly simulated, whatever the assumed particle habit; on the contrary, for larger radii, even if column simulations reveal high overestimations in simulated radiances, the strong habit sensitivity around 400 cm^{-1} leads to fairly better plate simulations (19/04 and 24/06 cases) and vice versa (27/09 case).

In Fig.5.3.18 selected spectra with high optical depths are plotted together vs wavenumber for selected radii, showing the modification of the spectral pattern of biases. Hence, grouping together clouds with similar thickness, the increasing of the effective radius leads to overestimations.

It's also interesting to study the spectral dependence of the sensitivity to habit changes: Fig.5.3.19 reports differences between simulations assuming column and plate habits, for five spectral intervals and sorted by increasing optical depth. From this plot, it's strongly evident the maximum radiance sensitivity to particle habit around 450 cm^{-1}, where scattering coefficient show an important peak, while at lower wavenumber and in the IR window, this sensitivity is attenuated or even suppressed. A similar wavenumber-sensitivity to the assumed habit is also found in Maestri et al. (2014).

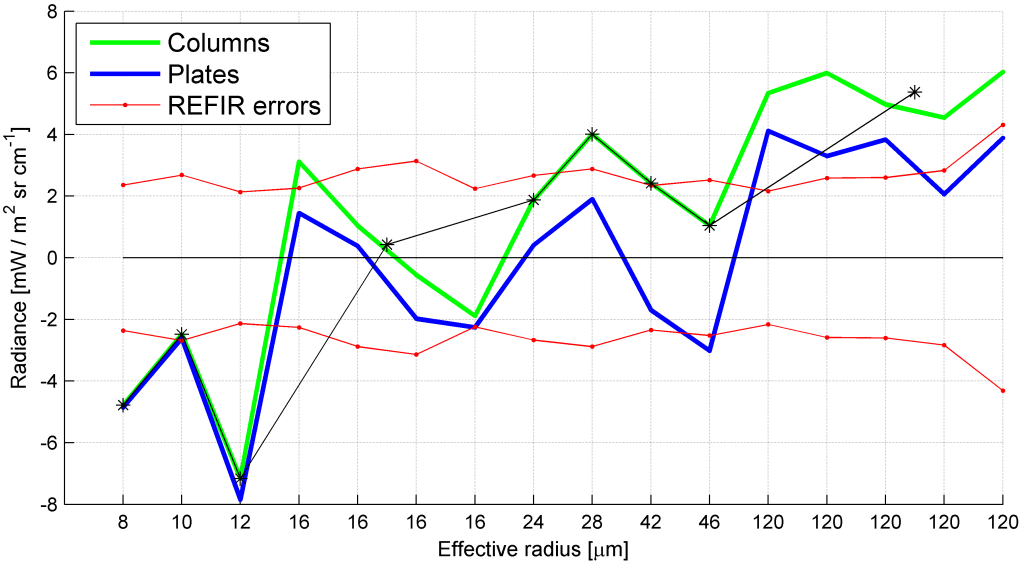


Figure 5.3.17: *Mean deviations on 410-500 cm^{-1} region of simulated spectra from measurements, relative to cloud with $OD < 0.7$, for plate and column habits, sorted by radius*

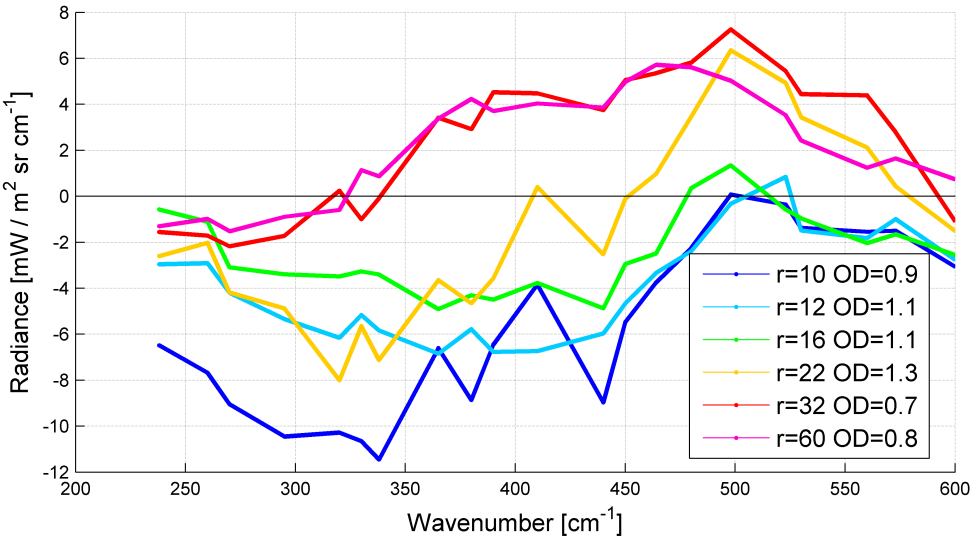


Figure 5.3.18: *Differences in the FIR region between simulated and observed spectra for cloud with $OD > 0.7$ and selected increasing radii.*

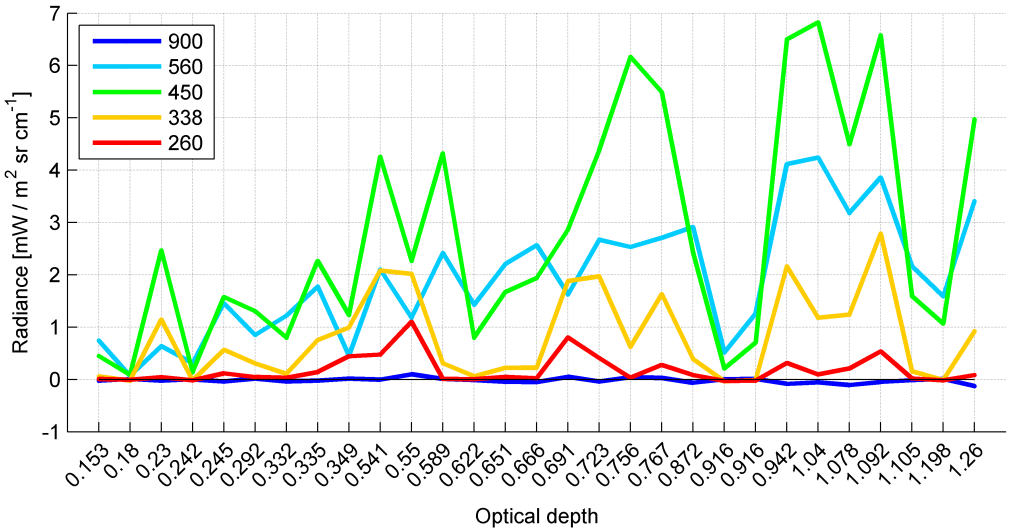


Figure 5.3.19: *Deviations of column crystal simulated radiances from plate ones, averaged over five selected intervals and sorted by increasing OD.*

A combined dependency of the magnitude of these differences to particle radius and OD can be evaluated: in the Fig.5.3.19, looking at 560 and 450cm^{-1} lines, a positive trend can be seen with maximum differences for thick clouds, but low values are found also for cases on the right side of the plot: these correspond to clouds with little ice crystals. Again Maestri et al. (2014), even if for only one cloud case with increasing thickness, have already pointed out this positive trend with OD.

This combined dependency, can be highlighted in Fig.5.3.20: it is a plot of averaged plate-column simulation differences around 450 cm^{-1} vs optical depth. It is found the positive trend with increasing OD, but it is differentiated for different effective radii: simulations for clouds with bigger particles show a greater slope with respect to the more flattened trend shown by smaller particles ones.

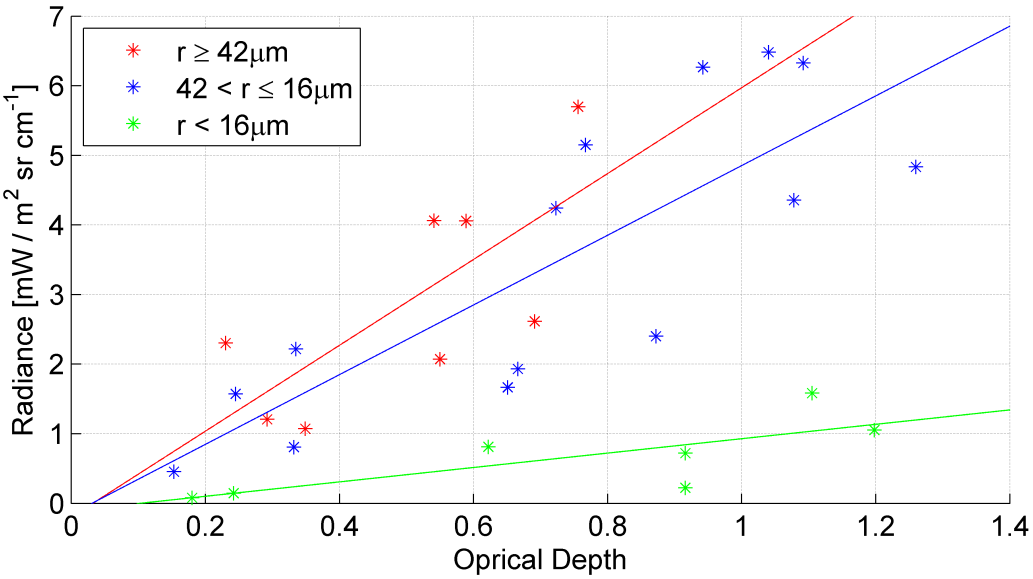


Figure 5.3.20: *Differences between plates and columns simulations averaged around 450 cm^{-1} vs optical depth, with cloud cases classified into three groups about particle size.*

As a further development of this work a new vertical layering of atmospheric profiles has been built, with a more detailed leveling for the troposphere, in order to describe better also high clouds. Moreover, the temperature of the lower level is set to the shelter one and the second level temperature (3 m above) is the value measured by a fixed weather station. From the third level, the temperature and humidity profiles are obtained using the methodology described in section 5.2.1.

A preliminary simulation has been performed with this new profile and the result is shown in Fig.5.3.21 for the 13/06/10 cloud case, assuming column crystals. Differences between the new and the old profile simulations are reported in Fig.5.3.22 together with deviations from measured spectrum of simulations performed using both profiles. As can be seen, the most important difference is related to the CO_2 band, now better described, including the central extreme peak. Also in the FIR region some differences appears, at most of about $0.5\,mW/\,m^2\,cm^{-1}\,sr$ and towards a general increasing of the radiance value. The latter modification improves the simulations in clear air but do not significantly impact the ones in presence of clouds.

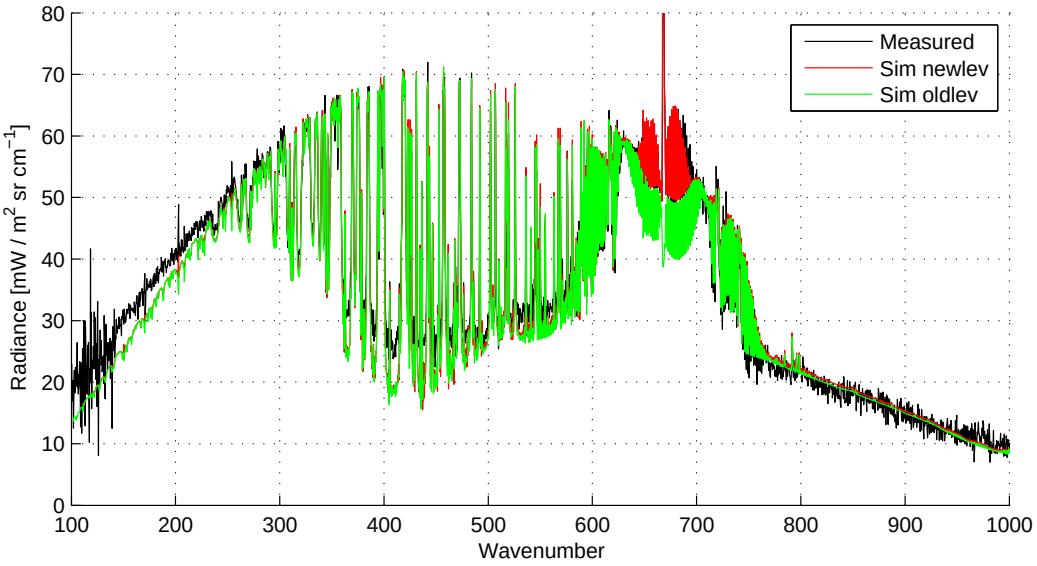
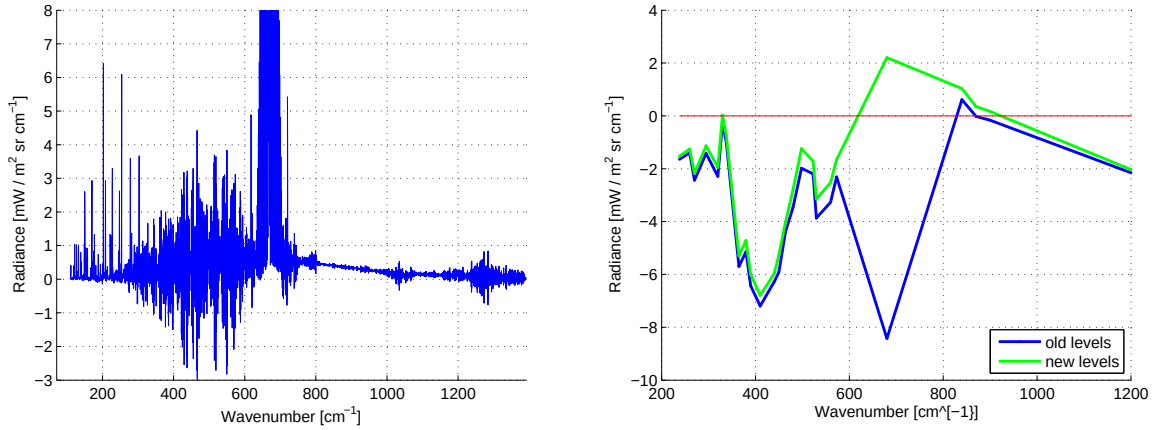


Figure 5.3.21: *Measured and simulated spectra using both new and old profiles superimposed; 10/06/13 case study, assuming column habit*



(a) **(b)**

Figure 5.3.22: *In panel (a), spectral differences between simulations using new and old atmospheric profiles; in panel (b) deviations from measured spectrum using both profiles.*

Conclusions

The relevance of the FIR for Earth's radiative budget considerations and for many remote sensing applications has been theoretically stressed in many works. Even if the lack of measurements in this spectral region, due to technological and experimental difficulties, would be filled only by means of a dedicated space mission, the PRANA project has been the first occasion to systematically observe atmospheric emission in the FIR, using the interferometer REFIR-PAD, with a focus on water vapor and Antarctic cloud spectral signatures.

In the present work, a preliminary selection of good quality spectra is performed, using, as thresholds, radiance values, with estimated errors, in few chosen spectral regions. As a result, thousands of good quality spectra covering the whole 2013 are obtained, in the presence of the most various meteorological conditions. Mean radiance in selected micro-windows is converted into BT and the differences between each pair of BTs are used as features for the description of each REFIR-PAD spectrum. A feature selection algorithm is implemented with the purpose to discriminate the presence, the type of cloud and the phase of the particulate. A supervised-learning approach is chosen and, for each month, a training set of spectra is manually collected using co-located Lidar quick-looks. Few BT differences are found significant for the discrimination between clear and cloudy cases but also between water and ice clouds. More difficult is the discrimination between precipitating and non-precipitating clouds, due to the complexity of radiative transfer processes and the variety of conditions and microphysical properties. The supervised classification step of the overall monthly datasets is performed using a SVM.

Finally, an accurate manual check of each non-precipitating cloud case is performed on the base of the obtained supervised classification and with the help of Lidar profiles. On these selected spectra, the retrieval methodology RT-RET is applied, using only a selection of spectral channels in the main atmospheric window region, assuming columns and plates as crystal habits. Retrieved effective radius, optical depth and IWC show neither seasonal trends nor evident correlations with different cloud typologies, due to the high variability among selected cases. Mean and median effective radii are also compared with values found in the literature (Mahesh et al. (2001) and Walden et al. (2003)): a possible coherence is found only if problematic retrieved radii equal to $120\mu m$ are neglected.

Simulations are performed using profiles of atmospheric and cloud parameters, the latter supplied by the retrieval scheme, for the whole spectral interval of the REFIR-PAD measurements. Significant differences from measured spectra are found for FIR dirty windows in many simulations, with strong variability in their spectral pattern. Averaging these differences over all the selected cases, a systematic under-estimation is found at the lower wavenumber end while, towards the CO_2 band, mean over-estimation can be seen, especially for simulations with columns. Considering the spectral interval around 450-500 cm^{-1}, where the cloud signal is particularly high and scattering processes are important, a high variability is observed and a significant result is found, specifically for clouds with high optical depth: small effective radii lead to simulations that under-estimate observations, with little sensitivity to particle habit, while larger radii lead to over-estimations, with an increasing spread between column and plate simulations. The spectral shape of simulated/observed radiance differences in the FIR shows an interesting shift with increasing radii among similar ODs. A deeper investigation is needed, in order to understand these first results, which are considered preliminary.

New simulations will be performed for another crystal habit, such as bullet rosettes, frequently reported (Lawson et al., 2006) in the presence of precipitating clouds. In addition a new vertical layering has been developed in order to take into account the optical path inside the instrumental shelter and to describe with more accuracy high clouds. In the end, a wider perspective would open, were Lidar optical depth profiles available. This would open to the possibility of handling more common precipitating clouds and seasonal trend considerations about microphysical properties would be more meaningful.

Abbreviations

AERI	=	Atmospheric Emitted Radiance Interferometer
AIRS	=	Atmospheric InfraRed Sounder
AOD	=	Absorption Optical Depth
AVHRR	=	Advanced Very High Resolution Radiometer
BASIL	=	BASILicata Raman Lidar system
BS	=	BeamSplitter
BT	=	Brightness Temperature
BTD	=	Brightness Temperature Difference
CAL	=	CALibration accuracy
CALIPSO	=	Cloud-Aerosol Lidar and Infrared Path-finder Satellite Observation
CBB	=	Cold BlackBody
CPI	=	Cloud Particle Imager
DISORT	=	Discrete Ordinates Radiative Transfer code
ECMWF	=	European Center for Medium-range Weather Forecasts
ECOWAR	=	Earth COoling by WAter vapouR
EU	=	European Union
FGSP	=	First Guess Spectral Properties
FFT	=	Fast Fourier Transform
FIRST	=	Far-InfraRed Spectroscopy of the Troposphere
FOV	=	Field Of View
FTS	=	Fourier Transform Spectrometer
GBMS	=	Ground-Based Millimeter-wave Spectrometer
GMT	=	Greenwich Mean Time
HBB	=	Hot BlackBody
HSR	=	High Spectral Resolution
ILS	=	Instrumental Line Shape
IR	=	InfraRed
IWC	=	Ice Water Content
IWP	=	Ice Water Path

LbLRTM	=	Line-By-Line Radiative Transfer Model
LIDAR	=	LIght Detection And Ranging
LW	=	Long Wave
MIR	=	Mid-InfraRed
ML	=	Mid-Latitude
MODIS	=	MODerate resolution Imaging Spectroradiometer
NESR	=	Noise Equivalent Spectral Radiance
NIR	=	Near InfraRed
OD	=	Optical Depth
OLR	=	Outgoing Long-wave Radiation
OPD	=	Optical Path Difference
PAD	=	Prototype for Applications and Development
PCA	=	Principal Component Analysis
PRANA	=	Proprietà Radiative del vapore Acqueo e delle Nubi in Antartide
PSD	=	Particle Size Distribution
PWV	=	Precipitable Water Vapor
RBB	=	Reference BlackBody
RBF	=	Radial Basis Function
REFIR	=	Radiation Explorer in the Far InfraRed
RS	=	RadioSonde
SGSP	=	Second Guess Spectral Properties
SNR	=	Signal to Noise Ratio
SPD	=	Scattering Property Database
SPS	=	South Pole Station
SR	=	Spectral Regression
SRF	=	Spectral Response Function
STD	=	Standard Deviation
SVM	=	Support Vector Machine
SW	=	Short Wave
TAFTS	=	Tropospheric Fourier Transform Spectrometer
TOA	=	Top Of the Atmosphere
ZPD	=	Zero Path Difference

Bibliografia

Radiation Explorer in the Far Infra-Red, Final Report: Executive Summary, ENV4-CT6 0344. 2000.

G. P. Anderson, J. H. Chetwynd, S. A. Clough, E. P. Shettle, and F. X. Kenizys. AFGL atmospheric constituent profiles (0-120km). Technical Report, U.S. Air Force Geophysics Laboratory. 1986.

M. P. Bailey and J. Hallett. A comprehensive habit diagram for atmospheric ice crystals: Confirmation from the laboratory, AIRS II, and other field studies. *Journal of the Atmospheric Sciences*, 66(9):2888–2899, 2009.

G. Bianchini and L. Palchetti. Technical Note: REFIR-PAD level 1 data analysis and performance characterization. *Atmospheric Chemistry and Physics*, 8:3817–3826, 2008.

G. Bianchini, L. Palchetti, and B. Carli. A wide-band nadir-sounding spectroradiometer for the characterization of the Earth's outgoing long-wave radiation. pages 63610A–63610A, 2006.

C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2010.

A. Bozzo, T. Maestri, R. Rizzi, and E. Tosi. Parameterization of single scattering properties of mid-latitude cirrus clouds for fast radiative transfer models using particle mixtures. *Geophysical Research Letters*, 35(16).

Deng Cai, Xiaofei He, Wei Vivian Zhang, and Jiawei Han. Regularized locality preserving indexing via spectral regression. In *Proceedings of the 16th ACM conference on Conference on information and knowledge management (CIKM'07)*, pages 741–750, 2007a.

Deng Cai, Xiaofei He, Wei Vivian Zhang, and Jiawei Han. Spectral Regression of Dimensionality Reduction. *Technical Report No. 2856*, 2007b.

B. Carli, A. Barbis, J.E. Harris, L. Palchetti, and D.Lastrucci. Design of an efficient broadband far infrared FT spectrometer. *Applied Optics*, 38(18):3945–3950, 1999.

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

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

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*, 136:718–739, 2010.

K.F. Evans and G.L. Stephens. A new polarized atmospheric radiative transfer model. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 46:413–423, 1991.

J.S. Foot. Some observations of the optical properties of clouds. II: Cirrus. *Quarterly Journal of the Royal Meteorological Society*, 114(479):145–164, 1988.

F. Di Giuseppe and R. Rizzi. Far Infrared scattering effects in cloudy sky. *Physics and chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 24:191–202, 1999.

T.C. Grenfell and S.G. Warren. Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation. *Journal of Geophysical Research*, 104:697–709, 2010.

P. Grenier, J-P. Blanchet, and R. Munoz-Alpizar. Study of polar thin ice clouds and aerosols seen by CloudSat and CALIPSO during midwinter 2007. *Journal of Geophysical Research*, 114, 2009.

J. Hansen and L.D. Travis. Light scattering in planetary atmospheres. *Space Science Reviews*, 16(4):527–610, 1974.

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.

A. W. Hogan. Summer ice crystal precipitation at the South Pole. *Journal of Applied Meteorology*, 14(2):246–249, 1975.

K. Kikuchi and A. W. Hogan. Properties of diamond dust type ice crystals observed in summer season at Amundsen-Scott South Pole Station, Antarctica. *Journal of the Meteorological Society of Japan*, 57:180–190, 1979.

M. D. King, Si-Chee Tsay, Steven E Platnick, Menghua Wang, and Kuo-Nan Liou. Cloud Retrieval Algorithms for MODIS: Optical Thickness, Effective Particle Radius, and Thermodynamic Phase. *MODIS Algorithm Theoretical Basis Document*, 1997, 1997.

T. Lachlan-Cope. Antarctic clouds. *Polar Research*, 29:150–158, 2010.

R.P. Lawson, B.A. Baker, P. Zmarzly, D. O'Connor, Q. Mo, J. Gayet, and V. Shcherbakov. Microphysical and optical properties of atmospheric ice crystals at South Pole Station. *Journal of Applied Meteorology and Climatology*, 45:1505–1524, 2006.

K. N. Liou. *An Introduction to Atmospheric Radiation*. Academic Press, 2002.

D. Lubin and D.A. Harper. Cloud radiative properties over th South Pole from AVHRR infrared data. *Journal of Climate*, 9:3405–3418, 1996.

T. Maestri. Il ruolo del vapor acqueo e delle nubi nel raffreddamento nel lontano infrarosso, thesis. 2000.

T. Maestri and R. E. Holtz. Retrieval of Cloud Optical Properties From Multiple Infrared Hyperspectral Measurements: A Methodology Based on a Line-by-Line Multiple-Scattering Code. *IEEE Transactions on geoscience and remote sensing*, 47, 2009.

T. Maestri and R. Rizzi. A study of infrared diabatic forcing of ice clouds in the tropical atmosphere. *Journal of Geophysical Research*, 4139, 2003.

T. Maestri, R. Rizzi, E. Tosi, P. Veglio, L.Palchetti, G. Bianchini, P. Di Girolamo, G. Masiello, C. Serio, and D. Summa. Analysis of cirrus cloud spectral signatures in the Far InfraRed. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 141:49–64, 2014.

A. Mahesh, V. P. Walden, and S. Warren. Ground-based remote sensing of cloud properties over the Antarctic Plateau. Part II: cloud optical depths and particle sizes. *Journal of Applied Meteorology*, 40:1279–1294, 2001.

D.H. Martin. Polarizing interferometric spectrometer for the near and sub-millimeter spectra. *Infrared and Millimeter Waves*, pages 65–148, 1982.

D.H. Martin and E.Pupplet. Polarized interferometric spectrometry for the millimeter and submillimeter spectrum. *Infrared Physics*, 10:105–109, 1969.

B. M. Morley, E.E. Uthe, and W. Viezze. Airborne LIDAR observation of clouds in the Antarctic troposphere. *Geophysical Research Letters*, 16:491–494, 1989.

C. Naud, J.E. Russel, and J.E. Harries. Remote sensing of cirrus properties in the far infrared. *Proceedings of SPIE*, 4168, 2001.

T. Ohtake. Atmospheric ice crystals at the South Pole in summer. *Antarctic Journal of the United States*, 13:174–175, 1978.

T. Ohtake and T. Yogi. Winter Ice crystal at South Pole. *Antarctic Journal of the United States*, 14:201–203, 1979.

L. Palchetti, G.Bianchini, A. Barbis, J.E. Harris, and D. Lastrucci. Design and mathematical modeling of the space-borne far-infrared Fourier transform spectrometer for REFIR experiment. *Infrared physics & technology*, 40, 1999.

P. Ricchiazzi, C. Gautier, and D. Lubin. Cloud scattering optical depth and local surface albedo in the Antarctic: Simultaneous retrieval using ground-based radiometry. *Journal of Geophysical Research*, 100:21091–21104, 1995.

R. Rizzi and T. Maestri. Some considerations on the infrared cloud forcing. *Journal of Geophysical Research*, 108, 2003.

R. Rizzi and L. Mannozzi. Preliminary results on the planetary emission between 100 and 600 cm^{-1}. *in REFIR-Radiation Explorer in the Far Infrared, Final Rep. ENV4-CT6 0344*, 1:77–87, 2000.

R. Rizzi, C. Serio, and R. Amorati. Sensitivity of broad-band and spectral measurements of outgoing radiance to changes in water vapor content. In *International Symposium on Optical Science and Technology*, pages 181–190, 2002.

L. S. Rothman, I. E. Gordon, A. Barbe, D. C. Benner, P. F. Bernath, M. Birk, V. Boudon, L. R. Brown, A. Campargue, J-P. Champion, et al. The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 110: 533–572, 2009.

K. Sassen. Depolarization of laser light backscattered by artificial clouds. *Journal of Applied Meteorology*, 13:923–933, 1974.

K. Sassen. Ice crystal habit discrimination with the optical backscatter depolarization technique. *Journal of Applied Meteorology*, 16:425–431, 1977.

K. Sassen. *Lidar Backscatter Depolarization Technique for Cloud and Aerosol Research, in Light Scattering by Nonspherical Particles: Theory, Measurements, and Geophysical Applications.* 2000.

R.S. Stone. Properties of austral winter clouds derived from radiometric profiles at the South Pole. *Journal of Geophysical Research*, 12:961–971, 1993.

S. Theodoris and K. Koutroumbas. *Pattern Recognition.* Academic Press, 2009.

S. Theodoris and K. Koutroumbas. *An Introduction to Pattern Recognition: A MATLAB Approach.* Academic Press, 2010.

P. Veglio and T. Maestri. Statistics of vertical backscatter profiles of cirrus clouds. *Atmospheric Chemistry and Physics*, 11:12925–12943, 2011.

V.P. Walden, S.G. Warren, and E. Tuttle. Atmospheric ice crystals over the Antarctic Plateau in winter. *Journal of Applied Meteorology*, 42:1391–1405, 2003.

S. Warren. Optical constants of ice from the ultraviolet to the micro-wave. *Applied Optics*, 23:1206–1225, 1984.

P. Yang, M. G. Mlynczak, H. Wei, D. P. Kratz, B. A. Baum, Y. X. Hu, W. J. Wiscombe, A. Heidinger, and M. I. Mishchenko. Spectral signature of ice clouds in the far-infrared region: single-scattering calculations and radiative sensitivity study. *Journal of Geophysical Research*, 108, 2003.

P. Yang, M. G. Mlynczak, H. Wei, D. P. Kratz, B. A. Baum, Y. X. Hu, W. J. Wiscombe, A. Heidinger, and M. I. Mishchenko. Scattering and absorption property database for non-spherical ice particles in the near through far infrared database. *Applied Optics*, 44: 5512–5523, 2005.

P. Yang, M. G. Mlynczak, H. Wei, D. P. Kratz, B. A. Baum, Y. X. Hu, W. J. Wiscombe, A. Heidinger, and M. I. Mishchenko. Spectrally Consistent Scattering, Absorption and Polarization Properties of Atmospheric Ice Crystals at Wavelengths from 0.2 to 100 μm. *Journal of the atmospheric sciences*, 70:330–347, 2013.

K.P. Zetie, S.F. Adams, and R.M. Tocknell. How does a Mach-Zehnder interferometer work? *Physics Education*, 35, 2000.

V. M. Zolotarev, B. A. Mikhilov, L. L. Alperovich, and S. I. Popov. Dispersion and absorption of liquid water in the infrared and radio regions of the spectrum. *Optics and Spectroscopy*, 27:430–432, 1969.

Acknowledgements

Desidero ringraziare innanzitutto i Professori Rizzi e Maestri per la grande professionalità, disponibilità e cortesia che mi hanno dimostrato in questi mesi e per l'aiuto fornito durante la stesura della tesi, sopportando da ultimo la mia carente scrittura inglese.
Esprimo la mia gratitudine verso il CNR di Firenze per aver fornito il dataset di misure analizzato in questo lavoro: in particolare ringrazio il Dott. Palchetti per la grande disponibilità e attenzione mostrata e il Dott. Del Guasta per aver messo a disposizione quick-looks e matrici Lidar, fondamentali per sviluppo della parte di retrieval e simulazione dei casi studio.
Un ringraziamento anche al Dott. Masiello dell'Università degli studi della Basilicata, per la collaborazione e la preparazione del dataset di reanalisi ECMWF.
Ringrazio inoltre tutti i professori che durante questi 5 anni mi hanno permesso di crescere sia dal punto di vista culturale che umano.

Ovviamente un particolare ringraziamento ai miei genitori che, con il loro sostegno, mi hanno permesso di raggiungere questo importante traguardo.

Desidero inoltre ringraziare mia sorella Marta, preziosa alleata nel risolvere numerose piccole e grandi questioni.

Grazie a tutti i miei amici, a casa come in studentato, che mi hanno sostenuto in questa avventura e che hanno dato sostanza e colore alle mie giornate, così come ai miei compagni di corso che mi hanno accompagnato in questi anni di studio, molti dei quali per me più veri amici che semplici compagni.