ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI CHIMICA INDUSTRIALE
DIPARTIMENTO DI CHIMICA FISICA ED INORGANICA
DOTTORATO IN SCIENZE CHIMICHE XIX CICLO
S.S.D. CHIM/02

PhD Thesis

Tomographic Inversion of MIPAS Measurements; Optimization of Retrieval Grids and Generation of Innovative Products

ENZO PAPANDREA

ADVISOR:
PROF.
MASSIMO CARLOTTI

CO-ADVISOR:
DR.
BIANCA MARIA DINELLI

COORDINATOR: PROF. VINCENZO BALZANI

BOLOGNA, MARCH 2007

Keywords

MIPAS, ENVISAT, Remote Sensing, Stratosphere/Troposphere, Spectroscopy, Trade-off Study, Carbon Dioxide (CO_2)

Contents

Overview

The main objective of this thesis is to perform a trade-off study between the accuracy and the horizontal resolution of the retrieved parameters on MIPAS observations. The results are compared with the work previously done by Luca Magnani in his PhD Thesis [1] on simulated data.

Spectroscopic data are acquired by the **M**ichelson **I**nterferometer for **P**assive **A**tmospheric **S**ounding (**MIPAS**). It is a high resolution Fourier Transform Spectrometer developed by the **E**uropean **S**pace **A**gency (**ESA**) for the detection of limb emission spectra in the upper atmosphere.

Two methods to vary the horizontal resolution have been tested. One method varies the separation between profiles in the retrieval grid. The second method exploits a variation of the measurements grid at the expenses of modifying spectral resolution.

Another objective of this study is try to retrieve in the stratosphere one of the most important greenhouse gases (the CO_2) with acceptable retrieval errors. For the purpose of this study the GMTR analysis system has been upgraded in order to take into account the CO_2 line mixing, by using the model and software routines presented in a recent paper [2]. Furthermore it has been upgraded with the capability to perform retrievals on a target-dependent vertical retrieval grid.

Below is reported the content of each chapter of this document.

Chapter 1: provides a brief introduction to the Earth's Atmosphere characteristics and composition. It also presents an overview of the spectroscopic techniques used to study the Atmosphere, in particular is described here the Remote Sensing.

Chapter 2: summarizes the significant contributions of **F**ourier **T**ransform **S**pectroscopy (**FTS**) in general and to atmospheric spectroscopy in particular. A review of the historical developments from the 19^{th} century work of Michelson to the experience of present day systems is presented.

Chapter 3: illustrates the so called *Inverse Problem*, that is the problem of finding the best representation of the required parameter from measurements that are a complicated function of the parameter. Its solution is required in the case of indirect measurements, such as spectral radiances, in which it is not possible to detect directly the quantity of interest (in our case the constituent concentration).

Chapter 4: describes the MIPAS experiment. The main scientific objectives of MIPAS, the description of the instrument and its measurement requirements and principles are presented in this chapter. A scientific code for Level 2 analysis of middle infrared emission spectra measured by MIPAS experiment, called **O**ptimized **F**orward **M**odel (**ORM**), is routinely used by ESA.

Chapter 5: describes the ORM limits and how they can be overcomed with the recently proposed **GMTR** [3] code that is the result of the merging of Geofit method [4] implemented with the **M**ulti **T**arget **R**etrieval (**MTR**) approach [5]. This method allows to model the horizontal variability and the simultaneous retrieval of more than one target and has been used for our results.

Chapter 6: summarizes the theoretical study of trade-off between the accuracy associated to the retrieved parameters and the spatial resolution obtainable with the GMTR analysis system. This study was performed on simulated data [6] as MIPAS spectra where not available when it has been done. Two methods to vary the horizontal resolution have been tested. One method varies the separation between profiles in the retrieval grid, while the second one exploits the different measurement grid of some MIPAS orbits acquired during August, 2004.

Chapter 7: a practical application of the optimized retrieval grid is the study of CO_2 distribution. This chapter describes the line mixing implementation [2] in the retrieval code and the consequent advantages that it carries out. It also describes the microwindows selection process for the simultaneous retrieval of pressure, Temperature and CO_2 and the subsequential creation of lookup tables with CO_2 line mixing and irregular grids for these microwindows. Finally, monthly mean CO_2 distribution maps are shown.

Chapter 1

Characteristics and Remote Sensing of the Earth's Atmosphere

The vertical structure of the Earth's atmosphere has been defined by scientists in a variety of ways depending on which physical property is used. The most common classification is based upon the temperature profile. The vertical distribution of temperatures in the atmosphere depends on the balance of heat fluxes driven, at each level, by radiative processes or heat transfer due to atmospheric motions. Furthermore, the distribution of some chemical species, generally determined by chemical and dynamical processes, affects the thermal structure of the atmosphere. The field of atmospheric science covers therefore a wide variety of disciplines involving radiation, chemistry, and dynamical processes. In this first chapter, a brief introduction to the basic physical and chemical properties of the atmosphere is given. Furthermore this chapter provides an introduction to Remote Sensing of the Earth's atmosphere. Remote Sensing is defined as the technique of acquiring information about a system without actually coming into physical contact with it. This is done by sensing reflected, absorbed or emitted energy and retrieving the desired information. In Remote Sensing the "required" quantity is in general not the "measured" quantity but some more or less complicated function of it. Information transfer through space is accomplished using spectroscopic techniques. The spectroscopy of the atmosphere consists in recording a portion of the electromagnetic spectrum to identify the presence and determine the amount of specific molecular species in the atmospheric region under investigation. The number and type of spectroscopic techniques currently in use for the Remote Sensing of the atmosphere is large and varied: observations are made at almost all wavelengths, both in emission and absorption, therefore only a brief overview will be summarized in the following sections.

1.1 Why do we study the Earth's atmosphere?

Only recently, it has been realized that the atmosphere is an important part of the Earth system, but for a long time we have had a little understanding of the structure of this gaseous envelope and of the fundamental role it plays in sustaining life on Earth. Until about 1950, it was assumed that the mean state of the atmosphere was stable, well balanced and not likely to be significantly affected by anthropogenic activity. With the development of new sounding methods and techniques, and the improvement in sensitivity of detectors and analytical capabilities, our knowledge on the atmospheric composition has not only become more detailed, but it has also revealed the complex interactions among different molecular species as well as their behavior in the presence of solar radiation.

At least 40 different molecular species have been identified as part of the atmospheric mixture and have been measured. These gases, together, control the global radiative budget and dynamical variability of the atmosphere. The climatic conditions for life to persist on Earth are linked to the way some gases act, for example as a "shield" against harmful (to most living things) solar ultraviolet radiation or as a "thermal blanket" trapping the heat and insulating the Earth from cold space or as a "detergent" of the lower atmosphere. Some other gases are present only as a result of human activities and have an effect which modify the natural balance that exists in the atmosphere.

The chemical composition of the Earth's atmosphere varies to a large degree with local physical conditions, and is therefore dependent on recurring events such as diurnal, seasonal or solar cycles. Furthermore, there are episodic phenomena (i.e. volcanic eruptions, biomass burning, …) that can perturb the regional or even the hemispherical concentrations of many telluric gases with respect to their global and natural values. Nowadays observational methods and techniques for fulfilling the investigation of the chemical composition and variability of the atmosphere (in terms of latitude, longitude, altitude and time dependence) have become more sophisticated and more sensitive, providing global coverage, simultaneous measurements of many species, absolute concentration determinations and temporal variations.

1.2 Composition and vertical structure

The atmosphere (from the Greek: *atmos* = vapor and *sphaira* = ball) is the thin layer of gas, held by gravitation, that surrounds a planet. Over 99% by volume of the Earth's atmosphere consists of molecular nitrogen (N_2, 78.1%), and molecular oxygen (O_2, 20.9%). These gases, as well as the noble gases (argon, neon, helium, krypton, xenon), posses very long lifetimes against chemical destruction and, hence, are relatively well mixed below approximately 90 km altitude. "Minor" constituents, such as CO_2, CO, CH_4, N_2O, O_3, H_2O, and over twenty "trace" species, account for the remaining 1% of the atmospheric volume. In addition, the atmosphere also contains various solid and liquid particles such as aerosols, water drops and ice crystals, which are highly variable in space and time. These particles suspended in the atmosphere play an important role in absorption and scattering of solar radiation and in the physics of clouds and precipitation. Even if most of the gaseous constituents are well mixed, the atmosphere itself is not physically uniform. It has a heterogeneous vertical structure where significant variations with altitude in composition, temperature, pressure and density are present.

The ratio of atmospheric constituents in the absence of sources or sinks is determined by either molecular diffusion or mixing due to fluid motion. Molecular diffusion attempt to order the atmosphere in such a way the lightest gases are present at the top of the atmosphere while the heaviest molecules occupy the bottom. Mixing due to fluid motion of the molecules is instead independent of molecular mass and, if this is the dominant process, the atmosphere's composition will be independent of height. At lower heights the mean free path between molecular collisions is so small that the time necessary for molecular diffusion is many orders of magnitude greater than that required for turbulent motions to homogenies them. However, at higher heights the mean free path increases to such an extent that molecular diffusion is the dominant mechanism.

Temperature varies greatly both vertically and horizontally. However, despite horizontal variations, its vertical structure is qualitatively similar everywhere, and so it is meaningful to think of a "typical" temperature profile. Figure 1.1 represents, for middle latitude conditions, the vertical temperature profile of the U.S. Standard Atmosphere (1976) [7]. Typically, the profile is divided into a set of layers on the basis of the different vertical temperature gradients, also known as *lapse rates* ($-dT/dz$). The altitudes at which these layers occur are not fixed but vary in time and location. As indicated in the figure, the atmosphere results neatly distinguished into four layers: *troposphere, stratosphere, mesosphere* and *thermosphere*, each characterized by substantially different chemical and physical processes. The boundaries between these layers are respectively called the *tropopause, stratopause, mesopause* and *thermopause*. Frequently, the different regions are referred as the *lower atmosphere* (troposphere), the *middle atmosphere* (stratosphere and mesosphere) and the *upper atmosphere* (above about 100 km).

The behavior of the temperature profile can be understood by considering the dominant processes occurring at specific heights. Much of the electromagnetic radiation emitted by the Sun is at wavelengths in or near the visible part of the spectrum. These are not strongly absorbed by the atmosphere and so, unless the presence of clouds to scatter the photons and reflect a portion of the radiation back to space, most of the energy reaches the ground where it is absorbed. The Earth's surface is then heated by the incoming solar radiation, and re-radiates the absorbed energy as a black-body spectrum. The heating of the lower atmosphere by the ground causes active convection and a decrease in temperature due to volume expansion. Molecules, like water vapor, absorb some of the outgoing terrestrial radiation in the infrared region of the electromagnetic

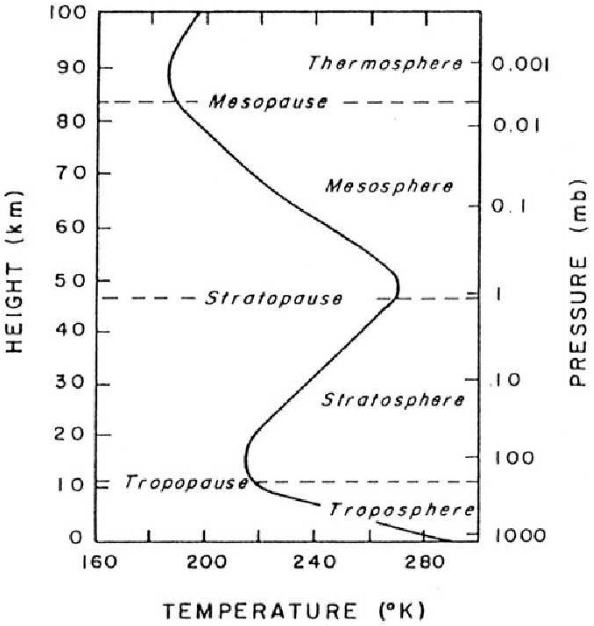

Figure 1.1: Vertical distribution of temperature and pressure between the surface and 100 km altitude as defined by the U.S. Standard Atmosphere 1976 model for mid-latitude conditions. The designation of each part of the atmosphere is shown.

spectrum. The energy associated with the absorbed radiation causes the molecules to vibrate and/or rotate with respect to their centers of mass.

The lower-most layer, the *troposphere*, ranges in thickness from the surface to the tropopause at the approximate altitude of 18 km over the tropics, 12 km at mid-latitudes, and from 6 to 8 km near the poles. Here the mean temperature decreases with increasing altitude with a typical lapse rate of $6.5°C/km$. This layer, which contains about 85-90% of the atmospheric mass, as well as nearly all of the water vapor and aerosols, is characterized by strong instability with significant vertical exchanges of energy and mass associated with convective activity. Much of the variability observed in the atmosphere occurs within this layer. Owing to the temperature structure of the troposphere, it is in this region that most of the weather related phenomena (i.e. storms, clouds, precipitations) essentially take place.

The layer above the tropopause, known as *stratosphere*, is the one in which the vertical temperature gradient reverses because of ozone and, to a lesser extend, molecular oxygen absorbtion of ultraviolet solar radiation according to the cyclic mechanism discovered by Chapman [8]. The negative temperature lapse rate suppresses the vertical motion and creates within this layer, which extends up to 50 km, very stable conditions dominated by radiative processes. The lack of convective air motion results in stratification (warmer air lies above cooler air, hence the name "layered sphere") and prevents the vertical mixing between the troposphere and the stratosphere. Since virtually no water can cross the cold tropopause, and since only small amounts of water are formed within the layer from oxidation of CH_4 molecules, the stratosphere is very dry (a few ppmV of H_2O). The ozone concentration profile peaks in this atmospheric layer at about 35 km when considering the relative concentration, i.e. volume mixing ratio (see fig. 1.2). The stratosphere contains about 90% of the atmospheric ozone, and most of the solar ultraviolet radiation is absorbed by this "layer". The importance of the ozone layer lies in the fact that the dissociation of ozone molecules helps the Earth to maintain its heat balance, and reduces the amount of harmful UV radiation that reaches its surface.

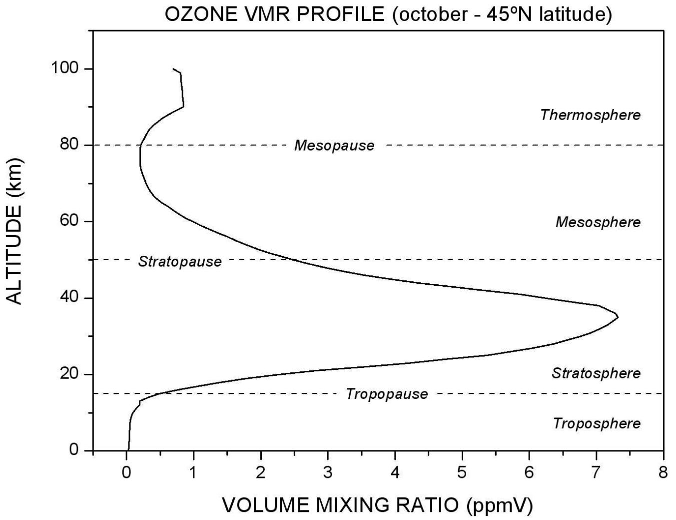

Figure 1.2: Vertical distribution of ozone between the surface and 100 km altitude as given by the IG2 climatology for mid-latitude conditions.

Like the troposphere, the temperatures in the *mesosphere*, from about 50 to about 90 km, declines with altitude. The lapse rate is again positive, approximately $2°C/km$. In this layer most chemicals are in an excited state, as they absorb UV energy from the Sun; since ozone concentration decreases with altitude, no ozone causes here heating. Dynamical instability occurs frequently in this layer and convection plays here a large role again.

Above about 90 km, in the region denoted as *thermosphere*, the temperature rises once again to reach maximum values that are strongly dependent on the amount of solar activity. It is in this layer that very short UV wavelengths photo-dissociate molecular oxygen into atoms and ions as well as molecular nitrogen at rather higher altitudes. Vertical exchanges associated with dynamical mixing become insignificant and molecular diffusion becomes an important process, producing separation of species according to molecular or atomic weight.

As one moves even farther from the Earth's surface, gravity decreases and it is not as strong as it is near the surface. All gaseous molecules in the atmosphere are in constant motion and, at a given temperature, lighter molecules have root mean square (*rms*) speeds higher than heavier molecules. If *rms* speeds are high enough, some molecules may escape to space. Long time ago, the escape to space of hydrogen atoms from the *exosphere*, the atmospheric region above about 500 km, has contributed to substantial changes in the chemical composition of the Earth's atmosphere and accounts for the relative lack of hydrogen in spite of its abundance in the Universe.

The vertical profile of pressure comes from the competition between gravity, which pulls molecules toward Earth's surface, and the molecular kinetic energy, which keeps the molecules moving. Expressed mathematically, this competition results in the law $p = p_{surface} \cdot exp(-mgz/k_B T)$, where $p_{surface}$ is the surface pressure, m is the average molecular weight of the molecules that compose air, g is the acceleration due to gravity, z is the height, k_B is the Boltzmann constant, and T is the temperature. On average the factor kT/mg is 7±1 km. Thus, the atmospheric pressure falls off exponentially with height by a factor of e^{-1} (2.7) every 7 km.

The Sun and the solid Earth are approximately black-bodies, with temperatures of typically 5800 K and 280 K, reaching their maximum at 0.50 μm and 10 μm respectively. The 5800 K black-body curve exceeds the 280 K curve at all wavelengths, as seen in fig. 1.3. The global mean temperature of the Earth-Atmosphere system is about 250 K, much lower than that of the Sun's photosphere. As a consequence, the atmospheric black-body signal is less and the wavelength for the intensity peak of the Planck function is longer. When the Sun-Earth distance is considered, as in fig. 1.4, only a minor overlap around 4 μm is noticeable. Hence, atmospheric radiation with wavelengths shorter than 4 μm origins likely from Solar irradiance and is usually studied through absorption and scattering while radiation with wavelengths longer than 4 μm likely originates from Telluric emittance and is usually studied through emission processes. The atmospheric attenuation is high below 3 μm and above 20 μm. In between, the atmosphere is more transparent with structures from absorbing species (fig. 1.4).

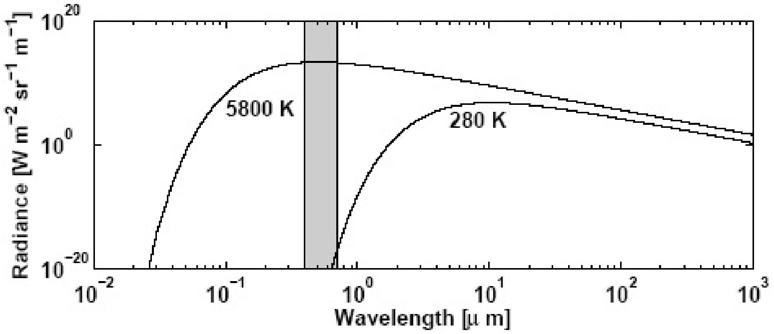

Figure 1.3: Planck black-body distributions corresponding to Sun (5800 K) and Earth (280 K) surfaces. The visible region (0.4–0.7 μm) is shaded in grey.

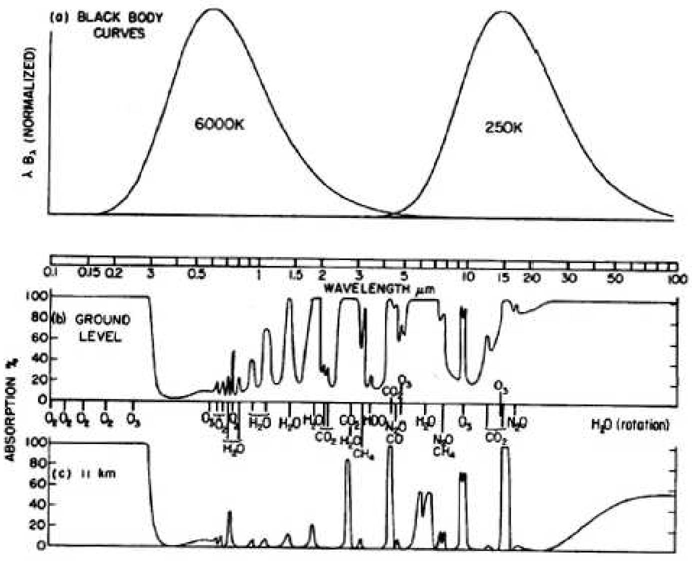

Figure 1.4: Normalized black-body distributions corresponding to Sun ($\sim$ 6000 K) and Earth's effective temperature ($\sim$ 250 K). The atmospheric attenuation at ground level is around 100 % below 3 μm and above 20 μm. In between, the atmosphere is more or less transparent with structures from absorbing species.

1.3 Observation techniques

A wide variety of apparatuses is used to observe and measure the Earth's atmosphere. The methods currently in use for studying the chemical state of the atmosphere can be distinguished into two basic and very different approaches, namely *in situ* and *remote Sensing* investigations.

In situ measuring method implies bringing the instrument directly into the location where the measurement is to be performed (co-location). Consequently, the determination of profiles of gas concentrations versus altitude requires in situ instruments to be airborne and operate at different heights. Typical platforms that have been used extensively over the last two decades include airplanes, balloons, and rockets. A detailed review of in situ methods and techniques can be found in Gille [9] and Ehhalt [10].

Whenever in situ measurements are insufficient to provide the desired information, Remote Sensing techniques are used. Remote Sensing methods, which are more specifically discussed in this chapter, imply that the gathering of information about features of the atmosphere, both chemical and physical, is achieved by sensing systems located at some distance from the air parcel or layer to be studied. In principle, these methods have two specific advantages: they sound large areas of the atmosphere and do not perturb the local state of the air parcel under investigation. Remote Sensing from space is the only possible solution for achieving global datasets with temporal and spatial resolution, hence it is used in many atmospheric fields. In situ measurements, with their high accuracy are however important in the validation of Remote Sensing missions.

Both Remote Sensing and in situ methods can be either *passive* or *active* techniques. The distinction depends on the possible type of sensor carried by the instrument. A passive sensing technique measures the interaction of the atmosphere with naturally generated radiation, recording that part of the electromagnetic spectrum where atmospheric constituents leave their specific absorption or emission fingerprints. Examples of passive experiments include interferometers, radiometers and photometers. Most methods are passive, but some, notably LIDAR (LIght Detection And Ranging) and GPS (Global Positioning System) occultation, are active sensing techniques that propagate their own electromagnetic radiation toward the target to be investigated and measure the intensity of the return signal.

While Remote Sensing is the only means to sound the atmosphere from space platform, thus allowing global coverage and geographic variability, it has, however, limitations in both spatial resolution and/or short term variability. For such detailed studies, undertaken locally, in situ methods as well as laser techniques operated from the ground (LIDAR [11] and DIAL [12] instruments, laser heterodyne radiometry and microwave sounding [13]) are better suited. Figure 1.5 gives a representation of the methods, modes and techniques currently in use for atmospheric composition studies. It should be noted that no unique method-technique-instrument of investigation has the virtue to provide the complete characterization of the atmosphere in all its dimensions and, therefore, the development of several complementary techniques remains essential.

1.4 Geometries and platforms of observation

Since Dobson [14] constructed his first ground based ultraviolet spectrometer to measure total ozone in the 1920's, a great variety of atmospheric Remote Sensing instruments and techniques have been designed. Many configurations for atmospheric Remote Sensing instruments, with specific advantages and disadvantages, can be identified.

Ground based instruments are maintainable and stable, providing long time series, however, they give poor spatial coverage and are generally sensitive to tropospheric clouds, turbulence and water vapor.

Balloons may be used to measure profiles of atmospheric geophysical parameters with high vertical resolution up to 35-40 km. Their major disadvantage is that their flight track cannot be actively controlled and they only provide a limited coverage, as a "snapshot" of the atmosphere.

Aircrafts can be used for controlled flights up to 20 km and can provide better coverage, but are not suitable for global measurements with spatial, temporal and vertical resolution. Remote Sensing systems have the advantage of being able to make multiple measurements from a range of heights simultaneously. However, the vertical resolution is usually quite coarse when compared with in situ measurements.

Satellites are suitable for global coverage measurements on a daily basis. Geostationary orbits with an altitude of 35,900 km permit continuous measurements with high temporal resolution but are limited to 60° coverage in longitude and may only be used efficiently up to latitudes of ±60°. Polar orbits with altitudes of 600-850 km give near global coverage. Satellites are expensive and have a lifetime limited to a few years, which makes long time series impossible. Multi-satellite missions with similar instruments may, of course, provide better temporal and spatial coverage or longer time series. Space-based platforms are becoming increasingly important for obtaining measurement on a synoptic or even global scale. A variety of instruments are employed, using both active and passive sensing, covering a wide range of wavelengths of the electromagnetic spectrum, and basing their sounding methodology on different physical properties such as scattering, emission, refraction, transmission or occultation. In addition, many different observation geometries are available.

The geometry with which the atmosphere can be observed is defined by the location of the observer and the orientation of the line of sight. The location, as already said, can be the ground (either from sea level or high mountain sites), an aircraft, a balloon, a rocket platform or a satellite. The line of sight can be oriented either in the vertical direction, *vertical sounding*, or in the horizontal direction, looking at the Earth's limb, *limb sounding*. Different combinations of location and line of sight are geometrically possible.

Vertical sounding can distinguish the contributions from different altitudes exploiting the fact that the atmospheric spectral features have a pressure (and, therefore, an altitude) dependent line shape. Very high spectral resolution instruments are able to record the detailed shape of a spectra. The signal in the far wings of the lines originates mainly from low altitudes, whereas the signal in the center of the lines originates from high altitudes. Therefore, by appropriate deconvolution methods, the distribution of the atmospheric constituents can be determined with a vertical resolution approximately equal to the atmospheric scale height. Vertical sounding can be made looking either in nadir and zenith directions. In the case of nadir vertical soundings, the atmosphere is observed from above, and the signal-to-noise is limited by the small contrast that exists between atmospheric and Earth's surface temperatures. Nadir sounding measurements are possible from aircrafts, stratospheric balloons and satellites. In the case of zenith vertical soundings, the atmosphere is observed from below, either in emission or in absorption mode. In this case the background signal, being a bright source or empty space, provides better temperature contrast. However, only relatively strong features that occur in a transparent atmospheric spectral interval can be measured and, therefore, high mountains, dry sites and aircraft platforms must be used in order to reduce water vapor attenuation. With the *limb* sounding technique, the observer is at high altitude and

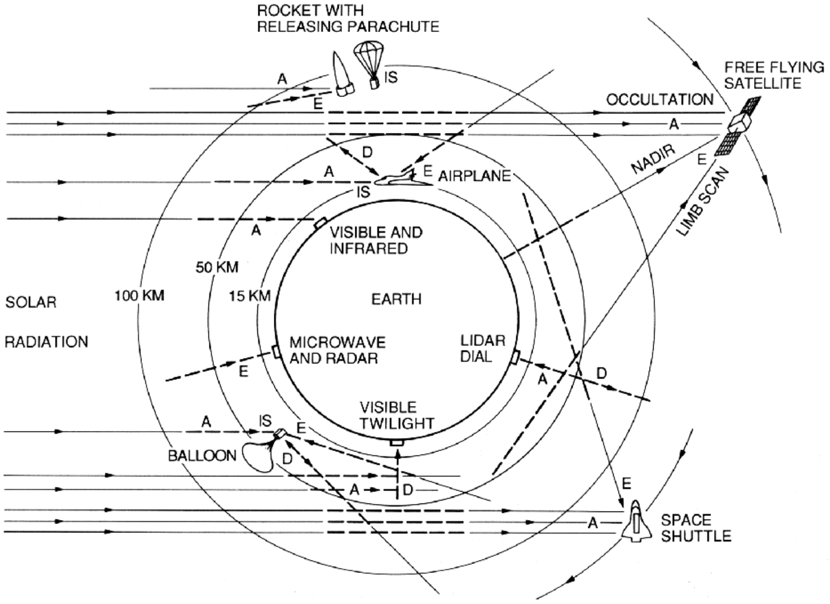

Figure 1.5: Representation of the most important methods, modes and observation techniques currently in use for atmospheric composition investigations. Letters *A*, *E*, and *D* assigned to rays reaching various platforms mean Absorption-, Emission-, and Diffusion-mode operation, respectively. The letters *IS* stand near those platforms from which in situ studies can be performed.

the line of sight is oriented below the horizontal direction, so that the Earth's surface is not reached and only the atmospheric limb is observed (at zenith angles greater than 90°). Limb measurements can be made both in emission and in absorption mode. Limb sounding measurements, in general, maximize the observed signal and provide a good vertical resolution in the determination of the vertical distribution, since, in a given spectrum, the largest contribution to the signal arises from molecules near the *tangent point* (minimum distance of the optical path from the Earth's surface). Observations at different zenith angles provide measurements with different tangent altitudes and deconvolving the data using an inverse procedure it is possible to retrieve the atmospheric composition as a function of altitude (more details on the inversion procedure will be presented in Chapter 2). The vertical distribution obtained in this way, of course, is limited to altitudes below the observation height, so that high altitude platforms are desired. Therefore, the emerging need to determine the vertical distributions of constituents in the upper atmosphere above the ceiling of stratospheric balloons (that fly at altitudes of up to about 40 km) provided in the past a strong incentive to develop high resolution spectroscopic instruments capable to make measurements from space.

1.5 Emission and absorption measurements

The spectrum of the Earth's atmosphere can be observed in emission using the black-body thermal emission of the atmosphere itself or in absorption using the Sun or some other natural source (i.e. stars) as the radiation source in the line of sight. Most of the high resolution spectroscopic observations, which have provided data for a large number of infrared active constituents of the atmosphere, were made in the absorption mode, because solar absorption methods have the advantage of using a very bright source providing spectra with high signal-to-noise ratio.

A major disadvantage of spectroscopic or radiometric absorption methods is the fact that the measurements are constraint both in space and time, because observations can only be made at rather restrictive times of the day. This aspect poses serious difficulties in the case of diurnally varying species. These restrictions do not apply of course to emission measurements, which can be made continuously at all times and in all possible directions, except toward the sun, providing the best possible diurnal and geographical coverage[1]. The emitting medium is a relatively weak source, however, since the signal-to-noise ratio is proportional to the signal intensity of the measurement and to the square root of the measurement time, the signal-to-noise enhancement provided by the intensity of the absorption measurements can be comparable to or smaller than the signal-to-noise enhancement provided by the long measurement time of emission measurements. On the other hand, absorption measurements are not dependent (to the first order) on a detailed knowledge of the temperature along the line of sight of the observations, as in the case of measurements made in emission. In the latter case, the errors resulting from uncertainties in temperature depend on the spectral region that is observed.

1.6 High altitude platform Far Infrared Experiments

Until the 1960s the far-infrared (10 cm^{-1} – 400 cm^{-1}, 1000 μm – 25 μm) was a relatively unexplored region of the electromagnetic spectrum. The reasons for this came from the initial unavailability of strong sources, the poor performance of detectors and the inefficiency of classical instruments designed to observe this region.

Better sensitivity became available as technological innovations were developed and grating instruments were extended to these long wavelengths. However, it was not until the practical realisation of multiplex instruments and the availability of very sensitive liquid helium-cooled detectors in the 1950's and 1960's that far infrared spectroscopy became a practical monitoring technique [15].

---

[1] For satellite measurements it should be noted, however, that in order to better attain best geographical coverage, emission measurements usually adopt sun-synchronous orbits which again are constraints to measurements at fixed time during the diurnal cycle.

Following the early activity of Gebbie in 1957, a number of groups carried out studies of the sub-millimetre emission spectrum of the lower atmosphere from aircraft and balloon. Measurements of the submillimeter far infrared emission spectrum were made by Harries in 1971 from the Comet 2E experimental aircraft [16, 17, 18] and in 1972 from the Concorde 002 aircraft [19, 20]. These measurements showed the large improvement in spectral resolution by using a more sensitive cooled detectors. Using a Michelson interferometer with a Golay cell detector, spectra might be obtained up to a maximum resolution of about 0.2 cm^{-1}, while using the liquid helium cooled detectors it was possible to reach an unapodized spectral resolution of 0.0625 cm^{-1}.

Kendall and Clark reported further measurements of stratospheric emission spectrum from balloon altitudes during one of the 1976 flights [21]. At higher altitudes accessible from balloon platforms, it became possible to detect weak lines and to use limb sounding methods to derive concentration height profiles for a number of species. Using a Michelson interferometer and cooled detectors, the data were measured at a resolution up to 0.08 cm^{-1} over the spectral range 30-110 cm^{-1}[22].

Further measurements of the stratospheric emission spectrum in the sub-millimetre spectral interval with a higher resolution were made by Carli et al. in 1979 from the **SIBEX** (**S**ubmillimiter **I**nfrared **B**alloon **EX**periment) balloon platform at about 38 km altitude [23]. The measured emission spectrum were obtained in the 7–85 cm^{-1} spectral range with a resolution of 0.0033 cm^{-1}. This improvement in resolution, made possible by the use of both a high efficiency polarising Michelson interferometer [24] and new sensitive bolometers, has provided very detailed information on the complex structure of the stratospheric sub-millimetre spectrum.

In 1978 M. J. Bangham carried out a number of line-by-line spectral emission computations, generally for the case of limb sounding geometry with an observer on a balloon or a spacecraft [25]MB78. An example of this work is given in fig. 1.6 (top). Only a part of this spectrum is compared with the measurement made in 1979 from SIBEX. This spectrum provides an example of the two features due to the two chlorine isotopes of HCl (bottom of fig. 1.6) that have been positively observed in the far infrared with a spectral resolution of 0.0033 cm^{-1}. Carli et al. measured further high resolution stratospheric spectra during two balloon flights in 1982 and 1983 that made it possible to extent the analysis in the far infrared region up to 190 cm^{-1} with the detection of several species such as ClO, OH and HBr [26]. Measurements at higher frequencies were obtained by Chance and Traub [27]. These measurements cover the 70–260 cm^{-1} range with a resolution of 0.032 cm^{-1} and made possible the study of a new spectral range and the identification of several lines due to important stratospheric species such as HOCl and H_2O_2.

Based on the experience acquired with the balloon experiments, in 1993–1995 a new instrument named **SAFIRE-A** (**S**pectroscopy of the **A**tmosphere using **F**ar **I**nfraRed **E**mission/**A**irborne) has been developed by Carli and al. [58] for the atmospheric com-

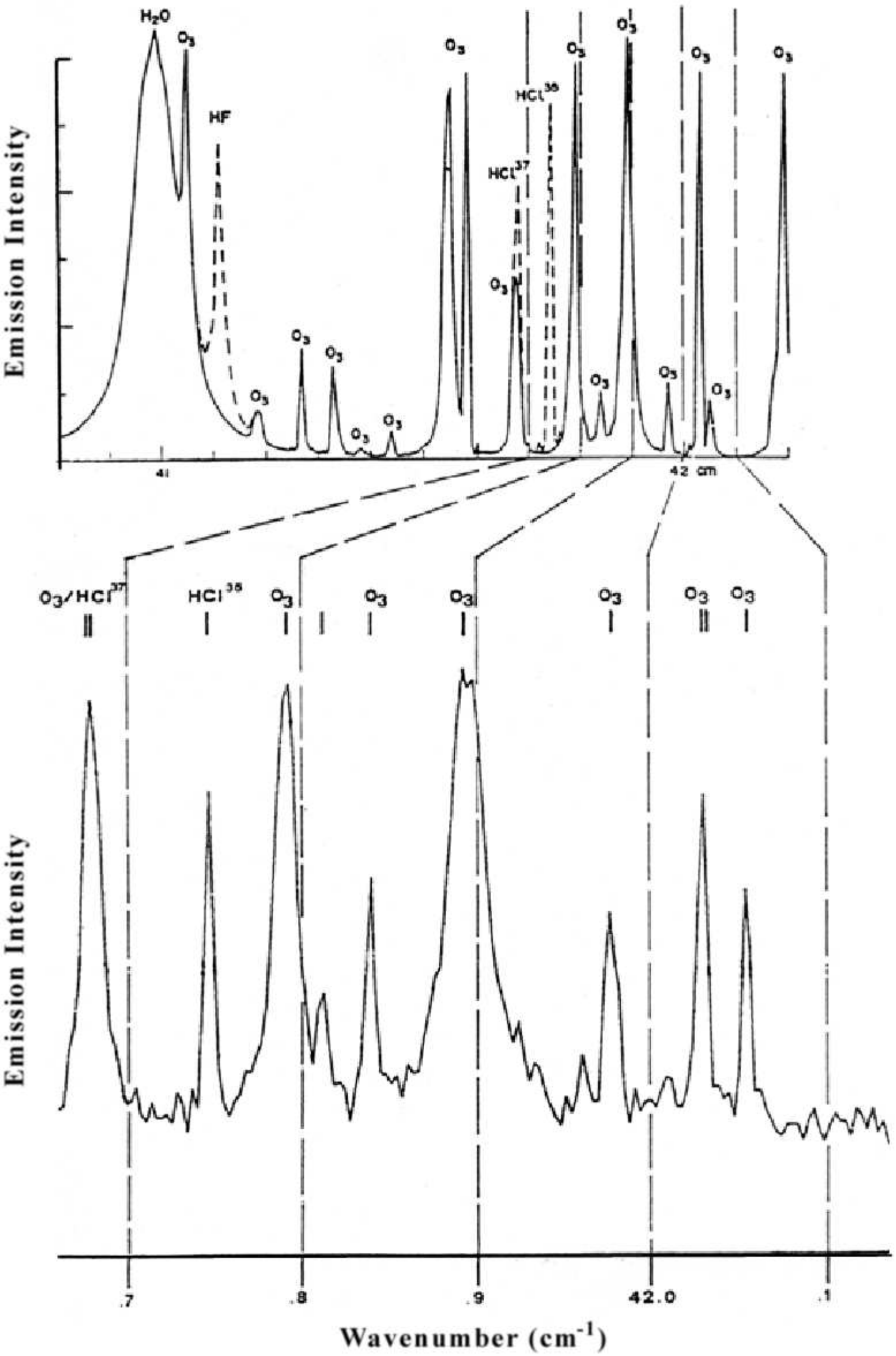

Figure 1.6: Comparison of a part of the calculated spectrum of atmospheric emissivity by Bangham (top) in the 41 cm^{-1} region in the case of limb view at 20 km of tangent height and the experimental spectrum observed by Carli et al. from SIBEX balloon at about 38 km with a tangent height of 28 km (bottom). In the calculated spectrum, the solid line is for water vapor, oxygen and ozone and the broken line shows the effect of adding HCl and HF with a relative concentration of 1 part in 109 by volume.

position through limb scanning emission measurements in order to exceed the limited location of balloon flights. SAFIRE-A is a polarising Fourier transform spectrometer that covers the spectral region from 10 to 250 cm^{-1} with a resolution of 0.004 cm^{-1} which successfully operated on a M-55 stratospheric research aircraft in mid latitude and polar regions during several measurement campaigns.

1.7 Satellite FTS Experiments

Remote Sensing from satellites is a powerful experimental technique used in global atmospheric research. In this field, as the need emerged of measuring the distributions of several trace species simultaneously, the requirement of instrument performance became more and more demanding. The progress made in the last decade with FTS spectrometers lead to the development of enhanced sensors. A selection of FTS instruments operating in the IR range previously or recently launched, currently flying, in preparation or considered for future use is summarized below.

After the first space-based 16T FTS developed by the Block Engineering Inc. and the IRIS series of long-duration instruments developed for use on the NIMBUS satellites, the use of infrared high-resolution FTSs to monitor the atmospheric limb was pioneered by the **ATMOS** (Atmospheric Trace MOlecule Spectroscopy) experiment [28], [29], [30]. ATMOS flew four times on the NASA Space Shuttle (1985, 1991, 1992 and 1994) and recorded a remarkable series of high resolution (±50 cm optical path difference) solar occultation spectra. The ATMOS experiment provided the first comprehensive set of simultaneous measurements of the distributions of minor and trace constituents of the atmosphere and demonstrated the power of the limb absorption technique as a tool for Remote Sensing from space.

On board the ADEOS (ADvanced Earth Observing Satellite) platform flew the Interferometric Monitor for Greenhouse gases (**IMG**) [31], a Michelson-type Fourier Transform Spectrometer developed by NASDA (National Space Development Agency of Japan). ADEOS was launched in 1996 and unexpectedly ended its operation about 8 months after because of the destruction of the solar paddle. IMG was a nadir looking sensor that measured the up-welling infrared radiation from the Earth (714–3030 cm^{-1}) with 0.1 cm^{-1} spectral resolution. The mission provided substantial information on the spatial distribution (horizontal and vertical) of greenhouse gases.

CRISTA (CRyogenic Infrared Spectrometers & Telescopes for the Atmosphere) is a limb-scanning satellite experiment, designed and developed to measure infrared (140–2500 cm^{-1}) emissions of the Earth's atmosphere [32]. The vertical resolution is about 2–3 km and the spectral resolution is 2 cm^{-1}. CRISTA is mounted on the ASTRO-SPAS satellite. It has successfully completed two missions: CRISTA-1 in 1994 and CRISTA-2 in 1997.

Currently in orbit there are other high-resolution FTS instruments: the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on the European Space Agency's ENVISAT-1 satellite, the Atmospheric Chemistry Experiment (ACE) developed by the Canadian Space Agency, and the Tropospheric Emission Spectrometer (TES) on NASA's Aura satellite.

MIPAS [33] is a high-resolution limb sounding Fourier Transform spectrometer, on board the ENVIromental SATellite (ENVISAT-1) successfully launched on March 2002. MIPAS measures atmospheric trace gases from the upper troposphere to the thermosphere and performs limb sounding observations in the middle infrared region (685–2410 cm^{-1}) with a spectral resolution of 0.025 cm^{-1}.

ACE-FTS [34] is the main instrument on the SCISAT-1 satellite launched in August 2003. It is a high spectral resolution (0.02 cm^{-1}) Fourier Transform Spectrometer operating in solar occultation mode in the mid-IR (750–4400 cm^{-1}). The primary goals of the ACE mission are to understand the chemical and the dynamic processes that control the distribution of ozone in the stratosphere and upper troposphere, particularly in the Arctic, to explore the relationship between atmospheric chemistry and climate change, to study the effects of biomass burning in the free troposphere, and to measure aerosols and clouds to reduce uncertainties in their effects on the global energy balance.

TES [35] has been launched on the EOS (Earth Observation System) Aura satellite in July 2004. It is a high-resolution infrared-imaging Fourier Transform spectrometer primarily aimed at investigating tropospheric chemistry including troposphere-biosphere interactions and troposphere-stratosphere exchange on global, regional and local scales. TES employs both the natural thermal emission of surface and atmosphere and reflected sunlight, thereby providing day-night coverage anywhere on the globe. It operates in a combination of limb and nadir modes, with spectral resolution of 0.1 cm^{-1} and 0.025 cm^{-1} respectively over the spectral range from 600 to 3450 cm^{-1}.

Additional FTS space-based instruments are planned for the near future. The ATmospheric RAdiation Spectrometer (**ATRAS**) is a proposed follow-on instrument of the ADEOS-IMG. Similar to the IMG, ATRAS will observe detailed infrared spectra for the Japanese space program [36].

The Infrared Atmospheric Sounding Interferometer (**IASI**) is a space born instrument mounted on the METOP (METeorological OPerational) satellite [37]. This European instrument has the mission objectives of operational meteorology and world wide climate monitoring. The IASI instrument covers the spectral range between 645 and 2740 cm^{-1} with a spectral resolution of 0.25 cm^{-1}. The program includes the development of three instruments mounted on three successive satellites scheduled for 2006, 2010 and 2015.

Among the numerous instruments described, our interest will focus on MIPAS which is an important European effort, extremely productive for the high resolution sounding of the atmosphere over the coming years, providing an unprecedented global coverage. Chapter 4 will be entirely dedicated to the MIPAS instrument.

Chapter 2

Fourier Transform Spectroscopy and its Atmospheric Applications

The purpose of this chapter is to summarize significant contributions of Fourier Transform Spectroscopy (FTS) to spectroscopy in general and to atmospheric spectroscopy in particular. The reviews of historical developments are presented from the 19th century work of Michelson to the experience of present day systems.

2.1 Michelson and his interferometer

Fourier transform interferometry initiated in 1880 when Dr. A. A. Michelson invented the interferometer [38, 39] which now bears his name. Michelson used the interferometer to prove, with greater precision than anybody before him, that the velocity of light was an universal constant and the ether did not exist, and for his discovery he received the Nobel Prize.

In the last decade of the nineteenth century, Michelson made many classic observations with his interferometer. Among the many applications he studied there were some attempts at spectroscopic analysis. At the time when the origins of hyperfine structure in the spectral lines were unexplained, he showed the presence of splitting in many apparently simple spectral lines, for example the doublet nature of the red Balmer line of hydrogen and the complex structure of the mercury green line. But, because of the lack of computers and electronic detectors, he could not perform Fourier transform calculations. The technical resources enabled him to measure the position of the nodes of the interference and to work with visibility curves rather than with complete interferograms. Since, as Lord Rayleigh [40] pointed out, unique spectral distributions can not be obtained from the visibility curves alone, Michelson was therefore not able to derive a spectrum.

To compute Fourier transforms, he invented a mechanical analog computer which could handle about 80 data points and which he termed "harmonic analyzer" (shown at the Smithsonian Museum in Washington). Nevertheless for some reason, he did not use the analog computer for **F**ourier **T**ransform **S**pectroscopy (**FTS**).

2.1.1 Michelson interferometer

Figure 2.1 is an optical diagram of the basic Michelson interferometer. The term "Michelson interferometer" is often used to refer to the particular two-beam arrangement with two plane mirrors [41, 42].

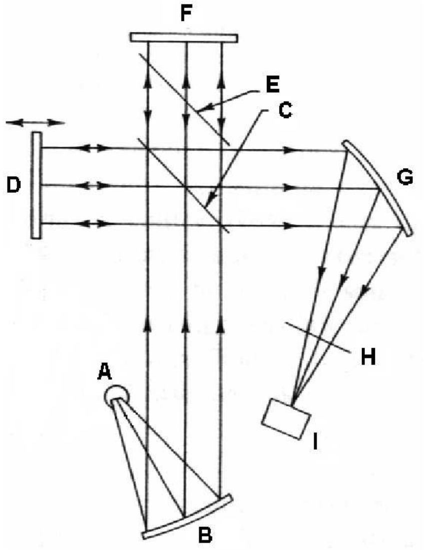

Figure 2.1: The two-beam interferometer invented by Michelson. A: source, B: collimator, C: beam-splitter, D: movable mirror, E: compensator, F: fixed mirror, G: focusing mirror, H: spectral filter, I: detector.

To understand the mechanics of the interferometer, let us start with the source, which is usually chosen according to the desired spectral region - mercury lamps for the far-infrared, glowers for the near-infrared, and a variety of lamps in the visible. Afterwards the radiation beam is collimated by a mirror or a lens, it is amplitude-divided at the beam-splitter. One part of the radiation beam goes through a compensator to a fixed mirror, which reflects the beam back through the compensator. The beam is then reflected by the beam-splitter toward a focusing mirror. The other part of the radiation beam is reflected by the beam-splitter, goes to and returns from a movable mirror, is transmitted through the beam-splitter, and goes to the focusing mirror. The compensator is sometimes introduced to keep equal the optical properties in the two arms. The focusing mirror merely focuses the recombined radiation beams on the detector. The path difference is twice the arm displacement from the balanced position of the movable mirror, that is called "origin" or "**Z**ero **P**ath **D**ifference" (**ZPD**) position. The detected signal as a function of the optical path difference is the *interferogram.*

A monochromatic source yields a cosine variation in the flux of the combined beams at the detector. The period of the cosine function is uniquely determined by the wavelength and the optical path difference for the radiation beams in the two arms. For a source of many frequencies, the interferogram is the sum of the fluxes of each wavelength pattern. Fourier analysis enables to convert the interferogram into a spectrum, i.e., signal versus frequency. That is, Fourier analysis of the interferogram picks out the pattern for each frequency and determines the magnitude of the flux at that frequency.

2.2 The work of Rubens and Wood

The first experimental observation of a full interferogram was not made until 1911, when it was recorded with a micro-radiometer by Rubens and Wood [43].

The interferogram was obtained with an interferometer consisting of two thin crystalline quartz plates mounted with their faces parallel to one another. Their separation was varied from 0 to 85 μm and was measured by counting the number of sodium D line fringes formed in the air gap between the plates. The interferogram was observed by selecting a path difference, detecting the radiant flux with a micro-radiometer, and then moving to the next path difference and repeating the procedure. The spectra were not obtained by a direct transformation of the interferogram, but by a trial and error method. The weakness of Rubens and Wood's method lays, like Michelson's method, in investigating the experimental results by comparing guessed plausible spectral distributions with the observed interferograms.

Because of the lack of adequate analytical methods there was a decline of interest in interference spectroscopy until the 1950s.

2.3 Foundations of Fourier Transform Spectroscopy

The next sketch of the history brings to the discussing subject that might be called the "renaissance" of the Michelson interferometer. Most of the foundations of Fourier spectroscopy were worked out in the period from 1951 to 1961 after two key contributions, namely, the discovery of the throughput or *étendue* advantage by Jacquinot [44] and of the so-called multiplex gain of FTS by Fellgett [45].

2.3.1 Jacquinot advantage

In spite of the use by Michelson and Rubens of the interferometer as a spectroscopic tool, one of its major advantages did not appear to have been noticed by these authors. Jacquinot pointed out in 1954 [46, 47] that an interferometer, being an instrument possessing circular symmetry, has an angular admission advantage over conventional grating spectrometers, which employ slits and consequently have not such symmetry.

All spectrometers must have a limited angular acceptance if they have non-zero resolving power. Jacquinot showed that when prism, grating, and interference spectrometers (e.g. Fabry-Perot, Amplitude and Polarization interferometers) were compared at *equal resolving power* and at *equal instrument aperture*, the radiant throughput of the interference spectrometer was much higher than that of the grating spectrometer (assuming that the instruments have the same entrance area).

For the interferometer, Jacquinot pointed the attention on the fact that the product of the infinitesimal area of any optical elements and the infinitesimal solid angle subtended by the source is a constant for the instrument from the source to the detector. This product was called the *étendue* or throughput.

Jacquinot also emphasized that the Fourier method of interferometric spectrometry combines this advantage with the advantage of wide spectral range, which is lacking in the classical interferometric method of Fabry and Perot [48, 49]. Jacquinot not only brought the throughput advantage to the attention of the scientific community, but he trained many fine students, among whom Drs. Pierre and Janine Connes.

2.3.2 Fellgett advantage

Fellgett put forward his ideas in 1951 in his doctoral thesis [50] and was the first person who performed a numerical transformation of an interferogram[51]. Fellgett subsequently constructed a Michelson-type interferometer with a resolving power of about 60 in the region 4000 to 8000 cm^{-1}.

Using a single detector, in an interferometer the whole of the spectral band is observed for the whole of the duration of the experiment, whereas in a grating (or any dispersive) spectrometer the spectral elements are observed sequentially for short periods which add up to give the total time of the experiment. Thus, the interferometer receives information about the *entire* spectral range during the *entire* scan, while the grating instrument receives information only in a narrow band at a given time. Fellgett gave the name *multiplex spectrometry* to the spectroscopic techniques in which all the spectral elements are simultaneously observed.

The signal-to-noise obtained in the recovered spectrum is a factor of primary importance which determines the quality of the spectroscopic measurement. The multiplex gain is a salient feature of Fourier spectroscopy, but the multiplex and the Fourier attributes are distinct and independent.

This multiplex advantage can be abated or even become a disadvantage, depending on the type of noise that perturbs the recorded signal. This advantage is gained only in the case when the noise is independent of the signal power (e.g. detector noise). Whenever the noise is proportional to the square root of the source intensity (e.g. photon noise), the multiplex advantage is just balanced by the increased noise (if all the spectral elements are of comparable intensity) and there is no net gain. Whereas, when the noise is linearly proportional to the source intensity (e.g. scintillation noise), multiplexing is transformed into a disadvantage when comparison is made with sequential measurements.

2.3.3 Other pioneering Fourier Transform Spectroscopists

The throughput and multiplex advantages created a new interest in Michelson's interferometer especially for infrared spectroscopy, and in the middle 1950s, several laboratories started to exploit this technique. The usefulness of the FTS method in the visible range was lessened somewhat by the frequent loss of the Fellgett advantage and by the availability of multi-detectors (photographic slabs).

Mertz devised an interferometer in the visible region of the spectrum [52]. The interferograms obtained were not Fourier transformed, the hope being that an empirical classification of star types could be made without such procedure. In 1954 Mertz also devised a polarization interferometer for the far infrared, but published details of this instrument much later [53]. Again, however, the interferograms were not transformed.

In the latter half of the 1950s, a number of workers took up interference spectroscopy using a variety of approaches. Strong and co-workers [54] constructed a lamellar grating interferometer for the far infrared. In 1956, Gebbie and Vanasse [55] published the first digitally computed far infrared spectrum showing water vapour absorption lines that were in satisfactory agreement with theory.

Further results followed with the application of the technique to absorption spectroscopy and to the observation of the far infrared solar spectrum. The Michelson-type interferometer, having a dielectric film to divide the beams now most widely used, was first described and built in 1959 by Gebbie [56]. Figure 2.2 shows the first spectra obtained with such a device with which Gebbie started sub-millimetre atmospheric spectroscopy [57].

Concurrent with these long-wave developments, progress was made with astronomical applications in the near infrared. J. Connes [55] obtained some preliminary observations of interferograms in 1958 and two years later published, with Gush [58, 59], interferometric spectra of the night sky in the region of 6000 cm^{-1} to 1040 cm^{-1}, with a resolving power of 1000. This improved resolution enabled the vibrational structure of the OH radical to be clearly resolved for the first time.

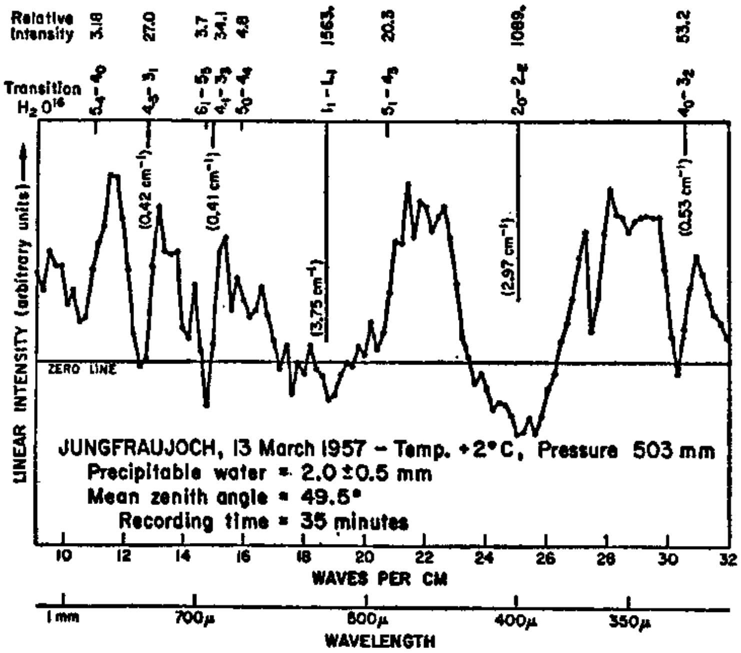

Figure 2.2: Historical experiment by Gebbie which shows the sub-millimetre spectrum of solar radiation modified by atmospheric transmission and absorption. The calculated positions and intensities of pure rotation lines of H_2O are given. The numbers in parentheses are calculated half-widths for selected lines.

2.4 Practical application of the fundamental principles

The era from 1951 to 1961, which could be called the "exploration period of FTS", ended with the thesis of J. Connes [60, 61, 62] in which all the main mathematical and technical aspects of FTS were comprehensively discussed. It remained to the following elaboration period to work for its practical use.

One of the major development of the 1960's was in the area of data handling and computational techniques and the corresponding revolution of computers. A disadvantage of FTS was that, to obtain a given resolution, the number of the interferogram data points to which a numerical transformation is applied increased in direct proportion to the resolving power. It is not coincidental that the resolving power of the measurements has been improved in parallel with the progress made in digital computer technology. There was a feature of electronic digital computers that was entirely different from mechanical devices [63, 64, 65]: that the scale of operations could be extended indefinitely by increasing the time taken to complete the task. This was the essential difference between a digital and an analog computer, and was central to the advantages of getting spectra by a Michelson interferometer with digital Fourier transformation. The advent of digital computer was a crucial ingredient at that time, but prior to 1966 the Fourier transformation was an enormously difficult task for an interferogram having many data points, since the classical algorithm required extremely time-consuming computations.

A major breakthrough was the discovery of the so-called "Fast Fourier Transform" (FFT) algorithm by Cooley and Turkey [66] in 1965. The number of arithmetic operations for this "Fast Fourier Transform" algorithm is reduced to $2Nln_2N$ compared to N^2 for the classical "Discrete Fourier Transform" algorithm; thus the Cooley-Turkey algorithm, without any approximation, reduces the computation time of Fourier transforms by several orders of magnitude. Moreover, at the *1970 International Conference on Fourier Spectroscopy in Aspen, Colorado*, J. Connes [62] presented several mathematical techniques for further reducing computation times. She noted the advantage of recording one-sided interferograms, by exploiting the fact that the interferogram is a real and symmetric function, and showed that further savings in time of up to a factor of four were possible.

Although Fourier transform spectrometers were the logical extension of Michelson's pioneering research, these instruments did not become really useful until the development of high speed electronic data processing systems. The introduction of minicomputers [67] has greatly improved the capabilities of the Fourier spectrometers. These minicomputers have made real-time analysis possible. In the 1960's, the Beckman Corporation, Grubb-Parsons, Digilab Inc., Coderg, and Idealab marketed Fourier transform spectrometers. All of these companies produced instruments which covered the range from 10 cm^{-1} (or lower) to several hundred wavenumbers.

2.4.1 Application to atmospheric monitoring

An initial non-laboratory application for FTS was atmospheric radiation measurements from balloon-borne platforms [68]. Since then, FTS instruments have been extensively used for balloon, aircraft and rocket-borne atmospheric research. The rapid progress of satellite technology stimulated interest in Remote Sensing of atmospheric parameters as a global sounding. In September 1962, the era of Remote Sensing measurements using space-borne Fourier transform systems began with the flight of the *Block Engineering 16T* on board the Discoverer satellite [69]. Until that time it had been thought that FTS instruments were too fragile and the optical alignment required could not be maintained through launch and into orbit.

In the 1960's many workers made a variety of astronomical measurements. Rotational spectra of gases on Venus, the spectra of CO_2 on Mars, the ice-water phase of the "frost" particles in the rings of Saturn and spectra of the Earth's atmosphere from satellite were observed. Pierre and Janine Connes with J. P. Maillard measured excellent spectra from the planets and published a great deal of the data in "Atlas des Spectres dans le Proche Infrarouge de Venus, Mars, Jupiter et Saturn" [70].

R. Hanel and his colleagues at the Goddard Space Flight Canter (NASA) packaged a

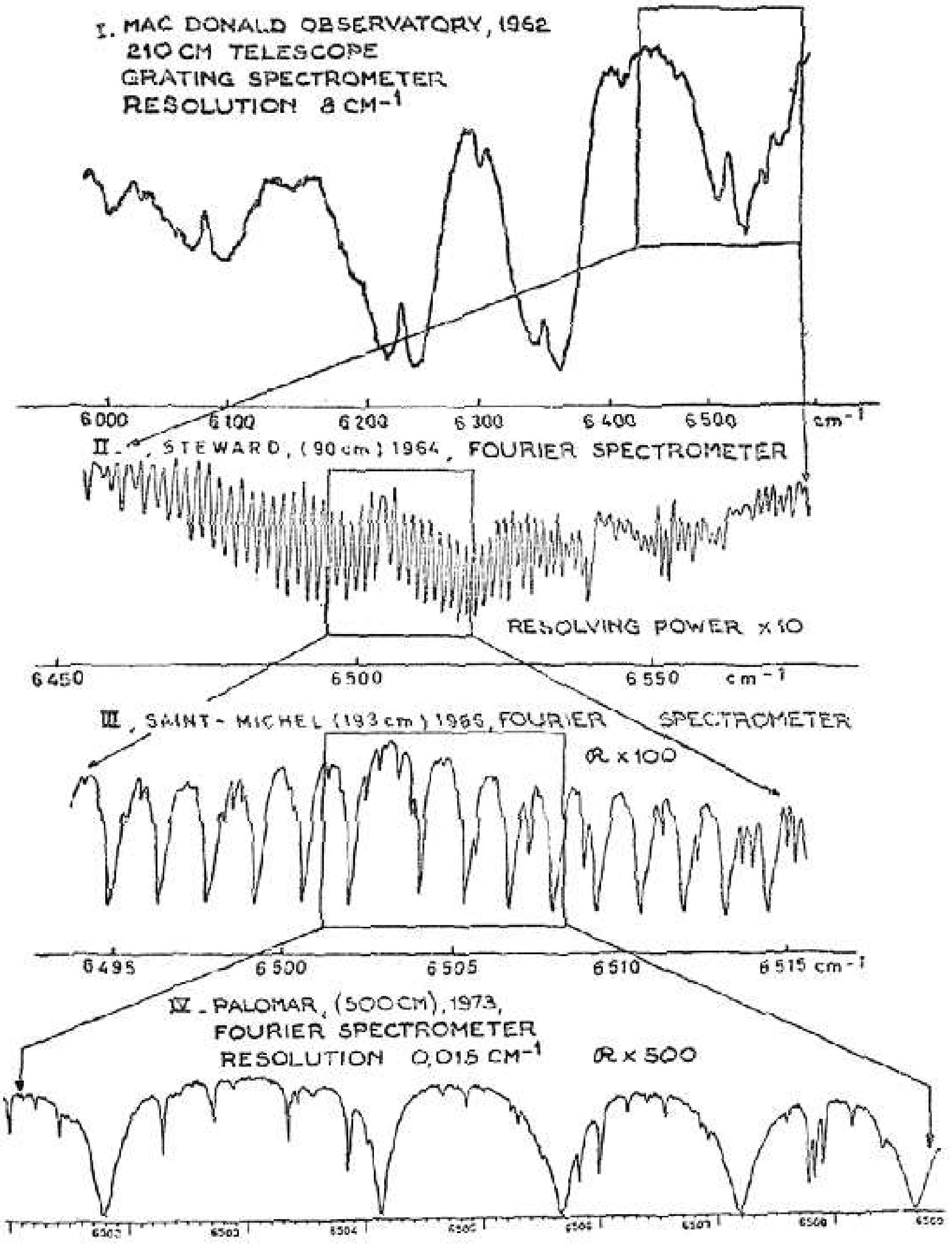

Figure 2.3: Improvements in the near-IR Venus spectrum due to Fourier spectroscopy; same type detectors (cooled PbS) with almost the same NEP used throughout. Curve I by Kuiper (1962); Curve II from Connes and Connes (1966); Curve III from Larson and Fink (1975); and Curve IV from P. Connes and Michel (1975).

Michelson interferometer (**IRIS**-B, **I**nfra**R**ed **I**nterferometer **S**pectrometer) for the Nimbus III satellite (launched in 1969). They obtained data for about three and one-half months, until the thermistor bolometer failed, from different parts of the Earth [71, 72]. The IRIS series represented the state of the art for that time. IRIS-B validated the use of FTS for remote atmospheric temperature sounding while IRIS-D (packaged for the Nimbus IV satellite launched in 1970) provided similar data operating continuously for a year [73]. An advanced version of the instrument (IRIS M) was flown on the Mariner 9 orbiter in 1971/72, allowing investigation of the infrared spectrum of Mars [74]. IRIS Voyager 2 (launched in 1977), which conducted a planetary tour, survived for over 12 years in space prior to leaving the solar system [75]. The fact that it functioned for so long after fabrication demonstrated that thermo-mechanical stability was a manageable FTS design issue. These space ventures have demonstrated the power of Remote Sensing with Fourier transform spectrometers. The wide spectral range at moderately high spectral resolution and the reliability achieved by these instruments have permitted scientific investigations which would otherwise have been impossible.

In the same years J. E. Chamberlain, J. E. Gibbs and H. A. Gebbie at National Physical Laboratory (NPL) in England started on developing Michelson interferometer equipment for molecular spectroscopy and atmospheric studies. One notable contribution was their work on asymmetric Fourier spectroscopy [76, 77], simultaneously developed by E. E. Bell and R. Sanderson at Ohio State University [78, 79], to obtain the complex indices of refraction of gases.

At the Air Force Cambridge Research Laboratory, E. V. Loewenstein, G. Vanasse, H. Sakai, R. Murphy and A. T. Stair Jr. made several notable contributions to theoretical and experimental Fourier transform spectroscopy. They studied signal-to-noise ratio problems and mathematical transforms in different mathematical spaces, i.e. Hilbert space, and improved the theory of measurements in Fourier transform spectroscopy [80]. They made a number of measurements on atmospheric gases in the laboratory and from airplanes, and studied molecular spectra, including relaxation processes which are associated with environmental problems - natural and man-made [81].

In 1975 P. Connes showed the fast development of higher and higher resolution in FTS astronomical observations. The observed spectra, fig. 2.3, of the Venus atmosphere over the last eleven years prove that the resolution has improved from 8 cm^{-1} in 1962 with gratings to 0.015 cm^{-1} in 1973 with FTS (more than 500 times). Four strong CO_2 Venusian bands are shown in I by Kuiper (1962) [82]; the rotational structure is resolved in II by Connes and Connes (1966) [83]; III shows lines from much weaker overlapping bands by Arson and Fink (1975) [84]; IV gives a good approximation of the true line profile by P. Connes and Michel (1975) [25].

Chapter 3

Inverse Methods

Indirect measurements are used whenever direct measurements are difficult or expensive, although indirect measurements often bring complex problems of interpretation. The measured quantity contains some information about the required quantities and the problem of trying to reconstruct these quantities is known as the Inverse Problem. There exists a large amount of literature on this subject, both from the mathematical point of view and applied to specific scientific problems. A very good textbook for Remote Sensing retrievals is [85]. The notation and concepts introduced in this chapter are inspired to those described there.

3.1 The problem

In accordance with convention, the collection of values to be retrieved is referred to as the state of the atmosphere that we usually denote with x. The *forward problem* is the mapping from the state of the atmosphere to the quantities that we are able to measure. The details of the forward problem is in our case given by the physical theory of the atmosphere. The forward mapping may be linear or non-linear and is denoted by F. In practice we are never able to make exact measurements and the data that we actually measure are a corrupted version of the error-free data obtained through the forward process from the state of the atmosphere. The difference between error-free data and the measured data, denoted by y, is called the *measurement noise* and is denoted by ε. Thus the mapping from the state of the atmosphere to actual data is given by the relation: $y = F(x) + \varepsilon$. The *inverse problem* is then the problem of finding the original state (or the quantities to be retrieved) x of the atmosphere given the measured data y and the knowledge of the forward problem F.

3.1.1 Ill-posed problems

There are several classifications of the forward problem depending on whether the state of the atmosphere and measured data are functions of a continuous or discrete variable, i.e. have infinite-dimensions or finite-dimensions. In most inverse problems the quantities to be retrieved are often functions of continuous variables such as time or (physical) space, so that the dimension of the state space is infinite. On the other hand, only a limited number of data can be measured, so that data space has finite dimensions. Thus most inverse problems are formally ill-posed. According to Hadamard (1923), the inverse problem of solving $F(x) = y$ for x given y is called ill-posed if it meets one or more of the following three conditions:

i the inverse of the forward operator F does not exist;

ii the inverse is not *unique*;

iii an arbitrary small change in the measured data can cause an arbitrary large change in the retrieved quantities.

In the case of a well-posed problem, relative error propagation from the measured data to the solution is controlled by the *conditional number*:

$$\frac{\|\Delta x\|}{\|x\|} \leq cond(F)\frac{\|\Delta y\|}{\|y\|}$$

where Δy is the variation of the measurements y and Δx the corresponding variation of the parameters x. Since the fraction error in the retrieved parameters depends on the conditional number multiplied by the fractional error in the measurements, small values of the conditional number are desirable. If $cond(F)$ is not too much greater than unity, the problem is said to be *well-conditioned* and the solution is stable with respect to small variations of the measurements. Otherwise the problem is said to be *ill-conditioned*. The separation between well-conditioned and ill-conditioned problems is not very sharp and the concept of "well-conditioned" problem is more vague than the concept of "well-posed".

3.2 Inverse problems

As discussed previously, there are problems in which the dimension of the state space is infinite and in this case the parameters can not be determined by the measurements because there exists an infinite number of solutions which satisfy the measurements. At this point, it is convenient to express the continuous function with a representation in terms of a finite number of parameters. Thus the mapping from the state vector $\mathbf{x}$ (with n elements) to the *measurement vector* $\mathbf{y}$ (with m elements) may be written as:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \varepsilon, \tag{3.1}$$

where we indicate with bold characters the vector quantities. After discretisation, the problem may either be over-constrained ($m > n$) or under-constrained ($m < n$). It is usually important to appreciate the degree of *linearity* of any given inverse problem (as discussed in Sect. 3.2.2). The near linear nature of any inverse problems has allowed the development of appropriate inverse methods based on linear theory.

If the non linearities are significant, a liberalization of the forward model about some reference state $\mathbf{x}_0$ is often an adequate approximation, and we obtain the following expression:

$$\mathbf{y} - \mathbf{F}(\mathbf{x}_0) = \frac{\partial \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}}(\mathbf{x} - \mathbf{x}_0) + \varepsilon = \mathbf{K}(\mathbf{x} - \mathbf{x}_0) + \varepsilon$$

which defines the $m \times n$ *weighting function matrix* $\mathbf{K}$, not necessarily square, in which each element is the partial derivative of a forward model element with respect to a state vector element, i.e. $K_{ij} = \partial F_i(x)/\partial x_j$. The term *weighting function* is peculiar to the atmospheric remote sounding literature, but it may also be called the Jacobian (it is a matrix of derivatives) or the kernel (hence $\mathbf{K}$).

3.2.1 Linear inverse problems

Let us consider first a linear problem in the absence of measurement errors. In this case the problem reduces itself to the solution of linear simultaneous equations:

$$\mathbf{y} = \mathbf{K}\ \mathbf{x}$$

and can have no solutions, one solution or an infinite number of solutions. The m weighting function vectors $\mathbf{K}_j$ will span some subspace of state space which will be of dimension not greater than m and may be less than m if the vectors are not linearly independent. The dimension of this subspace is known as the *rank* of the matrix $\mathbf{K}$, denoted by p, and is equal to the number of linearly independent rows (or columns). If $m < n$ the problem is under-constrained (and ill-posed) because the number of unknowns exceeds the number of simultaneous equations and the parameters can not be determined from the measurements. We can make the problem well determined by reducing the number of unknowns. If $m = n = p$, then $\mathbf{K}$ is square and in this case a unique solution can be found and is called the *exact solution*:

$$\mathbf{x}_e = \mathbf{K}^{-1}\ \mathbf{y}. \tag{3.2}$$

However, in the presence of noise, if the problem is ill-conditioned this solution may be unsatisfactory.

If $m > n$ the problem is described as over-constrained. In this case, the rank of $\mathbf{K}$ can be equal to or less than the number of unknowns, n. If $p < n$, this means that the measurements are "blind" to certain aspects of the unknowns: those components of $\mathbf{x}$ along the first p orthogonal base vectors of the state space are determined by the measured data, while the data tell nothing about all the components of x along the remaining n-p base vectors. This under-determined part of state space is called *null space* of $\mathbf{K}$.

For $m > p = n$ there is not a solution that can fit all the measurements and we have to use some criterion, such as least squares, to select one acceptable solution. In the least square method we look for a solution that minimizes the sum of the squares of the differences between the measurements and the forward model calculations made using the solution; these differences are called *residuals*, the sum of the squares is called the residual norm or χ^2. That is, we minimize:

$$(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}})^T(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}}) = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{K}\hat{\mathbf{x}} - \hat{\mathbf{x}}^T\mathbf{K}^T\mathbf{y} + \hat{\mathbf{x}}^T\mathbf{K}^T\mathbf{K}\hat{\mathbf{x}} \tag{3.3}$$

Equating to zero the derivative with respect to $\mathbf{x}$:

$$\frac{\partial}{\partial\hat{\mathbf{x}}}(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}})^T(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}}) = 2\mathbf{K}^T(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}}) = 0$$

$$\mathbf{K}^T\mathbf{K}\hat{\mathbf{x}} = \mathbf{K}^T\mathbf{y}.$$

These are known as the "normal equations" of the least squares problem. Since $\mathbf{K}^T\mathbf{K}$ is an $n \times n$ matrix, it is invertible if $p = n$. If the matrix is invertible, then we obtain a unique solution and find the best fit parameters $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = (\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\hat{\mathbf{y}}. \tag{3.4}$$

The matrix $(\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T$ is also known as the Moore-Penrose inverse of $\mathbf{K}$. If the rank p of $\mathbf{K}$ is less than n, there are an infinite number of solutions, all of which minimize the square of the residual norm.

All real measurements are subject to experimental error or noise, so that any practical retrieval must allow for this. For a proper treatment of experimental error we need a formalism in which to express uncertainty in measurements and the resulting uncertainty in retrievals, and with which to ensure that the latter is as small as possible. A good approximation for experimental error is to describe our knowledge of the true value of the measured parameter by a Gaussian or "normal" distribution P(y) with a mean ȳ and variance σ^2. When the measured quantity is a vector, as in our case, different elements of the vector may be correlated, in the sense that:

$$S_{ij} = \varepsilon\{(y_i - \overline{y}_i)(y_j - \overline{y}_j)\} \neq 0$$

where S_{ij} is called *covariance* of y_i and y_j , and ε is the expected value operator. These covariances can be assembled in a matrix, which we will denote by $\mathbf{S}_y$ for the covariance matrix of $\mathbf{y}$. Its diagonal elements are the variances of the individual elements of $\mathbf{y}$. A covariance matrix is symmetric and non-negative definite. The Gaussian distribution for a vector is of the form:

$$P(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{S}_y|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\mathbf{y}-\hat{\mathbf{y}})^T \mathbf{S}_y^{-1}(\mathbf{y}-\hat{\mathbf{y}})} \tag{3.5}$$

where $\mathbf{S}_y$ must be non singular.

In presence of Gaussian distributed noise with zero mean and covariance matrix $\mathbf{S}_\varepsilon$, the weighted least squares are used and the quantity to be minimized is, instead of expression (3.3):

$$\chi^2 = (\mathbf{y} - \mathbf{K}\hat{\mathbf{x}})^T \mathbf{S}_\varepsilon^{-1}(\mathbf{y} - \mathbf{K}\hat{\mathbf{x}}), \tag{3.6}$$

and equating to 0 the derivative of (3.6) with respect to $\hat{\mathbf{x}}$, gives:

$$\hat{\mathbf{x}} = (\mathbf{K}^T \mathbf{S}_\varepsilon^{-1}\mathbf{K})^{-1}\mathbf{K}^T \mathbf{S}_\varepsilon^{-1}\mathbf{y}.$$

The expectation value of the expression (3.6) is $m - n$. Therefore a quantity called χ^2-test, or simply χ-test, can be defined as:

$$\chi - test = \frac{\chi^2}{m - n} \tag{3.7}$$

and has an expectation value of 1. Therefore the deviation of χ-test from unity provides a good estimate of the agreement between the model and the observations.

Since ill-conditioned nature of the inverse problems, the mathematical solution often gives results that are unacceptable in the sense that they do not agree with our understanding and preliminary knowledge of the measured quantity. If this is the case, rather than looking for a measurement of the true state we must look instead for an *estimate* of the true state which is acceptably accurate or the best estimate in some statistical sense. In literature there are many statistical methods or probability techniques that tell how to combine the measurements with other information in order to select from the all possible solutions the best one.

A very powerful tool in probability theory is the Bayesian approach [86], in which we may have some prior understanding or expectation about some quantity and we want to update the understanding in the light of the new information.

Bayes' theorem

The Bayes' theorem shows the relationship between the conditional probability of **y** given **x**, P(**y**|**x**), and the conditional probability of **x** given **y**, P(**x**|**y**). For the vector case the Bayes' theorem states that:

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{x}|\mathbf{y})P(\mathbf{x})}{P(\mathbf{y})}, \tag{3.8}$$

where the left hand side of Eq. (3.8), P(**x**|**y**), is the quantity that we obtain when we update the *a priori* knowledge P(**x**) of the state with the measurement **y**. P(**x**|**y**) requires the knowledge of the forward model and the statistical description of the measurement error. The denominator, P(**y**), can be determined by normalization. Note that the Bayesian view is general. Given a measurement together with a description of its error statistics, a forward model describing the relation between the measurement and the unknown state, and any *a priori* information that might be available, it allows us to identify the class of possible states that are consistent with the available information, and to assign a probability density to them.

The linear problem with Gaussian statistics

As a simple example of Bayesian approach consider a linear problem in which all of the probability density functions *pdf* are Gaussian. The Gaussian distribution for a random vector **x** is obtained from Eq. (3.5). The maximum probability value for **x** is equal to the expected value $\hat{\mathbf{x}}$, because the *pdf* is symmetric about $\mathbf{x} = \hat{\mathbf{x}}$.

Anticipating that the linear problem has a linear solution, we expect that the linear solution is of the general form $\hat{\mathbf{x}} = \mathbf{x}_0 + \mathbf{G}\mathbf{y}$, where **G** is a $n \times m$ retrieval gain matrix and $\mathbf{x}_0$ is some constant offset. **G** has various names and in atmospheric literature is often called *contribution function* matrix. If there is an error ε in the measurement, then there will obviously be a corresponding error $\mathbf{G}_\varepsilon$ in the solution and thus the size of the gain provided by the matrix **G** gives a measurement of the ill-conditioning of the solution.

It is useful to use the scalar quantity (a *cost function*), -ln{P(**y**|**x**)} + constant, that is the logarithm of the Gaussian distribution (see Eq. (3.5)), to express P(**y**|**x**) as:

$$-2\ln P(\mathbf{y}|\mathbf{x}) = (\mathbf{y} - \mathbf{K}\mathbf{x})^T \mathbf{S}_\varepsilon^{-1} (\mathbf{y} - \mathbf{K}\mathbf{x}) + c_1, \tag{3.9}$$

where c_1 is a constant and $\mathbf{S}_\varepsilon$ is the measurement error covariance. Similarly, we can describe the *a priori* knowledge of **x** by a Gaussian *pdf*:

$$-2\ln P(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) + c_2, \tag{3.10}$$

where $\mathbf{x}_a$ is the *a priori* value of **x** and $\mathbf{S}_a$ is the associated covariance matrix.

Substituting Eqs. (3.9) and (3.10) in Eq. (3.8) we obtain for the a posteriori *pdf*:

$$-2\ln P(\mathbf{x}|\mathbf{y}) = (\mathbf{y}-\mathbf{K}\mathbf{x})^T\mathbf{S}_\varepsilon^{-1}(\mathbf{y}-\mathbf{K}\mathbf{x}) + (\mathbf{x}-\mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x}-\mathbf{x}_a) + c_3, \tag{3.11}$$

where c_3 is a constant that includes the cost function of the measurements (-ln P($\mathbf{y}$)). Since P($\mathbf{x}|\mathbf{y}$) is also a quadratic form in $\mathbf{x}$, it is possible to write it as a Gaussian distribution with expected value $\hat{\mathbf{x}}$ and covariance $\hat{\mathbf{S}}$:

$$-2\ln P(\mathbf{y}|\mathbf{x}) = (\mathbf{x}-\hat{\mathbf{x}})^T\hat{\mathbf{S}}^{-1}(\mathbf{x}-\hat{\mathbf{x}}) + c_4. \tag{3.12}$$

Equating terms in Eqs. (3.11) and (3.12) that are quadratic in $\mathbf{x}$, one obtains:

$$\mathbf{x}^T\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}\mathbf{x} + \mathbf{x}^T\mathbf{S}_a^{-1}\mathbf{x} = \mathbf{x}^T\hat{\mathbf{S}}^{-1}\mathbf{x}$$

from which it turns out that:

$$\hat{\mathbf{S}}^{-1} = \mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1}. \tag{3.13}$$

Likewise equating the term linear in $\mathbf{x}^T$ one obtains:

$$(-\mathbf{K}\mathbf{x})^T\mathbf{S}_\varepsilon^{-1}\mathbf{y} + \mathbf{x}^T\mathbf{S}_a^{-1}(-\mathbf{x}_a) = \mathbf{x}^T\hat{\mathbf{S}}^{-1}\hat{\mathbf{x}}.$$

Canceling the $\mathbf{x}^T$'s, because this expression must be valid for any value of $\mathbf{x}$, and substituting for $\hat{\mathbf{S}}^{-1}$ from Eq. (3.13) one obtains:

$$\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{y} + \mathbf{S}_a^{-1}\mathbf{x}_a = (\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1})\hat{\mathbf{x}}$$

and hence:

$$\hat{\mathbf{x}} = (\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1})^{-1}(\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{y} + \mathbf{S}_a^{-1}\mathbf{x}_a). \tag{3.14}$$

Note that if we assume that the *a priori* knowledge of $\mathbf{x}$ is uniform, i.e. P($\mathbf{x}$)=C, where C is a constant, $\mathbf{S}_a^{-1}$ is equal to zero and the expected value $\hat{\mathbf{x}}$ and the corresponding covariance $\hat{\mathbf{S}}$ become:

$$\hat{\mathbf{x}} = (\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{y} \quad \text{and} \quad \hat{\mathbf{S}}^{-1} = \mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}.$$

This is the weighted least squares solution that is similar to the solution expressed by Eq. (3.4) in which $\hat{\mathbf{S}}^{-1}$ is used instead of the Moore-Penrose inverse matrix. For the well-posed problem, an exact solution is possible, so that there exists a matrix $\mathbf{G}$ such that $\mathbf{KG} = \mathbf{I}_m$, the identity matrix (e.g. we could choose $\mathbf{G} = \mathbf{K}^T(\mathbf{KK}^T)^{-1}$). If we insert $\mathbf{KG}$ before $\mathbf{y}$ in Eq. (3.14), we obtain:

$$\hat{\mathbf{x}} = (\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1})^{-1}(\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}(\mathbf{Gy}) + \mathbf{S}_a^{-1}\mathbf{x}_a).$$

We can see that this represents a weighted mean of the *a priori* $\mathbf{x}_a$ and *any* exact retrieval $\mathbf{x}_e = \mathbf{Gy}$ with matrix weights $\mathbf{S}_a^{-1}$ and $\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}$ respectively.

3.2.2 Non-linear inverse problems

A non-linear problem may be thought simply as a problem in which the forward model is non-linear, but there may be non-quadratic terms due to *a priori* constraint that would lead to a non-linear problem even if the forward model were linear. Any non-Gaussian probability density function *pdf* as *a priori* information will lead to a non-linear problem. We can make a qualitative classification of the linearity of inverse problems as follows:

- *Linear*: when the forward model can be put in the form $\mathbf{y} = \mathbf{Kx}$ and any *a priori* is Gaussian; very few practical problems are truly linear.
- *Nearly linear*: problems which are non-linear, but for which a linearisation about some *a priori* state is adequate to find a solution.
- *Moderately non-linear*: problems where linearisation is adequate for the error analysis, but not for finding a solution. Many problems are of this kind.
- *Grossly non-linear*: problem which are non-linear even within the range of the errors.

Much of what has been described so far for linear problems applies directly to moderately non-linear problems when they are appropriately linearised. The main difference is that there is no general explicit expression for optimal solutions in the moderately non-linear case, as there is from linear and nearly linear problems, so that it must be found numerically or iteratively.

The forward model is now a non-linear mapping from the state space into measurement space. The inverse mapping from measurement space into state space will map the *pdf* of the measurement error into a *pdf* in state space. If the problem is not worse than moderately non-linear, and the measurement error is Gaussian, then the retrieval error will be Gaussian. In the non-linear case it may no longer be possible to write down an explicit solution that must be found numerically or iteratively. For non-linear problems we can consider either the maximum *a posteriori* approach or the equivalent least squares method. The Bayesian solution for the linear problem, Eq. (3.11), with or without the *a priori* information respectively, can be modified for an inverse problem in which the forward model is a general function of the state and the measurement error is Gaussian:

$$\begin{aligned} &[\mathbf{y}-\mathbf{F}(\mathbf{x})]^T\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x})]+[\mathbf{x}-\mathbf{x}_a]^T\mathbf{S}_a^{-1}[\mathbf{x}-\mathbf{x}_a] \\ &\qquad or \\ &[\mathbf{y}-\mathbf{F}(\mathbf{x})]^T\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x})]. \end{aligned} \tag{3.15}$$

3.3 Newton and Gauss-Newton methods

The degree of difficulty in solving a non-linear problem depends on the degree of linearity of the forward model $\mathbf{F}(\mathbf{x})$. Newtonian iteration is a straightforward numerical method for finding the zero of the gradient of a given cost function, such as Eq. (3.15). For the general vector equation $\mathbf{g}(\mathbf{x})=\mathbf{y}-\mathbf{F}(\mathbf{x})=0$, the Newton's iteration can be written:

$$\mathbf{x}_{i+1}=\mathbf{x}_i-[\nabla_x\mathbf{g}(\mathbf{x}_i)]^{-1}\mathbf{g}(\mathbf{x}_i), \tag{3.16}$$

where $\mathbf{x}_i$ is the initial guess of $\mathbf{x}$ and the inverse is the inverse of the matrix:

$$\begin{aligned} \nabla_x\mathbf{g}&=\mathbf{S}_a^{-1}+\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}-[\nabla_x\mathbf{K}^T]\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x})] \\ &\qquad or \\ \nabla_x\mathbf{g}&=\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}-[\nabla_x\mathbf{K}^T]\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x})]. \end{aligned} \tag{3.17}$$

The function $\nabla_x\mathbf{g}$ is the second derivative of the cost function, Eq. (3.15), known as the *Hessian*. The Hessian involves both the Jacobian $\mathbf{K}$, the first derivative of the forward model, and $\nabla_x\mathbf{K}^T$, the second derivative of the forward model. The latter term is a complicated object and problems for which this term can be ignored are called *small residual* in the numerical methods literature. Ignoring this term, one obtains the Gauss-Newton method:

$$\begin{aligned} \mathbf{x}_{i+1}&=\mathbf{x}_i+(\mathbf{S}_a^{-1}+\mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}_i)^{-1}[\mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}(\mathbf{y}-\mathbf{F}(\mathbf{x}_i))-\mathbf{S}_a^{-1}(\mathbf{x}_i-\mathbf{x}_a)] \\ &\qquad or \\ \mathbf{x}_{i+1}&=\mathbf{x}_i+(\mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}_i)^{-1}\mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x}_i)] \end{aligned} \tag{3.18}$$

where $\mathbf{K}_i=\mathbf{K}(\mathbf{x})|_{\mathbf{x}_i}$ and Eq. (3.18) represents the iterative solution in a non-linear problem with and without the *a priori* information.

3.3.1 Levenberg-Marquardt method

Both Newton's method and Gauss-Newton will find the minimum in one step for a cost function that is exactly quadratic in $\mathbf{x}$ (linear problem) and will get close if the cost function is nearly quadratic. However, these methods can give bad results far from the true minimum if the true solution is sufficiently far from the currently assumed solution. In these cases, the residual may even increase rather than decrease. For the non-linear least squares problem, Levenberg [87] proposed the iteration:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_i + \gamma_i \mathbf{I})^{-1} \mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i)] \tag{3.19}$$

where γ_i is chosen at each step as the value that minimises the cost function. Marquardt [88] simplified the choice of γ_i, by not searching for the best γ_i at each iteration, but by starting a new iteration step as soon as a value is found for which the cost function is reduced. An initially arbitrary value of γ is updated at each iteration. A simplified version of Marquardt's method is given by Press and al. [89]:

- if χ^2 increases as a result of a step, increase γ, do not update $\mathbf{x}_i$ and try again,
- if χ^2 decreases as a result of a step, update $\mathbf{x}_i$ and decrease γ for the next step.

The factor by which γ is increased or decreased is usually empirically determined.

3.4 Regularization methods

In some case, the *a priori* knowledge is not provided by a measurement of the unknown but by some general understanding of its shape and behavior.

We may regard solving the inverse problem as a competition between two conflicting desires: firstly the desire to minimize the residual and secondly the desire for the solution to have a small sum-square norm or to be smooth or similar to what we believe the answer should be. One way of selecting a solution from several feasible reconstructions is to introduce a second function $\Omega(\mathbf{x})$ representing our aversion to a particular reconstruction. For example, we can decide that the solution of minimum norm should be selected from the feasible set; this can be done by choosing $\Omega(\mathbf{x}) = \|\mathbf{x}\|^2$. Sometimes, we have a preference for reconstructions that are close to some default solution or *a priori* knowledge $\mathbf{x}^\infty$. This may be appropriate if we have historical information about the quantity. This can be done by choosing $\Omega(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}^\infty\|^2$. More generally, it may not be the norm of $\mathbf{x} - \mathbf{x}^\infty$ which needs to be small, but a linear operator acting on this difference. Introducing the operator $\mathbf{L}$ for this purpose, we can set:

$$\Omega(\mathbf{x}) = \|\mathbf{L}(\mathbf{x} - \mathbf{x}^\infty)\|^2 = (\mathbf{x} - \mathbf{x}^\infty)^T \mathbf{L}^T \mathbf{L} (\mathbf{x} - \mathbf{x}^\infty). \tag{3.20}$$

The matrix $\mathbf{L}$ is of size $p \times n$, where $p \leq n$ (n: dimension of unknown space, p: rank of $\mathbf{K}$). Typically, $\mathbf{L}$ is the identity matrix or a banded matrix approximation to the $(n-p)^{th}$ derivative. There are many ways of balancing the conflicting requirements of minimizing the residual and the $\Omega(\mathbf{x})$ function and these lead to a variety of regularization methods.

3.4.1 Twomey-Tikhonov regularization

The first methods applied to the retrieval problem, in which error sensitivity and constraints were considered, were published at about the same time by Twomey [90] and by Tikhonov [91]. Both methods consider the minimization of a weighted sum of the square of the residual norm and the square of the departure of the solution from some *a priori* $\mathbf{x}_a$. This is perhaps the most common and well-known of regularization schemes:

$$(\mathbf{y} - \mathbf{K}\mathbf{x})^T(\mathbf{y} - \mathbf{K}\mathbf{x}) + \lambda^2(\mathbf{x} - \mathbf{x}_a)^T\mathbf{L}^T\mathbf{L}(\mathbf{x} - \mathbf{x}_a), \tag{3.21}$$

where a whole family of solution is parameterized by the weighting factor λ^2. λ is called the *regularization parameter*. A formal solution to the problem may be found, minimizing the function (3.21):

$$\hat{\mathbf{x}} = (\mathbf{K}^T\mathbf{K} + \lambda^2\mathbf{L}^T\mathbf{L})^{-1}(\mathbf{K}^T\mathbf{y} + \lambda^2\mathbf{L}^T\mathbf{L}\mathbf{x}_a)$$

For values of the regularization parameter at one end of its range, i.e. if λ is very large, the solution is usually smoother, more similar to the *a priori* solution and less affected by noise on the data because we effectively ignore the measured data (and any noise on the data) completely. Whereas for values of this parameter at the other end, i.e. if λ is small, the solution can be very sensitive to noise as it is primarily determined by the requirement of minimizing the data residual. Of course, if λ is reduced to zero, the problem reduces to the least-squares case considered earlier.

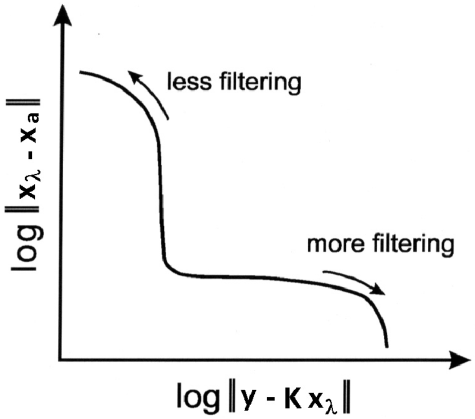

Figure 3.1: The generic form of the L-curve.

Perhaps the most convenient graphical tool for setting the regularization parameter λ is the "L-curve". When we plot $\log \|\mathbf{y} - \mathbf{K}\mathbf{x}_\lambda\|$ versus $\log \|\mathbf{x}_\lambda - \mathbf{x}_a\|$ we get the characteristic L-shaped curve with a corner separating vertical and horizontal parts of the curve (as shown in fig. 3.1). The regularization is a trade-off between the residual norm $\|\mathbf{y} - \mathbf{K}\mathbf{x}\|$ and the solution semi-norm $\mathbf{L}\|\mathbf{x} - \mathbf{x}_a\|$. In the vertical part of the curve the solution semi-norm is a very sensitive function of the regularization parameter because the solution is undergoing large changes with λ in an attempt to fit the data better. On the horizontal part, the solution is not changing by very much as λ is changed. However, the data residual is increasing sharply with increasing λ and so it is desirable to choose a solution which lies not too far from the corner.

Chapter 4

The MIPAS Experiment and the Optimized Retrieval Model

The ***M****ichelson* ***I****nterferometer for* ***P****assive* ***A****tmospheric* ***S****ounding (****MIPAS****) is a high resolution Fourier Transform-type Spectrometer that operates onboard the* ***ENVI****ronmental* ***SAT****ellite (****ENVISAT****) since March 1*st*, 2002. ENVISAT is an advanced polar-orbiting Earth observation satellite which provides measurements of atmosphere, oceans, lands, and ice over a period of five years. The ENVISAT payload consists of a set of ten instruments that operate over a wide range of the electromagnetic spectrum, from centimeter waves to the ultraviolet. MIPAS is one of the three atmospheric chemistry sensors and observes atmospheric constituents using a passive Remote Sensing technique. In this chapter the main aspects on the MIPAS experiment will be outlined. More details are available in Endemann et al. [92] or in ESA web page: http://envisat.esa.int. In the second part of this chapter the data processing scheme and the scientific code for Level 2 near real time analysis of middle infrared emission spectra measured by MIPAS experiment is presented. This code is called* ***O****ptimized* ***R****etrieval* ***M****odel (****ORM****) and was developed in an ESA-supported study "Development of an optimized algorithm for routine p, T and VMR retrievals from MIPAS limb emission spectra" (ESA contract 117171 95/NL/CN) [93]. The study was led by IROE-CNR (Firenze - Italy) in collaboration with Department of Physical Chemistry of the University of Bologna (Italy), Oxford University (United Kingdom), Leicester University (United Kingdom), IMK (Karlsruhe - Germany) and LPMA (Paris - France). Since middle infrared emission spectra are strongly sensitive to the temperature and limb measurements are strongly affected in general by the observation geometry (pressure at tangent altitudes), for a correct interpretation of the retrieval of the atmospheric constituents a good knowledge of pressure and temperature is required. Therefore the ORM has been developed to perform first the retrieval of pressure and temperature and then the retrieval of volume mixing ratio profiles of six key species (*H_2O*,* O_3 *,* HNO_3 *,* CH_4 *,* N_2O *and* NO_2*).*

4.1 MIPAS objectives

The MIPAS experiment is based on a Fourier Transform Spectrometer designed to sound the atmospheric limb emission in the middle and upper atmosphere [94]. MIPAS observes the Earth's horizon, measuring a wide spectral interval in the infrared region where many of the atmospheric trace-gases have important emission features. MIPAS observations are independent from sun position, then allowing continuous measurements, i.e. also during night or polar winter. The combination of high spectral resolution, full coverage of the mid-infrared region, high sensitivity, and full global and seasonal coverage, provide unprecedented insights into chemistry and dynamics of the atmosphere.

The primary scientific objective of MIPAS is to advance the understanding of the chemistry of the stratosphere and to provide information on species whose temporal changes affect the climate (i.e. the whole nitrogen family). In addition to stratospheric chemistry, contributions are also expected to research fields such as global climatology, atmospheric dynamics, and tropospheric chemistry.

4.2 The MIPAS instrument

4.2.1 Measurement requirements

MIPAS measures sequences of spectra at different limb heights to allow the retrieval of concentration profiles. The instrument is designed for limb-scanning observations giving maximum sensitivity and good vertical resolution [92], [94]. The instantaneous field of view (IFOV) is about 3 km high and 30 km wide at tangent altitude to collect sufficient atmospheric radiance.

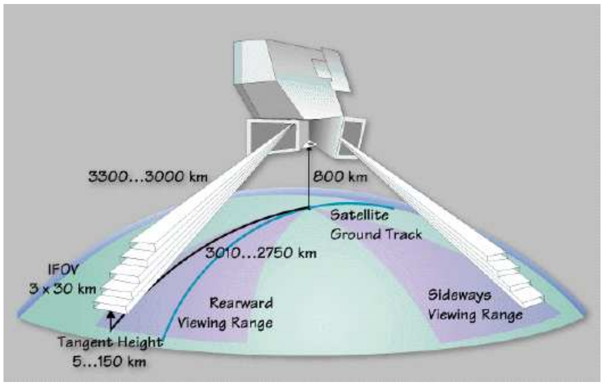

Figure 4.1: MIPAS observation geometries: the instrument is capable to perform measurements in two pointing directions.

As represented in fig. 4.1, the instrument performs measurements in two pointing modes: *rear-ward* and *sideways*. The rear-ward viewing range is used for all nominal measurements, whereas the sideways viewing range, across the orbit plane, is important for the study of specific events or atmospheric phenomena, like volcano eruptions, trace-gas concentrations above major air traffic routes, or concentration gradients. The polar orbit and the pointing capabilities of MIPAS allow a global geographical coverage, concentrated on the polar regions. A daily ground track coverage of MIPAS is shown in fig. 4.2: in the case of a nominal measurement viewing scenario the instrument performs 14 orbits with about 75 elevation scan sequences per orbit.

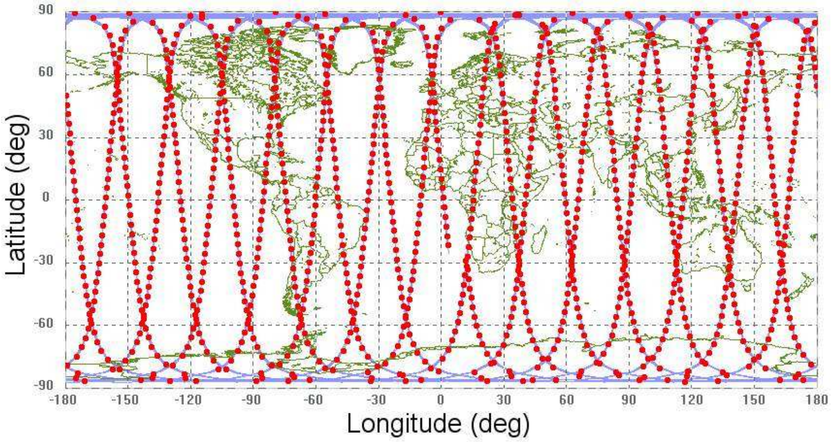

Figure 4.2: Example of the daily geographical coverage of MIPAS (14 orbits with 75 elevation scan sequences per orbit).

For most of the measuring time, MIPAS has been operated in the "nominal" observation mode, measuring consecutive backward-looking limb-scans with the lines of sight approximately lying in the orbit plane. One typical elevation limb-scanning sequence of MIPAS consists of 17 spectra and takes 75 s (a complete high resolution spectrum is measured in about 4 s). A nominal tangent elevation scan starts at 68 km height and descends in 8/5 km steps to 42 km and, from 42 km to 6 km, in 3 km steps. However MIPAS is sufficiently flexible to perform any elevation scan sequence within the height range from 5 to 150 km with variable step sizes. In the nominal observation mode the spectral resolution, as defined by the sampling theorem, is equal to 0.025 cm^{-1} and corresponds to a full width at half maximum (FWHM) of the unapodised spectrum better than 0.035 cm^{-1}.

Depending on the scientific objective, the "special" observation modes may differ from the "nominal" mode in the adopted spectral resolution, the altitude coverage, the vertical/horizontal sampling steps, or the azimuth direction of the line of sight, as summarized in table 4.1.

MIPAS spectral coverage is set from 685 to 2410 cm^{-1} (corresponding to wavelengths from 4.15 to 14.6 μm). This range covers almost the complete thermal infrared region

Observation Mode	.	Scientific Objective	Pointing direction	Coverage	Altitude Range (km)	Height Resolution (km)	Horizontal Spacing (km)
Nominal		Stratospheric chemistry and dynamics	rear	Global	6–68	3-5-8	530
Polar Winter Chemistry	S1	Polar Chemistry and Dynamics	rear	Global	7–55	2-10	420
Tropospheric Stratospheric Exchange	S2	Exchange between Stratosphere and Troposphere, Troposphere Chemistry	rear	Global	5–40	1.5-10	420
Impact of Aircraft Emissions	S3	Study of major air traffic corridor	side	Primarily North of 25° lat.	6–40	1.5-10	330
Stratospheric Dynamics	S4	Small scale structures in the middle Atmosphere	rear	Global	8–53	3	390
Diurnal Changes	S5	Diurnal changes near the terminator	side	Near the terminator	15–60	3	480
UTLS	S6	Upper Troposphere / Lower Stratosphere	rear	Global	6–35	2-7	120
Upper Atmosphere	S7	Upper atmosphere	rear	Global	20–160	3-8	800

Table 4.1: MIPAS nominal and special observation modes (before March 24, 2004).

in which emission lines of most atmospheric species are present.

A good radiometric sensitivity is essential to allow detection of weak atmospheric signals. The radiometric sensitivity is expressed by the *Noise Equivalent Spectral Radiance* (NESR), which characterizes the instrument noise on the measurement in terms of incident radiance. For MIPAS the required sensitivity must be better than 50 $nW \cdot cm^{-2} \cdot sr^{-1} \cdot cm$ at the long wavelengths (low frequencies), decreasing to 4.2 $nW \cdot cm^{-2} \cdot sr^{-1} \cdot cm$ at the short wavelengths (high frequencies). These NESR values are based on pre-launch test results and have been now updated with in-flight measurement values. The requirement on radiometric accuracy is equally stringent: a good absolute knowledge of the received radiance is required for an accurate determination of the atmospheric temperature, which is a key parameter in the data retrieval.

4.2.2 Measurement principles

The interferometric design allows all the incoming radiation to be directed onto the detectors along the full duration of the interferometric sweep. This provides a clear advantage over spectrometers which use dispersive elements and scanning devices to send the individual spectral elements sequentially to the detector. Furthermore, with regard to a dispersive spectrometer, an interferometer can accept incoming light covering a much larger solid angle, then increasing the signal-to-noise ratio. Another advantage of a Fourier Transform spectrometer lies in the fact that a single detection element can be used to record a broadband spectrum (limited only by the spectral response) with high spectral resolution, while for dispersive spectrometers a large array detector must be employed. The disadvantage of the interferometer is that the modulated output has to be sampled at regular intervals along the optical path, increasing the mechanical complexity of the instrument. Nevertheless the use of an FTS instrument is the only way to obtain complete high resolution IR observations of an extended source, such as the Earth's limb, with good sensitivity. A Fourier Transform interferometer is the heart of MIPAS [95]: it is the only means currently available to obtain simultaneous (all related species at the same time) high-resolution infrared spectra of the atmosphere with a global scale.

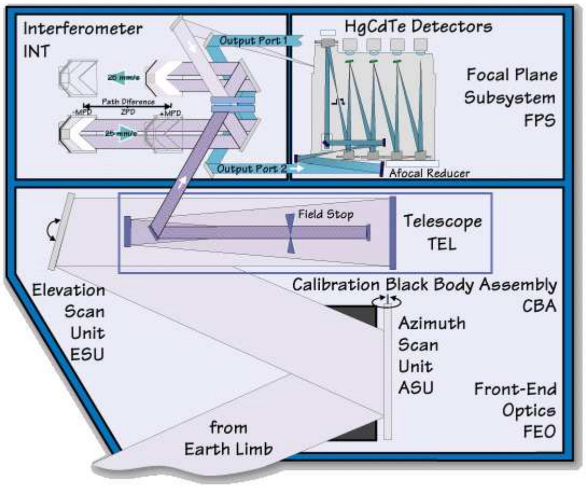

Figure 4.3: Schematic view of the optical layout of MIPAS.

Figure 4.3 shows the optical layout of MIPAS, indicating the path of the incoming light from the entrance to the detector elements. The atmospheric radiance enters through the front-end optics; here the first optical element is the azimuth scan unit (ASU) that allows the line of sight (LOS) selection within the two field of view (FOV) directions. In the azimuth scan unit there is an internal black-body source used for the in-flight gain calibration of the instrument. From the entrance aperture in the azimuth scan unit the beam is reflected to the elevation scan unit (ESU) which determines the tangent altitude of each measurement. The angle covered by this mirror is less than 3° which is sufficient to encompass limb heights between 5 and 250 km; the highest values will be used for measurements of cold space as part of the calibration procedure. From the ESU the radiation enters the telescope which reduces the dimension of the beam to match the input dimension of the interferometer. A field stop in the focal plane of the front-end telescope defines the instantaneous field of view (IFOV) of MIPAS. The atmospheric radiance is then directed to the Michelson interferometer.

In order to meet the radiometric and spectrometric performance requirements, as well as the lifetime requirement of five years of continuous operation in space, a symmetrical dual slide interferometer with dual input and output ports was selected. From the input port the atmospheric radiance collected by the front-end optics enters the interferometer, whereas a second input port is closed with a cold black-body to suppress disturbing radiance. The two output beams leaving the interferometer are reduced in size by two small telescopes and directed to the focal plane subsystem which houses the Hg-Cd-Te photoconductive detectors. All optics and detectors in the focal plane subsystem are cooled to 70 K in order to reduce their thermal emission.

MIPAS spectral coverage is divided into five spectral bands. Each detector is optimized for highest sensitivity in a particular band. Table 4.2 summarizes the spectral range of these bands and their NESR values.

MIPAS band	Spectral range (cm^{-1})	NESR ($nW cm^{-2} sr^{-1} cm$)
a	685.0 – 970.0	50
ab	1020.0 – 1170.0	40
b	1215.0 – 1500.0	20
c	1570.0 – 1750.0	6
d	1820.0 – 2410.0	4.2

Table 4.2: Spectral coverage and NESR requirements for each band of MIPAS.

Within the interferometer, the incoming radiation is divided by a beam-splitter into two beams of similar intensity. These are directed onto two moving retro-reflectors and returned on to the beam-splitter which now acts as a beam-combiner. The beams are superimposed and interfere with the resulting intensity at the output ports varying as a function of optical path difference. The interference-modulated signal constitutes an interferogram which is the Fourier Transform of the spectrum of the incoming radiation. The original spectrum can be reconstructed from the recorded interferogram by an inverse Fourier Transform.

The spectral resolution of a Fourier Transform Spectrometer is determined by the Maximum Path Difference (MPD) achievable in the particular interferometer. To meet the mission objectives, the MIPAS MPD is equal to about 20 cm, or to a mechanical motion of each corner cube over a 50 mm long path. In the dual-slide design of MIPAS, the lengths of both arms of the interferometer are simultaneously modified in a “push-pull” arrangement moving the retro-reflectors. To record a useful interferogram, the modulated output has to be sampled at very regular optical path difference intervals (the required sampling accuracy for MIPAS is about 30 nm). This is done with the help of a laser beam transmitted in the same optical set up, which is used to trigger the sampling electronics of the detector at very precise path difference positions.

4.3 Current status of the MIPAS instrument

MIPAS started working successfully since April 2002. The measurement campaign proceeded with only few interruptions due to various types of anomalies experienced by the interferometer unit. All these anomalies pointed to a degradation of the interferometer subsystem, but MIPAS scientific return was not affected: the generated products were still meeting the engineering and scientific requirements.

On March 26^{th}, 2004 MIPAS regular operations were suspended for a serious anomaly related to an unexpected behavior of the interferometric mirror slides. An exhaustive series of tests was carried out in order to identify the cause. The analyses highlighted a combination of different effects, the most important being a mechanical degradation of the interferometer slides. This anomaly did not allow to operate the instrument in its original configuration.

During the unavailability period, MIPAS was tested in several different operational modes in order to assess the safer one with respect to the instrument health and to continue in the production of data with the best quality, i.e. “2RR” double-slide moving mirrors with a reduced resolution of 0.0625 cm^{-1}, and the “1RR” single-slide moving mirror with resolution of 0.05 cm^{-1}. Amongst the possible configurations, the MIPAS Science Team decided to operate the instrument for the future in the more stable double-slide mode in which the spectral resolution is to be kept at 0.0625 cm^{-1}, that is 41% of the full spectral resolution used in the original nominal configuration. The instrument is now operated in discontinuous way and the data quality is marginally affected. The consequences of MIPAS operating on a campaign basis are that:

i currently the Near-Real Time (NRT) data are no longer generated, no systematic operational Off-Line (OFL) processing is performed while the algorithms are being adapted to the new observation modes;

ii any request for special measurements (i.e. in conjunction with airborne or ground based campaigns) have to be planned with great advance.

Although it was evident that the anomaly was caused by a mechanical problem, the precise cause is still not fully understood. An empirical understanding is that the interferometer performance improves after a long period of interruption.

After the detailed anomaly review, the mission has restarted with an in-flight re-characterization. From January 10^{th}, 2005 MIPAS is back in operation working with the measurement scenario of 3 days-on and 4 days-off cycle. This scenario allows a global coverage to be obtained in the three days of operations, while the four days switch-off of the instrument is request for relaxing the interferometer slides system.

All MIPAS "nominal" and "special" observation modes have been necessarily modified by the management team, defining a new operational scenario.

The configuration tested for the "nominal" mode ranges from 6 to 70 km at variable steps. A limb scan is now composed by 27 sweeps with an along track sampling around 400 km.

Other observation modes include:

UTLS-1: The Upper Troposphere Lower Stratosphere (**UTLS-1**) mode that enables 2D retrievals starts around 5 km and reaches the height of 46-53 km with 18 sweeps per scan and an along track sampling around 270 km. In this "special" mode MIPAS is operated continuously for 3 days.

UTLS-2: A second Upper Troposphere Lower Stratosphere (**UTLS-2**) mode has been tested. It ranges in altitude from 12 to 42 km with 11 sweeps per scan and an along track sampling around 165 km (horizontal over-sampling), with MIPAS operated for 1 day.

MA: In the Middle Atmosphere (**MA**) mode the scan is composed of 29 sweeps (18-102 km at 3 km steps) with MIPAS operated for 3 days close to the Solstice or the Equinox.

AE: The Aircraft Emissions (**AE**) mode, formerly referred to as S3, covers the altitude range from 8 to 38 km in 11 sweeps per scan with MIPAS operated for 2 days.

UA: In the Upper Atmosphere (**UA**) mode, formerly referred to as S6, MIPAS measures limb radiances from 42 to 102 km at 3 km steps (the lowest height is still not fixed) and from 102 to 172 km at 5 km spacing steps with totally 35 sweeps per scan and MIPAS operated for 1 day.

4.4 Data processing

MIPAS interferometric measurements are transformed into calibrated spectra of atmospheric radiance. The spectra are fed to an inversion model to compute vertical profiles of atmospheric geo-physical parameters. From the most raw interferograms to the final atmospheric profiles, there is series of necessary processing steps. The processing chain is separated in two major parts: a *space segment* and a *ground segment*. The flowchart in fig. 4.4 illustrates the flow of MIPAS data processing.

Data acquired in orbit by the instrument are transmitted to ESA ground stations where undergo further stages of processing, resulting in higher level data products. The ground processing is divided into two major processing phases: Level 1b and Level 2 processing. The goal of the *Level 1b processing* is to decode the instrument source packets and transform the interferograms into calibrated and geolocated spectra of atmospheric radiance. In the subsequent *Level 2 processing* phase, the calibrated spectra are processed by ESA to retrieve atmospheric pressure at tangent altitudes and vertical distribution of temperature and VMR of relevant atmospheric constituents. All the resulting data products are delivered in a format specific for the ENVISAT Ground Segment (Payload Data Segment - PDS) and are identified by labels indicating the step in the processing chain.

MIPAS Level 0 data correspond to the data stream received directly from the instrument without any further processing.

MIPAS Level 1A data are intermediary data, not archived or distributed, used as starting point for the subsequent processing stage.

MIPAS Level 1b data consist of formatted, geolocated, radiometrically and spectrally calibrated radiance spectra, also annotated with quality indicators.

MIPAS Level 2 data comprise retrieved profiles of pressure, temperature and VMR of H_2O, O_3, N_2O, CH_4, HNO_3, and NO_2.

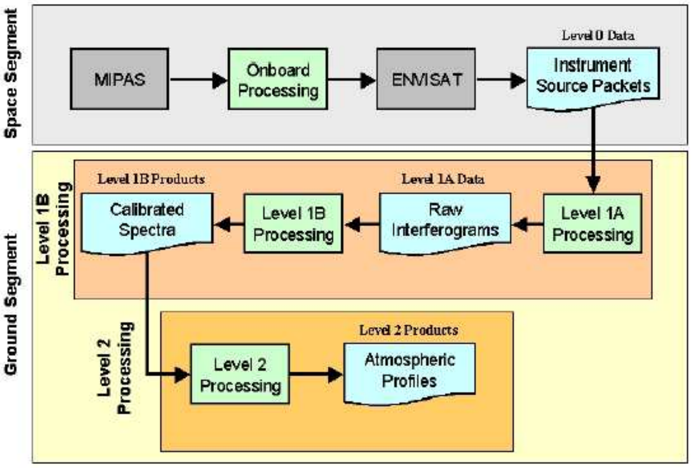

Figure 4.4: MIPAS data processing flowchart.

4.5 Optimized Retrieval Model (ORM)

4.5.1 General features

The objective of the ORM [93] was to process continuously and in near real time (less than three hours from the measurements to the results) MIPAS measurements. To this purpose it was necessary to develop physical and mathematical optimization of the retrieval scheme and to optimize the trade-off between run time and accuracy of the retrieval. The desirable performances of the retrieval code had the following targets for the overall error budget, including systematic errors:

- noise error on temperature within 2 K,
- noise error on tangent pressure within 3%,
- noise error on VMR of the target species within 5%, at all altitudes covered by the typical MIPAS scan.

The preliminary objective for the run time of the total retrieval, that is the p, T and VMR retrieval of the six MIPAS target species for a limb scanning sequence of 17 limb views, was a maximum of 1 hour on a Sun Sparc 20 workstation. The redundancy of information provided by the broad-band and high resolution spectral measurements of MIPAS and the multiplicity of unknowns lead to the adoption of the following main features.

The retrieval is based on the use of selected spectral intervals called microwindows. These spectral intervals are not wider than 3 cm^{-1} and are chosen in order to contain the best information on the target parameters and to avoid the analysis of spectral regions that are either "blind" to the unknown or affected by systematic errors (e.g. regions characterized by uncertain spectroscopic data, interference by non-target species, non local thermal equilibrium, and line-mixing effects). By analysing the sensitivity of the radiance to target parameters changes, list of appropriate microwindows were compiled for the retrieval of the six key species as well as pressure and temperature. A microwindows database has been created with respect to minimization of retrieval errors [96, 97, 98].

The retrieval is based on the non-linear least squares criterion without use of *a priori* information and a Global fit approach [99]. In the global fit the vertical distribution of each geophysical parameter is retrieved simultaneously at all altitudes using the whole limb scanning sequence. Compared with the onion peeling method [100], that determines the gas amounts sequentially starting from the top of the atmosphere, the global fit approach avoids the error propagation on the retrieved amounts of gas and allows the use of retrieved altitudes that are no longer strictly connected with the geometry of the observed spectra. In addition, it permits the full exploitation of hydrostatic equilibrium condition and allows several optimisations in the radiative transfer modeling.

The unknowns of the problem are retrieved sequentially: first, temperature and tangent pressure are retrieved simultaneously (p, T retrieval), then the VMR of the target species are retrieved individually in sequence according to the degree of their reciprocal spectral interference. The reason for this approach is that a simultaneous retrieval of all the species would require too large amount of computer memory. Simultaneous p, T retrieval exploits the hydrostatic equilibrium assumption, that provides a relationship between temperature, pressure and geometrical altitude. The sequence of the target species retrieval is: H_2O, O_3, HNO_3, CH_4, N_2O, NO_2. Besides temperature, tangent pressure and VMR profiles, sampled at tangent pressures, the fitted quantities are the parameters of atmospheric continuum and instrument zero-level calibration correction.

The atmospheric continuum is the spectral intensity contribution of all emission sources that are frequency independent within a microwindow. This contribution is caused by the wings of far lines (the most important being those of H_2O), the pressure broadened bands of O_2 and N_2 and the emission of aerosols. In the retrieval model the atmospheric continuum is assumed to be constant within the microwindows and to be both altitude and microwindow dependent.

The instrument zero-level calibration correction accounts for an additive, microwindow dependent, offset (but tangent altitude independent) that is caused by internal emission of the instrument, scattering of light into the instrument or third order nonlinearity of the detectors.

4.5.2 Retrieval scheme

The problem of retrieving the vertical distribution of a physical or chemical quantity from a limb scanning observation is a typical inverse non-linear problem (see Chapter 3). Figure 4.5 illustrates the MIPAS retrieval algorithm scheme. Starting from some first-guess values of the unknown parameters and using data on observation geometry and instrumental characteristics, the forward model program computes the simulated spectra, which are compared with the measured spectra provided by the MIPAS Level 1b processor. The MIPAS Level 1b processor converts the scene interferograms into fully calibrated radiance spectra, using pre-processed radiometric offset, gain calibration data and spectral axis correction parameters. The quadratic summation of the differences between the simulated and measured spectra provides the value of the chi-square that has to be minimized. A new profile is generated by modifying the first-guess with the correction provided by Eq. (3.19), i.e. the correction to the assumed value of the considered parameter in order to obtain its correct value. Convergence criteria are applied in order to establish if the minimum of the chi-square function has been reached. If the convergence criteria are fulfilled the procedure stops. If the convergence is not reached the improved profile is used as a new guess for generating simulated spectra that are again compared with the measured ones. The main components of the retrieval algo-

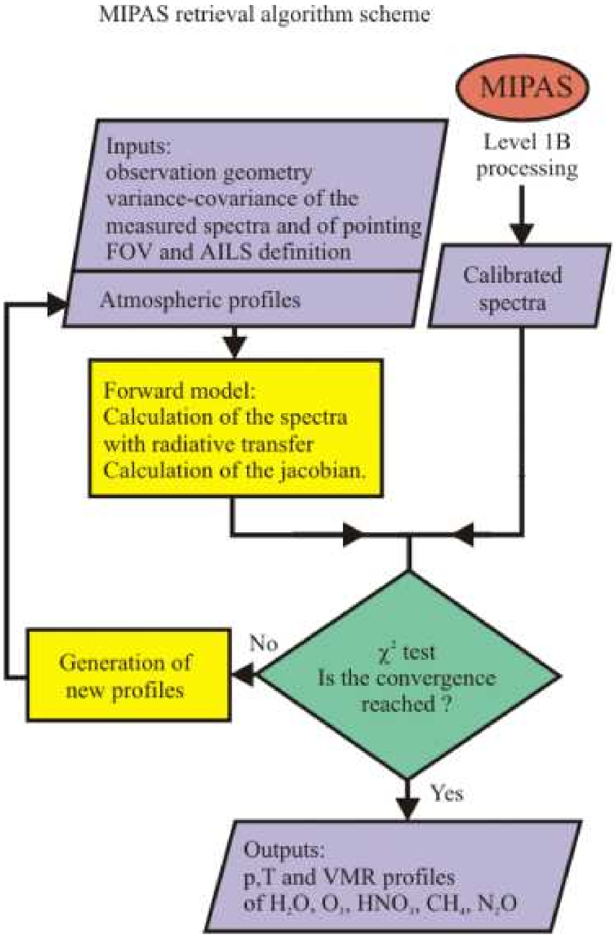

Figure 4.5: MIPAS retrieval alghorithm sequence.

rithm are the forward model, the Jacobian calculation and the convergence criteria, that will be each briefly described in the following subsections.

Optimised Forward Model

The Forward model simulates the spectra measured by the instrument in the case of known atmospheric composition. The signal measured by the spectrometer is equal to the atmospheric radiance which reaches the spectrometer modified by the instrumental effects, due to the finite spectral resolution and the finite field of view of the instrument. These instrumental effects are taken into account by convolving the atmospheric limb radiance with respectively the Apodized Instrument Line Shape (AILS) and the MIPAS Field Of View (FOV) function (defined in the following subsections). The atmospheric radiance that reaches the instrument when it is pointing to the limb at tangent altitude zt is calculated by means the radiative transfer equation (the Beer-Lambert law for a non-homogeneous medium):

$$L(\tilde{v}, z_t) = \int_{x_i}^{x_0} \left[B(\tilde{v}, x)k(\tilde{v}, x)\eta(x)\right] \times \exp\left[-\int_{x}^{x_0} k(\tilde{v}, x')\eta(x')dx'\right] dx \tag{4.1}$$

where x is the position along the line of sight between the observation point x_0 and the point x_i at the farthest extent of the line of sight, $B(\tilde{v}, x)$ is the source function, $k(\tilde{v}, x)$ is the absorption cross section, and $\eta(x)$ is the number density of the absorbing molecules. The exponential term represents the transmittance between x and x_0. In the case of local thermodynamic equilibrium $B(\tilde{v}, x)$ is equal to the Planck function. The simulated measured spectrum is equal to the convolution of the atmospheric radiance $L(\tilde{v}, z_t)$ with the $AILS(\tilde{v})$ and the $FOV(z_t)$:

$$S(\tilde{v}, z_t) = \int\int L(\tilde{v} - \tilde{v'}, z_t - z'_t) \times AILS(\tilde{v'})d\tilde{v'}FOV(z'_t)dz'_t \tag{4.2}$$

The computation of Eqs. (4.1) and (4.2) requires many operations that must be repeated for several variables, each with numerous values. The first objective of the ORM is to optimize the sequence of operations in order to avoid repetition of the same calculations and to minimize the number of memorized quantities. A second objective is to perform the calculations with the required accuracy. The implemented main optimizations are described as follows.

Definition of the frequency grid Limb radiance spectra contain spectral features varying from the narrow, isolated, Doppler-broadened lines at high altitudes, to wide, overlapping, Lorentz-broadened lines at low altitudes. The simplest procedure to resolve the sharp lines at high altitude is to choose a wavenumber grid fine enough that the narrowest line is adequately sampled (0.0005 cm^{-1} is required). The choice of a "fine grid" implies a large number of calculations that is prohibitive for most applications and represents an excessive amount of computer time and storage. Since not all the points of the fine grid are equally important for the reconstruction of the spectral distribution, a subset of spectral grid points, denoted *irregular grid*, is sufficient. Full radiative calculations need to be only performed for this subset of points and the remaining spectrum is obtained using an interpolation scheme (a linear interpolation is chosen).

The irregular grid is specific of each microwindow and of the microwindow boundaries and depends on the spectral convolution required by the Apodized Instrument Line Shape (AILS).

Definition of the atmospheric layering The non-homogeneous nature of the atmosphere along a radiation path is treated by sub-dividing the atmosphere into a set of layers. In this way the calculation of the Radiative Transfer integral (4.1) becomes a summation over the constituent layers. The layer boundaries should be chosen in such a way that the gas within the layer may be considered homogeneous and well represented by appropriate Curtis-Godson [101] parameters for pressure and temperature . An optimal set of boundaries are calculated based on a maximum allowed variation of temperature and average Voigt half width of a reference line across a layer. Since the transmittance weighting functions for limb geometries peak near the tangent point, it is important that a fine layer structure is used. Within each layer a series of paths are defined along the actual ray trajectory for each of the different gases comprising the layer. For each ray-path (i.e. combination of layer and limb view), the Curtis-Godson quantities are calculated. In the approximations of flat layers and straight line of sight, the same values of equivalent pressure and temperature are obtained independently of the limb angle between the line of sight and the vertical direction (secant law approximation). This is an useful approximation because its implementation allows the calculation and storage of less equivalent pressures and temperatures and therefore less cross-sections.

For a spherically symmetric atmosphere the transmittance in paths symmetrically located at either sides of the tangent point is identical and path parameters are only required for one of the two paths.

Computation of absorption cross sections The computation of cross-sections in Eq. (4.1) is a very time consuming part of the forward model, due to the large number of spectral lines to be considered, the high spectral resolution required and the number of p, T combinations for which they have to be computed. An alternative to the explicit **L**ine-**B**y-**L**ine (**LBL**) calculation of absorption cross-sections, in which the computation is performed for every line at every wavenumber grid point with a preselected line spectroscopic database, is the use of pre-computed Look-Up Tables (LUTs).

The basic idea of LUTs [102] is to pre-calculate, for each frequency grid point, the absorption cross sections of each gas for a set of different pressures and temperatures within the range of the atmospheric variability. These data are stored in files, which are read at the beginning of each retrieval, and then they are interpolated to the equivalent pressure and temperature of the atmospheric paths.

The use of LUTs is advantageous only if they can be stored in random access memory (RAM), and the feasibility of this depends on the amount of memory required.

The combined use of irregular grids and LUTs means that the absorption cross-sections need only be reconstructed at a subset of fine grid points. Therefore, LUTs and irregular grids used in combination allows to reduce significantly the computing time required by forward model.

Apodized Instrument Line Shape convolution The Instrument Line Shape (ILS) function is, by definition, the response of the spectrometer to a monochromatic radiance. In the case of a perfect Fourier transform spectrometer, the ILS is equal to the convolution of the sinc function, associated with the finite spectral resolution of the instrument, with a term due to the finite angular aperture. If the angular aperture is circular, this term is equal to a rectangular function shifted in wavenumber and with a width that varies linearly with the wavelength [42]. In a real instrument alignment errors and irregular angular aperture lead to a more elaborate ILS.

The ILS, which is an input of the forward model, is assumed to be independent of the tangent altitude and does not include the instrument responsivity and phase error corrections, since retrieval is performed using calibrated and phase-error corrected spectra provided by Level 1b processing.

Measured spectra are apodized with the Norton-Beer [103] strong function before entering the retrieval process in order to reduce the interference of far lines inside the individual microwindows. Apodization reduces the amplitude of the side lobes of the ILS and accordingly the size of the spectral interval in which the atmospheric spectrum must be calculated. When the spectrum is convoluted with the apodization function, correlation between different spectral points is introduced which appear as off-diagonal terms in the variance covariance matrix of the observations.

The Apodized Instrument Line Shape (AILS) is obtained by convoluting the measured ILS with the apodization function. The high resolution spectrum computed with the radiative transfer model is convoluted with the AILS function. The code is optimised in order to provide in a single operation both this convolution process and the resampling of the spectrum from the irregular grid of the radiative transfer model to the coarse grid of the measured spectral data.

Instrumental Field of View convolution The Field of View (FOV) of an instrument is the angular distribution of the instrument sensitivity. In the case of a satellite measurement, like MIPAS, a linear relationship exists between viewing angle and tangent altitude; therefore the FOV can be defined by a distribution as a function of altitude.

The effect of the spread of the FOV in the altitude domain is taken into account performing, for each spectral frequency, the convolution of the FOV pattern with the atmospheric limb radiance. This convolution requires the forward model calculation for a number of lines of sight that span the vertical range of the FOV pattern.

To reduce the number of computations, in the ORM an analytical convolution is performed. An analytical expression for the spectrum with generic tangent altitude is determined interpolating a polynomial through spectra calculated at contiguous tangent pressures in the range of the FOV pattern. In altitude ranges in which the radiance profile varies rapidly with tangent altitude, additional spectra are simulated at intermediate tangent levels to maintain the numerical accuracy of the convolution.

The simultaneous computation of the whole sequence of limb scanning spectra, which is required by the global fit approach, allows a simple and efficient computation of the FOV convolution and avoids the reiterated computation of spectra with adjacent tangent altitudes.

Jacobian calculation

Another important part of the retrieval code is the fast determination of the derivatives of the radiance with respect to the retrieval parameters. In most cases the computation of the forward model and its derivatives will take far longer than the linear algebra. In most circumstances it is preferable to evaluate the algebraic derivative of the forward model code rather than perturb the forward model for each element of the state vector and recompute the forward model several times.

In the ORM, whenever possible (this is the case of tangent pressure, volume mixing ratio, atmospheric continuum and instrumental offset), derivatives are computed analytically in the sense that analytical formulae of the derivatives are implemented in the program. When the calculation of sufficiently precise analytical derivatives requires computations as time consuming as the calculation of spectra (this is the case of temperature), an optimised numerical procedure is implemented. The temperature derivatives are determined in a "fast numerical" way, i.e. the derivatives are computed in parallel with the spectra in order to exploit the common calculations.

Convergence criteria

Convergence criteria are needed to establish when the minimum of the chi-square function has been reached. The convergence criteria adopted in the ORM are a compromise between the required accuracy of the parameters and the computing speed of the algorithm.

The adopted convergence criteria are based on three conditions:

- Condition on *linearity*: at the current iteration the relative difference between the actual chi-square and the expected value of chi-square computed in the linear approximation must be less than a fixed threshold,
- Condition on *attained accuracy*: the relative correction that has to be applied to the parameters for the subsequent iteration is below a fixed threshold. Different thresholds are used for the different types of parameters depending on their required accuracy. Furthermore, whenever an absolute accuracy requirement is present for a parameter, the absolute variation of the parameter is checked instead of the relative variation. The non-target parameters of the retrieval, such as continuum and instrumental offset parameters are not included in this check,
- Condition on *computing time*: the maximum number of iterations must be less than a given threshold,

The convergence is reached if either condition on linearity of the inversion problem or condition on attained accuracy is satisfied. If only the condition on computing time is satisfied, the retrieval is considered unsuccessful.

4.5.3 Algorithm performance and validation

The run-time performances of the Optimised Retrieval Model was tested using different computers. Tests were performed on simulated observations using two different sets of microwindows: a preliminary standard set and a set which optimises the trade-off between accuracy and run-time performance. In these tests the initial guess profiles of the retrieval were chosen sufficiently close to the reference profiles (the ones used to simulate the observations), so that convergence is reached in only one iteration. In these conditions, the run-time required to perform p, T and VMR retrieval of the MIPAS target species from a limb-scanning sequence of 16 limb-views is only 51 seconds on a Digital DEC-SERVER with 600 MHz CPU. Considering that the measurement time per scan is 75 seconds and that more than one computer can be used for the operational analysis data, the run-time requirements are fully satisfied also for retrievals that need more than one iteration.

To validate the approximations implemented in the OFM, a parallel activity was performed at Oxford for the development of a line-by-line forward model code based on GENLN2 [104]. This code was compared with several existing codes and was elected as Reference Forward Model (RFM) [105].

The retrieval accuracy of the ORM code was validated by performing retrievals from spectra generated by its own forward model and by the RFM. Results indicate that both forward model error, i.e. error due to imperfect modeling of the atmosphere, and convergence error, i.e. error due to the fact that the inversion procedure does not find the real minimum of the chi-square function, are much smaller than the measurement error due to radiometric noise.

In absence of MIPAS measurements, measurements obtained by instruments similar to the ENVISAT version of MIPAS have been used to validate the ORM: the balloon instrument MIPAS-B2 [106] (same instrument but different geometry) and the Space Shuttle instrument ATMOS (satellite geometry but different mode of observation).

4.5.4 Error budget

The main error sources that affect the accuracy of the retrieved profiles are:

- noise error, due to the mapping of radiometric noise in the retrieved profiles,
- systematic error, due to incorrect input parameters,
- retrieval errors.

The amplitude of noise error was evaluated performing test retrievals with observations generated starting from assumed atmospheric profiles (reference profiles) and perturbed with random noise of amplitude consistent with MIPAS noise specification.

The noise error obtained in these test retrievals is consistent with the accuracy requirements at most altitudes covered by the standard MIPAS scan.

Systematic errors, such as spectroscopic errors or errors due to imperfect knowledge of the VMR profiles of non-target species, are taken into account in the definition of the optimum size of each microwindow and in the selection of the optimal set of microwindows.

Among the systematic errors there is the effect of temperature error on VMR retrievals that is determined using tabulated propagation matrices which estimate the effect for different measuring conditions. Preliminary results indicate that temperature error can be a significant component of the error budget.

Systematic errors have a major role in the total error budget and their accurate characterisation is an important requirement that can only be met with real data.

Chapter 5

The Geo-Fit Multi-Target Retrieval

In this chapter is described how the limit of the horizontal homogeneity of ORM can be overcomed with the new ***G****eo-fit [4]* ***M****ulti-***T****arget* ***R****etrieval [5] model (****GMTR*** *[3]) that implements the Geo-Fit two-dimensional inversion for the simultaneous retrieval of several targets.*

5.1 Geo-Fit Rationale

In the analysis of data from a limb-scanning experiment, the assumption of horizontal homogeneity of the atmosphere can be avoided if each limb observation contributes to determining the unknown quantity at a number of different “locations” among those spanned by its line of sight [4]. However, the attempts to derive, from a single limb scan, atmospheric parameters at different locations along the lines of sight, usually fail because of an ill-conditioned problem (the retrieval of a horizontal gradient can be attempted in these cases).

In the MIPAS experiment a solution to the problem regarding the analysis of observations taken along the orbit track can be found by exploiting the fact that limb-scanning measurements are continuously repeated along the plane of the orbit. Indeed, this repetition makes it possible to gather information about a given location of the atmosphere from all the lines of sight that cross that location regardless of which sequence they belong to. Since the loop of cross-talk between nearby sequences closes when the starting sequence is reached again at the end of the orbit, in a retrieval analysis the full gathering of information can be obtained by merging in a simultaneous fit the observations of a complete orbit. One of the Earth’s poles is the optimal choice as a starting and ending point of the analyzed orbit because, in this case, the cross-talk loop is closed when the same air mass is observed. The use of any other latitude as a starting and ending point would face the problem that, because of the Earth’s motion, a different longitude is observed at the beginning and at the end of the orbit.

Since MIPAS Level 1b files correspond to orbits that do not originate at the North

Pole, a geo-fit analysis requires the preliminary construction of a full orbit of measurements by combining the measurements stored in two Level 1b files. An important feature of the geo-fit is that the retrieval grid is independent from the measurement grid, which is the grid identified by the tangent altitudes of the measurements. Therefore the atmospheric profiles can be retrieved with horizontal separations different from those of the measured limb scans.

5.2 Multi-Target Retrieval Rationale

In the task of determining the altitude distribution of atmospheric constituents, the knowledge of pressure and temperature distributions is necessary for the determination of all VMRs. In the data-analysis process, the usual approach is the preliminary retrieval of pressure and temperature (exploiting the assumption of known CO_2 VMR) followed by the sequential retrieval of the target VMRs. A drawback of this approach is that the retrieval errors affecting pressure and temperature profiles do propagate into the retrieved VMR values. Moreover, molecular species with a “rich” spectrum (such as water and ozone) may also propagate their measurement error in the other products because their spectral features often “contaminate” the frequency intervals analyzed for the retrieval of other species. The error propagation process can be minimized with a careful choice of both the analyzed spectral intervals and the sequence of the retrievals. Nevertheless, the error propagation cannot be completely avoided and its assessment requires some postprocessing operations. A strategy that minimizes this source of systematic errors is represented by the simultaneous retrieval of all the quantities whose correlation in the observed spectra is the cause of the error propagation [5]. This strategy is referred to as multitarget retrieval.

The main advantages of MTR are the following:

i No systematic error propagation due to “interfering” species.

ii The error due to the cross-talk between different target quantities is properly represented in the covariance matrix of the retrieved parameters.

iii The selection of the spectral intervals to be analyzed is no longer driven by the necessity to reduce the interferences among the target species.

iv The information on pressure and temperature can be gathered from the spectral features of all target species and not only from CO_2 lines.

5.3 Use of Microwindows

Similarly to ORM the new retrieval model works on a limited number of observations. Actually, the redundancy of information coming from MIPAS measurements makes it possible to select a set of narrow (less than 3 cm^{-1} wide) spectral intervals containing the best information on the target quantities, while the spectral regions containing little or no information can be ignored [107]. The use of narrow spectral intervals, called "microwindows" (MWs), allows one to limit the number of analyzed spectral elements and to avoid the analysis of spectral regions that are characterized by uncertain spectroscopic data, interference by nontarget species, or are influenced by unmodeled effects, such as nonlocal thermodynamic equilibrium (NLTE) or line mixing. The use of MWs instead of broad spectral intervals also allows one to keep the inversion problem within acceptable dimensions from the computational point of view.

The MWs selected for a given retrieval target can be used at all the measured limb views within a scan or only at a subset of them. The optimal set of MWs is arranged in the so-called occupation matrix (OM), which is a logical matrix with as many rows as the limb views of the analyzed observations and as many columns as the MWs used in the inversion. The entry (i, j) of the OM defines whether the MW j at altitude i is used in the retrieval.

The algorithm used for the MW selection is described in Ref. [107]; it operates the selection of the spectral elements with the aim of maximizing the information content [85] of the used measurements with respect to the target parameters.

5.4 Retrieval Algorithm

The parameters of the retrieval are derived using a nonlinear least-squares fit based on the Gauss-Newton method. The mathematics of this method, applied to the retrieval of atmospheric parameters, can be found in Refs. [93] and [108]. Here we recall the basic equations and provide some details about their implementation in the new analysis scheme. The Gauss-Newton iterative solution expression is

$$\mathbf{y} = (\mathbf{K}^T \mathbf{S}_n^{-1} \mathbf{K})^{-1} \mathbf{K}^T \mathbf{S}_n^{-1} \mathbf{n} \tag{5.1}$$

in which $\mathbf{y}$ is the vector containing the corrections to the adopted state vector [85], $\mathbf{n}$ is the vector containing the differences between each observation and the corresponding simulation, $\mathbf{S}_n$ is the variance-covariance matrix (VCM) associated with vector $\mathbf{n}$, and $\mathbf{K}$ is the matrix (usually denoted as Jacobian matrix) containing the derivatives of the observations with respect to the retrieved parameters.

If the geo-fit analysis is operated to retrieve the altitude distribution of a single target quantity, the entries of the state vector are: (i) values of the target quantity at all the geolocated retrieval grid points and (ii) atmospheric continuum cross sections, at the same location of the target quantities, for the central frequencies of the analyzed MWs. A conventional order is assigned to the elements of the state vector within the orbit plane. If the geo-fit analysis is operated to perform a multitarget retrieval, the target quantities are sequentially ordered within the state vector. The results of an MTR analysis are, of course, not related to the storage order that, however, must be consistent within the arrays of Eq. (5.1).

In accordance with the geo-fit rationale, the analyzed spectral points are ordered in vector $\mathbf{n}$ starting from the first limb scan (of the examined orbit) measured beyond the North Pole. The next limb scan that runs through the North Pole terminates the set of analyzed observations.

In matrix $\mathbf{K}$, each column refers to a retrieved parameter while each row refers to a spectral element of the analyzed observations and contains the derivatives of that observation with respect to all the retrieval parameters. In each row of matrix $\mathbf{K}$, a subset of contiguous entries corresponds to a given retrieval target. Within each subset only minor sections contain entries that are different from zero: indeed a given observation depends only on a limited set of retrieval parameters [which are those located around the portion of the atmosphere sounded by the instrument field of view (FOV) relating to the considered observation]. Therefore with a suitable layout of the parameters, the matrix $\mathbf{K}$ can be compressed in such a way that only the sections of a row containing entries other than zero are calculated and stored. Auxiliary vectors provide the information about the position of the calculated entries in the expanded matrix $\mathbf{K}$. The derivatives that enter in matrix $\mathbf{K}$ are calculated using their analytical expression for all the targets (including the atmospheric continuum).

The entries of matrix $\mathbf{S}_n$ are defined on the basis of the noise level of the measured spectra, the correlation among the spectral points introduced by the zero filling when it is used in the Fourier transform of the interferograms, and the correlation between the spectral points introduced by the apodization process. Observations belonging to different microwindows are assumed to be uncorrelated; matrix $\mathbf{S}_n^{-1}$ is then block diagonal and only a minor portion of each row has entries different than zero. Therefore a compression can be applied to the rows of this matrix with a procedure similar to that used for matrix $\mathbf{K}$. The compression of matrices $\mathbf{K}$ and $\mathbf{S}_n^{-1}$ enables a dramatic reduction of the memory requirements and avoids products with zero entries, hence reducing by orders of magnitude the number of operations needed for the arrays algebra involved in Eq. (5.1).

Equation (5.1) provides the solution that minimizes the 2 (cost) function defined as

$$\chi^2 = \mathbf{n}^T \mathbf{S}_n^{-1} \mathbf{n} \tag{5.2}$$

The quantity χ_R^2 defined as

$$\chi^2_R = \frac{\chi^2}{(m-n)} \tag{5.3}$$

(in which m is the number of observations and n is the number of retrieved parameters) has an expectation value [109] of 1: Therefore the deviation of χ^2_R from unity quantifies the capability of the forward model to reproduce the observed spectra and provides a good estimate of the quality of the retrieval.

The errors associated with the solution of the inversion procedure are characterized by the VCM of $\mathbf{y}$ given by

$$\mathbf{V}_y = (\mathbf{K}^T \mathbf{S}_n^{-1} \mathbf{K})^{-1} \tag{5.4}$$

The matrix $\mathbf{V}_y$ maps the experimental random errors (represented by $\mathbf{S}_n$) onto the uncertainty of the values of the retrieved parameters; in particular, the square root of the diagonal elements of $\mathbf{V}_y$ provides the estimated standard deviation (ESD) of the corresponding parameter. Conventionally the ESD indicates the precision of the quantity it refers to.

5.5 The Retrieval Grid

The spatial resolution and the precision of the retrieved profiles are generally negatively correlated [6] and dependent on the retrieval grid. When horizontal homogeneity is assumed, the problem of choosing the retrieval grid consists of the selection of the altitudes at which the parameters are retrieved. Since the weighting functions of the observations generally peak at the tangent altitudes [108], a common choice is to let these altitudes coincide with the retrieval grid.

In the case of geo-fit, a two-dimensional retrieval grid must be adopted; the first dimension is the distance of the retrieval points from the Earth's surface, the second dimension is an angular (polar) coordinate identifying the position of the retrieval grid points in the orbit plane. The location of the tangent points of the observations could still be the leading criterion for the choice of the retrieval grid. However, even in the case of identical limb scans, because of the inhomogeneity of the atmosphere, the ray tracing (see *Ray Tracing* in 5.6.1) leads to different tangent altitudes because of the different refractive indices of the atmosphere sampled by the lines of sight. The resulting spread of the tangent altitudes would make difficult the interpolation process needed for both the construction of the atmospheric fields and the representation of the results. Therefore in the case of two-dimensional retrievals, a uniform retrieval grid defined at fixed altitudes is a more suitable choice. Regarding the polar coordinate of the retrieval grid points, an important feature of the geo-fit approach (see 5.1) is that the retrieval grid is fully independent from the measurement grid; therefore a user-defined retrieval grid that corresponds to the desired horizontal sampling can be chosen.

Besides this opportunity, a natural choice for the horizontal grid would be to use the polar coordinate of the tangent points of the measurements. However, due to the motion of the satellite and to the finite time interval needed to acquire the measurements, the tangent points of the individual views of a limb scan are not vertically aligned. In order to ease the interpretation of the retrieval results and to avoid the interpolation processes that may be required after the retrieval, it is preferable to have the retrieval grid points vertically aligned.

On the basis of the above considerations the adopted strategy for the retrieval grid is as follows:

i In the vertical domain the retrieval grid points are placed at the "nominal" (i.e., commanded) altitude of the tangent points of the limb observations [except for the single scan analysis (see *Geo-Fit and MTR Functionalities* in 5.6.2)].

ii In the horizontal domain the user can choose to place the retrieval grid points: (a) at the polar coordinate of the tangent altitude of the observations, (b) vertically aligned at the average polar coordinate of the tangent points of the individual scans, or (c) vertically aligned at user-defined arbitrary positions.

5.6 Forward Model

5.6.1 Geometrical Aspects

Reference Frames

The MIPAS Level 1b products are geolocated using two different reference frames; the satellite coordinates are provided with respect to the World Geodetic System 1984 (WGS84) (Ref. [110]) while the tangent points are identified using their altitude above sea level as well as their latitude and longitude. In the geo-fit analysis scheme, two 2D reference frames are adopted, both lying in the orbit's plane that is in the plane defined by the Earth's center and two arbitrary positions of the satellite. The first reference frame is a Cartesian (xy) frame whose origin is placed at the Earth's center; the positive x semiaxis joins the origin with the North Pole while the positive y semiaxis lies at 90° from the x positive semiaxis in the direction of the satellite motion. The second reference frame is a polar coordinates frame with the origin at the Earth's center; the polar axis coincides with the positive x semiaxis of the Cartesian frame while the polar angle increases in the direction of the satellite motion. The polar angle of this reference frame is referred to as "orbital coordinate" (OC).

Earth Model and Discretization of the Atmosphere

Whenever the atmosphere is assumed to be horizontally homogeneous, its discretization is generally operated only in the vertical domain using altitude levels that define horizontally homogeneous layers. If the horizontal variability of the atmosphere has to be modeled, a second discretization must be operated in the horizontal domain. In the GMTR, the horizontal discretization is built using segments perpendicular to the Earth's geoid and extended up to the boundary of the atmosphere. The spherical approximation of the Earth's shape suggests representing the altitude levels with concentric circumferences and the vertical elements with a set of evenly spaced radii that originate at the Earth's center. The result of the two-dimensional discretization is like a spider's web in which the intersections between levels and radii identify points denoted as "nodes". Furthermore, consecutive levels and radii in this web define plane figures, denoted as "cloves".

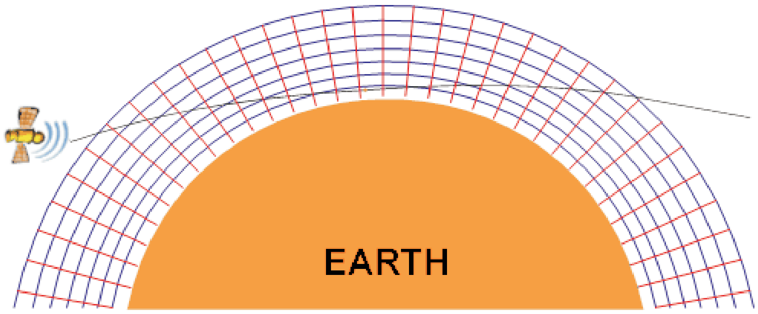

Figure 5.1: Two-dimensional discretization of the atmosphere and ray tracing.

A sketch of this two-dimensional discretization of the atmosphere is reported in fig. 5.1. If the geo-fit approach is implemented in an operational analysis system, the spherical approximation is no longer acceptable and the Earth's shape must be represented by an ellipsoid according to the WGS84 model. In this case the altitude levels also follow the elliptic shape. The definition of concentricity does not apply to the ellipses; however, in the Earth's dimensions scale, the objective of building two evenly spaced levels can still be attained, with an approximation of a few centimeters, using ellipses that differ in both semiaxes by the desired altitude separation. For what concerns the horizontal domain, the solution adopted in the case of spherical symmetry cannot be applied because a radius displaced by more than 10° from the ellipse axes would sensibly deviate from the vertical direction at the Earth's surface. However, the process of building a discretization along the horizontal domain can still start from a set of evenly spaced radii that originate at the Earth's center. These radii identify a set of points on the surface of the Earth's ellipsoid; from these points, segments perpendicular to the Earth's surface are originated and extended up to the atmospheric boundary.

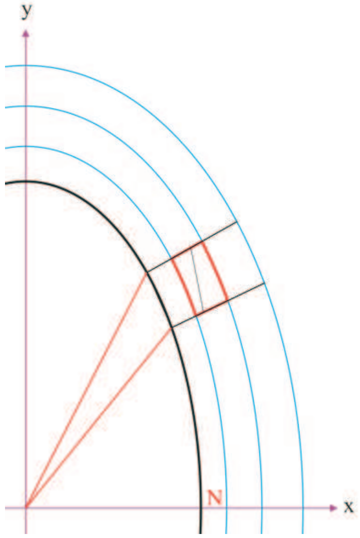

Figure 5.2: Elements of a two-dimensional discretization of the atmosphere with elliptical geometry. Three levels and two radii of the atmospheric discretization are represented with a highly distorted scale. The clove delimited by red contours is crossed by the path of a line of sight. The Cartesian reference frame and the position of the North pole (N) are also reported.

The adopted solution is illustrated, on a highly distorted scale, in fig. 5.2 in which three levels and two radii of the atmospheric discretization are represented. In this figure the path of a line of sight crossing the clove defined by red contours is also reported (see *Ray Tracing* in 5.6.1). Again, the elliptical geometry is such that a segment perpendicular to the Earth's surface is not perpendicular to the ellipses adopted for the altitude levels. However, this deviation is also of a minor entity in the Earth's scale because in the most unfavorable situation (at 45° from the ellipse axes) it translates into a displacement of a few hundreds of meters on a level at 100 km altitude.

Ray Tracing

If the horizontal variability of the atmosphere has to be modeled, the ray tracing procedure must be consistent with the two-dimensional discretization. In this case, the ray tracing starts from the coordinates that define the position of the satellite and the elevation angle corresponding to the considered limb view. These two geometrical elements define a segment (representing the considered ray) that extends from the satellite to the boundary of the atmosphere. From this point on, while crossing the atmosphere, the propagating ray will intersect levels and radii of the two-dimensional discretization. At each intersection, the direction of propagation of the ray is modified by the effect of refraction, in agreement with the Snell's law applied to the boundary between the two cloves involved in that intersection. The resulting shape of a line of sight is therefore piecewise linear with each segment representing the optical path (p) within a clove. In fig. 5.1 we show, on a highly distorted scale, an example of the ray tracing of a line of sight in the case of two-dimensional discretization of the atmosphere. In the geo-fit analysis, the ray tracing, and therefore the radiative transfer, is operated for the line of sight corresponding to the tangent altitude of the observation, and for two additional lines of sight that delimit the instrument FOV. For each spectral point, the spectrum is then represented as a function of the elevation angle by means of a second-order polynomial satisfying the three simulated lines of sight. This polynomial is analytically convolved with the FOV response function to provide the simulation of the observed limb radiance.

Pointing Angles Calibration

As described in 5.6.1 (*Ray Tracing*), the entire optical path of a line of sight depends on the physical properties of the atmosphere that is encountered. Actually, the refractive index assigned to each clove of the atmospheric discretization depends on the local pressure and temperature. Therefore a ray tracing operated on the adopted (see *Ray Tracing* in 5.6.1) atmospheric model (starting from the coordinates of the satellite and using the nominal value of the elevation angle) in general will not lead to the value of the tangent altitude supplied in the MIPAS Level 1b data. Furthermore, in a geo-fit analysis the tangent points are all represented in the orbit plane while, in practice, they sometimes deviate from this plane. For these reasons the GMTR operates a calibration of the pointing that, using an iterative procedure, derives the value of the elevation angle that leads to the tangent altitude reported in the Level 1b data. For this purpose, the ray tracing is repeated by adjusting the elevation angle until the nominal value of the tangent altitude is reached within a predefined tolerance.

Radiative Transfer

Definition of Atmospheric Fields The initial status of the atmosphere is defined by an input model that provides, in the whole orbit plane, the distributions of pressure, temperature, VMR of the atmospheric constituents, and atmospheric continuum cross section. The profiles cover the altitude range of the observations up to a maximum conventional value for the upper boundary of the atmosphere.

Values of physical and chemical quantities are calculated at the position of all the nodes of the two-dimensional discretization, starting from the input atmospheric model. Mixing ratios and temperatures are obtained by linear interpolation with respect to both the radial and the OC values. Linear interpolation is also used to derive the pressure values in the horizontal domain; in the vertical domain, an exponential variation of pressure with altitude is assumed. Finally, atmospheric continuum is obtained using the same interpolation laws as for pressure.

Average values of the refractive index (used for the ray tracing discussed in 5.6.1) and of the atmospheric continuum are associated with each clove of the two-dimensional discretization of the atmosphere starting from the quantities that have been calculated at the nodes that identify the clove itself. Other atmospheric properties are associated with each individual path p (within a clove) that results from the ray tracing of a line of sight. These quantities, calculated for each spectroscopically active gas (g) at the frequency of the simulation, are

i slant column ($\mathrm{col}_{g,p}$),

ii equivalent pressures (p_{eq}) and temperatures (T_{eq}), namely the Curtis-Godson [101] quantities,

iii absorption cross section $c_{g,p}$ calculated for the equivalent values of pressure and temperature.

The calculation of the above quantities is carried out assuming local thermodynamic equilibrium in the sounded atmosphere. For this reason the new retrieval scheme is not suitable for the analysis of daytime MIPAS measurements recorded within the "upper atmosphere" observation modes that are affected by NLTE processes.

Radiative Transfer Equation The radiative transfer is calculated for each line of sight with the classical equation referring to the case of no symmetry around the tangent point. For the reader's convenience we recall that the signal S reaching the spectrometer is computed by approximating the radiative transfer path integral as a summation over the N paths constituting the line of sight:

$$S = \sum_{p=1}^{N} B_p \left[1 - e^{\tau_p}\right] \left(\prod_{k=p+1}^{N} e^{-\tau_{k,p}} \right) \tag{5.5}$$

in which B_p is the Planck function computed for the equivalent temperature of the atmospheric mixture in the corresponding clove, and τ_p is the "single path" optical depth given by the following expression:

$$\tau_p = \sum_{g=1}^{N_{gas}} c_{g,p} \cdot col_{g,p} + c_{cont_p} \cdot col_{air_p} \tag{5.6}$$

In the second term of Eq. (5.6), c_{cont_p} represents the contribution to the "single-path" optical depth of all emission sources characterized by a constant amplitude in the considered MW (e.g., the contribution of atmospheric continuum), col_{air_p} is the air column for the path p.

Look-Up Tables and Irregular Frequency Grids In the radiative transfer calculation, the absorption cross sections $c_{g,p}$ in Eq. (5.6) are derived from precalculated look-up tables (LUTs). The use of LUTs enables one to save computing time with respect to the explicit line-by-line calculation; it requires one to precalculate for each frequency grid point the absorption cross sections of each gas for a set of different pressure and temperature pairs within the range of the atmospheric variability. These data (stored in files) are accessed and interpolated to the equivalent pressure and temperature of the atmospheric paths p_{eq}, T_{eq} pairs. Since the frequency grid at which the cross sections have to be calculated is rather fine ($5\times\ 10^{-4}$ cm^{-1}), the amount of data to be stored is large. To reduce the amount of data contained in the look-up tables (and consequently the time required to read them) a compression procedure has been adopted based on the matrix singular value decomposition, as described in Refs. [102, 93].

Limb radiance spectra contain spectral features ranging from the narrow, isolated, Doppler-broadened line centers (from high altitudes) to wide, overlapping, Lorentz-broadened line wings (from low altitudes). However, it is possible to identify a minimum subset of spectral grid points (irregular grid) that are sufficient to reconstruct the full radiance spectra over the whole range of tangent altitudes and atmospheres. Full radiative transfer calculations are then only required for this subset of points; the remaining grid points are obtained using a predetermined interpolation scheme. Typically, only 5-10% of the full resolution grid is required for the reconstruction of the spectra. The combined use of irregular grids and look-up tables results in a very efficient optimization of the computing time.

The algorithms described in Refs. [111, 112] for the automated selection of auxiliary data have been updated to support multitarget retrievals and retrievals that do not assume the horizontal homogeneity of the atmosphere. These algorithms have then been used to generate a full set of LUTs and irregular grids for a large set of atmospheric targets and observation modes.

5.6.2 Functionalities of the Retrieval Model

The new retrieval model is designed to allow flexible implementation of its functionalities. In this section we report the main options available to the user when the functionalities are exploited in a retrieval session. The optimal combination of functionalities can be selected by also taking into account the available computing resources.

Geo-Fit and MTR Functionalities

The geo-fit functionality can be fully exploited by operating retrieval analyses on along-track looking observations acquired in a full orbit. Alternatively, the retrieval can be operated on a subset of observations that cover a segment of the orbit. The subset can be reduced to a single limb scan; in this case the atmosphere is considered horizontally homogeneous and the vertical retrieval grid is defined by the tangent altitude of the observations.

The MTR functionality enables one to retrieve simultaneously more than one target quantity. Assuming the availability of auxiliary data, the only limitation to the number of targets is determined by the available computing resources (RAM memory, in particular). On the other hand, the system can also be operated to retrieve profiles of a single molecular species, assuming as known all the other quantities necessary in the forward model calculations. Pressure and temperature profiles can only be retrieved simultaneously.

Sequence of Retrievals: Cascade Mode

The execution of the retrieval can be reiterated on different target species within a sequence controlled by an external shell script. This functionality (denoted as "cascade mode") permits one to operate retrievals by using the results of the previous analyses. For this purpose the analysis program stores the atmospheric field of the retrieved target(s) (that is, the value that the quantity assumes at all the nodes of the atmospheric discretization) in temporary files. When a step of the retrieval sequence is run in "cascade mode", the program checks for the existence of these files and (if any) it reads and uses the previously retrieved atmospheric fields. Any combination of MTR and sequential retrievals can be implemented; however, in a typical retrieval sequence, the MTR of pressure, temperature, water, and ozone is followed by one or more "cascade" retrievals of "minor" atmospheric species. In the retrieval chain each step may have its own retrieval grid, independent of the previous steps.

Selection of Data from Michelson Interferometer for Passive Atmospheric Sounding Level 1b Files

As stated in section 5.1, the standard geo-fit analysis operates on spectra relating to a complete orbit originating at the North Pole; the line of sight of the observations lies in the orbit plane (along track measurements). Since the first limb scan reported in the Level 1b files is usually quite far from the North Pole, the analysis program joins two Level 1b files of two consecutive orbits to produce the set of measurements to be processed. However, in the case of measurement gaps, data from a single Level 1b file can also be processed. Furthermore, if measurements pointing across the orbit track are encountered in the Level 1b file(s), they are automatically identified and can be processed separately using the single limbscan option.

Exploitation of the Payload Data Segment Products

The initial guess of the target profiles and the profiles of the nontarget species are taken from a climatological database specifically developed for MIPAS data analysis. However, if atmospheric fields retrieved by the ESA ground processor are available, it can be preferable to use them to define the initial guess atmosphere because they represent an atmospheric status that is closer to the reality than the one provided by the climatological database. Alternatively, the user can choose to start from a different set of atmospheric profiles (e.g., those retrieved from an adjacent orbit) by addressing the software to the specific location where it can be read.

Retrievals are sensitive to the tangent altitudes that define the line of sight of the observations; good knowledge of these quantities is therefore crucial. Engineering estimates of the tangent altitudes are reported in the Level 1b data. However, if the tangent altitude estimates derived by the ESA Level-2 processor are available, these can be used in place of the Level 1b engineering estimates.

Chapter 6

Trade-off Studies between spatial resolution and retrieval uncertainties

In this chapter are reported some tests performed in order to investigate the trade-off between the accuracy associated to the retrieval parameters and the spatial resolution. Two methods to vary the horizontal resolution have been tested. One method varies the separation between profiles in the retrieval grid. For this purpose has been used the orbit 2081 of the 24th July, 2002. The second method exploits a variation of the measurements grid at the expenses of spectral resolution. The orbit selected for this scope is the 12858 of the 15th August 2004. In the last part of this chapter are shown the results obtained performing the retrieval of an orbit acquired with a very fine measurement grid (orbit 7023 of the 4th July 2003, Special Mode S6).

6.1 Trade-off studies varying the retrieval grid

As reported in section 5.5, a useful advantage of the system code is the possibility to choice a user-defined retrieval grid, so that a trade-off study is allowed.

In the following sections is described the trade-off study between the accuracy of the retrieved parameters and the horizontal resolution corresponding to a given retrieval grid. This study is performed on real data and is compared with the work previously done by Luca Magnani in a PhD Thesis [1] on simulated data.

The analyses have been made on the well-characterized orbit 2081 of the 24th July, 2002.

6.1.1 Tests and comparisons

In this set of tests have been adopted several retrieval grids that differ each other for the spatial separation between adjacent retrieved profiles.

In the GMTR code the file "*in_grid.dat*" is used if the system is asked to operate the retrieval on a user-defined horizontal retrieval grid and contains the values of the orbital coordinate where profiles are retrieved.

Some tests have been therefore performed using different "*in_grid.dat*" input files. For each target the number of retrieved profiles and their spatial separation is reported in table 6.1.

N° of retrieved profiles	24	36	48	**72**	80	90	100	120	144	180
Spatial separation (deg.)	15.0	10.0	7.5	**5.0**	4.5	4.0	3.6	3.0	2.5	2.0
Spatial separation (km)	1667	1111	833.5	**555.6**	500.1	444.5	400.1	333.4	277.8	222.3

Table 6.1: Number of retrieved profiles and spatial separation between two consecutive profiles in the executed tests. The value corresponding to the nominal conditions are reported in bold.

A first test has been performed on p, T, H_2O and O_3. First temperature and pressure profiles were retrieved simultaneously (pT retrieval), then the target species VMR profiles were individually retrieved according to the order of their reciprocal spectral interference (H_2O followed by O_3). All the targets have been retrieved using the retrieval grids reported in table 6.1.

Figure 6.1 reports the χ-test values at convergence for the pT retrievals as a function of the angular separation. The data point referring to the nominal retrieval grid is indicated with an arrow.

In this figure the behavior of the χ-test shows a minimum in proximity of the retrievals performed with the nominal retrieval grid. The increase of χ-test value when the horizontal resolution improves was predictable, since the value of equation (3.7), for similar values of χ^2, increases as the number of retrieval parameters increases. The explanation of the increase is different when the spatial resolution gets worse in comparison to the nominal case. This behavior is due to the fact that, when the distance among the retrieved profiles increases, the analysis system is not able to reproduce the horizontal structures of the atmosphere within two adjacent retrieved profiles. Therefore in this case the system is forced to approximate the atmospheric field with a gradient then providing worse fits.

Figure 6.2 reports the χ-test values at convergence for H_2O and O_3. For these target species it has been possible a comparison between the data based on real atmospheric spectra (left panel) and the data based on simulated spectra (right panel) formerly derived by Luca Magnani in his PhD thesis [1].

For ozone the χ-test seems to show a minimum in proximity of the retrievals performed with the nominal retrieval grid in both cases. Nevertheless in the real data case it is less pronounced and for high spatial resolution the curve is nearly flat.

For water vapor the curve of the real case is similar to that of ozone but with lower

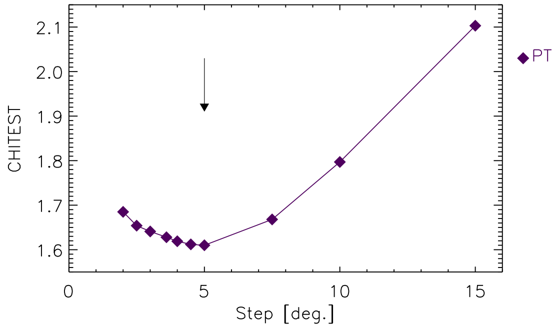

Figure 6.1: χ-test for p, T. The data point referring to the nominal retrieval grid is indicated with an arrow.

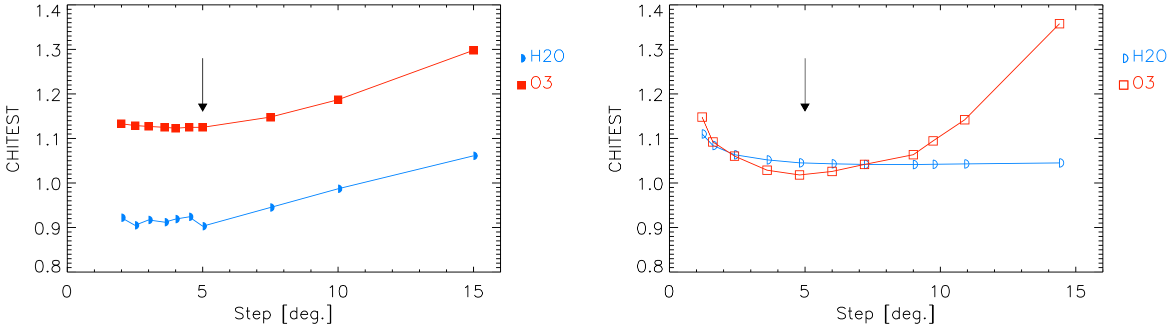

Figure 6.2: χ-test for H_2O and O_3, real (left) and simulated (right) data. The arrows have the same meaning of fig. 6.1.

values, while simulated data have a flat behavior for low spatial resolution. This behavior can be explained by the fact that the distribution of the water vapor obtained with real atmospheric spectra show high resolution structures that where not present in the model used in simulated spectra. Values below unit are caused by an overestimation of the noise level in the orbit file.

Figure 6.3 reports the ESD values (5.4) (expressed in percentage) of the real case for the p, T, H_2O and O_3, averaged on the whole orbit at all the altitudes where the retrieval has been performed: in the panels different altitudes are represented with different colors and different symbols. In the computation of the mean value the too noisy points (values outside the mean $\pm 3\ \sigma$) have been discarded.

A decreasing behavior of the ESDs, for increasing profiles separation, can be clearly appreciated for all the target species. This is because when the spatial resolution improves (and therefore the number of retrieved profiles increases), the number of retrieved parameters and the number of degrees of freedom increases as well .

Figure 6.4 reports the p and T ESD values averaged also in altitude and fig. 6.5 the same kind of plots for ozone and water vapor so as to allow a direct comparison with

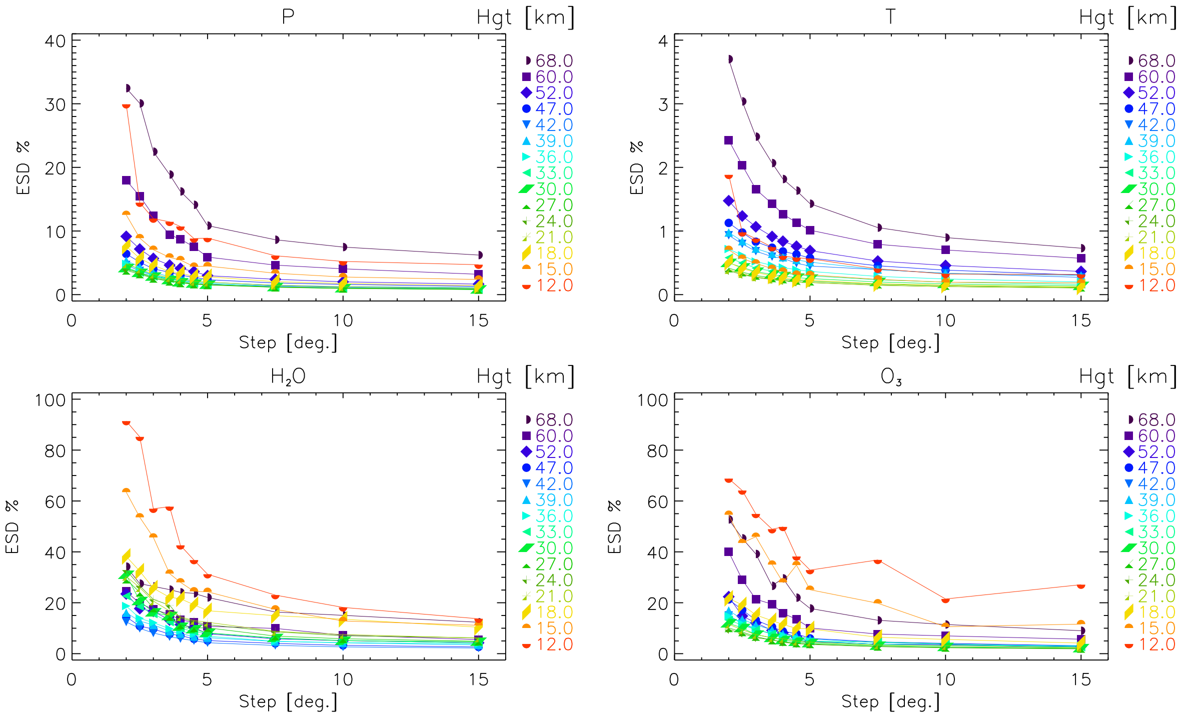

Figure 6.3: ESDs for p (top left), T (top right), H_2O (bottom left), O_3 (bottom right). The legend at the left side of each panel indicates the altitude where the values have been computed.

simulated data.

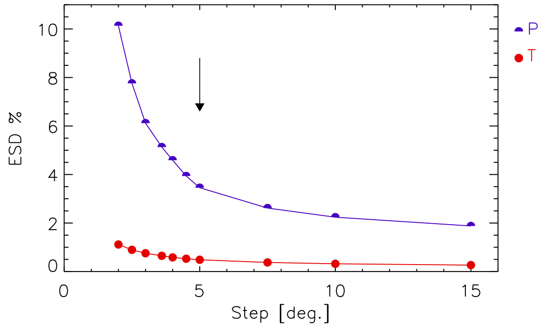

Figure 6.4: ESDs for p and T. The arrow has the same meaning of fig. 6.1

The comparison evidences a very good agreements between real and simulated data. In both plots the water vapor ESDs are larger than those of ozone, due probably to larger water vapor's uncertainty below tropopause.

Pressure, temperature, water vapor and ozone are correlated, so it would be useful [5] to retrieve them together in a Multi Target Retrieval in order to avoid the systematic error arising from the non correct modeling of the other species. Therefore dedicated

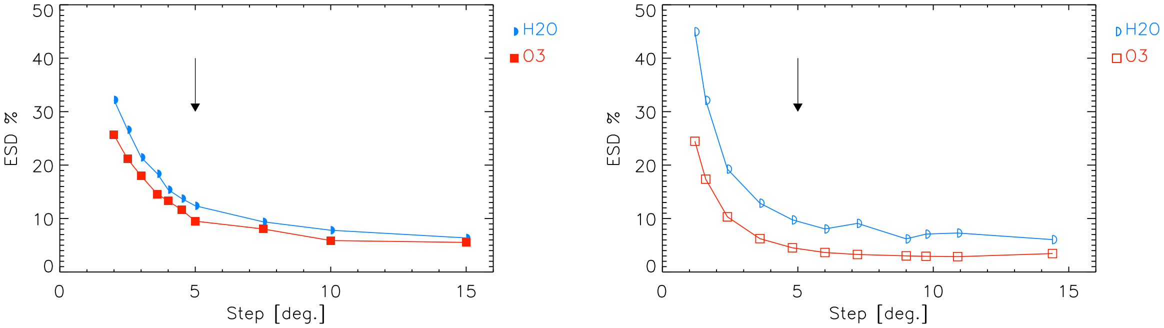

Figure 6.5: ESDs for H_2O, O_3 averaged in altitude, real (left) and simulated (right) data. The arrows have the same meaning of fig. 6.1

MWs for their simultaneous fit have been selected and used in another test performed on the same retrieval grids.

Figure 6.6 reports the χ-test values at convergence for the simultaneous p, T, H_2O and O_3 retrieval.

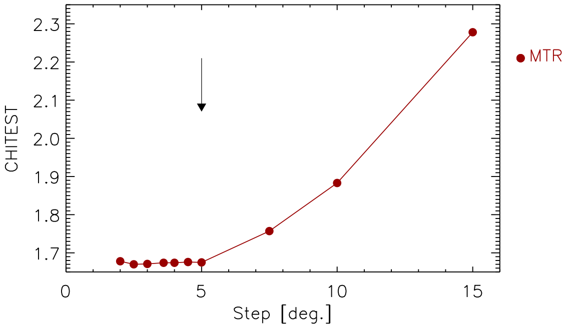

Figure 6.6: χ-test for p, T, H_2O and O_3 retrieved simultaneously. The arrow has the same meaning of fig. 6.1

The χ-test shows a flat behavior when we improve the spatial resolution (in this case the direct increase of the χ-test due to eq. (3.7) is well-balanced by an improvement of the fit) and an increase when we degrade the retrieval grid, as in the previous case. Figure 6.7 reports the ESD values for the p, T, H_2O and O_3 MTR, averaged on the whole orbit at all the altitudes where the retrieval have been performed.

In this second test a decreasing behavior of the ESDs with increasing profiles separation can be appreciated for all the target species, confirming the trend seen before. Furthermore lower ESD values can be noticed in comparison with the test performed without the MTR functionality can also be noticed. In fig. 6.8 are shown the corresponding ESDs averaged over all altitudes.

Other tests have been carried out on the remaining key species routinely retrieved

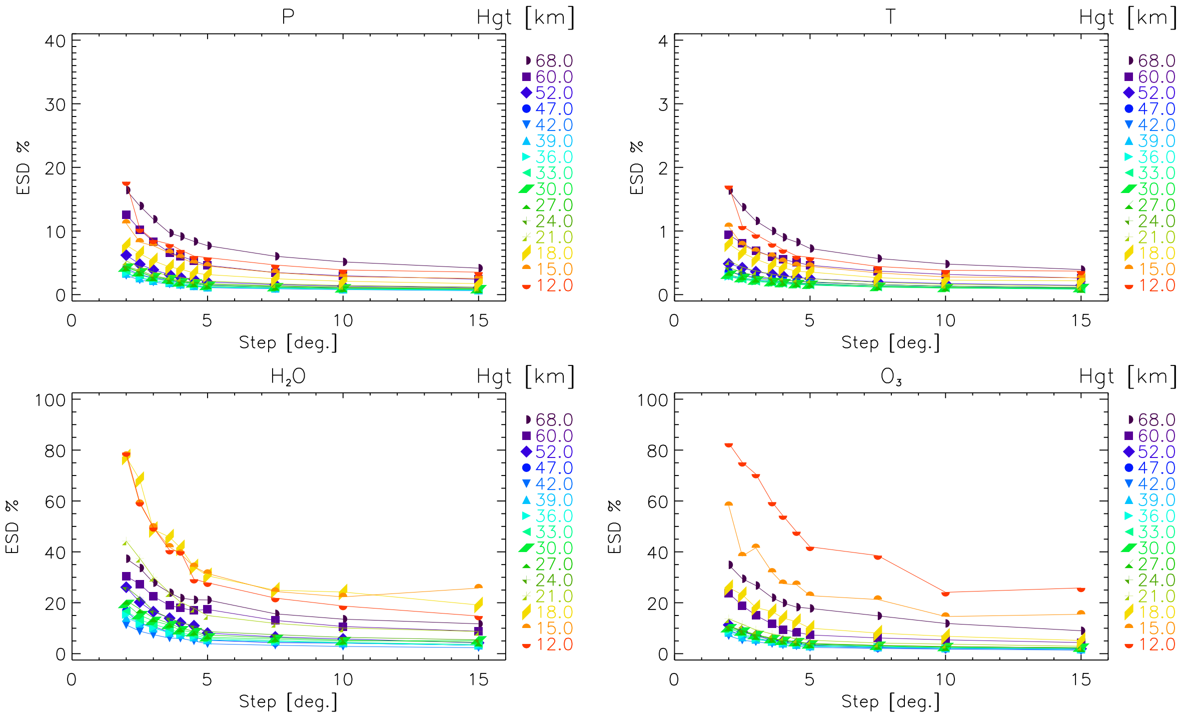

Figure 6.7: ESDs for p (top left), T (top right), H_2O (bottom left), O_3 (bottom right) retrieved simultaneously. The legend at the left side of each panel indicates the altitude where the values have been computed.

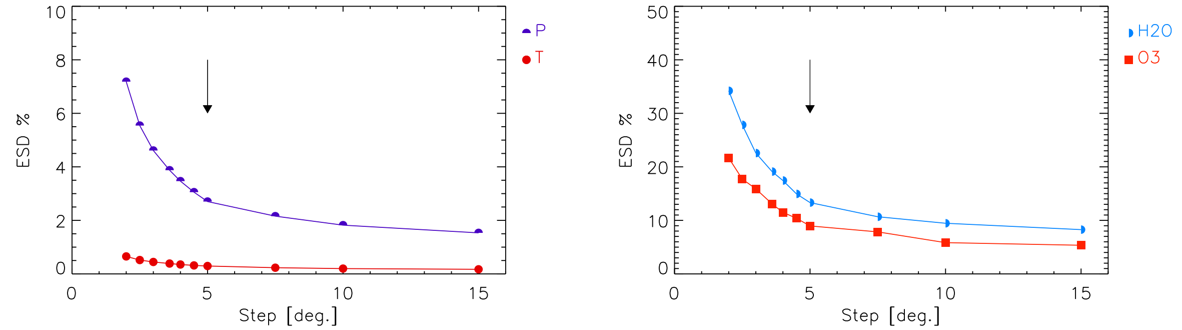

Figure 6.8: ESDs for p, T (left) and H_2O, O_3 (right) averaged in altitude. The arrows have the same meaning of fig. 6.1

by ESA: HNO_3, CH_4, N_2O and NO_2. These species have been retrieved sequentially in the specified order assuming, for the species already retrieved, the values computed previously.

Figure 6.9 reports the χ-test values at convergence for these tests (left panel corresponds to real data, right panel to simulated data). NO_2 results are reported in a separated plot (6.10) as we do not have simulated data for a direct comparison for this specie.

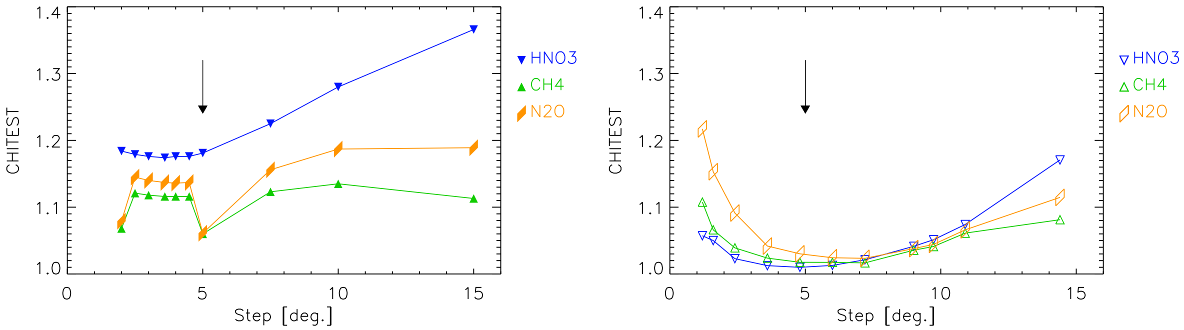

Figure 6.9: χ-test for HNO_3, CH_4 and N_2O, real (left) and simulated (right) data. The arrows have the same meaning of fig. 6.1

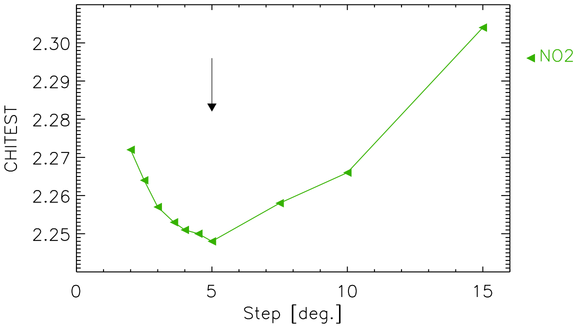

Figure 6.10: χ-test for NO_2. The arrow has the same meaning of fig. 6.1

In general, results obtained with real data show larger values than those obtained with simulated data. A discontinuity is observed in the results obtained with spatial resolution higher than nominal case for the species N_2O and CH_4 in the left panel of fig. 6.9. This may be explained by a different degree of convergence reached for different spatial resolutions. For HNO_3 the minimum is reached for values of the spatial resolution slightly higher than the nominal case. NO_2 (fig. 6.10) shows a clear minimum in coincidence with the nominal retrieval grid.

Figure 6.11 reports the ESD values for HNO_3, CH_4, N_2O and NO_2, averaged on the whole orbit at all the altitudes where the retrieval has been performed.

Generally we can state that an increase of the number of retrieved profiles (operation that leads to an improvement in the spatial resolution) results in a retrieval precision which is still acceptable up to a factor 2 of improvement. This is a confirmation of the conclusion made with the simulated data.

A detailed overview of the quality of the performed retrievals can be obtained with 2-D maps. These maps, shown in figs. 6.14, 6.16 and 6.18 report the retrieved values of all the target species as a function of the altitude (vertical scale) and the orbital coordinate (horizontal scale). In definition of the orbital coordinate the North Pole is at

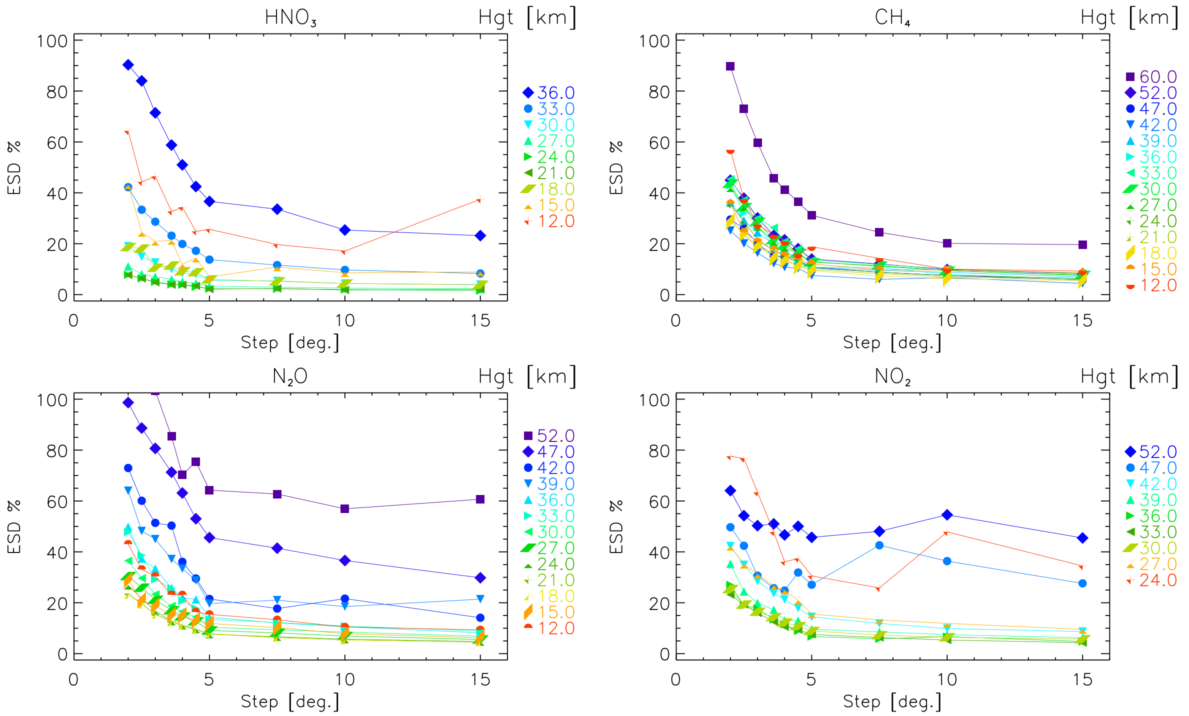

Figure 6.11: ESDs for HNO_3 (top left), CH_4 (top right), N_2O (bottom left), NO_2 (bottom right).

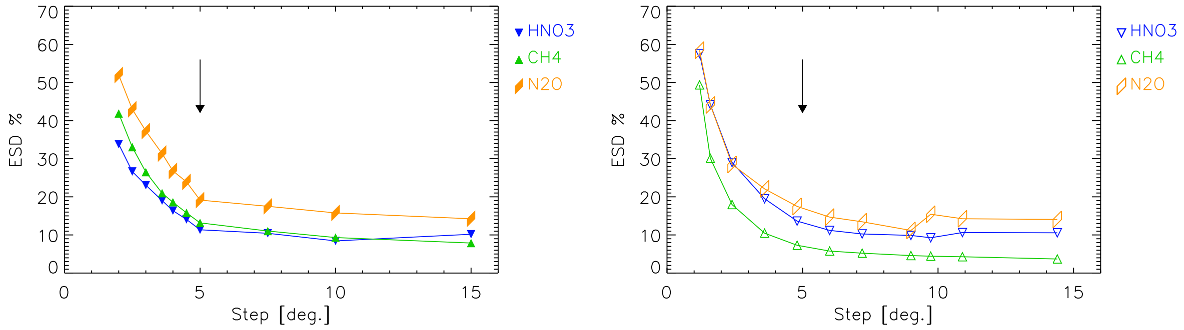

Figure 6.12: ESDs for HNO_3, CH_4, N_2O averaged in altitude, real (left) and simulated (right) data. The arrows have the same meaning of fig. 6.1

0°. The south pole is at 180° and the equator at 90° and 270°. In all maps the vertical white strips cover cloudy regions and the retrieval grid is displayed with white circles. The profiles's separation is reported in the figure captions.

Equivalent maps for the percentage ESDs are shown in figs. 6.15, 6.17 and 6.19. It can be seen that the most evident effect if we try to use a too fine retrieval grid is the presence in the maps of some oscillations. This is caused mainly by the fact that MIPAS observations do not contain enough information to retrieve profiles with horizon-

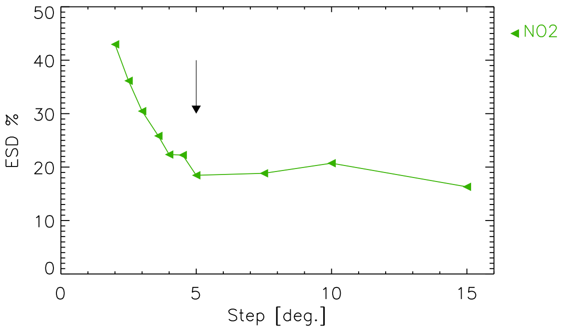

Figure 6.13: ESDs for NO_2. The arrow has the same meaning of fig. 6.1

tal separation smaller than the nominal separation between the measured scans. A less important contribution to that behavior is given by the fact that the retrieved profiles are periodically closer to the tangent points of the observations, where the maximum of the information is concentrated.

This effect can be observed for all the species considered and its size varies slightly from specie to specie.

Looking at the ESDs maps a general consideration can be made: as the spatial resolution improves (fig. 6.18) the precision drops in all the orbit and for all the altitudes, fig. 6.19.

If we degrade the horizontal resolution (fig. 6.14) an overall improvement in the ESDs (fig. 6.15) happens but, on the other hand, we lose some details of the retrieved quantity specially when a strong horizontal gradient is present.

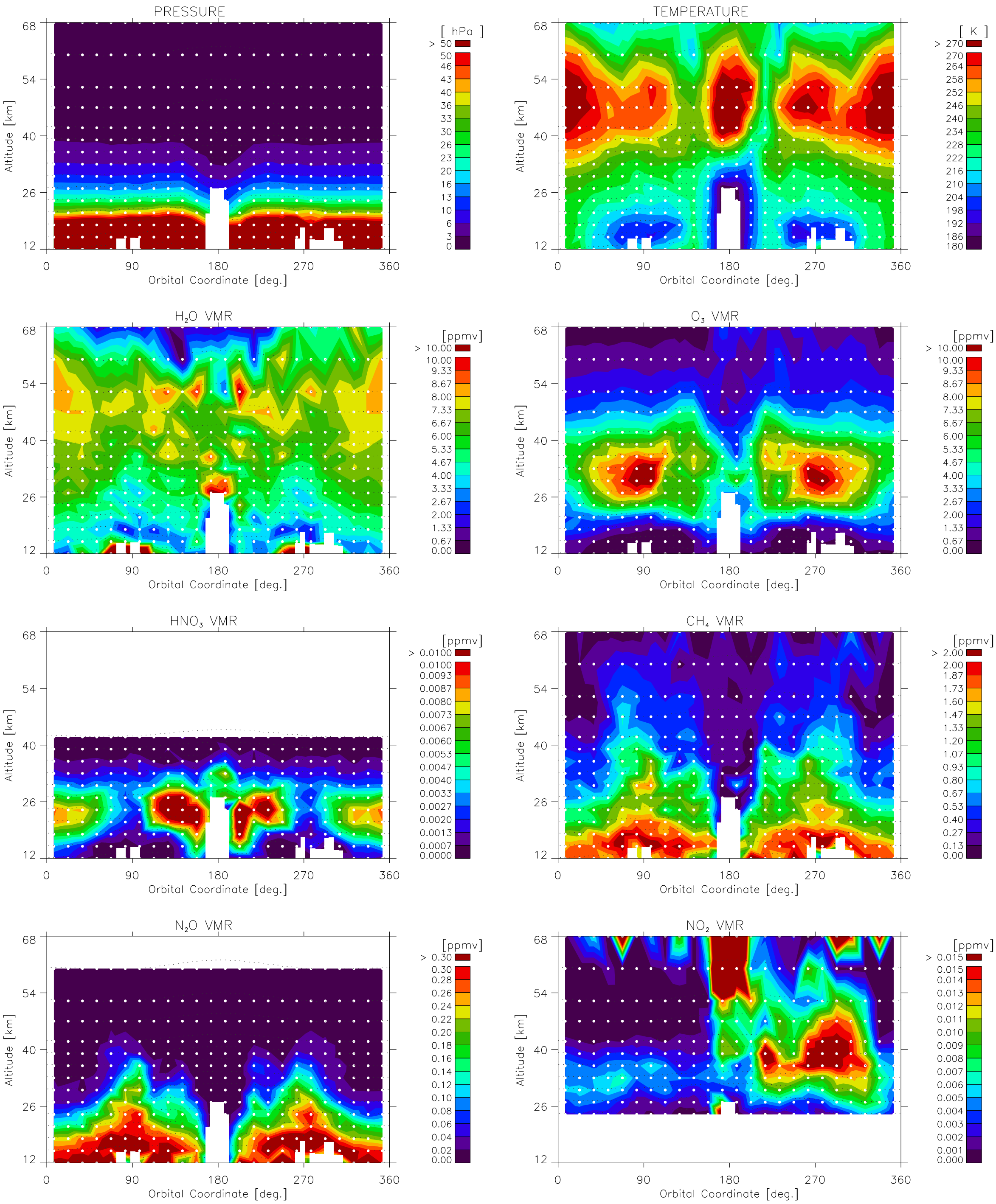

Figure 6.14: Set of maps that report the values of p, T and VMRs for a retrieval grid spaced by 15.0° between two consecutive profiles.

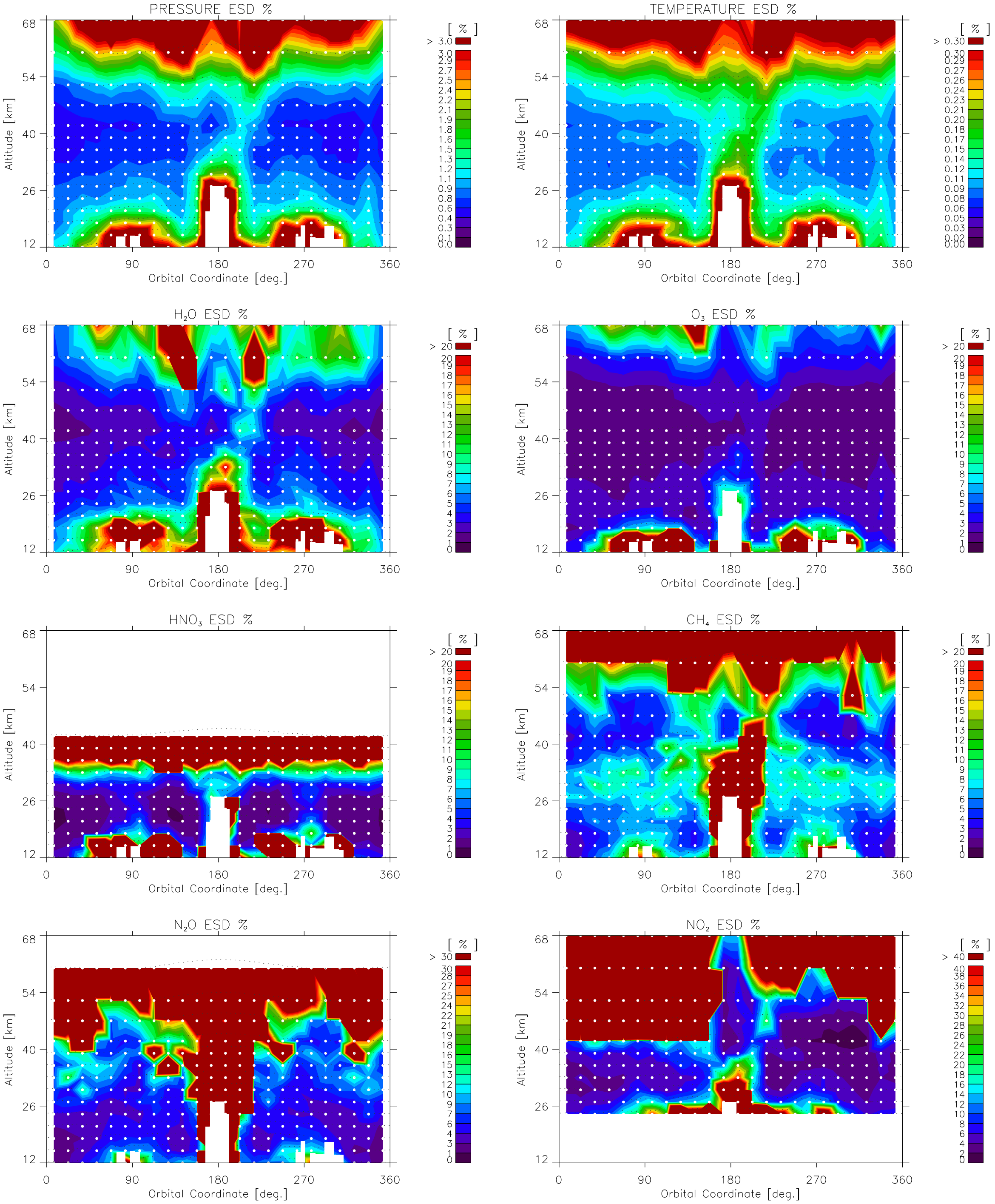

Figure 6.15: Set of maps that report the values of ESDs for a retrieval grid spaced by 15.0° between two consecutive profiles. The target specie is reported on top of figures.

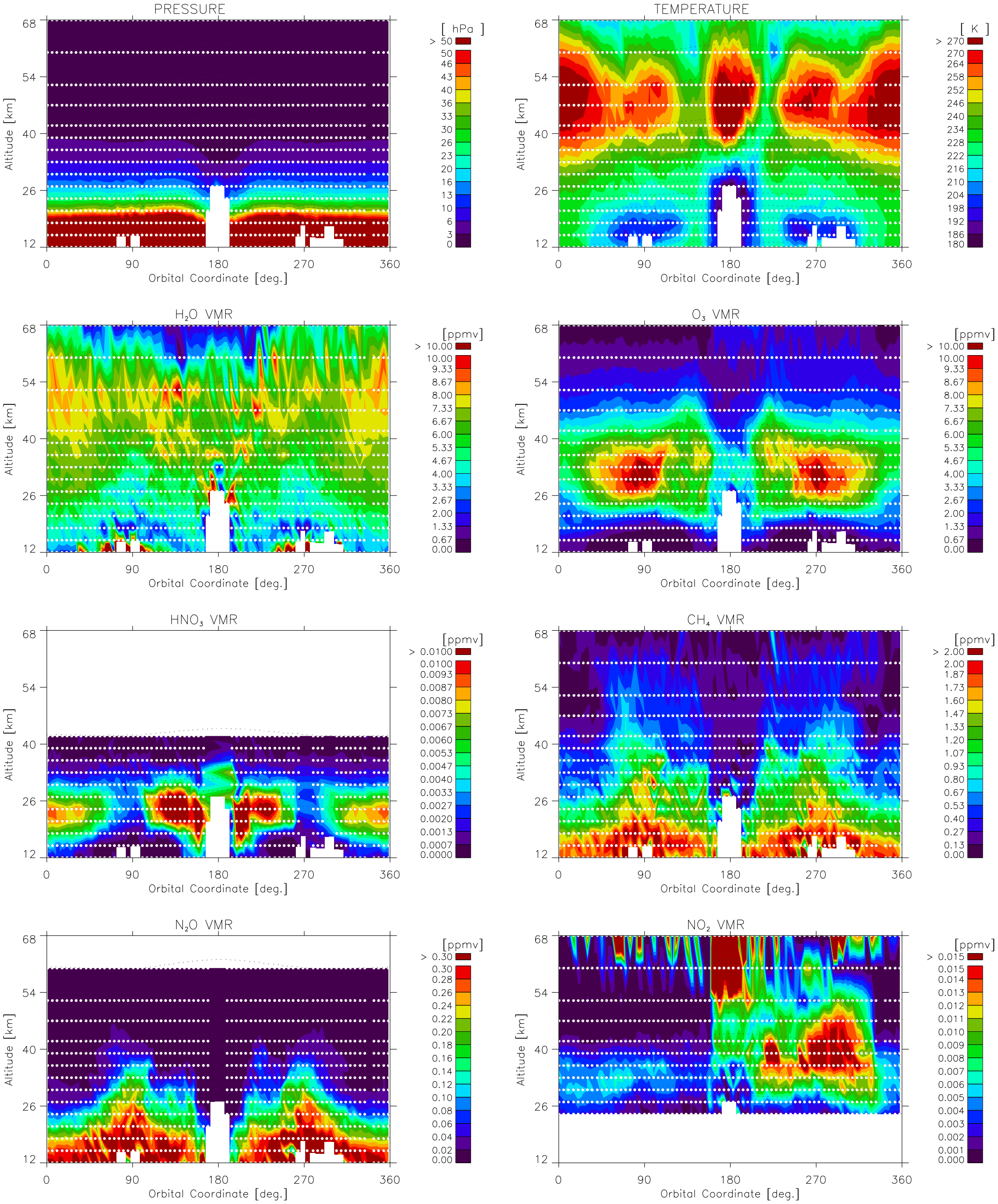

Figure 6.16: Set of maps that report the values of p, T and VMRs for a retrieval grid spaced by 5.0° between two consecutive profiles (nominal retrieval grid).

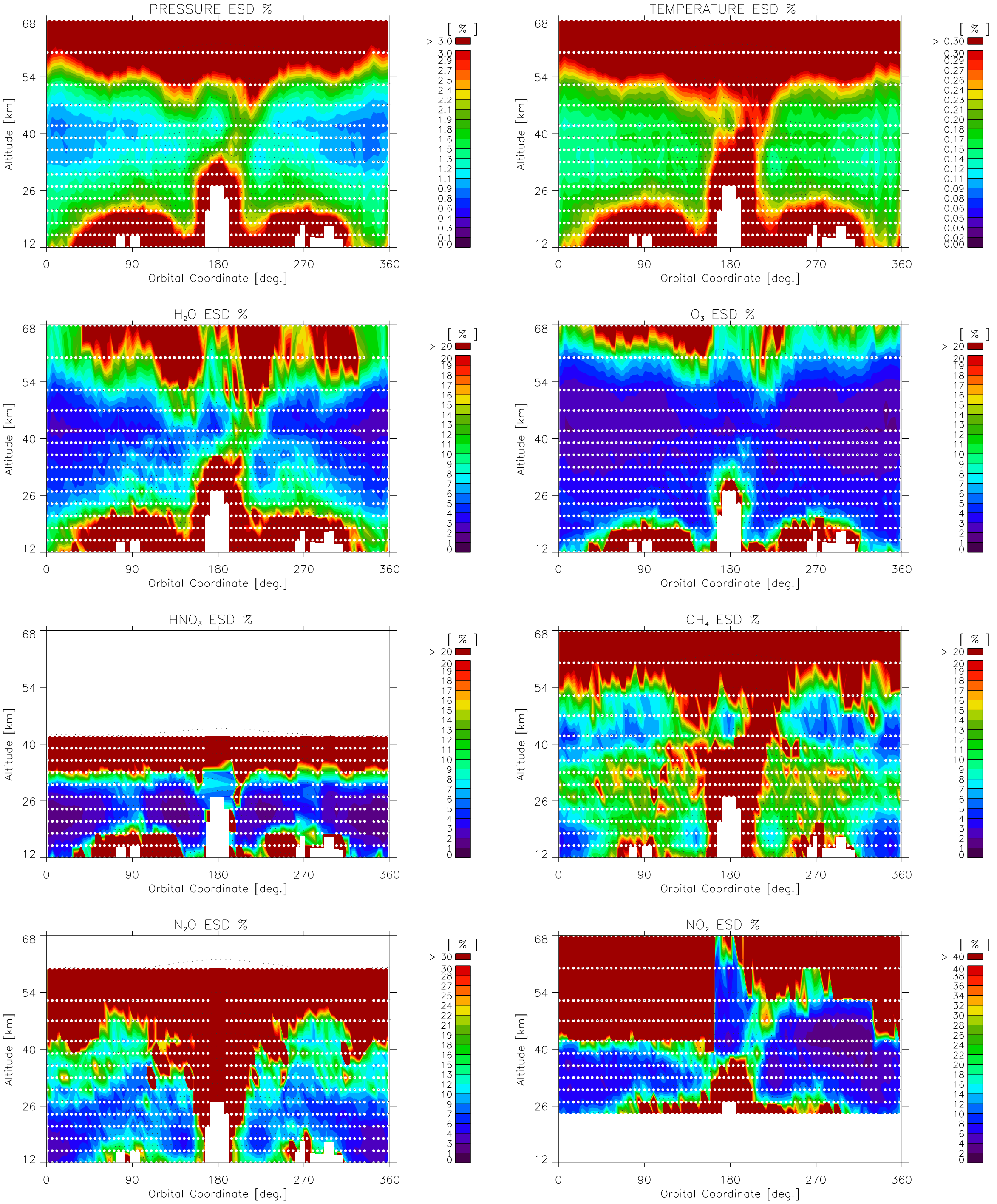

Figure 6.17: Set of maps that report the values of ESDs for a retrieval grid spaced by 5.0° between two consecutive profiles (nominal retrieval grid).

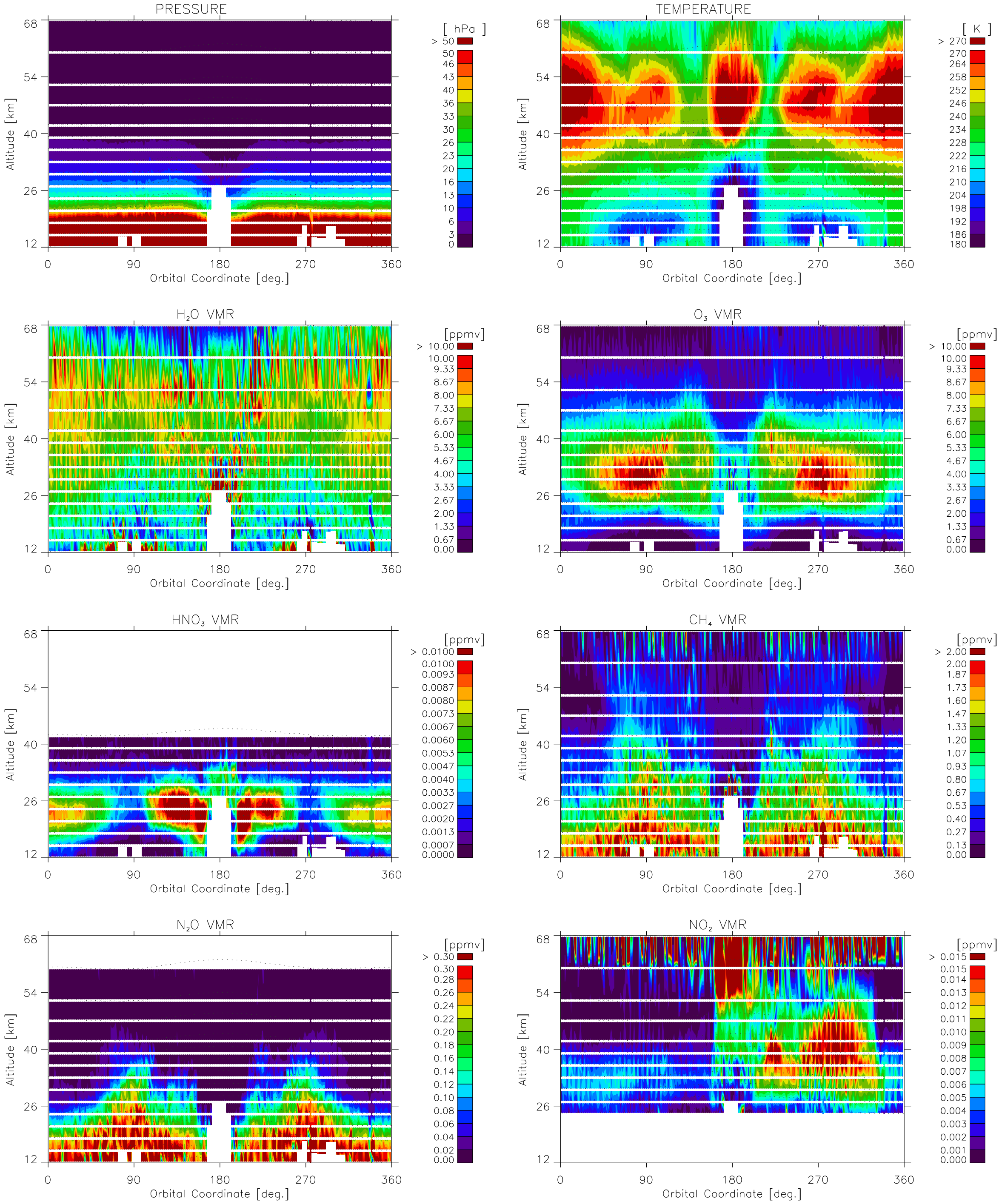

Figure 6.18: Set of maps that report the values of p, T and VMRs for a retrieval grid spaced by 2.0° between two consecutive profiles.

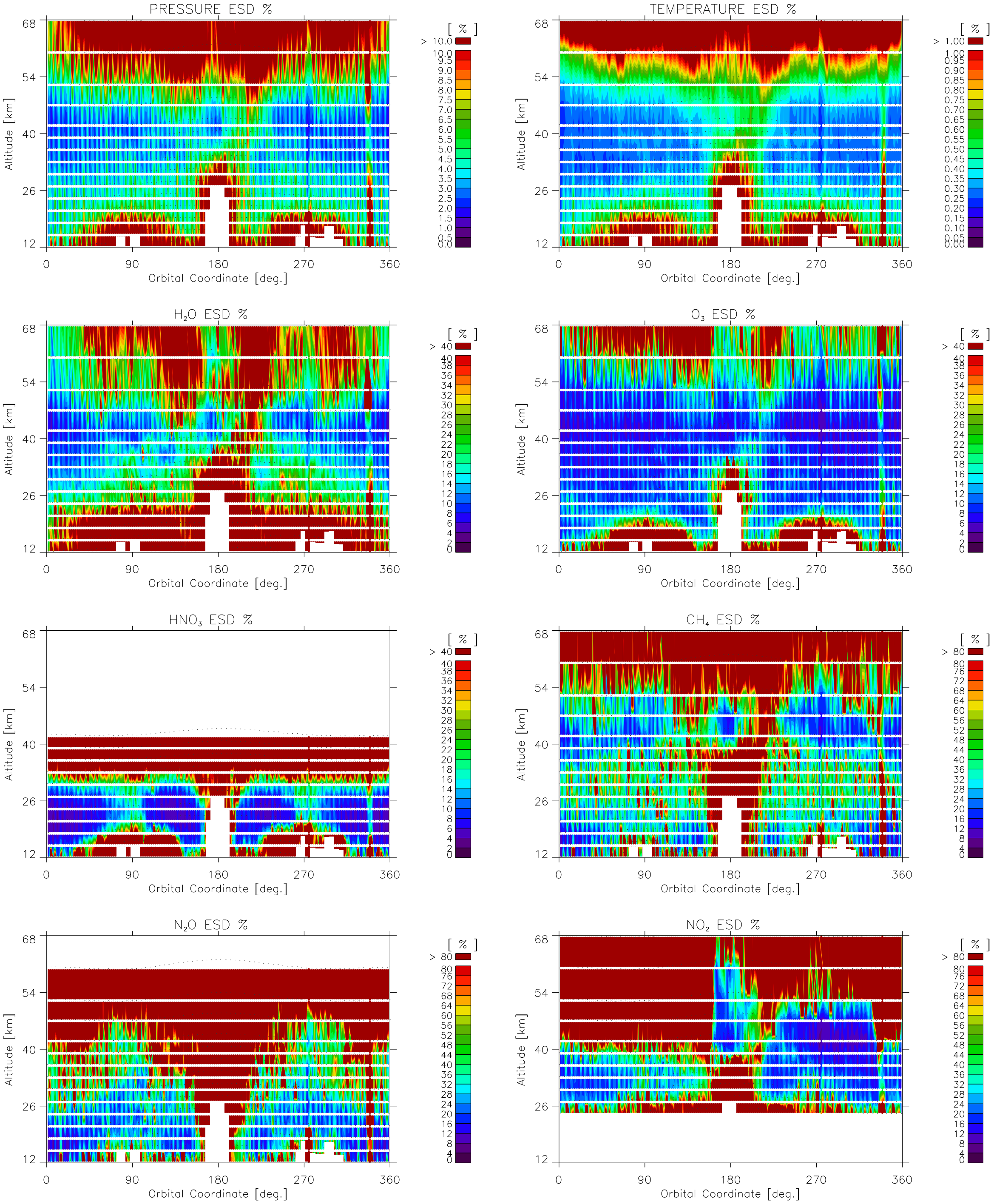

Figure 6.19: Set of maps that report the values of ESDs for a retrieval grid spaced by 2.0° between two consecutive profiles.

6.2 Trade-off studies varying the measurement grid

A second possibility to perform a trade-off study between spatial resolution and retrieval precision is to vary the measurements grid modifying the spectral resolution.

The MIPAS interferometer is designed to acquire interferograms with a constant speed of the interferometric mirror. This speed is determined by the sensitivity of the detectors and cannot be increased beyond the value used in the nominal measurement scenario. With this constraint, the possible strategies to increase the number of measured limb-scans along the orbit consist in:

i limiting the altitude range of the acquired limb-scans,

ii decreasing the vertical sampling of the limb-scans,

iii decreasing the spectral resolution (i.e. the maximum path difference of the measured interferograms).

In all cases the time interval required for the measurement of a limb-scan turns-out to be reduced with respect to the nominal measurement scenario: in cases [*i*] and [*ii*] because fewer sweeps per limb-scan are measured, in case [*iii*] because the time interval required to measure the individual interferograms is reduced compared to the case of maximum spectral resolution. Within the unconstrained least squares approach, strategy [*i*] is not practical because the retrieval of profiles from limb-scans covering a limited altitude range was proven to be affected by large systematic errors due to the assumption of the profile shape above the uppermost and below the lowermost (due to the finite instrument FOV) retrieval altitudes [113, 114]. Strategy [*ii*] gains horizontal resolution at the cost of vertical resolution. Therefore only strategy [*iii*] allows a complete comparison with the results obtained varying the retrieval grid.

MIPAS has not been built for this kind of experiment but an unfortunate event that occurred to the interferometer changed the original plans. After two years of nearly continuous measurements (from July 2002 to March 2004) at the end of March 2004 the instrument was stopped due to problems with the mirror drive of the interferometer. After that, series of tests have been performed with reduced maximum optical path difference (8 cm maximum OPD), corresponding to both a lower spectral resolution (0.0625 cm^{-1} instead of 0.025 cm^{-1}) and a shorter measurement time (1.8 sec/sweep instead of 4.5 sec/sweep).

In particular, during August 2004 MIPAS operated with the new resolution and the old measurement scenario (17 sweeps/sequence coincident with the same tangent altitudes). Since the low resolution (LR) spectra are acquired in only 40% of the time of high resolution spectra this allows limb-scans to be obtained every 200 km rather than every 500 km.

Since the beginning of 2005 the scan pattern has been changed to increase the number of tangent heights per profile from 17 to 27, while reverting to the original horizontal resolution.

6.2.1 Tests

The tests performed in August 2004 allow us to study the impact of a finer measurement grid on the retrieved values and ESDs. The orbit number 12858 of the 15th August has been selected for our purposes.

Figure 6.20 gives the ESDs quantifier obtained by the retrieval of pressure, temperature, water vapor and ozone in comparison to the same quantifiers obtained by the analysis of the full resolution orbit 2081.

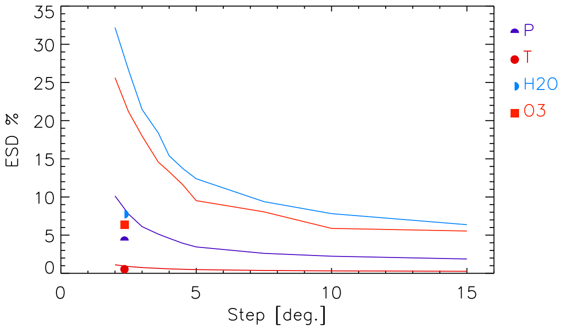

Figure 6.20: ESDs quantifiers for p, T, H_2O, O_3. Symbols correspond to the LR data

The corresponding pressure, temperature, water vapor and ozone VMRs and ESDs maps are showed in comparison with orbit 2081 through figs. 6.21, 6.22. Each of these two panels contains: the maps for the orbit 12858 retrieved with its nominal retrieval grid (profiles spaced by 2.35°) on the left-hand side and the maps for the orbit 2081 retrieved the nominal retrieval grid of the orbit 12858 on the right-hand side.

For the evaluation of the results shown in figs. 6.20, 6.21 and 6.22 it must be recalled that different observations have been analyzed by the two retrievals because the selection of the optimal spectral intervals to be analyzed depends on the spectral resolution of the measurements; however the total number of analyzed spectral points is comparable in the two retrievals.

Apart from these considerations, fig. 6.22 clearly shows that, in order to gain horizontal resolution, the strategy based on the reduction of spectral resolution (and then of the measurement time) is advantageous with respect to the strategy that acts on the retrieval grid. This outcome confirms the indication reported in [1] that was obtained with simulated observations and making use of several assumptions. A prominent feature in the right-hand panel of figs. 6.21 and 6.22 is the fine color alternation that appears, at

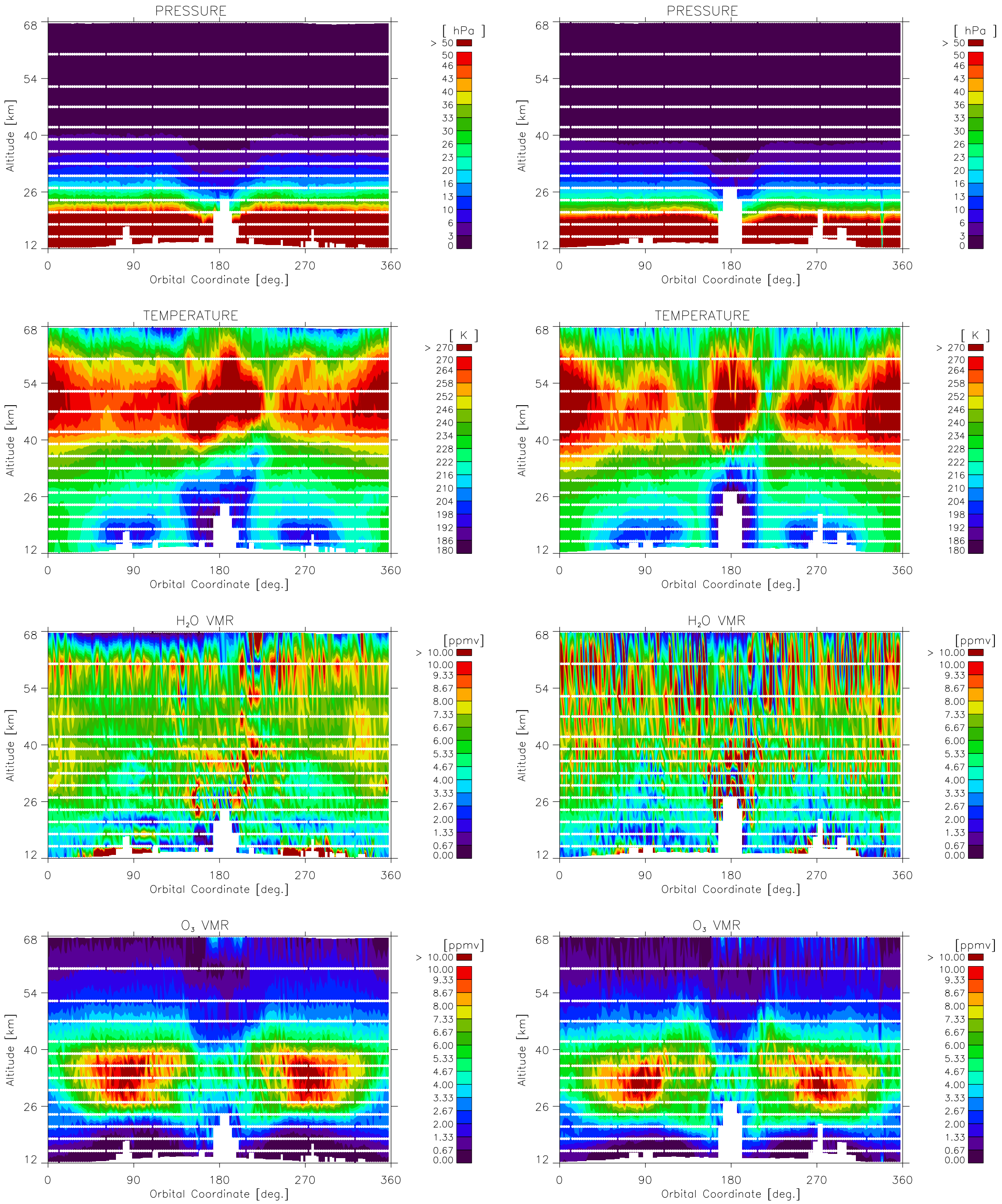

Figure 6.21: Set of maps that report the values of p, T and O_3, H_2O VMRs of orbit 12858 (left) and orbit 2081 retrieved with the nominal retrieval grid of orbit 12858 (right).

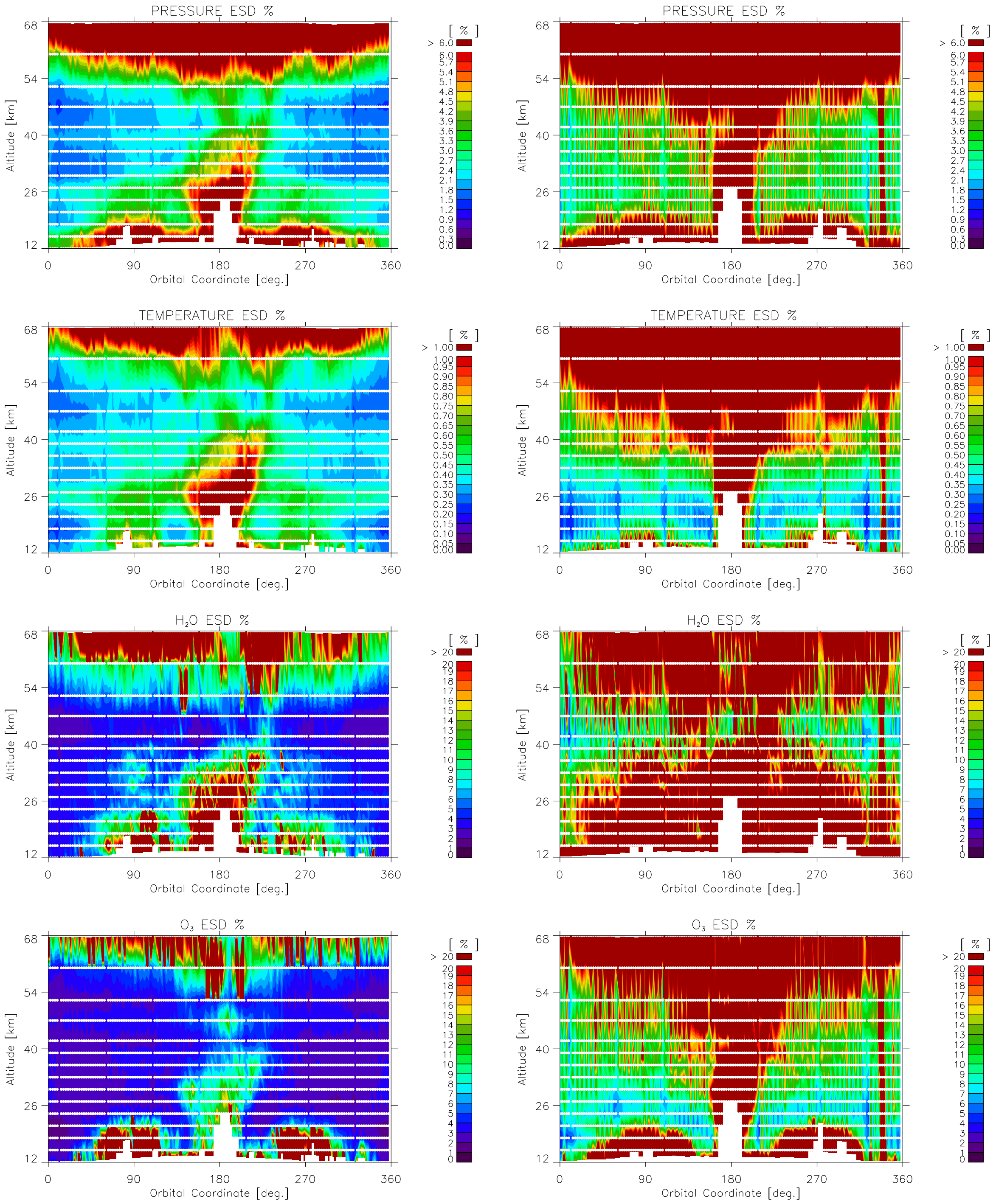

Figure 6.22: Set of maps that report the values of p, T, O_3 and H_2O ESDs of orbit 12858 (left) and orbit 2081 retrieved with the nominal retrieval grid of orbit 12858 (right).

most altitudes, along the orbital coordinate. The horizontal oscillation, although present, is much less pronounced in the left-hand panel of these panels.

This behavior is the result of two concurrent effects:

1. the random coincidence between the OC of the retrieved profiles with the OC of the observations where the information is largest (that is where the weighting functions peak).
2. the anti-correlation between values at the same altitude, that increases when the adjacent profiles get closer. The first effect is expected to be minimum with the reference retrieval grid (as in the case of the left panel of fig. 6.22) and maximum when this same grid is increased by a factor 2 (rather close to the case of the right panel of fig. 6.22).

6.3 S6 Observation mode

A even finer measurement grid can by found in the S6 observation mode of the original measurements configuration that has been designed to study important dynamical processes in the upper troposphere (UT) and lower stratosphere (LS) in greater detail. S6 is achieved by scanning from 6 to 35 km tangent height (in 2 km steps up to 24 km) and with a spacing along-track of about 112 km, at the expenses of reducing spectral resolution to 0.1 cm^{-1}.

The UT/LS region is characterized by strong horizontal and vertical gradients in temperature and constituents which can be modelled only by adopting a 2-D representation of the atmosphere in the retrieval model. Therefore the GMTR is suitable to perform an analysis on this observation mode and to draw some conclusions about a so thick measurement grid.

Figure 6.23 show the results obtained for a Multi Target Retrieval of pressure, temperature, water vapor and ozone with the nominal retrieval grid of the Special Mode S6 (profiles spaced by 1.4°). The horizontal oscillations, well visible in the figure, are clearly unphysical and appear both in the VMR maps and in the corresponding ESDs maps whose values raise by effect of the strong correlations. The size of the oscillations can be evaluated in figs. 6.24 where the solid line reports the O_3 VMR (left) and ESD (right) values at the altitude of 27 km limited (for clarity) to an orbit segment going from the north pole to the equator.

It is worth to point out that the instabilities induced by the horizontal correlations appear and can be evaluated only by a 2-D retrieval algorithm. In this respect a direct analogy is with the oscillation of the altitude profiles when an atmospheric over-sampling has been operated in the vertical domain.

The strategy of thinning out the retrieval grid has been studied in order to avoid unphysical oscillations in the retrieved horizontal distribution of atmospheric parameters.

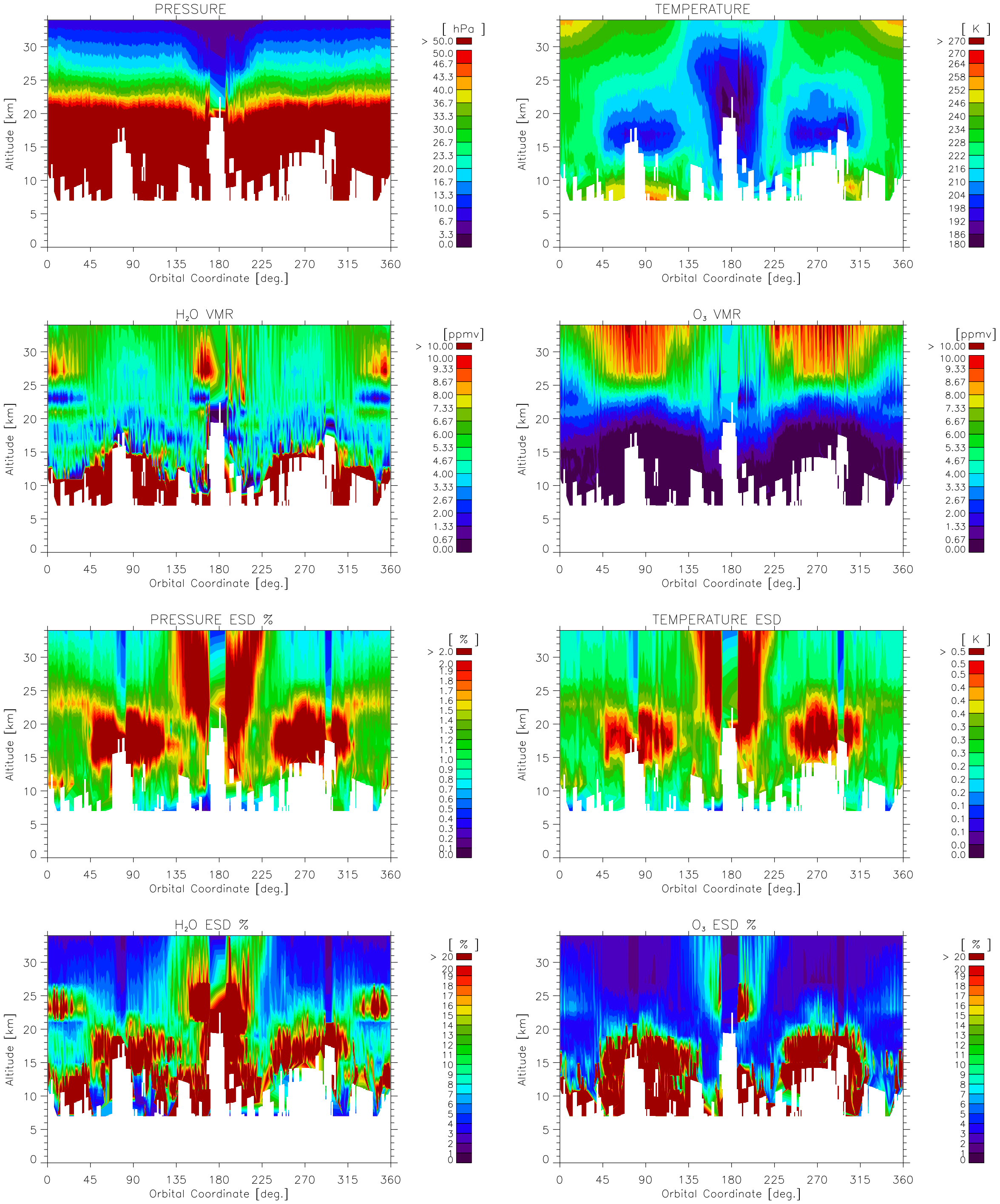

Figure 6.23: Set of maps that report the values of p, T, VMRs and ESDs for a retrieval grid spaced by 1.4° between two consecutive profiles (nominal retrieval grid).

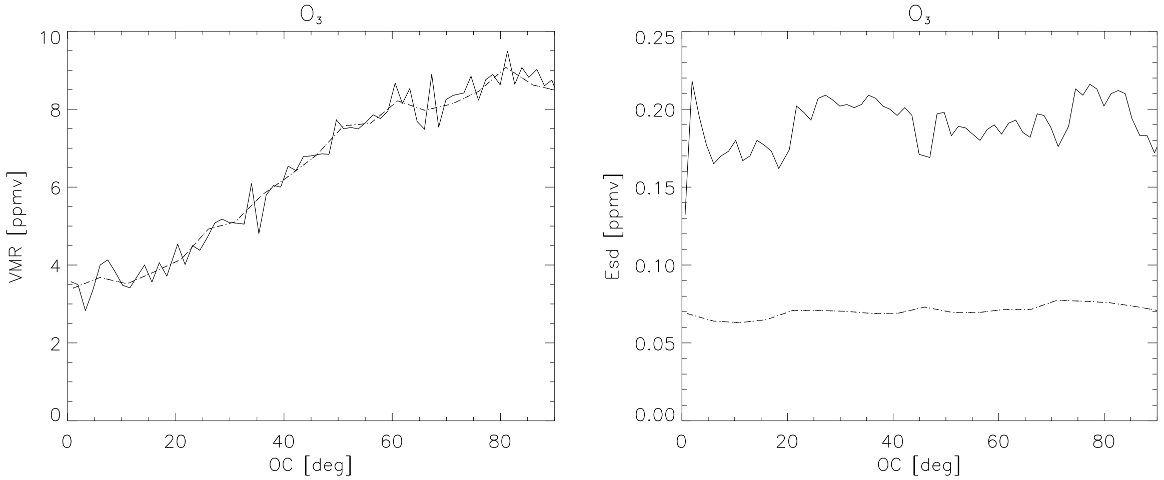

Figure 6.24: VMR (left) and ESD (right) values for an orbit segment going from the north pole to the equator at 27 km tangent height. The solid lines report a retrieval performed with a grid spaced by 1.4° between two consecutive profiles (nominal retrieval grid) while the dashed-dotted lines a retrieval grid spaced by 5.0° between two consecutive profiles.

For this purpose we use the behavior of the horizontal distribution at 27 km, reported in fig. 6.24, that is representative of the general issue. In this figure is also reported the smoothed distributions obtained by retrieving ozone profiles separated by 5.0° with dotted-dashed line. Furthermore, this figure shows that the strategy of widening the retrieval grid leads to a rather constant precision of the retrieval products.

Chapter 7

CO_2 Distribution in Stratosphere from MIPAS Measurements

Global warming is the observed increase in the average temperature of the Earth's atmosphere and oceans in recent decades and its projected continuation into the future. Global average near-surface atmospheric temperature rose $0.6 \pm 0.2\,^{\circ}C$ in the 20^{th} century. The main cause of the human-induced component of warming is the increased atmospheric concentration of greenhouse gases such as carbon dioxide (CO_2), which leads to warming of the surface and lower atmosphere by increasing the greenhouse effect. Greenhouse gases are released by activities such as the burning of fossil fuels, land clearing, and agriculture. For this reason, a monitoring of CO_2 concentration from satellite measurements would be very useful. Furthermore CO_2 spectral signatures are often used to retrieve the p-T profiles that are afterward used to retrieve all the target species. Nevertheless a retrieval with acceptable errors (comparable with those obtained with other instruments) is difficult to be performed. With some opportune expedients we aim to lower as much as possible the retrieval errors. This chapter describes the strategy adopted to perform the CO_2 retrieval from MIPAS measurements: the line mixing implementation [2] in the GMTR retrieval code and the use of the target-dependent retrieval grid are important features for this study. This chapter also describes the microwindows selection process for the simultaneous retrieval of pressure, temperature and CO_2 with the MW-MAKE program [115] and the subsequential creation of lookup tables and irregular grids. Finally we show the results obtained processing all the full resolution MIPAS mission.

7.1 Reasons to investigate the CO_2 distribution in the Atmosphere

7.1.1 Environmental relevance

When sunlight reaches Earth's surface some is absorbed and warms the earth. Because the Earth's surface is much cooler than the sun, it radiates energy at much longer wavelengths than the sun. Some of these longer wavelengths are absorbed by greenhouse gases in the atmosphere before they are lost to space. The absorption of this longwave radiant energy warms the atmosphere (the atmosphere also is warmed by transfer of sensible and latent heat from the surface).

Greenhouse gases also emit longwave radiation both upward to space and downward to the surface. The downward part of this longwave radiation emitted by the atmosphere is the "greenhouse effect". The term is inappropriate, as this process is not the primary mechanism that warms greenhouses. Without the greenhouse effect life on this planet would probably not exist as the average temperature of the Earth would be a chilly -18° Celsius, rather than the present 15° Celsius.

The major natural greenhouse gases are water vapor, which causes about 36-70% of the greenhouse effect on Earth (not including clouds); carbon dioxide, which causes 9-26%; methane, which causes 4-9%, and ozone, which causes 3-7%. Note that it is not possible to state that a certain gas causes a certain percentage of the greenhouse effect, because the influences of the various gases are not additive.

The major atmospheric constituents (N_2 and O_2) are not greenhouse gases, because homonuclear diatomic molecules (e.g. N_2, O_2, H_2) neither absorb nor emit infrared radiation as there is no net change in the dipole moment of these molecules.

All of the major greenhouse gases have increased in concentration since the beginning of the Industrial Revolution. As a result of these higher concentrations, scientists predict that the greenhouse effect will be enhanced and the Earth's climate will become warmer. Predicting the amount of warming is accomplished by computer modeling. Computer models suggest that a doubling of the concentration of the main greenhouse gas, carbon dioxide, may raise the average global temperature between 1 and 3° Celsius.

A number of gases are involved in the human activities causes enhancement of the greenhouse effect; in particular the increase of carbon dioxide in the atmosphere is constant. Prior to 1700, carbon dioxide concentration was about 280 parts per million while in the year 2006 it reached about 385 parts per million.

7.1.2 Retrieval relevance

Since the spectroscopic parameters of CO_2 are quite accurate and its VMR is well known, the CO_2 spectral signatures are often used to retrieve the pressure and temperature profiles (and in some cases the pointing) from atmospheric measurements in the infrared region.

This is also the case for the inversion of MIPAS measurements, operated with both the ORM and the geo-fit algorithms. Both ORM and geo-fit derive pressure and temperature analyzing a subset of the MIPAS spectrum selected with the purpose to provide the best compromise between “high” sensitivity to the target parameters and “low” contribution to the systematic errors.

As the target species are retrieved from pressure and temperature a underestimate in the assumption of the CO_2 constancy could lead to systematic error in all the subsequent target species retrieval.

Despite its importance, the literature is poor of data about CO_2 distribution in the stratosphere, therefore a strategy that allows to retrieve the CO_2 from satellite measurements with low errors would be very useful: as we will see in section 7.3.5 with all the possible strategies up to now the size of the systematic errors is still too high to draw some absolute conclusions. Nevertheless if we compare CO_2 distributions measured at different times and/or locations a similar impact of the systematic errors is expected on the different CO_2 retrievals.

The following sections will describe the strategy adopted for the CO_2 retrieval from MIPAS measurements step by step.

7.2 Line Mixing implementation in the GMTR

In this section it is described why is the line-mixing effect important to be implemented in the analysis codes which operate nearby the 15 μm spectral region.

Among the possible spectral intervals, the intense CO_2 Q branches in the 15 μm region are promising due to both their high signal-to-noise ratio and their high sensitivity to pressure and temperature. In particular, the wings of the intense Q branches are attractive because of their high sensitivity to pressure coupled with the weak dependence on the instrument line-shape function. Indeed, in the line wings the absorption cross-section is proportional to p^2 and varies slowly with frequency. Furthermore, the CO_2 line wings have particularly narrow weighting functions, so their use is expected to enhance the vertical resolution of the retrieved profiles. Despite the above-mentioned arguments in favor of these spectral features, CO_2 Q branches are generally avoided for the retrieval of p and T. The reason for excluding Q branches stands in the complexity of the models that are necessary for these spectral features that are strongly affected by line-mixing effects.

In the past, the implementation of line-mixing models required a considerable computational effort for the calculation of the absorption cross-sections. Cross-section look-up-tables (LUTs) are today often adopted in the retrieval algorithms; they significantly reduce the CPU demand making possible the use of refined and complex models for the computation of CO_2 line-mixing. Line-mixing effects are evident not only in the Q branches of CO_2 but also in the line wings of P- and R-type transitions; the necessity of globally improving the modeling of these spectral features is now recognized as recently pointed out by the analysis of the far wings of the CO_2 $\nu 2$ band.

Recently, a model has been developed for the line-mixing effect in P, Q, and R transitions of all the vibrational bands of CO_2. This model, based on the energy corrected sudden approach, reproduces the observations with very good accuracy as demonstrated by comparisons of predictions with measured laboratory and atmospheric spectra and F. Niro et al. [2] have implemented the line-mixing model of in the operational geo-fit analysis system. This new feature allows retrieval of p and T profiles exploiting entire Q branches as well as line wings of individual CO_2 transitions.

7.3 Microwindow Selection

7.3.1 Introduction

As satellite spectra typically provide thousands of radiance measurements every second, a retrieval scheme including all the radiance points is currently not compliant with memory and computer speed resources. Consequently, attention has been focused on methods for determining the optimum subsets of such spectra which contains most of the potential information. This subsets are called Microwindows (MWs).

One approach is to simulate the propagation of random noise though the retrieval and select measurements which maximize the information content or degrees of freedom of the signal [116, 117], or which best satisfy other criteria [118]. While this is reasonable if the retrieval errors are predominantly due to random errors and/or errors in the *a priori* estimate, it does not allow for other (systematic) sources.

An alternative approach [98, 119] address this problem by selecting microwindows which minimize the total error, including both random and systematic components. However, this is only achieved by approximating the profile retrieval as a set of independent, single-layer retrievals. As a result, the effect of inter-level correlations an the actual retrieval is ignored. There is also the practical difficulty of consolidating the microwindows derived independently for each profile level into single microwindows applicable over the whole profile.

Microwindows for the operational MIPAS retrieval were selected using a third approach which includes simulating a full profile retrieval including the propagation of systematic error terms [107]. As such, it can be considered either as an extension of the first ('random error') method to include systematic errors, or as a multi-layer version of the second ('single-layer') method.

As well as selecting microwindows, this scheme also allows existing microwindows to be ordered according to different criteria and provides a complete error analysis of the resulting retrieval products.

7.3.2 Retrieval Model

A linear approximation can be used to relate the retrieved state vector $\mathbf{x}$ to set of measurements $\mathbf{y}$ via the *Gain Matrix* $\mathbf{G}$ [85]

$$\mathbf{x} = \mathbf{G}\mathbf{y} \tag{7.1}$$

$$\mathbf{G} = \mathbf{S}_a\mathbf{K}^T(\mathbf{S}_y + \mathbf{K}\mathbf{S}_a\mathbf{K}^T)^{-1} \tag{7.2}$$

where $\mathbf{S}_a$ is the*a priori* covariance, $\mathbf{S}_y$ is the measurement noise covariance, $\mathbf{K}$ is the *Jacobian Matrix* $K_{ij} = \partial y_i/\partial x_j$. Here it is assumed that spectral selection only depends on vertical domain so that only a one-dimension retrieval is modelled.

The total error covariance is the sum of the random error plus systematic error terms

$$\mathbf{S}_x^{tot} = \mathbf{S}_x^{rnd} + \mathbf{S}_x^{sys} \tag{7.3}$$

Two further assumptions are made:

- Assume that systematic error source i can be split into independent components $\mathbf{S}_y^i$

$$\mathbf{S}_x^{tot} = \mathbf{G}\mathbf{S}_y\mathbf{G}^T + \sum_i \mathbf{G}\mathbf{S}_y^i G^T \tag{7.4}$$

- Assume that each indipendent systematic error component is fully correlated $\mathbf{S}_y^i = (\mathbf{G}\delta\mathbf{y}^i)(\mathbf{G}\delta\mathbf{y}^i)^T$

$$\mathbf{S}_x^{tot} = \mathbf{G}\mathbf{S}_y\mathbf{G}^T + \sum_i (\mathbf{G}\delta\mathbf{y}^i)(\mathbf{G}\delta\mathbf{y}^i)^T \tag{7.5}$$

Thus systematic errors are represented in terms of *Error Spectra* $\delta\mathbf{y}^i$

7.3.3 Figure of Merit

Some *ad hoc* scalar parameter H is required to define the quality of a retrieval, conveniently expressed in ‘bits’

$$H = -\frac{1}{2}log_2F \tag{7.6}$$

where F is some scalar function of the retrieval covariance (smaller $F \Rightarrow H \Rightarrow$ better retrieval).

For the specific case where F is the ratio of the determinants of the retrieved/*a priori* covariances

$$F = \frac{|\mathbf{S}_x^{tot}|}{|\mathbf{S}_a|} \tag{7.7}$$

then H is the *Shannon Information Content.*

However, for MIPAS operational microwindows a different functional forma has been used

$$F = \frac{\prod_j (S_{x_{jj}}^{tot} + 2S_{x_{jj}}^{sys})}{\prod_j S_{a_{jj}}^{tot}} \tag{7.8}$$

Compared with the Shannon Information Content, it only minimize diagonal elements rather than full matrix and it additional weight against systematic errors.

Evaluating the Figure of Merit using only the diagonal elements of the covariance matrix is considered to be preferable, especially when considering microwindows for joint-retrievals of two more species, since the correlation information represented by the off-diagonal elements is often of no practical benefit to users of the data. However, when considering microwindows for species with low S/N, the second feature (weighting against systematic errors) may be dropped.

7.3.4 Error Sources/Estimation

In principle, any non-retrieved parameter which contributes significantly either to the forward model or to the instrument model uncertainty.

Those used for MIPAS, with revisions based on experience based on in-flight data, are:

Random Due to the propagation of instrument noise through the retrieval. Based on in-flight values of NESR from orbit 2081, allowing for apodisation.

NONLTE Non-LTE error. Due to assumption of local thermodynamic equilibrium when modelling emission in the MIPAS forward model. Based on calculation using vibrational temperatures.

SPECDB (formerly referred to as HITRAN) Spectroscopic database error. Due to uncertainties in the strength, position and width of infrared emission lines.

GAIN Radiometric Gain Uncertainty. Due mostly to the non-linearity correction in bands A, AB and B. A uniform value of ± 2% has been assumed for all bands.

SPREAD Uncertainty in width of apodised instrument line shape (AILS). A value of 2% has been assumed based on observed 2nd derivative signatures in the residual spectra during the first 6 months of operation.

SHIFT Uncertainty in the spectral calibration. the designed specification of ±0.001 cm^{-1} has been used, and is consistent with the 1^{st} derivatives signatures in the residual spectra.

CO_2MIX CO_2 line-mixing. Due to the neglecting line-mixing effects in the retrieval forward model (only affects strong CO_2 Q branches in the MIPAS A and D bands).

CTMERR Uncertainty in gaseous continua. Assumes an uncertainty of ± 25% in the modelling of continuum feature of H_2O (mostly), CO_2, O_2 and N_2.

GRA Horizontal gradient effect. Due to retrieval assuming a horizontal homogeneous atmosphere for each profile. Error is calculated assuming a ±1K/100km horizontal temperature gradient. This could be removed for 2-D dimensional retrievals (like GMTR).

HIALT Uncertainty in high-altitude column. Retrieval assumes a fixed-shape of atmospheric profile above the top retrieved level. Effect is calculated assuming 'true' profile can deviate by climatological variability. this has been added since launch.

TEM Temperature propagation error. temperature and pressure are retrieved first, this represents the contribution of a nominal 1K temperature error into the constituent retrievals.

PRE Pressure propagation error. As with temperature, effect of a nominal 2% pressure retrieval uncertainty.

species Uncertainties in assumed profiles of contaminant species. For most species this is the climatological 1σ variability.

7.3.5 Microwindow Selection for p, T and CO_2

The high correlation between pressure, temperature and CO_2 make opportune to retrieve them simultaneously so to remove the propagation of systematic errors into CO_2.

Ten microwindows have been selected for the purpose using the MWMAKE code [107] that operates the selection by minimizing the propagation of both random and systematic errors in the retrieval products, as previously described in details. The spectral distribution and the selected altitude range of the ten microwindows along with the list of absorbers that have spectral features within the microwindow are in table 7.1. Figure 7.1 gives a graphical overview of the microwindows distribution: eight are located in MIPAS band A and two in MIPAS band AB. In the same figure is also reported the CO_2 and the total spectral radiance at 21 km tangent altitude (note the presence of a double y axis: left axis indicates the validity altitude of the microwindows while the right axis indicates the radiance (in log scale) of the spectrum).

MW	Spectral range (cm^{-1})	Alt. range (km)	Absorbers list (HITRAN code)
1	729.275–732.275	15–60	2 1 3 4 10 12 62 11 23 26 27 29
2	1032.700–1035.700	21–68	2 1 3 4 6 11
3	791.275–793.050	6–33	2 1 3 10 12 61 11 63 18 23 26 27 60 29 56
4	1159.775–1161.975	6–33	2 1 3 4 6 12 9 11 52 56
5	686.100–689.100	52–68	2 1 3 4 10 12 23
6	967.375–969.800	6–30	2 1 3 4 10 12 11 29 30
7	774.250–774.450	6–9	2 1 3 10 12 61 11 63 18 23 26 27 60 29 56
8	767.900–768.150	6–12	2 1 3 10 12 61 11 63 18 23 26 27 60 29 56
9	955.500–958.500	6–36	2 1 3 4 10 12 11 29 30
10	750.325–751.100	6–15	2 1 3 4 10 12 62 61 11 23 26 27 60 29

Table 7.1: List of microwindows selected and their characteristics (spectral and altitude coverage and HITRAN code of the absorbers within the microwindow).

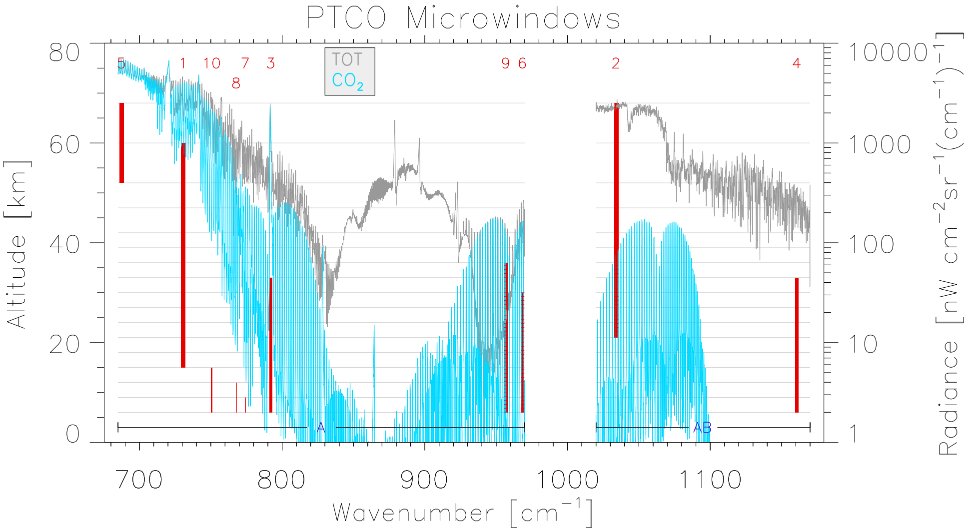

Figure 7.1: Spectral locations and tangent altitude range of the p, T, CO_2 microwindows (left axis). Total (gray) and CO_2 (azure) emission features in the MIPAS spectrum calculated for 21 km tangent altitude (right axis).

The corresponding CO_2 error budget is reported in the fig. 7.2. The legend of this figure shows, in decreasing order of magnitude, the components (explained in detail in section 7.3.4) that contribute to the systematic error (represented by the black line).

The error budget analysis evidences that large components of the systematic error come from the uncertainty on the O_3 and H_2O distributions.

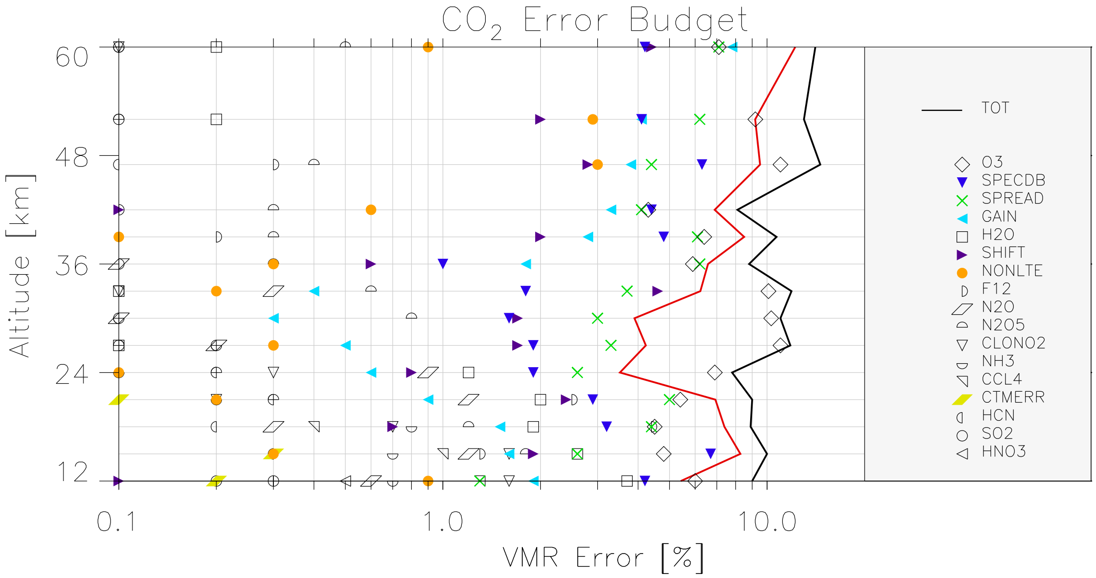

Figure 7.2: CO_2 error budget for the ten MWs selected. The various symbols represent the different components of the systematic error, in approximate order of importance. The black line represents the systematic error and the red line the same error without the H_2O and O_3 error components.

7.4 Analysis and Results

In order to minimize the O_3 and H_2O uncertainty p, T, H_2O and O_3 have been preliminarily retrieved: the resulting 2-D fields of H_2O and O_3 have then been used in the analysis for the simultaneous determination of p, T and CO_2, so that in the computation of the error budget the contributions of H_2O and O_3 can be neglected. Exploiting the possibility of the GMTR to operate on a target-dependent vertical retrieval grid a degraded vertical retrieval grid has been used for the CO_2 in order to lower the random error.

An *a posteriori* analysis has verified negligible differences in the p and T distributions determined in the two retrieval steps. As an example temperature maps derived in the two steps from orbit 7179 (15^{th} july 2003) are shown in fig. 7.3. The left panel of the figure refers to the temperature computed in the first retrieval step (p, T, H_2O and O_3 joint retrieval) while the right panel refers to the second step (p, T and CO_2 joint retrieval).

The accuracy obtained for the individual CO_2 VMR values is comparable with the one obtained for other well-determined MIPAS targets (such as O_3). However, for most applications, the accuracy requirements for CO_2 are stronger than those obtained. Considering the low variability of CO_2 with respect to both location and time, the random error component can be cut down with an averaging process and, in this case, the total

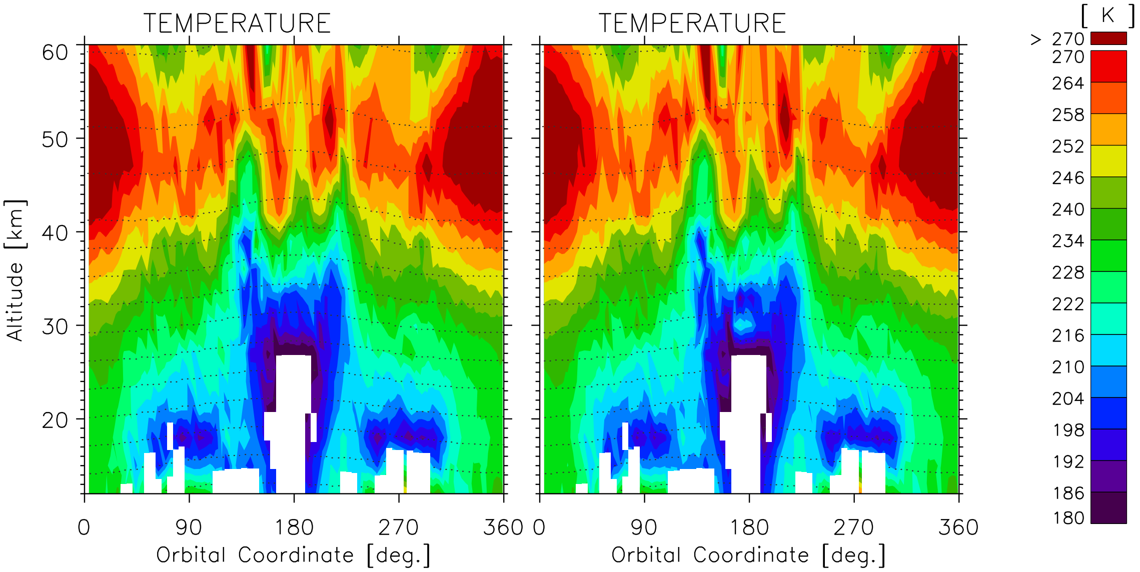

Figure 7.3: Temperature retrieved with an MTR p, T, H_2O, O_3 (left), and with an MTR p, T, CO_2 (right), orbit 7179

error can be approximated to the systematic error.

Monthly averages have been operated for all the full resolution MIPAS mission, within latitudinal bands of about 10 deg and altitude intervals of about 6 km (that correspond to the altitude grid adopted for the retrieval of CO_2).

The random errors associated to the values shown in these maps are bounded far below 0.1 % (more than 1000 VMRs have been averaged for each latitude/altitude element) so that they becomes negligible with respect to the systematic components. The total error budget for these maps is then represented by the red line in fig. 7.2 where, in addition to the random component, also the contributions of H_2O and O_3 have been neglected.

All the full resolution MIPAS mission (from july 2002 to march 2004) has been precessed. The CO_2 VMR maps (fig. 7.4 for) show a minimum around 36 km that extends over all the latitudes. Looking at the pressure maps (fig. 7.5) this minimum is clearly associated to a pressure level ($\sim$ 5 hPa); indeed, in the Antarctic winter (June-July-August maps) the minimum strip follows the shape of pressure levels that because of the winter vortex bend down while approaching the South Pole.

In general, the vertical and horizontal variability structures appear much larger than expected. However, the size of the systematic errors is such to limit the validity of possible deductions. On the other hand, the overall variability shows features that are different in the two months and do not show any correlation with T, H_2O, and O_3 distributions. This suggests that the observed variability exists independently from the effect of systematic errors. Analyses of the type shown in this study acquire value when

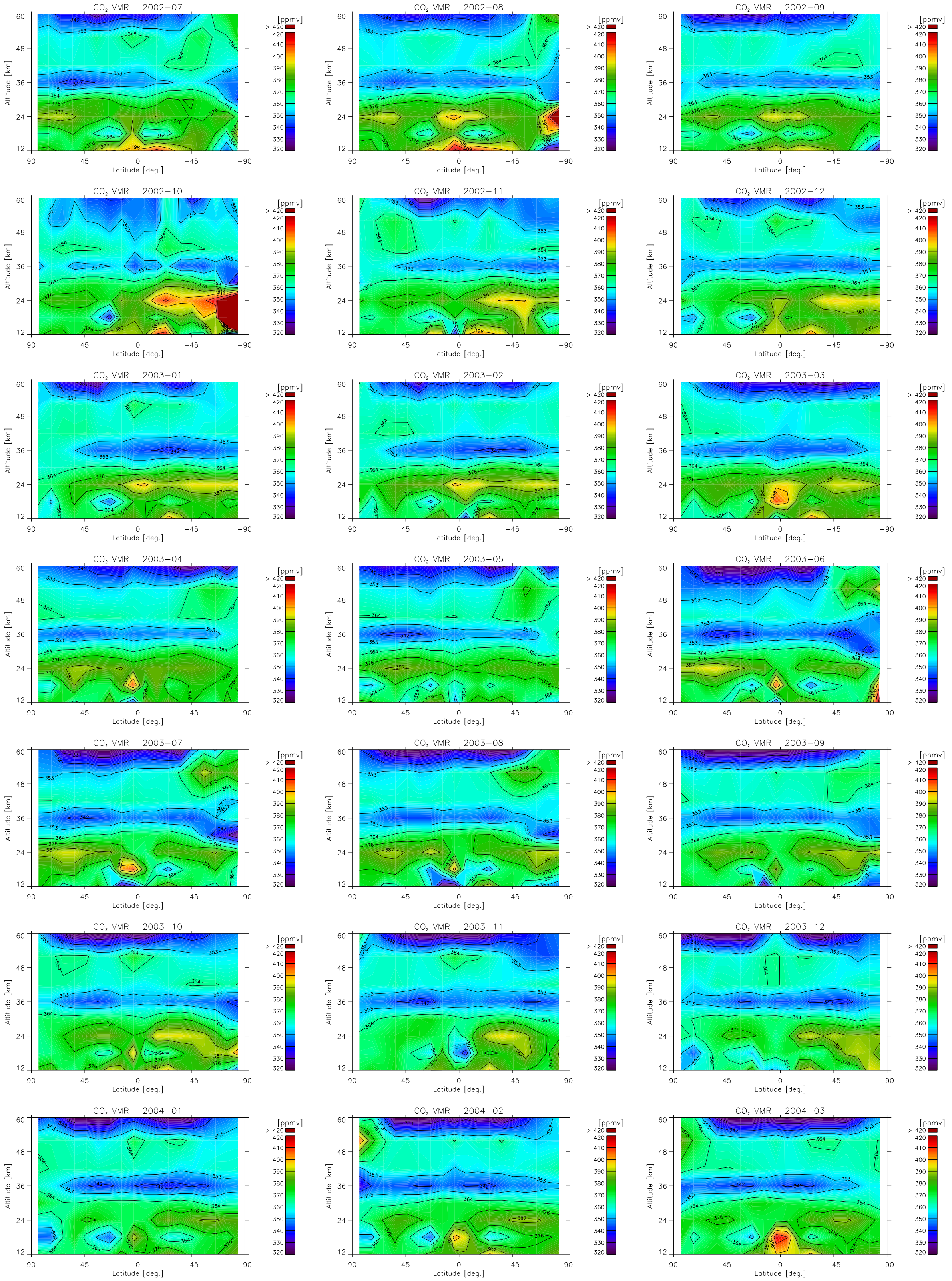

Figure 7.4: CO_2 VMR maps: 2002, 2003 and 2004.

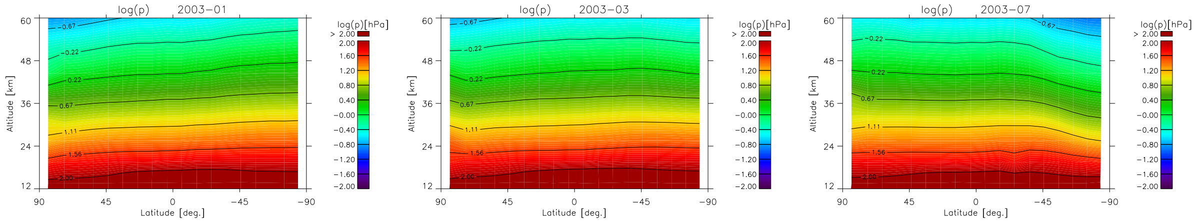

Figure 7.5: Pressure maps for January (left), March (center) and July (right) 2003.

comparing CO_2 distributions measured at different times and/or locations because, in this case, a similar impact of the systematic errors is expected on the different CO_2 retrievals.

The accuracy of the CO_2 products can be improved by including in the multi-target retrieval both H_2O and O_3. For this purpose a new set of MWs has been selected but not yet tested. On a different direction, an activity has started within the MIPAS community aiming at a more accurate evaluation and, possibly, a reduction of the systematic error components.

Conclusions

This thesis reports the results of a study aimed at the assessment of the trade-off between the accuracy of the retrieved parameters and the horizontal resolution of level 2 products derived from real MIPAS observations. A two-dimensional retrieval algorithm has been used to determine the trade-off of pressure, temperature, ozone, water vapor, nitric acid, methane, nitrous and nitric oxide. The results of this study have been compared with those obtained with simulated observations.

Two different strategies have been considered to change the horizontal resolution of the retrieval products:

1. Nominal MIPAS measurements have been analyzed on a horizontal retrieval grid different from the measurement grid. The analysis on real measurements confirms that the horizontal resolution of the retrieved profiles can be improved generally by a factor two or three (it depends on the target) with respect to the nominal value determined by the atmospheric sampling interval.
2. Obtain finer horizontal sampling of the atmosphere at the expenses of spectral resolution. This strategy has been tested for one measurement scenario using MIPAS observations acquired with spectral resolution reduced by about 60% but identical limb-scanning sequence.

The comparison of the two strategies has shown that, in order to gain horizontal resolution, the second is advantageous with respect to the strategy that acts on the retrieval grid. This outcome confirms the indication reported in Ref. [6] that was obtained with simulated observations and making use of several assumptions.

A further step in the exploitation of optimized retrieval grids has been carried out on a different important topic. An appropriate retrieval grid was identified to retrieve the carbon dioxide distribution in the stratosphere from MIPAS measurements.

CO_2 is an important greenhouse gas. Furthermore its spectral signatures are often used to retrieve the p-T profiles that are afterward used to retrieve all the target species. Despite its importance, the literature is poor of data about CO_2 distribution in the stratosphere. Indeed a retrieval with errors that are acceptable for the use of CO_2 as a reference gas is difficult to be performed: therefore a strategy that allows to retrieve the CO_2 from satellite measurements with low errors would be very useful. A strategy has been defined to reduce as much as possible the retrieval errors.

For this purpose some GMTR capabilities have been exploited:

i The MTR functionality allowed to perform a joint retrieval of pressure, temperature and CO_2. Indeed the high correlation between pressure, temperature and CO_2 makes opportune to retrieve them simultaneously in order to remove the propagation of p and T systematic errors into CO_2 VMR.

ii The line mixing modelling, implemented in the code, made feasible the use of the intense CO_2 Q branches in the 15 μm region that provide a high signal-to-noise ratio and a high sensitivity to pressure and temperature.

iii The possibility to use a target-dependent retrieval grid has been useful in order to adopt an optimal retrieval grid that reduces the ESDs.

All the full resolution MIPAS mission (from July 2002 to March 2004) has been precessed in order to retrieve CO_2 fields. Monthly averages have been operated on the results in order to reduce the random errors to negligible values. The averages have been computed within latitudinal bands of 10 deg at altitude intervals separated by 6 km.

The CO_2 VMR showed a minimum around 36 km that extends over all the latitudes that is clearly associated to a pressure level ($\sim$ 5 hPa); indeed, in the Antarctic winter (June-July-Augusts) the minimum strip follows the shape of this pressure level that, because of the polar vortex, bend down while approaching the South Pole. The study evidenced that, in general, the CO_2 vertical and horizontal variability structures appear much larger than expected.

The size of the systematic errors is such to limit the validity of possible deductions. On the other hand, the overall variability shows features that are different for each month and do not show any correlation with T, H_2O, and O_3 distributions. This suggests that the observed variability exists independently from the effect of systematic errors.

A future work aimed to improve these results could be carried out on the basis that the accuracy of the CO_2 products can be improved by including in the multi-target retrieval both H_2O and O_3.

Acknowledgements

There are lots of people I would like to thank for a huge variety of reasons.

Firstly, I would like to thank my Supervisor, Prof. Massimo Carlotti who has been an excellent advisor for my PhD. I would also like to thank my co-advisors, Dr. Bianca Maria Dinelli and Dr. Marco Ridolfi for the support they gave me during all the course of my PhD.

Thanks also to all the rest of the people who has been or are part of the group: Dr. Gabriele Brizzi (Gabri), Dr. Luca Magnani (the IDL expert...), Dr. Fabrizio Niro, Dr. Elisa Castelli, Dr. Valentina Mauri, Dr. Norma Caballero, Dr. Didem Alpaslan and all the academic staff (in particular Dr. Franca Tullini for her nice culinary presents) of the Dipartimento di Chimica Fisica ed Inorganica of the Facoltà di Chimica Industriale.

A special thank for the hospitality received during my five-months period in Oxford to my landlady Miss Miranda Jones and to Dr. Anu Dudhia, Dr. Don Grainger, Dr. Chiara Piccolo and all the other people of the Atmospheric, Oceanic and Planetary Physics of the University of Oxford.

Last but not least all the people who shared the same room at the office: starting from the first row, Dr. Francesca Ospitali (Francy), Dr. Roberta Michelazzi, Dr. Ivano Bilotti (Il Bombarolo), Dr. Antonio Monari (Dr. House or The Shark, as you like), Dr. Giampaolo Venditti (Zzampa, notice that the double 'zz' in the name is not a mistake) and the 'adopted' gambler tutor Dr. Mirko Minzoni.

Finally, I have to say 'thank-you' to all my friends, my girlfriend and my family, for all the encouragement they give me every day.

Is that everyone?

Enzo,
March 2007.

List of Tables

List of Figures

Bibliography

[1] L. Magnani. *Doctoral Thesis,* Assessment of the performance of the Level 2 data analysis system of MIPAS–ENVISAT experiment. Trade-off studies between spatial resolution and retrieval/observational parameters. (2004). 5, 77, 78, 93

[2] F. Niro, G. Brizzi, M. Carlotti, E. Papandrea, and M. Ridolfi. Precision improvements in the Geo-fit retrieval of pressure and temperature from MIPAS limb-observations by modeling CO_2 line-mixing. *J. Quant. Spectrosc. Radiat. Transfer* **103**, 14–26, doi: 10.1016/j.jqrst.2006.07.012 (2007). 5, 6, 99, 102

[3] M. Carlotti, G. Brizzi, E. Papandrea, M. Prevedelli, M. Ridolfi, B. Dinelli, and L. Magnani. GMTR: Two-dimensional geo-fit multitarget retrieval model for Michelson Interferometer for Passive Atmospheric Sounding/Environmental Satellite observations. *Appl. Opt.* **45**, 716–727 (2006). 6, 63

[4] M. Carlotti, B. M. Dinelli, P. Raspollini, and M. Ridolfi. Geo-Fit approach to the analysis of satellite limb-scanning measurements. *Appl. Opt.* **40**, 1872–1885 (2001). 6, 63

[5] B. M. Dinelli, D. Alpaslan, M. Carlotti, L. Magnani, and M. Ridolfi. Multi-target retrieval (MTR): the simultaneous retrieval of pressure, temperature and volume mixing ratio profiles from limb-scanning atmospheric measurements. *J. Quant. Spectrosc. Radiat. Transfer* **84**, 141–157 (2004). 6, 63, 64, 80

[6] M. Ridolfi, L. Magnani, M. Carlotti, and B.M. Dinelli. MIPAS-ENVISAT limb-sounding measurements: trade-off study for improvement of horizontal resolution. *Appl. Opt.* **3** No. 31, 1–11 (2004). 6, 67, 111

[7] United States Committee on Extension to the Standard Atmosphere, *U.S. Standard Atmosphere 1976,* National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, United States Air Force, Washington D.C. (1976). 9

[8] S. Chapman. A theory of upper-atmospheric ozone. *Mem. of the Roy. Met. Soc.* **4**, 26 (1930). 10

[9] J. Gille. ed. *Instruments and techniques for stratospheric research,* Atmospheric Technology No. 9 - Spring 1978, N.C.A.R., Bouldeder, Co. (1978). 13

[10] D. H. Ehhalt. In situ observations. *Phil. Trans. Roy. Soc. Lond.* **A296**, 175 (1980). 13

[11] G. Megie and J. Pelon. Active optical sounding of ozone and minor constituents in the middle atmosphere: a review of ground-based LIDAR measurements. *Handbook for MAP* **13**. 14

[12] E. V. Browell et al. NASA multipurpose airborne DIAL system and measurements of ozone and aerosol profiles. *Appl. Opt.* **22**, 522 (1983). 14

[13] J. W. Waters et al. *J. Quant. Spectrosc. Radiat. Transfer* **32**, 407 (1984). 14

[14] G. M. B. Dobson and D. N. Harrison. Measurements of the amount of ozone in the Earth's atmosphere and its relation to other geophysical conditions. *Proc. R. Soc. Lond.* **A110**, 660–693 (1926). 14

[15] L. Genzel and K. Sakai. *J. Opt. Soc. Am.* **67**, 871 (1977). 17

[16] J. E. Harries. *J. Opt. Soc. Am.* **67**, 880 (1977). 18

[17] J. E. Harries and W. J. Burroughs. *Nature* **227**, 824 (1970). 18

[18] W. J. Burroughs and J. E. Harries. *Quart. J. Roy. Met. Soc.* **97**, 516 (1971). 18

[19] J. E. Harries, N. R. W. Swann, J. E. Beckman, and P. A. R. Ade. *Nature* **236**, 159 (1972). 18

[20] J. E. Harries. *Nature* **241**, 515 (1973). 18

[21] D. J. W. Kendall and T. A. Clark. *Infrared Phys.* **18**, 803 (1978). 18

[22] D. J. W. Kendall. *Doctoral Thesis,* Univ. of Calgary, Canada (1979). 18

[23] B. Carli, F. Mencaraglia, and A. Bonetti. *Int. J. of Infrared and Millimeter Waves* **1**, 263 (1980). 18

[24] D. H. Martin and E. F. Puplett. *Infrared Phys.* **10**, 105 (1969). 18

[25] P. Connes and G. Michel. *J. Opt. Soc. Am.* , 2067 (1975). 18, 31

[26] B. Carli et al. *Int. J. of Infrared and Millimeter Waves* **6** No. 2, 149 (1985). 18

[27] K. V. Chance, J. C. Brasunas, and W. A. Traub. *Geophys. Res. Lett.* **7**, 704 (1980). 18

[28] I. R. Able, B. R. Breckinridge, and J. Pritchard. Optical design of the ATMOS Fourier Transform spectrometer. *Optical Systems Engineering,* Proc. SPIE **193**, 12 (1979). 20

[29] C. B. Farmer, O. F. Raper, and F. G. O'Callaghan. Final report on the first flight of the ATMOS instrument during the Spacelab 3 Mission, April 29 through May 6, 1985. *JPL Publication* 87–32, Jet Propulsion Laboratory, Pasadena, Calif., October (1987). 20

[30] M. R. Gunson et al. The Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment: Deployment on the ATLAS Space Shuttle missions. *Geophys. Res. Lett.* **23**, 2333–2336 (1996). 20

[31] H. Shimoda and T. Ogawa. Infrared Spaceborne Remote Sensing II. *M. S. Scholol,* ed., Proc. SPIE **2209**, 14 (1994). 20

[32] D. Offermann et al. The CRyogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) experiment and middle atmosphere variability. *J. Geophys. Res.* **104**, 16311 (1999). 20

[33] M. J. Endmann, G. Lange, and B. Fladt. in *Space Optics 1994: Earth Observation and Astronomy,* M. G. Cerutti-Maori and P. Roussel, eds., Proc. SPIE **2209**, 36 (1994). 21

[34] M. A. Soucy et al. ACE-FTS instrument design and on-orbit status. *Appl. Opt., submitted for publication* (2005). 21

[35] R. Beer and T. A. Glavich. Advanced Optical Instrumentation for Remote Sensing of the Earth's Surface from Space, G. Duchossois, F. L. Herr and R. J. Zander, eds. Proc. SPIE **1129**, 42 (1989). 21

[36] T. Ogawa et al. *Adv. Space Res.* **14** No. 1, 25 (1994). 21

[37] P. Javelle and F. Cayla. Space Optics 1994. *Earth Observation and Astronomy,* M. G. Cerutti-Maori and P. Roussel, eds., Proc. SPIE **2209**, 14 (1994). 21

[38] A. A. Michelson. *Philos. Mag.* **31**, 338 (1891). 23

[39] A. A. Michelson. *Philos. Mag.* **34**, 280 (1892). 23

[40] Lord Rayleigh. *Philos. Mag.* **34**, 407 (1892). 23

[41] J. Chamberlain. *The principles of Interferometric Spectroscopy.* John Wiles and Sons, New York, (1978). 24

[42] R. J. Bell. *Introductory Fourier Transform Spectroscopy.* Academic Press, New York and London, (1972). 24, 59

[43] H. Rubens and R. W. Wood. *Philos. Mag.* **21**, 249 (1911). 25

[44] P. Jacquinot and J. C. Dufour. *J. Rech. C.N.R.S.* **6**, 91 (1948). 25

[45] P. B. Fellgett. *J. de Physique et le Radium* **19**, 187 and 236 (1958). 25

[46] P. Jacquinot. *J. Opt. Soc. Amer.* **44**, 761 (1954). 25

[47] P. Jacquinot. *Rep. Progr. Phys.* **23**, 267 (1960). 25

[48] C. Fabry and A. Perot. *Ann. Chim. Phys. (7)* **12**, 459 (1897). 26

[49] C. Fabry and A. Perot. *Ann. Chim. Phys. (7)* **16**, 115, 289, 331 (1899). 26

[50] P. B. Fellgett. *Doctoral Thesis (unpublished),* Univ. of Cambridge(1951). Proceding of Ohio State Symposium on Molecular Spectroscopy (1952). 26

[51] P. B. Fellgett. *J. de Physique et le Radium* **28**, C2 – 165 (1967). 26

[52] L. Mertz. *J. de Physique et le Radium* **19**, 233 (1958). 27

[53] L. Mertz. *Opt. Commun.* **6**, 354 (1973). 27

[54] J. Strong and G. A. Vanasse. *J. Opt. Soc. Amer.* **49**, 844 (1959). 133. 27

[55] H. A. Gebbie and G. A. Vanasse. *Nature* **178**, 432 (1956). 27

[56] H. A. Gebbie. in *Symposium on Interferometry* , National Physical Laboratory, HMSO, London (1959). 27

[57] H. A. Gebbie. *Phys. Rev.* **107**, 1194 (1957). 27

[58] J. Connes and H. P. Gush. *J. de Physique et le Radium* **20**, 915 (1959). 27

[59] J. Connes and H. P. Gush. *J. de Physique et le Radium* **21**, 645 (1960). 27

[60] J. Connes. *J. de Physique et le Radium* **19**, 197 (1958). 28

[61] J. Connes. *Thesis,* Univ. of Paris (1960). 28

[62] J. Connes. *Rev. Opt. Theor. Instrum.* **40**, 45, 116, 171, 231 (1961). 28, 29

[63] J. L. Pritchard, A. Bullard, H. Sakai, and G. Vanasse. *J. de Physique* **28**, supplement to No. 3–4, C2–67 (1967). 28

[64] R. F. Edgar, B. Lawrenson, and J. Ring. *J. de Physique* **28**, supplement to No. 3–4, C2–73 (1967). 28

[65] J. E. Hoffman and G. Vanasse. *J. de Physique* **28**, supplement to No. 3–4, C2–79 (1967). 28

[66] J. W. Cooley and J. W. Turkey. *Math. Comput.* **19**, 267 (1965). 29

[67] H. Yoshinaga et al. *Appl. Opt.* **5**, 1159 (1966). 29

[68] M. J. Persky. *Report No. AFCRL 63–439, Block Associates Inc.* (1963). 29

[69] L. C. Block and A. S. Zachor. *Appl. Opt.* **3**, 209 (1964). 29

[70] J. Connes, P. Connes, and J. P. Millard. *J. Phys.* **28**, 120 (1967). 29

[71] R. A. Hanel and L. Chaney. *NASA Goddard SFC, Rep. X-620–66–476* (1966). 31

[72] R. A. Hanel et al. *Appl. Opt.* **9** No. 8, 1767 (1970). 31

[73] R. A. Hanel et al. *NASA Goddard SFC, Rep. X-622–71–272* (1971). 31

[74] R. A. Hanel et al. *Appl. Opt.* **11** No. 11, 2625 (1972). 31

[75] R. A. Hanel et al. *Appl. Opt.* **19** No. 11, 1391 (1980). 31

[76] J. E. Chamberlain, J. E. Gibbs, and H. A. Gebbie. *Nature* **198**, 877 (1963). 31

[77] J. E. Chamberlain, J. E. Gibbs, and H. A. Gebbie. *Infrared Phys.* **9**, 185 (1969). 31

[78] E. E. Bell. *Jpn. J. Appl. Suppl. I 4* , 412 (1965). 31

[79] E. E. Bell. *Infrared Phys.* **6**, 57 (1966). 31

[80] H. Sakai. in *Aspen International Conference on Fourier Spectroscopy 1970* (Ed. G. A. Vanasse, A. T. Stair, Jr., and D. J. Backer), AFCRL-71-0019, Special Report No. 114, 19 (1971). 31

[81] R. Murphy and H. Sakai. in *Aspen International Conference on Fourier Spectroscopy 1970* (1970). 31

[82] G. P. Kuiper. *Commun. Lunar Planet Lab. Univ. Ariz.* **1**, 83 (1962). 31

[83] J. Connes and P. Connes. *J. Opt. Soc. Am.* **56**, 896 (1966). 31

[84] U. Fick, H. P. Larson, and R. F. Poppen. *Astrophys. J.* **187**, 407 (1974). 31

[85] C. D. Rodgers. Inverse Methods for Atmospheric Sounding, Theory and Practice. Ed. F. W. Taylor (University of Oxford), World Scientific Publishing Co. Pte. Ltd. (2000). 33, 65, 103

[86] D. S. Sivia. Data Analysis: A Bayesian Tutorial, Clarendon, Oxford, UK (1998). 37

[87] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* **2**, 164 (1944). 42

[88] D. W. Marquardt. An Algorithm for the Least-Squares Estimation of NonLinear Parameters. *J. Soc. Appl. Math.* **11**, 431 (1963). 42

[89] Press et al. Numerical recipes: the art of scientific computing, Second Edition, Cambridge University Press (1995). 42

[90] S. Twomey. On the numerical solution of Fredholm integral equation of the first kind by the inversion of the linear system produced by quadrature. *J. Ass. Comput. Mach.* **10**, 97 (1963). 43

[91] A. N. Tikhonov. On the solution of incorrectly stated problems and a method of regularization. Dokl. Acad. Nauk SSSR **151**, 501 (1963). 43

[92] M. Endemann. MIPAS instrument concept and performance. *in Proceedings of the European Symposium on Atmospheric Measurements from Space,* ESA Earth Science Division ed. MIPAS instrument concept and performance, in Proceedings of the European Symposium on Atmospheric Measurements from Space, ESA Earth Science Division ed. (European Space Agency, ESTEC) **1**, 29 (1999). 45, 46

[93] M. Ridolfi et al. Optimized forward model and retrieval scheme for IMPAS near-real-time data processing. *Appl. Opt.* **39** No. 8, 1323 (2000). 45, 54, 65, 73

[94] ESA. *European Space Agency (ESA), Envisat-1: Mission and system summary,* Eur. Space Agency Sci. Tech. Rep., ESA/SP-1229, 85 pp (2000). 46

[95] P. Burkett, F. Fergg, and H. Fischer. A compact high resolution Michelson interferometer for passive atmospheric sounding (MIPAS). *IEEE Trans. Geosci. Remote Sensing* **GE-21**, 345 (1983). 49

[96] A. Dudhia, V. L. Jay, and C. D. Rodgers. Microwindows selection for High Spectral Resolution Sounders. *Appl. Opt.* **41**, 3665–3673 (2001). 54

[97] G. Echle, T. v. Clarmann, A. Dudhia, M. Lopez-Puertas, F. J. Martin-Torres, B. Kerridge, and J.-M. Flaud. Spectral microwindows for MIPAS-ENVISAT data analysis. in *Proceedings of the European Symposium on Atmospheric Measurements*

from Space, ESA Earth Science Division, ed. (European Space Agency - ESTEC) **2**, 481 (1999). 54

[98] T. von Clarmann and G. Echle. Selection of Optimized Microwindows for Atmospheric Spectroscopy. *Appl. Opt.* **37**, 7661 (1998). 54, 102

[99] M. Carlotti. Global-fit approach to the analysis of limb-scanning atmospheric measurements. *Appl. Opt.* **27** No. 15, 3250 (1988). 54

[100] A. Goldman and R. S. Saunders. Analysis of atmospheric infrared spectra for altitude distribution of atmospheric trace constituents. *J. Quant. Spectrosc. Radiat. Transfer* **21**, 155 (1979). 54

[101] J. T. Houghton. The Physics of Atmospheres. 2nd Ed, CUP, Cambridge (1986). 57, 72

[102] L. L. Strow, H. E. Motteler, R.G. Benson, S.E. Hannon, and S. De Souza-Machado. Fast Computation of monochromatic infrared atmospheric transmittances using compressed look-up tables. *J. Quant. Spectrosc. Radiat. Transfer* **59**, 481 (1998). 58, 73

[103] R. H. Norton and R.Beer. New apodizing functions for Fourier spectrometry. *J. Opt. Soc. Am.* **66**, 259 (1976). and errata corrige *J. Opt. Soc. Am.* **67**, 419 (1977). 59

[104] D. P. Edwards. GENLN2: A general line-by-line atmospheric transmittance and radiance model. Version 3.0 description and users guide. Report NCAR/TN-367+STR, Natl. Cent. for Atmos. Res., Boulder, Colo. (1992). 61

[105] P. Edwards. High Level algorithm definition document of the MIPAS Reference Forward Model, ESA Report PO-TN-OXF-GS-0004. (1997). 61

[106] B. M. Dinelli, E. Battistini, and B. Carli. MIPAS-B2 data analysis, TN-ISM-0002, ESA Contract No 11717/95/NL/CN. (2000). 61

[107] A. Dudhia, V. L. Jay, and C. D. Rodgers. Microwindow Selection for High-Spectral-Resolution Sounders. *Appl. Opt.* **41**, 3665–3673 (2002). 65, 102, 105

[108] M. Carlotti and B. Carli. Approach to the design and the data analysis of the limb scanning experiment. *Appl. Opt.* **33**, 3237 (1994). 65, 67

[109] P. R. Bevington and D. K. Robinson. Data Reduction and Error Analysis for the Physical Sciences. *3rd ed. McGraw-Hill* (2003). 67

[110] *World Geodetic System 1984 (WGS84), http://www.wgs84.com/ (UK), http://Earthinfo.nga.mil/GandG/puborder.html (USA)* . 68

[111] P. Morris. Generation of compressed cross-section look-up tables for NRT MIPAS retrievals", ESA report PO-TN-OXF-GS-0011. (1997). 73

[112] R. Wells. Generation of optimized spectral grids", Technical report for ESA study 11886/96/NL/GS, document: PO-TN-OXF-GS-0010. (14th July 1997). 73

[113] P. Raspollini et al. Level 2 near real-time analysis of MIPAS measurements on ENVISAT: updated performance assessment. *EGS-AGU-EUG joint assembly, Nice* (11 April 2003). 92

[114] A. Dudhia. MIPAS Microwindow Error Analysis. *http://www.atm.ox.ac.uk/group/mipas/err/* (2002). 92

[115] *http://www-atm.physics.ox.ac.uk/RFM/mwmake.html* . 99

[116] C. D. Rodgers. Information content and optimisation of high spectral resolution measurements. in *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II,* P. B. Hays and J. Wang (Eds.), Proc. SPIE 193 **2830**, 136–147 (1996). 102

[117] C. D. Rodgers. Information content and optimisation of high spectral resolution measurements. *Adv. Space Research* **21**, 362–367 (1998). 102

[118] F. Rabier et al. Channel selection methods for infrared atmospheric soundings interferometer radiances. *Q. J. R. Meteorol. Soc.* **128**, 1–15 (2001). 102

[119] G. Echle et al. Optimized spectral microwindows for atmospheric MIPAS-Enviat data analysis. *Appl. Opt.* **39**, 5531–5540 (2000). 102