University of Bologna

Faculty of industrial Chemistry

Atmospheric Physical Chemistry group of the
Department of Inorganic chemistry and physics

Master Thesis
Advanced Spectroscopy in Chemistry

Performance of a new set of Micro-windows for the retrieval of Pressure, Temperature, Water and Ozone from MIPAS observations

NAVE ANDY

Tutor:
Prof.
MASSIMO CARLOTTI

Co-Tutor:
Dr.
ENZO PAPANDREA

Academic Year 2011/2012

Summary:

The framework of this thesis is the study of the atmosphere thanks to the MIPAS instrument (Michelson Interferometer for Passive Atmospheric Sounding) on board of ENVISAT (Environmental satellite). MIPAS is a spectrometer designed by the European Space Agency to measure broad-band spectra of the atmosphere from space in the mid-infrared range. However, small spectral intervals are used to retrieve the altitude distribution (profile) of several targets (level 2 product). These spectral intervals, called Micro-Windows (MWs), are selected in order to maximize the information content and minimize the total error of the level 2 products at all the retrieval altitudes.

The first objective of this thesis was to compare the performance of two sets of MWs; one provided by the University of Bologna (8 MWs) and the other by the Oxford University (10 MWs) for the simultaneous retrieval of Temperature, Pressure, Water and Ozone. For this purpose I retrieved altitude profiles on simulated observation from the two sets of MWs for the above geophysical targets and for the other main targets (Nitric Acid, Methane, Nitrous Oxide and Nitrogen Dioxide) that are determined downstream in the retrieval chain making use of the previously retrieved profiles. The retrieval has been performed with the analysis code, called GMTR, that is based on the Geo-fit approach (tomographic retrieval) along with the Multi-Target Retrieval (MTR) functionality (interfering species retrieved simultaneously).

The strategy of simulated retrievals permits to know the reference atmosphere so that it becomes possible to compare the retrieved profiles with the reference profiles. Thanks to specific quantifiers, that I defined for the purpose, I plotted and compared easily the performance of the two sets of MWs. The result was that the Oxford set of MWs has a better overall performance.

The two sets of MWs were then compared using a recently developed diagnostic tool called Information Load (IL). I computed and compared the IL values of the two sets of MWs for the four main targets (Temperature, Pressure, Ozone and Water). The IL distributions miss a full coverage of the atmosphere for both sets of MWs; therefore I decided to merge the best elements of each set with the aim of obtaining a more uniform IL coverage. In order to select an optimal set I computed the IL values for each individual MW. A combined set of MWs was then created trying to cover all the atmosphere with the highest possible amount of IL. After several attempts I selected a set of 10 MWs (combined set). The IL distributions of the combined set are better than those of the original sets for all targets excepted for Pressure around 50 km of altitude.

I tested the performance of the combined set of MWs with the strategy of retrievals on simulated observations: by comparing the performance of the profiles obtained using the combined set with the performance of the corresponding profiles obtained from the Bologna and Oxford sets, I verified an overall better efficiency of the combined set with respect to both the other two at low altitudes and equal or worst efficiency at high altitudes.

I compared the performance of the three sets of MWs by retrieving the eight MIPAS main targets from a statistically meaningful set of real observations. In this case the estimated standard deviation (ESD) of the retrieved values (calculated by the retrieval algorithm) provides the comparison criterion: I found lower ESDs for the combined set of MWs excepted for Pressure above 40 km.

Finally I considered the systematic error components (that are provided for each set of MWs by the selection algorithm) and I calculated an estimate of the total error for the three sets: the combined set of MWs has slightly lower total error at low altitudes and generally higher error at high altitudes.

A further meaningful element to be kept in consideration is the computing time: the Oxford set of MWs (that was a generally better efficiency than the Bologna set) requires between 13 and 15 hours of CPU time to analyse on orbit of MIPAS measurements, the Bologna set requires between 3 and 5 hours while the CPU time for the combined set is between 9 and 11 hours.

Contents:

Introduction:

The framework of this thesis is the study of the atmosphere. In this scope spectroscopic instrument placed on board of artificial satellites are recording spectra of the atmosphere. These spectra are mainly used to retrieve altitude distribution of several targets (Temperature, Pressure and the main chemical components of the atmosphere). To avoid the redundancy of information and to save computer time, the retrieval process extracts narrow spectral intervals from the spectra. The retrieval process is currently using a set of spectral intervals selected by the University of Bologna. A new set of spectral intervals has been provided by Oxford University. The objectives of this thesis are the assessment of benefits and drawbacks of these two sets of spectral intervals and the merging of these two sets to create a "combined" set which should combine the benefits and discards the drawbacks.

Outline:

In Chapter 1, I describe the Environment Satellite and the inboard spectroscopic instrument, the Michelson Interferometer for Passive Atmospheric Sounding. This Spectroscopic instrument has the main assignment to extend our knowledge of atmospheric chemistry.

In Chapter 2, I introduce the so called Inverse Problem that is the problem of finding the best representation of the required parameters from measurements that are a complicated function of the parameters.

In Chapter 3, I briefly describe the retrieval process which retrieves from the MIPAS recorded spectra, the altitude profiles of Temperature, Pressure, and the main chemical components of the atmosphere.

In Chapter 4, I describe the diagnostics tools which are needed and will be used in the subsequent chapters to complete the analysis of the spectral intervals of the two provided sets, i.e. the retrieving of altitude profiles on simulated observations, the Information Load and the Systematic and Random errors.

In Chapter 5, I describe and use the quantifiers which are employed to compare the retrievals on simulated observations for Temperature, Pressure, Ozone, Water, Nitric Acid, Methane, Nitrous Oxide and Nitrogen Dioxide of the two sets of spectral intervals.

In Chapter 6, I compare the Information Load values retrieved thanks to the two sets of spectral intervals. Then I retrieve the Information Load for each individual spectral interval to select the most efficient and so create a more efficient, combined set of MWs.

In Chapter 7, I compare the combined set of spectral intervals in front of the two primary sets using retrievals on simulated and real observations. Then I finish the comparison looking at the systematic and total errors.

1 The MIPAS experiment

1.1 ENVISAT

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) is on board of the ENVIronmental SATellite (ENVISAT), it is an advanced polar-orbiting Earth observation satellite developed by the European Space Agency (ESA) that was successfully launched on a nearly polar orbit on 1 March 2002. The ENVISAT payload consists of a set of ten instruments (whose MIPAS) that operate over a wide range of the electromagnetic spectrum, from centimetre waves to the ultraviolet. This satellite provides measurements of atmosphere, oceans, lands, and ice and contributes significantly to extend our knowledge of climate and morphology of our planet. [1]

1.2 The MIPAS instrument

The MIPAS instrument is based on a Fourier Transform spectrometer that performs limb sounding observations of atmospheric emission in the middle and upper atmosphere. 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, 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. [2], [3]

1.2.1 The characteristic of the MIPAS instrument

The MIPAS instrument is similar to an ordinary Michelson Interferometer, except some modifications that were made to adapt it to operate on board of a satellite with a high sensibility. Some of these modifications are:

• The use of two movables mirrors, instead of one, to reduce the dimension of the instrument.

• These mirrors are moving in opposite direction in two parallel pathway, rather than perpendicular, to not confer to the orbital station a rotational moment when occurs the reversal of movement of the two reflective elements (see Figure 1.1).

• The use of a laser beam to obtain a high precision in the displacements of the moving elements

• Maintain all the optical elements at a temperature of 200 Kelvins, to reduce their thermal emission

• The use of a series of eight detectors Hg:Cd:Te, cooled at 70 Kelvins.

• The use of the radiation emission by a black body to a known temperature, for the radiometric calibration.
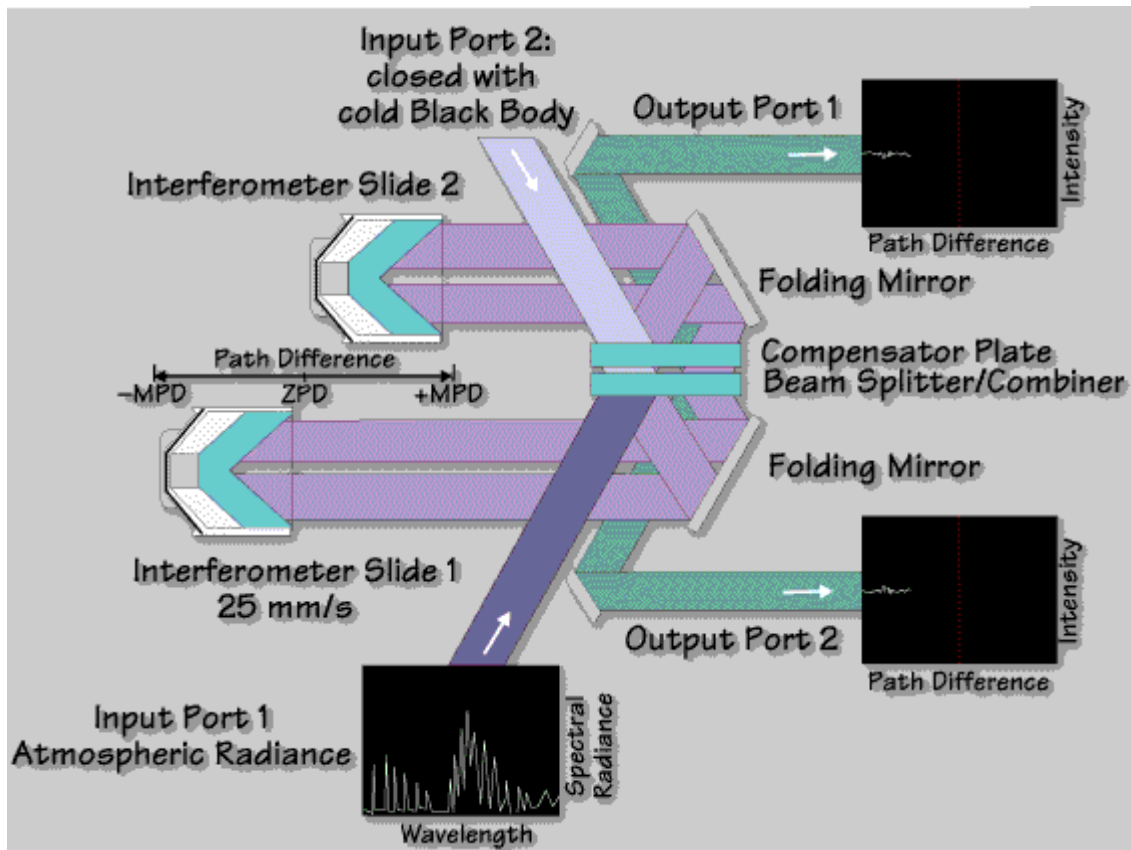
Figure 1.1: Schema of the two moving mirrors of the instrument MIPAS

1.2.2 Measurement description

The main purpose of MIPAS is the measurement of the distribution in altitude (profile) of Temperature and Pressure, as well as the volume mixing ratio (VMR)[1] of numerous species like water, ozone, nitric acid, methane, nitrous oxide, and nitrogen dioxide.

To obtain these results, MIPAS measures the emission of the atmosphere using the technique of "limb scanning", which consist in repeated measurement with different geometry of observation which not encounter the Earth's surface. Figure 1.2 shows a scheme describing the technique of limb scanning for a satellite. In this example, the orbital station flies at a altitude H_f, and performs with different angle ϑ_i, various measurements, which of them is characterise by the tangent altitude H_i, which is the minimum altitude reached by the line of sight.

---

[1] *The VMR (volume mixing ratio) is define as the ratio between the number of molecules of a tested species and the total number of molecules present in the same volume.*
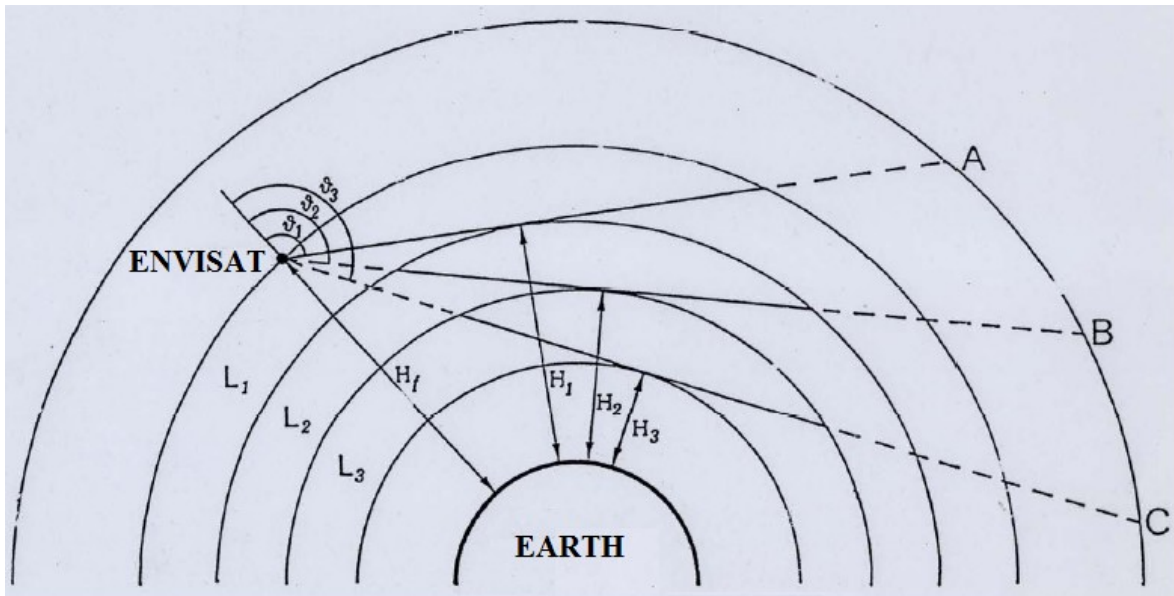
Figure 1.2: Example of limb scanning geometry

A, B, C Line of sight
H_f = *Flight Altitude*
H_i = *Tangent Altitude*
L_i = *Stratification of the atmosphere*
ϑ_i = *Zenith Angle*

1.2.3 The characteristic of the original MIPAS experiment

MIPAS provides series of spectra at different tangent altitude, the system operating the instrument MIPAS allows to scan between the altitude 5 to 150 km. Normally, a sequence of limb-scanning provides 16 spectra, takes 75 seconds (corresponding to a movement of the satellite of about 530 km), varying the tangent altitude from 6 to 68 km, where the studied species have the most important abundance.

Depending on the scientific objective, the "special" observation modes may differ from the "nominal" mode (see table 1.1) as these parameters: the adopted spectral resolution, the altitude coverage, the vertical/horizontal sampling steps, or the azimuth direction of the line of sight.

For one scan, the field of view is only of 3 km in altitude, to obtain a good vertical resolution, and is large of 30 km, to receive enough energy.

The usual measurements are in the opposite direction of the satellite movement. In special cases, it is possible to carry out the measurements in the perpendicular direction of the satellite flight, in the opposite direction of the sun. This mode is very useful to observe particular phenomena, as volcanic eruption, control of the pollution on the major air routes, or support eventual campaign of regional measurements.

The distance between the instrument and the tangent altitude is of about 3300 km. To obtain measurements on the desired line of view is necessary to effectuate an accurate calibration of the pointing system, which will takes the stars position as reference. The spectral interval of this instrument covers the mid-infrared: from 685 cm^{-1} to 2430 cm^{-1}. One complete spectrum is obtained in 4.5 seconds, with a resolution of 0.035 cm^{-1}. Of course, it is also possible to effectuate measurements with a reduce spectral resolution, in a reduce time of acquisition. In Figure 1.3 are reported the mains parameters of the MIPAS observation.

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	Polar Chemistry and Dynamics	rear	Global	7-55	2-10	420
Tropospheric Stratospheric Exchange	Exchange between Stratosphere and Troposphere, Troposphere Chemistry	rear	Global	5-40	1.5-10	420
Impact of Aircraft Emission	Study of major air traffic corridor	side	Primarily North of 25° latitude	6-40	1.5-10	330
Stratospheric Dynamics	Small scale structures in the middle Atmosphere	rear	Global	8-53	3	390
Diurnal Changes	Diurnal changes near the terminator	side	Near the terminator	15-60	3	480
UTLS	Upper Troposphere / Lower Stratosphere	rear	Global	6-35	2-7	120
Upper Atmosphere	Upper atmosphere	rear	Global	20-160	3-8	800

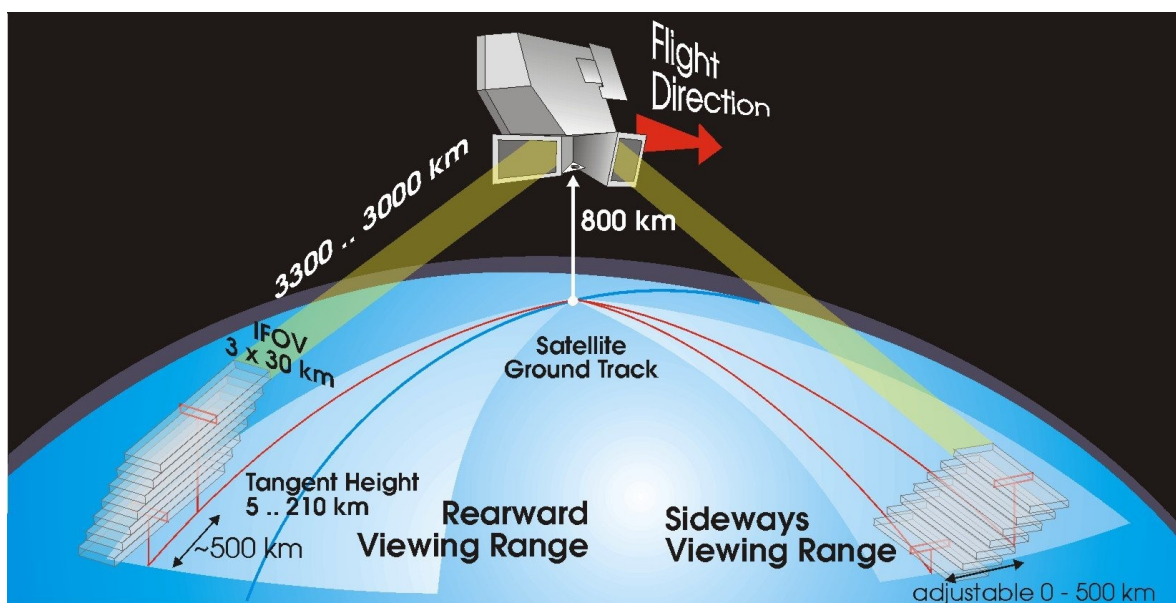Table 1.1: MIPAS nominal and special observation modes (before March 24, 2004)

Figure 1.3: The observation parameters of MIPAS

1.2.4 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 Figure 1.4 illustrates the flow of MIPAS 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 spectra of atmospheric radiance, with correction of the Doppler effect due to the movement of the orbital station . This process has also for purpose to obtain the tangent altitude and the geolocation of the data. 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.
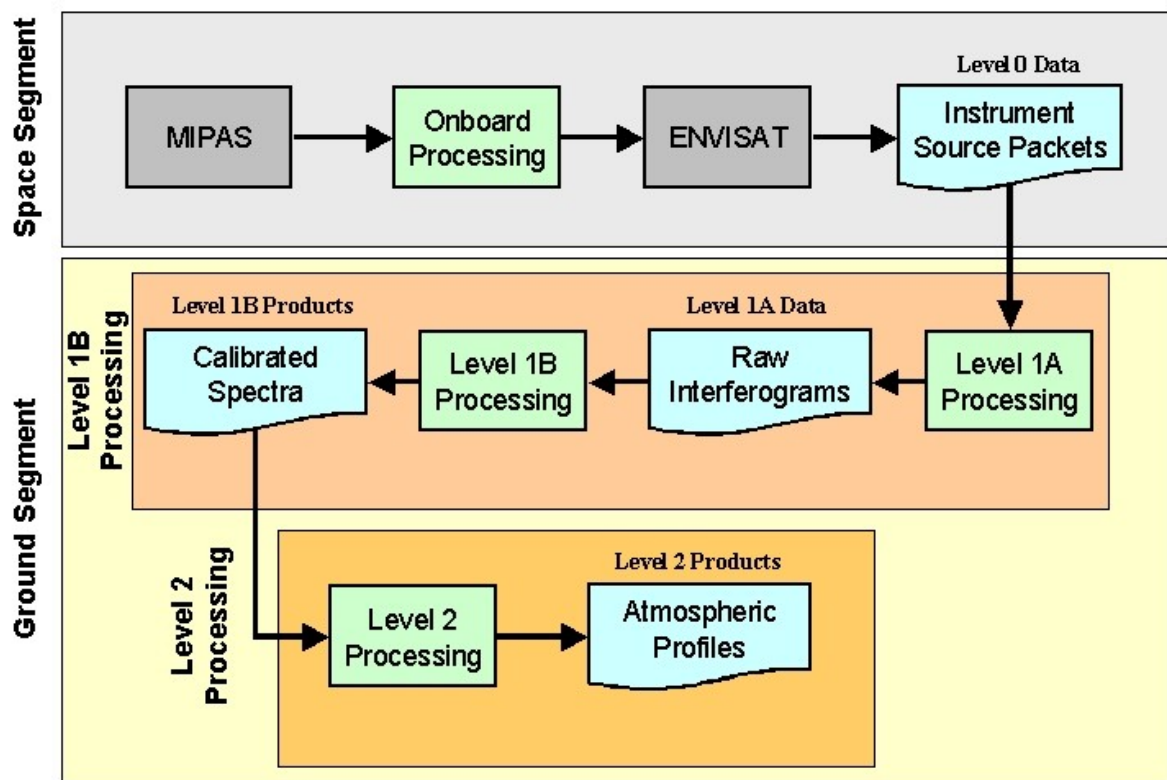

Figure 1.4: MIPAS data processing flowchart.

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, HNO_3, CH_4, N_2O and NO_2.

1.2.5 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 26th, 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 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 a 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 was operated in discontinuous way and the data quality was marginally affected. In Table 1.2, is described the new parameters associated to the observation mode, modified to match the reduced resolution.

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.

From January 2005 to December 2007, MIPAS was back in operation working with the measurement scenario of 3 days-on and 4 days-off. 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.

Between January 2005 and December 2007, the instrument operated on a 35% duty cycle but since December 2007 has been operating full-time with a standard 10 day sequence of 8 days nominal mode, 1 day MA, 1 day UA. [4], [5]

Observation Mode	Acronym	Description	Coverage	Altitude Range (km)	Scans Numbers	Horizontal Spacing (km)
Nominal	NOM	Stratospheric chemistry and dynamics	Global	5-77	27	410
Upper Troposphere / Lower Stratosphere	UTLS-1	Primary UTLS mode	Global	5.5-55	19	290
Upper Troposphere / Lower Stratosphere	UTLS-2	Test mode for 2-D retrieval	Global (limited numbers of orbits)	12-42	11	180
Middle Atmosphere	MA	Exchange between Upper Atmosphere and stratosphere	Global	18-102	29	430
Middle/Upper atmosphere in summer	NLC	Detection of Noctilucent clouds in the Polar Summer	North and South pole during polar summer	39-102	25	375
Upper Atmosphere	UA	Measurement high altitude NO and temperature	Global	42-172	35	515
Aircraft Emission	AE	Study of major air traffic corridor	Primarily North of 25° latitude	7-38	12	n.a.

Table 1.2: MIPAS nominal and special observation modes (after December 1, 2007)

The 8 April 2012, the contact with ENVISAT has been lost and all attempt to re-establish the communication has been unsuccessful. The ENVISAT satellite has already doubled its expected operations lifetime.

2 The Retrieval Problem

As I explained in chapter 1, the main objective of the MIPAS experiment is to obtain the altitude profiles of several targets relevant to the study of the atmosphere. The problem of retrieving the altitude profiles from the measured spectra is known as the Inverse Problem [3]. 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 [6]. The notation and concepts introduced in this chapter are inspired to those described there.

2.1 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.

2.1.1 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". 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":

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

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".

2.2 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 x (with n elements) to the "measurement vector" y (with m elements) may be written as:

$$\mathbf{y} = \mathbf{F(x)} + \varepsilon, \tag{2.1}$$

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 Section 2.2.2). 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:

$$\mathbf{y} - \mathbf{F(x_0)} = \frac{\partial \mathbf{F(x)}}{\partial \mathbf{x}}(\mathbf{x} - \mathbf{x_0}) + \varepsilon = \mathbf{K}(\mathbf{x} - \mathbf{x_0}) + \varepsilon$$

which defines the m×n "weighting function matrix" 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 **K**).

2.2.1 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:

$$x_e = \mathbf{K}^{-1}\,y, \qquad\qquad (2.2)$$

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 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:

$$\chi^2 \equiv (\mathbf{y} - \mathbf{K}\hat{x})^T (\mathbf{y} - \mathbf{K}\hat{x}) = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{K}\hat{x} - \hat{x}^T\mathbf{K}^T\mathbf{y} + \hat{x}^T\mathbf{K}^T\mathbf{K}\hat{x} \quad (2.3)$$

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

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

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

These are known as the "normal equations" of the least squares problem. Since $\mathbf{K}^T\mathbf{K}$ is an n × 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 $\hat{\mathbf{x}}$:

$$\hat{x} = (K^T K)^{-1} K^T \hat{y}. \qquad (2.4)$$

The matrix $(K^T K)^{-1} K^T \hat{y}$ is also known as the Moore-Penrose inverse of K. If the rank p of 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 S_y for the covariance matrix of y. Its diagonal elements are the variances of the individual elements of y. A covariance matrix is symmetric and non-negative definite. The Gaussian distribution for a vector is of the form:

$$P(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}|S_y|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\mathbf{y}-\hat{\mathbf{y}})^T S_y^{-1}(\mathbf{y}-\hat{\mathbf{y}})} \qquad (2.5)$$

where S_y must be non singular.

In presence of Gaussian distributed noise with zero mean and covariance matrix S_ε, the weighted least squares are used and the quantity to be minimized is, instead of expression (2.3):

$$\chi^2 = (\mathbf{y} - K\hat{\mathbf{x}})^T S_\varepsilon^{-1}(\mathbf{y} - K\hat{\mathbf{x}}), \qquad (2.6)$$

and equating to 0 the derivative of (2.6) with respect to $\hat{x}$, gives:

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

The expectation value of the expression (2.6) is $m - n$. Therefore a quantity called χ^2-test, or simply χ-test, can be defined as:

$$\chi - test = \frac{\chi^2}{m - n} \qquad (2.7)$$

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 [7] 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.

2.2.2 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 [7], 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:

$$[\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a]$$

or $\quad$ (2.8)

$$[\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})].$$

2.3 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 $\mathbf{F(x)}$. Newtonian iteration is a straightforward numerical method for finding the zero of the gradient of a given cost function, such as Eq. (2.8). For the general vector equation $\mathbf{g(x) = y - F(x) = 0}$, the Newton's iteration can be written:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\nabla_x \mathbf{g}(\mathbf{x}_i)]^{-1} \mathbf{g}(\mathbf{x}_i), \tag{2.9}$$

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

$$\nabla_x \mathbf{g} = \mathbf{S}_a^{-1} + \mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K} - [\nabla_x \mathbf{K}^T] \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F(x)}]$$

$$\text{or} \tag{2.10}$$

$$\nabla_x \mathbf{g} = \mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K} - [\nabla_x \mathbf{K}^T] \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F(x)}].$$

The function $\nabla_x \mathbf{g}$ is the second derivative of the cost function, Eq. (2.8), 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:

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

$$\text{or} \tag{2.11}$$

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

where $\mathbf{K}_i = \mathbf{K(x)}|\mathbf{x}_i$ and Eq. (2.11) represents the iterative solution in a non-linear problem with and without the a priori information.

2.3.1 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 [8] proposed the iteration:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_i + \gamma_i \mathbf{I})^{-1} \mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i)] \tag{2.12}$$

where γ_i is chosen at each step as the value that minimises the cost function. Marquardt [9] 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. [10]:

- if χ^2 increases as a result of a step, increase γ, do not update x_i and try again,

- if χ^2 decreases as a result of a step, update x_i and decrease γ for the next step.

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

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

3 MIPAS Retrieval Algorithm

3.1 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 2). Figure 3.1 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. (2.12). 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 [3].
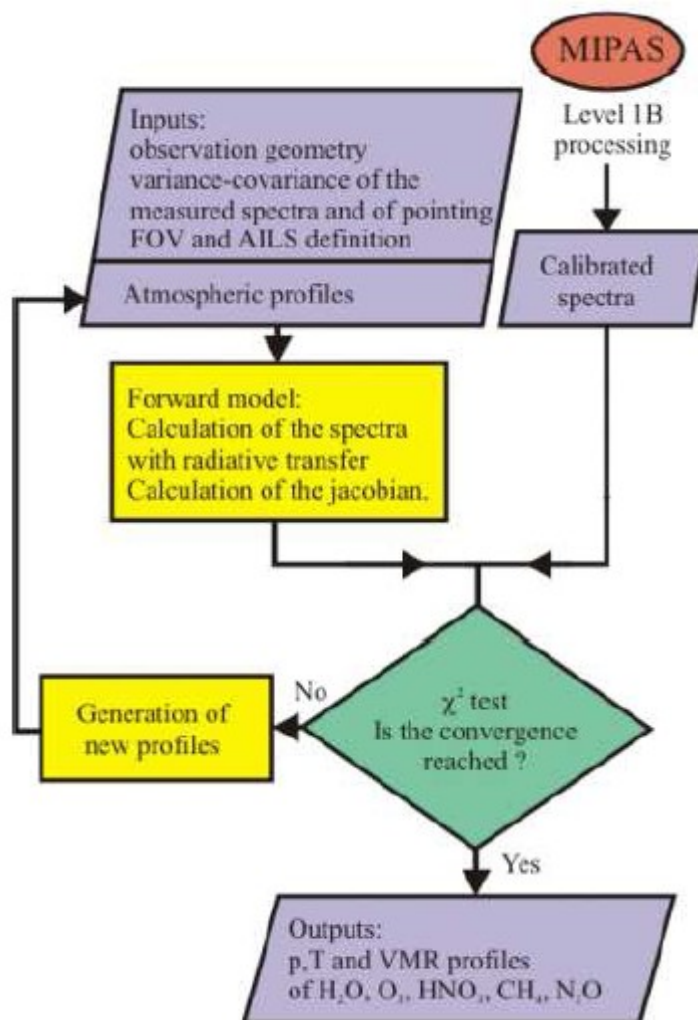
Figure 3.1: MIPAS retrieval algorithm sequence

3.1.1 Main Components of the Retrieval System

● **Initial guess atmosphere**

The initial guess of the target profiles and the profiles of the non-target 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 [11].

● **The Forward Model**

The Forward model simulates the spectra measured by the instrument in the case of known atmospheric composition. The forward model is capable of both simulating observations along a full orbit and accounting for the horizontal inhomogeneity of the atmosphere. This model, when it is calculating the signal that reaches the spectrometer, must be able to evaluate all the physical and chemical conditions that are encountered along the line of sight of a given observation. 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 [3].

● **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 retrieval process, 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 [3].

● **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 software are a compromise between the required accuracy of the parameters and the computing speed of the algorithm. The adopted convergence criteria are based on the conditions:

• 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 one of the conditions is satisfied. If only the condition on computing time is satisfied, the retrieval is considered unsuccessful [3].

3.2 Retrieval Characteristics

3.2.1 Use of Micro-Windows (MWs)

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 [13]. The use of narrow spectral intervals, called "micro-windows" (MWs), allows one to limit the number of analysed spectral elements and to avoid the analysis of spectral regions that are characterized by uncertain spectroscopic data, interference by non-target species, or are influenced by unmodeled effects, such as non-local 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 analysed 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 [13]; it operates the selection of the spectral elements with the aim of maximizing the information content [14] of the used measurements with respect to the target parameters [11].

3.2.2 The Geofit Multi-Target Retrieval

The Geofit Multi-Target Retrieval (GMTR) is a new retrieval model implements the geo-fit two-dimensional inversion for the simultaneous retrieval of several targets including a set of atmospheric constituents that are not considered by the ground processor of the MIPAS experiment. The detailed descriptions of the GMTR can be found in [16,17]. This retrieval model is implemented in an optimized computer code that is distributed by the European Space Agency as "open source" in a package that includes a full set of auxiliary data for the retrieval of 28 atmospheric targets. The GMTR have the following characteristics [11]:

● **Multi-target Retrieval**

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 analysed for the retrieval of other species. The error propagation process can be minimized with a careful choice [13] of both the analysed spectral intervals and the sequence of the retrievals. Nevertheless, the error propagation cannot be completely avoided and its assessment requires some post-processing 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. This strategy is referred to as multi-target retrieval (MTR). 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 analysed 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 [11].

● **Geofit 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. 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, see chapter 2, (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 analysed 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, [15] 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 [11].

● **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 the typical retrieval sequence considered in my work, 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 [11].

3.3 Computing Details

The new analysis system is organized into two separated modules: a preprocessor and a retrieval module. The preprocessor sets up a suitable environment for the retrieval module; its main functionalities consist of:

i. Combining (if available) two Level-1b files to compose a full pole-to-pole orbit and extracting the subset of spectral data points to be analysed.

ii. Mapping the geometrical coordinates from the three-dimensional reference frame into the two-dimensional reference frame in the orbit plane.

iii. Selecting the auxiliary information from the targets database.

iv. Building the initial atmospheric model.

The characteristics of the input data and the evolution of the retrieval analysis is controlled through a "settings file" (common to the two modules) whose entries are defined by the user. In the "settings file" each entry follows a key-word that is associated to that entry. This strategy makes irrelevant the order in which the entries appear within the file so that the user can organize the file at his convenience.

Since the retrieval module is independent from the preprocessing module, the user is allowed to run the retrieval with different settings on a fixed set of input files generated by a run of the preprocessor module.

Theses different modules consists of many blocks which are written mainly in standard FORTRAN 77. The only exceptions are some functions required to read Level 1b and Level 2 products that are written in C and called by FORTRAN routines [11,18].

4 Diagnostics Tools

In this chapter, I describe the diagnostics tools used in the experimental part to perform the analysis of the two sets of MWs.

4.1 Simulation of data

The GMTR software can be used to generate simulated retrievals (retrieval analyses operated on simulated observations). This option may be useful to test and select the most efficient initial parameters, e.g. the selection of an optimised set of MWs. The steps to generate the simulated retrievals are:

1. The creation of synthetic spectra for the whole orbit using the 2-D forward model which uses a set of reference profiles that describe the horizontal structures of the atmosphere. A special option is used to run only the forward model and then directly stop the GMTR software.

2. Random perturbations are introduced into the synthetic spectra, in order to start from an initial guess atmosphere different from the reference atmosphere and so avoid the convergence in only one iteration.

3. The software is operated normally using the spectra created by the forward model as if they were recorded by the MIPAS experiment.

This strategy provides the reference profiles which correspond to the initial atmosphere. The initial atmosphere corresponds to the atmosphere that the retrieval process should retrieve if the retrieval process were perfect. The comparison of the retrievals of the two sets of MWs is considerably facilitated since we can compare the results of each of them with the reference profiles. Indeed, the use of simulated retrievals is suitable to assess the real errors that are introduced when, for example, using different sets of MWs [11].

4.2 Information Load

For the analysis of limb-scanning observations with the 2-D approach the atmosphere is partitioned on both the vertical and the horizontal domains [19], [16]. This discretization leads to a web-like picture in which consecutive altitude levels and vertical radii define plane figures denoted as "cloves" (see Fig. 4.1). If the simultaneous analysis of several observation geometries is considered, we can assign to each clove the Information Load (IL) scalar quantifier (Ω) defined as [20]:

$$\Omega(q,h) = \left[\sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} \left(\frac{\partial S_{ijk}}{\partial q_h} \right)^2 \right]^{1/2}$$

(Eq. 4.1)

$\Omega(q,h)$ = IL of clove h with respect to atmospheric parameter q.
S_{ijk} = spectral signal of observation geometry i at frequency j of the analysed MW k.

l = number of observation geometries that go through clove h,
m = number of MWs in observation geometry i,
n = number of spectral points in MW j.

Eq. (4.1) can be written as:

$$\Omega(q,h) = \left[(\mathbf{k}^{\mathrm{T}}\mathbf{k})_h \right]^{1/2}$$

(Eq. 4.2)

where $\mathbf{k}$ is a vector containing the derivatives of all the observations that depend on the value of q in clove h. If the observations are affected by different noise levels (e.g. occur in different spectral bands) it is suitable to use the weighted IL (WΩ) defined as:

$$W\Omega(q,h) = \left[(\mathbf{k}^{\mathrm{T}}\mathbf{S}^{-1}\mathbf{k})_h \right]^{1/2}$$

(Eq. 4.3)

where $\mathbf{S}$ is the variance-covariance matrix of the observations relative to all the spectral points that contribute to the information load in clove h.

The uncertainty on the value of the target quantity q in clove h is connected with the quantity $1/W\Omega$ (see [20]).

A map of the WΩ quantifier, calculated for each clove of the 2-D atmospheric discretization permits to evaluate the distribution of the information load with respect to the geophysical parameter q. In these maps:

● The values of WΩ measure the amount of information contained in the corresponding cloves.

● The spatial distribution of WΩ indicates the regions of the atmosphere where the information is gathered from when operating a retrieval analysis.

The IL analysis provides a tool to compare the performance of different observation strategies and/or of different sets of spectral intervals. The WΩ maps also indicate the optimal strategy to select the geo-location of the retrieval grid.

In the 2-D context the observation geometries that look at clove h may come from different limb-scans while in the 1-D context they are a subset of the observation geometries of the analysed limb-scan.

The IL will be computed thanks to the software called "GMTR_INFO", derived from the GMTR software (Chap. 3)
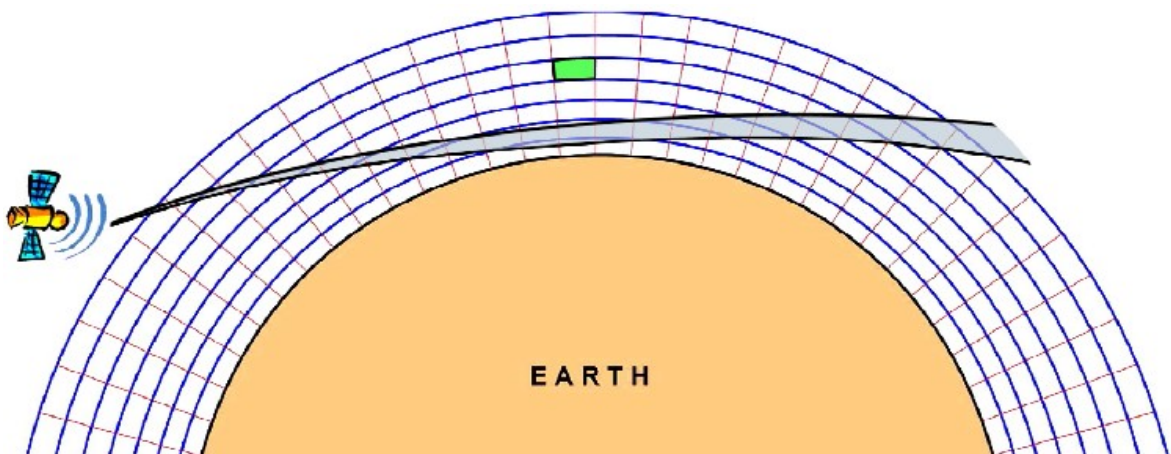
Figure 4.1: 2D discretization of the atmosphere and ray-tracing of the limb-view. A sample "clove" is high-lighted in green

4.3 Systematic and Random errors

The measurement are affected by errors that can be divided into two types of errors, indeed the total error is the square summation of the random and systematic errors. The random error is due to the propagation of instrument noise through the retrieval. Random errors are inconsistent with the repetition of measurement, they are scattered about the true value and tend to lack arithmetic meaning. Random errors are caused by unpredictable fluctuations due to interference of the environment with the measurement process.

The MW selection algorithm provide an estimate of the systematic error, which is the root sum square of individual systematic error sources:

• Non Local Thermodynamic Equilibrium (NLTE) error is due to assumption of local thermodynamic equilibrium when modelling emission in the MIPAS forward model. Based on calculations using vibrational temperatures supplied by M.Lopez-Puertas, IAA, Granada.

• Spectroscopic Database Errors (SPECDB formerly referred to as HITRAN) is due to uncertainties in the strength, position and width of infrared emission lines. Based on estimates supplied for each molecule/band by J.M.Flaud, LPM, Paris.

• Radiometric Gain Uncertainty (GAIN) is due mostly to non-linearity correction in bands of the recorded spectra.

• The SPREAD is the uncertainty in width of apodised instrument line shape (AILS). A value of 0.2% has been assumed based on likely variations in apodised instrument line shape from modelled.

• The SHIFT is the uncertainty in the spectral calibration. The design specification of $\pm 0.001 cm^{-1}$ has been used, and is consistent with the 1st derivatives signatures in the residual spectra.

• CO_2 line-mixing (CO2MIX) is due to neglecting line-mixing effects in the retrieval forward model (only affects strong CO_2 Q branches in the MIPAS spectral bands).

• The CTMERR is the uncertainty in gaseous continua. Assumes an uncertainty of $\pm 25\%$ in the modelling of continuum features of H_2O (mostly), CO_2, O_2 and N_2.

• Horizontal gradient effects (GRA) is due to the assumption in the retrieval process of a horizontally homogeneous atmosphere for each profile. Error is calculated assuming a $\pm$ 1K/100km horizontal temperature gradient.

• Uncertainty in high-altitude column (HIALT) is due to the assumption in the Retrieval process of a fixed-shape of atmospheric profile above the top retrieval level. Effect is calculated assuming `true' profile can deviate by climatological variability.

• The propagation of Pressure and Temperature (PT) random covariance into VMR retrieval.

• The uncertainties in assumed profiles (profiles supplied by J.Remedios, U.Leicester) of contaminant species (i.e. CH4, H2O, HNO3, N2O, NO2, O3) [species]. For most species this is the climatological 1-σ variability.

The definition of 'systematic error' here includes everything which is not propagation of the random instrument noise through the retrieval. However, to use these errors in a statistically correct manner for comparisons with other measurements is not straightforward. Each systematic error has its own length/time scale: on shorter scales it contributes to the Bias and on longer scales contributes to the Standard Deviation (SD) of the comparison.

Fortunately, two of the larger systematic errors (PT and SPECDB) can be treated properly:

• The propagation error (PT) is uncorrelated between any two MIPAS profiles (since it is just the propagation of the random component of the PT retrieval error through the VMR retrieval) so contributes to the SD of any profile comparison.

• Spectroscopic database errors (SPECDB) are constant but of unknown sign, so will always contribute to the Bias of any comparison, but note that the magnitude of these errors is very uncertain.

Of the other significant errors, the calibration-related errors (GAIN, SHIFT, SPREAD) should, in principle, be uncorrelated between calibration cycles however analysis of the residuals suggests that these errors are almost constant so could be included in the Bias.

The high altitude column (HIALT) and contaminant gas errors ([species]) are likely to be correlated over small areas (1000km) or times (weeks), hence contribute to the Bias for localised comparisons, but as the comparison datasets are extended these errors will contribute more to the SD.

Line mixing errors (CO2MIX) are also contribute towards the Bias but in principle the sign of these errors is known (unlike spectroscopic errors) so this bias could be removed. NLTE errors should also, in principle, contribute a known Bias but these are highly variable (especially diurnally) so care has to be taken to make sure that representative conditions for the comparison are used. [21]

Performance

of the

Micro-Window Sets

5 Comparison of the two sets of Micro-Windows using retrievals on simulated observations

The MIPAS instrument is recording spectra which are then used to retrieve the altitude distribution of different targets (Chap 1). This duty is achieved thanks to the GMTR software (Geofit Multi-Target Retrieval), described in Chap. 3. To avoid the redundancy of information and the interference between species the GMTR software extracts, from the spectra, narrow spectral intervals, called Micro-Windows (MWs) (see Sub-Sect. 3.2.1). The Oxford University and the University of Bologna have provided different sets of MWs. The aim of this chapter is to compare the efficiency of these two sets of MWs. I am comparing the two sets of MWs using retrieval on simulated observations, in order to compare the retrieved profiles w.r.t. the reference profiles (Sect. 4.1).

The orbits 31277 and 31278 of February, 23th 2008 are used to calculate the simulated observations in chapters 5 and 7 and to compute the IL values in chapter 6.

5.1 Comparison of the simulated retrievals

As described in Sub-Sect. 3.3.3 the GMTR software can be use to generate simulated retrievals. In order to compare the retrieved profiles from Oxford University (OX) and Bologna University (BO) sets of MWs, I used the MAPPA tool to convert the profiles over the full orbit into maps. MAPPA is a graphical software written in IDL language which plots all the retrieved profiles corresponding to one species into a single map, as a function of altitude and Orbital Coordinate (OC), that is the angular coordinate θ originating at the North pole and spanning the orbit plane over its 360° extension.
The values plotted on the maps are calculated using three equations:

$$DIFF = |ret - ref| \tag{5.1}$$

$$Ratio_{ref} = |ret - ref| / ref \tag{5.2}$$

$$Ratio_{esd} = |ret - ref| / esd \tag{5.3}$$

where "ret" corresponds to the profiles retrieved from the OX and the Bo sets of MWs, "ref" corresponds to the reference profiles and "esd" corresponds to the values of the Estimated Standard Deviation (ESD), as provided by GMTR. The purpose of these quantifiers will be described in Sect. 5.2.1. In Figs. 5.1 and 5.2 are plotted DIFF values for Temperature as a function of altitude and OC for OX and Bo sets of MWs respectively. In Figs. 5.1 and 5.2 I can localise very precisely where the divergence occurs between the retrievals from BO and OX sets of MWs. But to compare the performance of the two sets of MWs it is preferable to have an overall picture of the divergence between the retrievals. In Sect. 5.2, I describe the employed strategy.
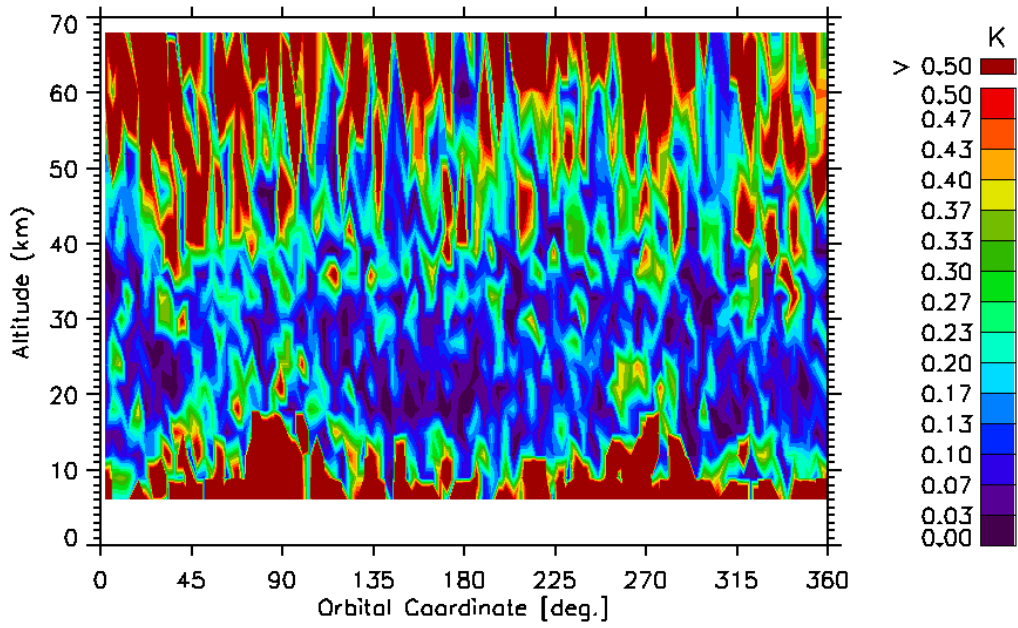
Figure 5.1: DIFF values for Temperature as a function of altitude and OC for the OX set of MWs
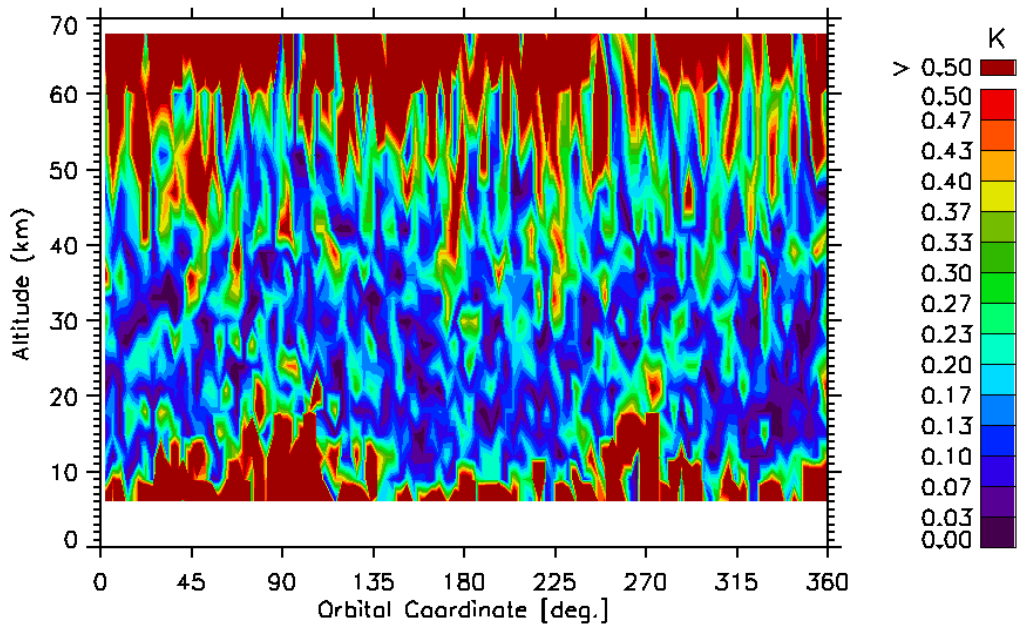
Figure 5.2: DIFF values for Temperature as a function of altitude and OC for the BO set of MWs

5.2 Comparison of the simulated retrievals using Quantifiers

To facilitate the comparison of the retrievals, I will introduce in sub-section 5.2.1, several quantifiers which refer to a single profile obtained with a full coverage orbit average.

5.2.1 Quantifiers Descriptions

a) AVE[DIFF] is the average value over the full orbit, at each altitude, of the absolute difference between retrieved and reference profile :

$$AVE[DIFF] = \frac{\sum_{i=1}^{N} |ret_i - ref_i|}{N} \qquad \text{(Eq. 5.4)}$$

where N represents the number of values over the full orbit at one given altitude.

AVE[DIFF] represents the average difference (without sign) between the retrieved values and the reference values. As explained in Sub-Sect. 3.3.3, the reference profiles represent what would be the retrieved profiles in the case of a perfect retrieval process, so AVE[DIFF] represents the "real" error of the retrieval process. On the other hand, the performance of the retrieval process is linked to the selection of the MWs, i.e. a better choice of MWs will lead to a better efficiency of the retrieval process. Therefore comparing the AVE[DIFF] values arising from the two sets of MWs, I can state which of the two sets of MWs is the most efficient.

For each target, the AVE[DIFF] values of the OX and BO sets of MWs will be plotted as a function of altitude, in black and in blue respectively. To abridge the description of these graphs, a divergence between the two curves will be directly connected to a better or a worse retrieval process which is caused by one of the two sets of MWs.

b) AVE[Ratio$_{ref}$] is the percentage of AVE[DIFF] w.r.t. the retrieved profile:

$$AVE[Ratio_{ref}] = \frac{\sum_{i=1}^{N} (|ret_i - ref_i|/ref_i)}{N} \times 100 \qquad \text{(Eq. 5.5)}$$

AVE[Ratio$_{ref}$] indicates the magnitude of the average difference w.r.t. the magnitude of the reference profiles (e.g. if the average difference is high w.r.t. the magnitude of the reference values, the result of this equation will be significantly high). Plotted values on the AVE[DIFF] graphs will be respectively revealed or hidden depending on their

relevance.

For each target, the AVE[Ratio$_{ref}$] values of the OX and BO sets of MWs will be plotted as a function of altitude, in black and in blue respectively. This quantifier will help us to judge the improvement or worsening magnitude (slight, intermediate or large).

c) AVE[Ratio$_{esd}$] is the average of DIFF divided by the values of the ESD:

$$AVE[Ratio_{esd}] = \frac{\sum_{i=1}^{N} (|ret_i - ref_i|/esd_i)}{N} \qquad \text{(Eq. 5.6)}$$

AVE[Ratio$_{esd}$] indicates the magnitude of the average difference w.r.t. the magnitude of the ESD. The ESD represents the estimated random error of the retrieved values. If this ratio is lower than three, i.e. the average difference is not three times greater than the ESD, then ESD and AVE[DIFF] are considered consistent. Moreover, the difference represent the real error therefore the closer the average value of the oscillations of this quantifier is to one, the better the ESD is reliable.

For each target, the AVE[Ratio$_{esd}$] values of the OX and BO sets of MWs will be plotted as a function of altitude, in black and in blue respectively. To abridge the description of these graphs, I will directly link an average value of the AVE[Ratio$_{esd}$] oscillations nearer to one to a better consistency between the real and estimated error of one of the two sets of MWs.

d) AVE[ESD] is the average of ESD:

$$AVE[ESD] = \frac{\sum_{i=1}^{N} esd_i}{N} \qquad \text{(Eq. 5.7)}$$

AVE[ESD] indicates the estimation of the software on the value of the error.

For each target, the AVE[ESD] values of the OX and BO sets of MWs will be plotted as a function of altitude, in black and in blue respectively.

As described in Sub-Sects. 3.3.1 and 3.2.2, I used the "cascade" mode to retrieve first the four mains targets simultaneously: Temperature, Pressure, Ozone and Water and then Nitric Acid, Methane, Nitrous Oxide and Nitrogen Dioxide. The magnitude of the peaks on the recorded spectrum will depend of the excitation of the molecules, therefore the results of the four last targets is dependant of the result of the mains targets (mainly temperature). Although the tested MWs only concern the retrieval of the four mains targets, I will also report the plotted quantifiers of the four others species in order to identify some eventual side-effects.

In order to compare the efficiency of the two sets of MWs, I present from Sub-Sects 5.2.2 to 5.2.9 the plotted quantifiers of the set of MW provided by Oxford University (OX) and the set of MWs provided by the University of Bologna (BO) as a function of altitude on the same Figure.

5.2.2 Temperature

In Figs. 5.3 and 5.4 the $Ratio_{ref}$ values are plotted as a function of altitude and OC for OX and BO sets of MWs respectively. In Figs. 5.5 and 5.6 the $Ratio_{esd}$ values are plotted as a function of altitude and OC for OX and BO sets of MWs respectively. In Figs. 5.1, 5.2 (from Sect. 5.1), 5.3 and 5.4 it seems that the red part at high altitudes is larger for DIFF and $Ratio_{ref}$ retrieved with the OX than with the BO set of MWs, which indicates a worse retrieval at high altitude caused by the OX set of MWs. In contrary, it seems that the red part at low altitudes is smaller for DIFF and $Ratio_{ref}$ retrieved with the OX set of MWs rather than with the BO set of MWs, which indicates a better retrieval at low altitude thanks to the OX set of MWs. In Figs. 5.5 and 5.6 I observe larger and more red spots for $Ratio_{esd}$ from profiles retrieved with the OX set of MWs.
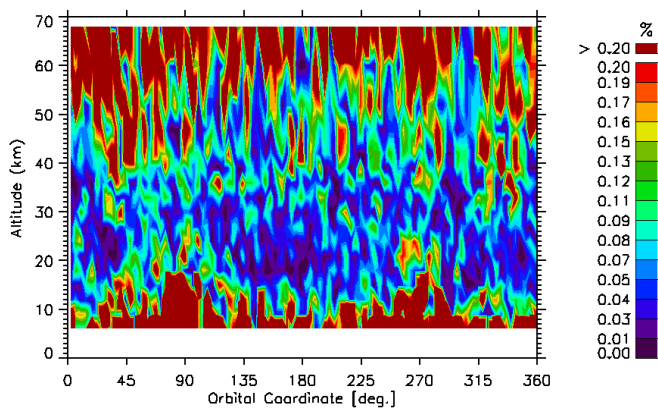
Figure 5.3: $Ratio_{ref}$ values for Temperature as a function of altitude and OC for the OX set of MWs
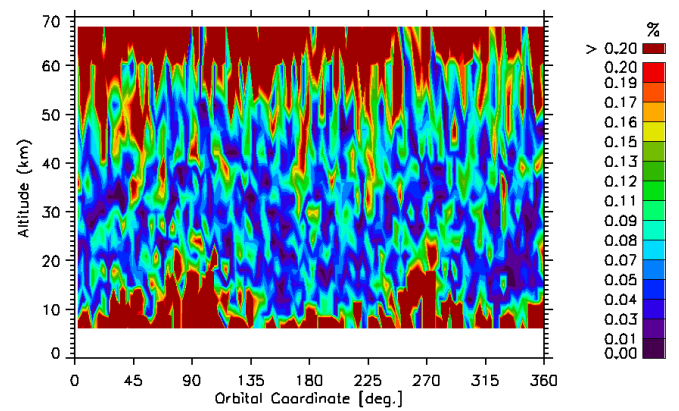
Figure 5.4: $Ratio_{ref}$ values for Temperature as a function of altitude and OC for the BO set of MWs
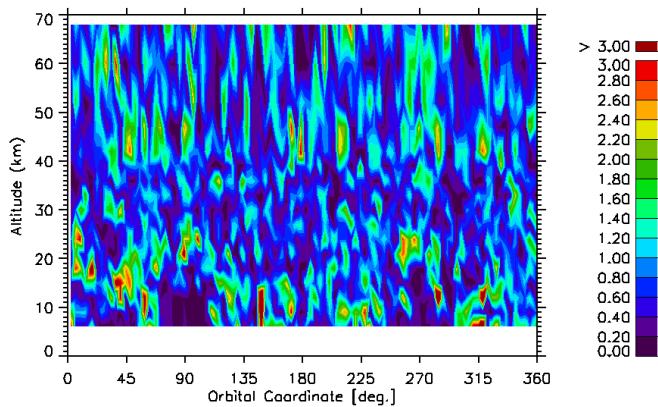
Figure 5.5: $Ratio_{esd}$ values for Temperature as a function of altitude and OC for the OX set of MWs
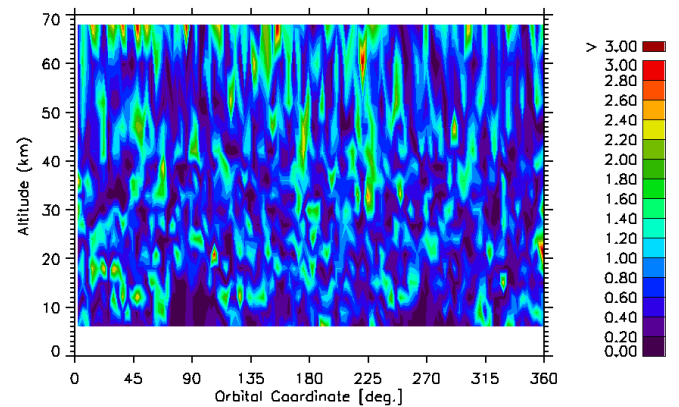
Figure 5.6: $Ratio_{esd}$ values for Temperature as a function of altitude and OC for the BO set of MWs

Figs. 5.7, 5.8, 5.9 and 5.10 report respectively AVE[DIFF], AVE[Ratio_ref], AVE[Ratio_esd] and AVE[ESD] values for Temperature as a function of altitude. In Fig. 5.7 and 5.8, I observe for the OX set of MWs a better performance at low altitudes (15 to 36 km) and a worse performance at higher altitudes (39 to 60 km). Fig. 5.9 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.7) and OX (0.85) sets of MWs which indicates a better consistency between the real and estimated error for the OX set of MWs. In Fig. 5.10 I observe for the OX set of MWs significantly lower AVE[ESD] values, from 18 to 33 km, and slightly higher AVE[ESD] values, from 33 to 50 km.
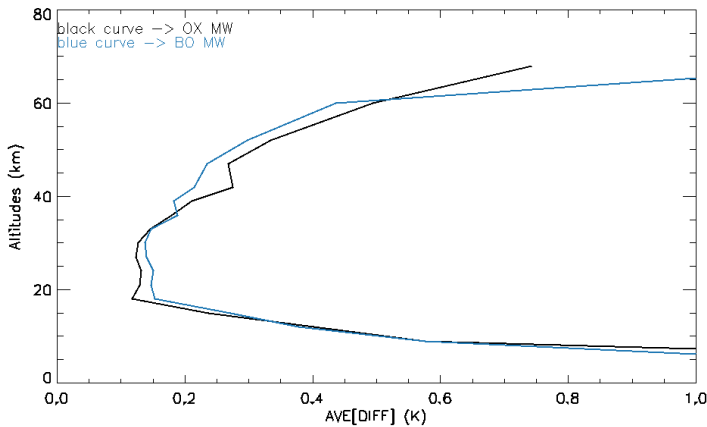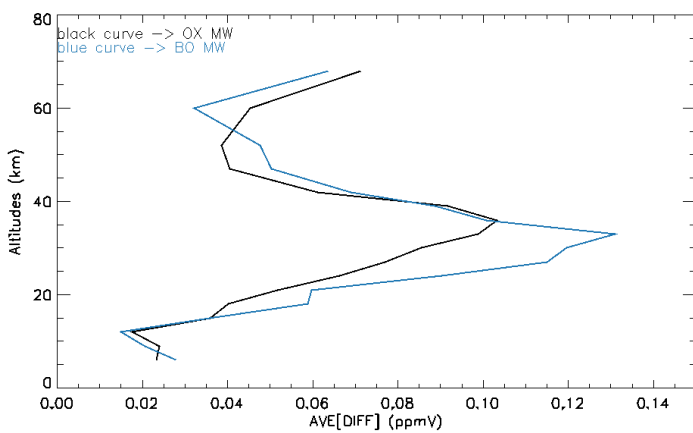

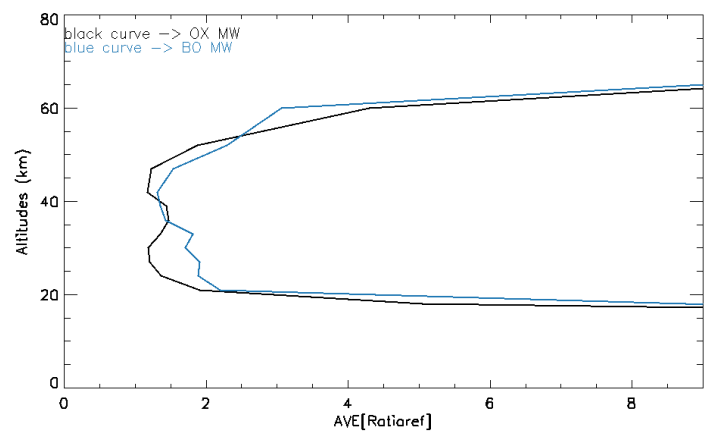Figure 5.7: AVE[DIFF] values for Temperature as a function of altitude for the two sets of MWs
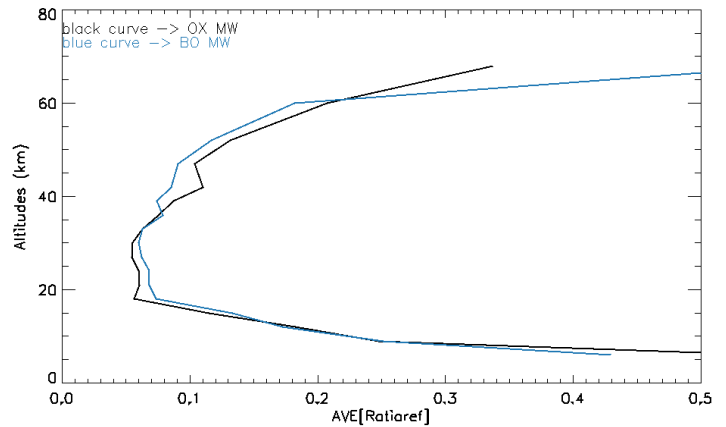
Figure 5.8: AVE[Ratio_ref] values for Temperature as a function of altitude for the two sets of MWs
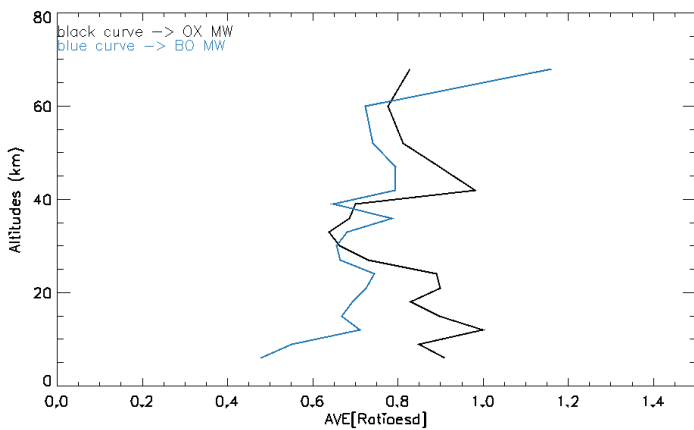
Figure 5.9: AVE[Ratio_esd] values for Temperature as a function of altitude for the two sets of MWs
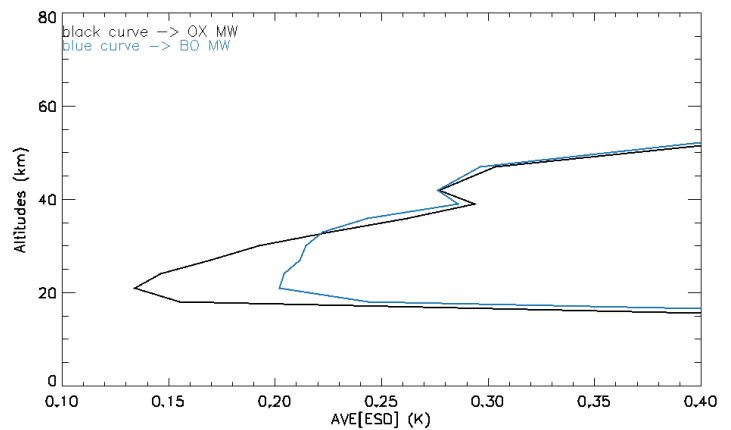
Figure 5.10: AVE[ESD] values for Temperature as a function of altitude for the two sets of MWs

I remark that it is more suitable and easier to comment and compare on Figs. 5.7 to 5.10 the efficiency of the two sets of MWs, than on Figs. 5.1 to 5.6. Therefore for the other targets the maps will not be introduced except if they can be used to localise precisely the improvement or the worsening of one of the quantifier.

37

5.2.3 Pressure

Figs. 5.11, 5.12, 5.13 and 5.14 report respectively ln(AVE[DIFF]), AVE[Ratio_ref], AVE[Ratio_esd] and ln(AVE[ESD]) values for Pressure as a function of altitude. In the case of Pressure and Water, I used a logarithmic scale for AVE[DIFF] and AVE[ESD] since these targets have a exponential decay w.r.t. altitude. In Fig. 5.11 and 5.12, I observe for the OX set of MWs better performance from 9 to 33 km and worse performance from 40 to 47 km. Fig. 5.13 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.75) and OX (0.9) sets of MWs which indicates a better consistency between the real and estimated error for the OX set of MWs. In Fig. 5.14 I observe for the OX set of MWs slightly lower AVE[ESD] values from 6 to 55 km and slightly higher AVE[ESD] values from 55 to 68 km.
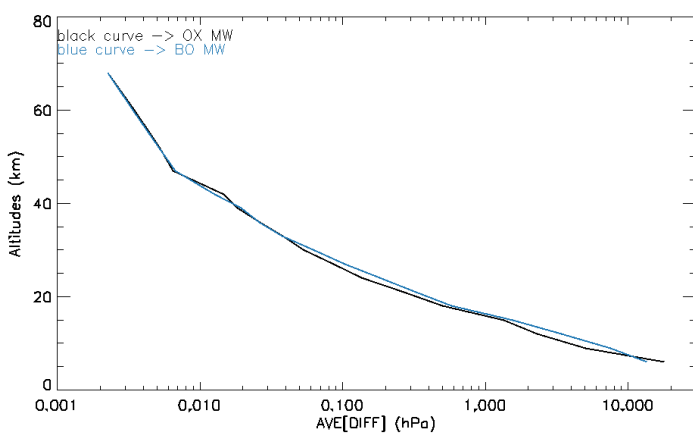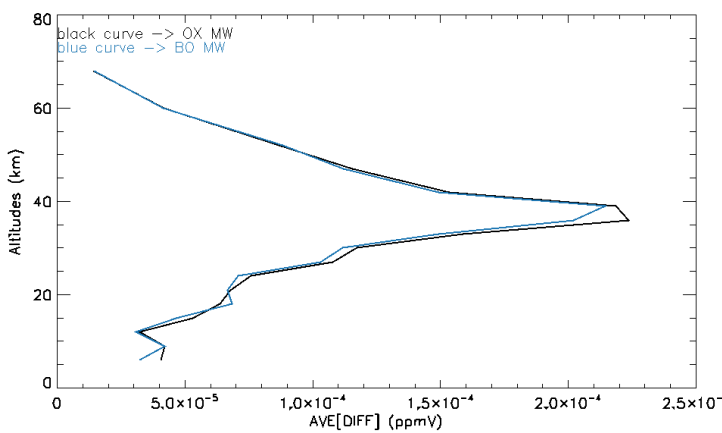

Figure 5.11: AVE[DIFF] values for Pressure as a function of altitude for the two sets of MWs
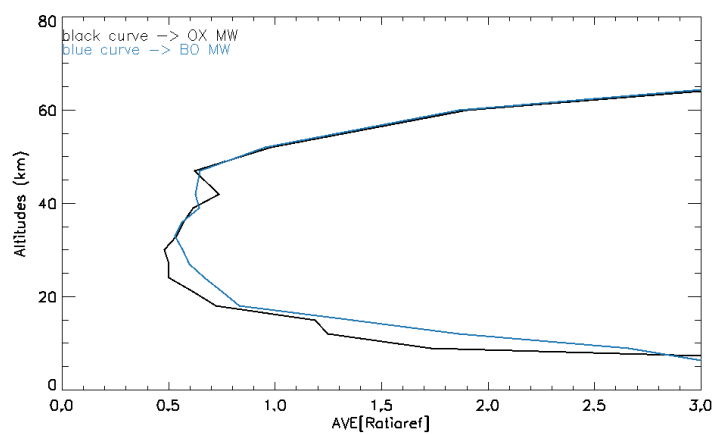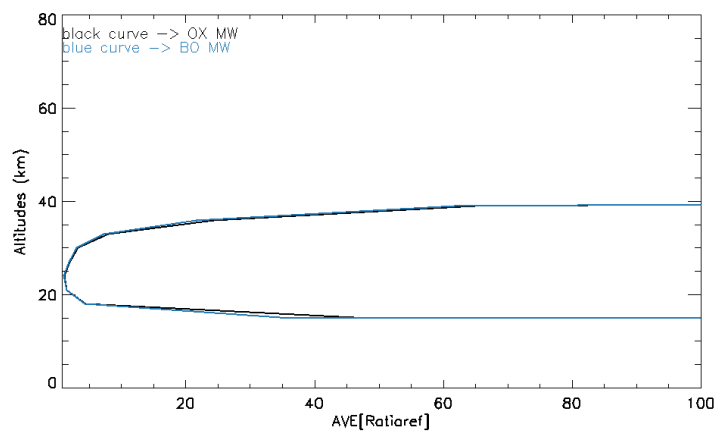

Figure 5.12: AVE[Ratio_ref] values for Pressure as a function of altitude for the two sets of MWs
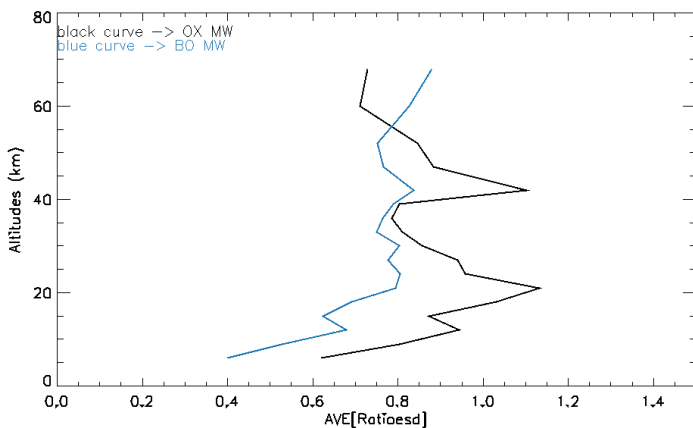
Figure 5.13: AVE[Ratio_esd] values for Pressure as a function of altitude for the two sets of MWs
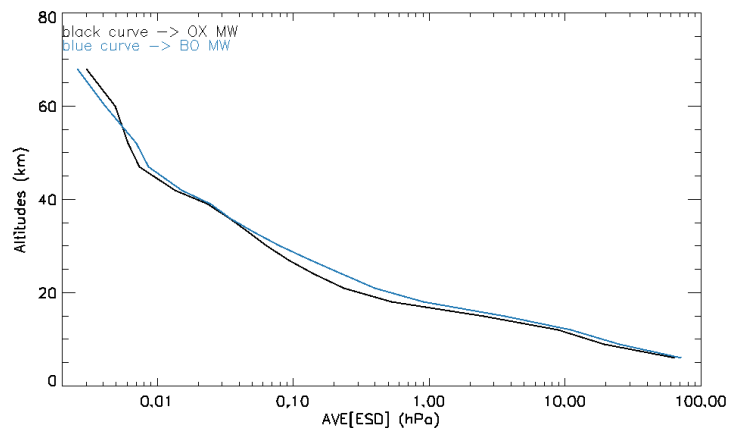
Figure 5.14: AVE[ESD] values for Pressure as a function of altitude for the two sets of MWs

5.2.4 Water
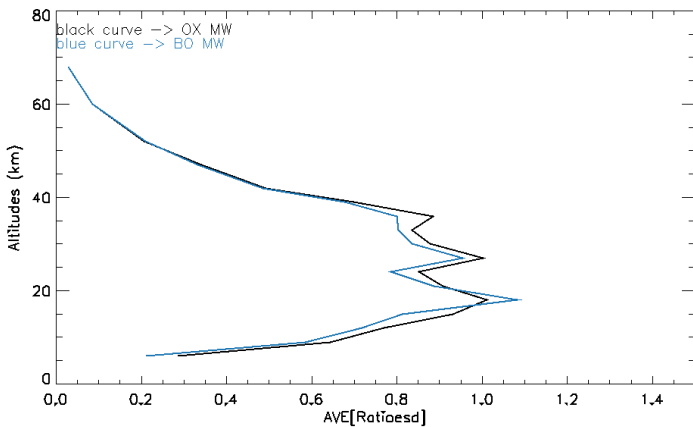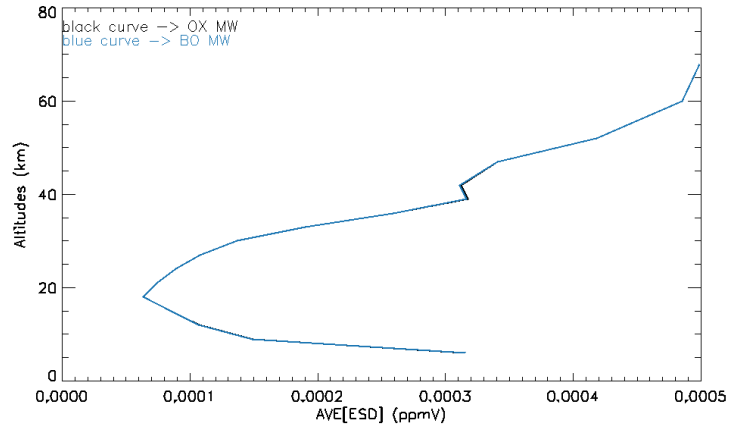
Figs. 5.15, 5.16, 5.17 and 5.18 report respectively ln(AVE[DIFF]), AVE[Ratio$_{ref}$], AVE[Ratio$_{esd}$] and ln(AVE[ESD]) values for Water as a function of altitude. In Fig. 5.15 and 5.16, I observe for the OX set of MWs better performance around 21 km, from 37 to 45 km and from 50 to 55 km and worse performance from 15 to 18 km, from 24 to 36 km and from 55 to 68 km. Fig. 5.17 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.75) and OX (0.8) sets of MWs which indicates a slightly better consistency between the real and estimated error for the OX set of MWs. In Fig. 5.18 I observe for the OX set of MWs lower AVE[ESD] values from 21 to 24 km and from 40 to 55 km and higher AVE[ESD] values from 25 to 39 km and from 55 to 68 km.
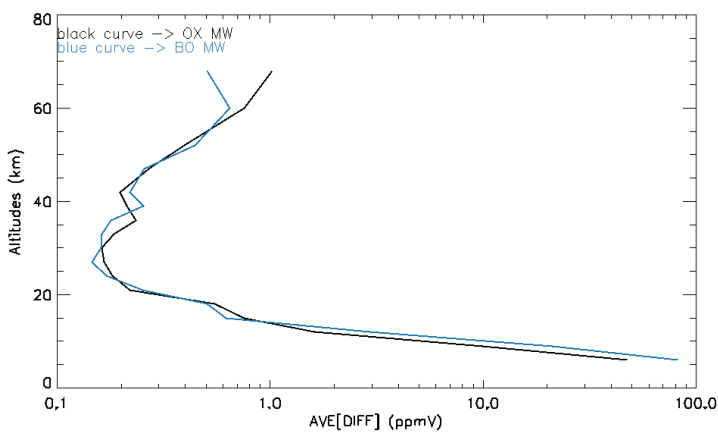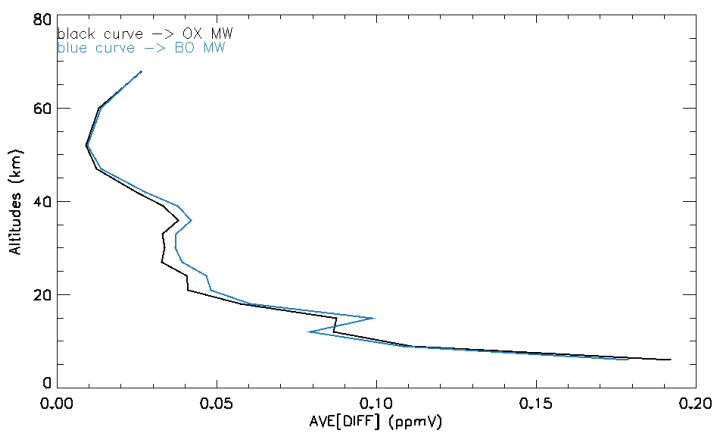

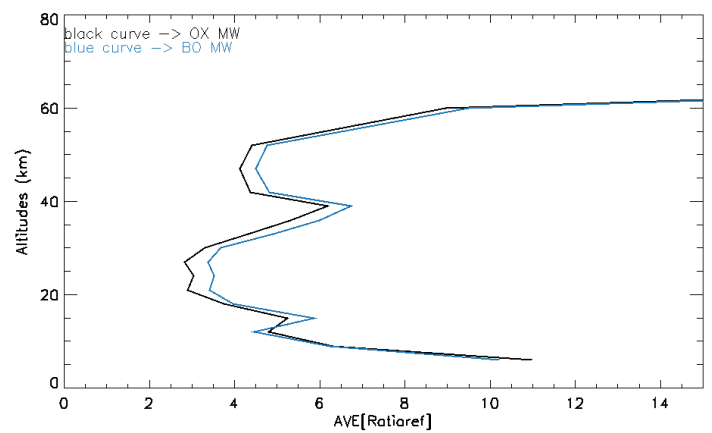Figure 5.15: AVE[DIFF] values for Water as a function of altitude for the two sets of MWs
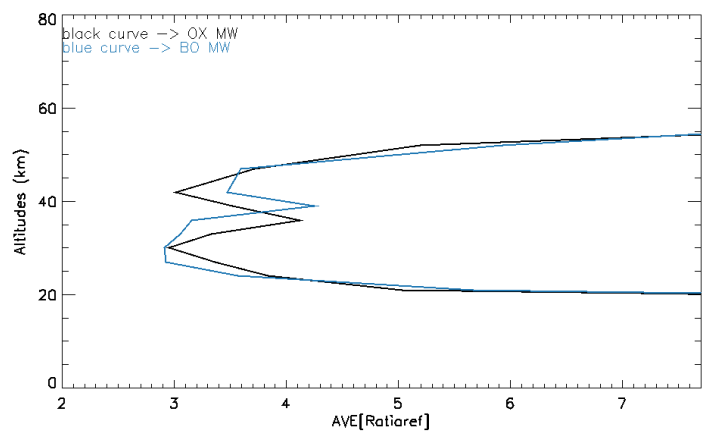
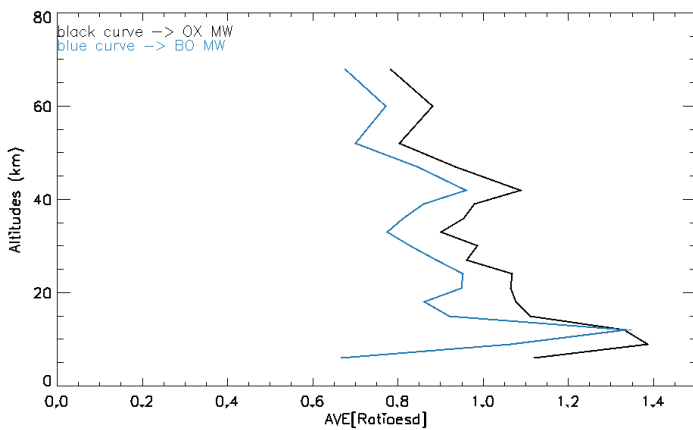Figure 5.16: AVE[Ratio$_{ref}$] values for Water as a function of altitude for the two sets of MWs
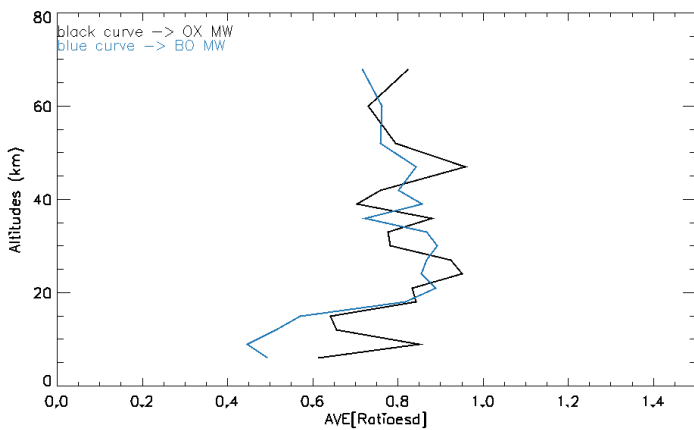
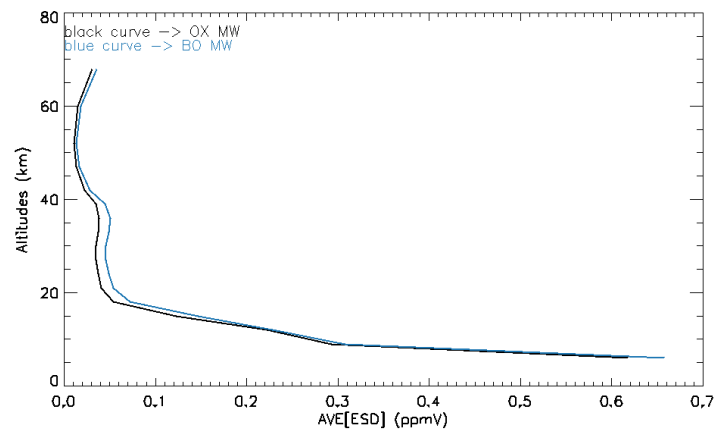Figure 5.17: AVE[Ratio$_{esd}$] values for Water as a function of altitude for the two sets of MWs
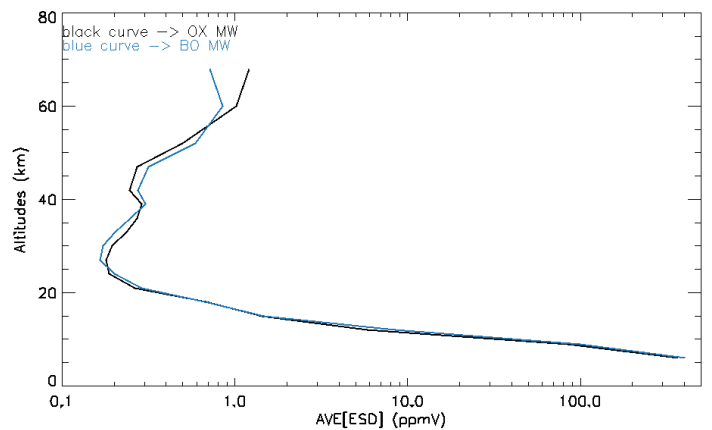
Figure 5.18: AVE[ESD] values for Water as a function of altitude for the two sets of MWs

5.2.5 Ozone

Figs. 5.19, 5.20, 5.21 and 5.22 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Ozone as a function of altitude. In Fig. 5.19 and 5.20, I observe for the OX set of MWs better performance from 15 to 35 km and from 40 to 55 km and worse performance from 9 to 12 km, from 35 to 40 km and from 55 to 68 km. Fig. 5.21 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.8) and OX (0.8) sets of MWs which indicates that the consistency between the real and estimated error is almost the same for the two sets of MWs. In Fig. 5.22 I observe for the OX set of MWs significantly lower AVE[ESD] values from 12 to 35 km.
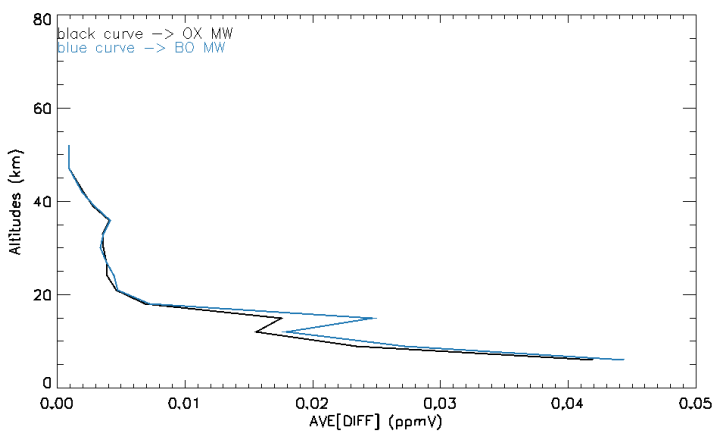
Figure 5.19: AVE[DIFF] values for Ozone as a function of altitude for the two sets of MWs

Figure 5.20: AVE[Ratio$_{ref}$] values for Ozone as a function of altitude for the two sets of MWs
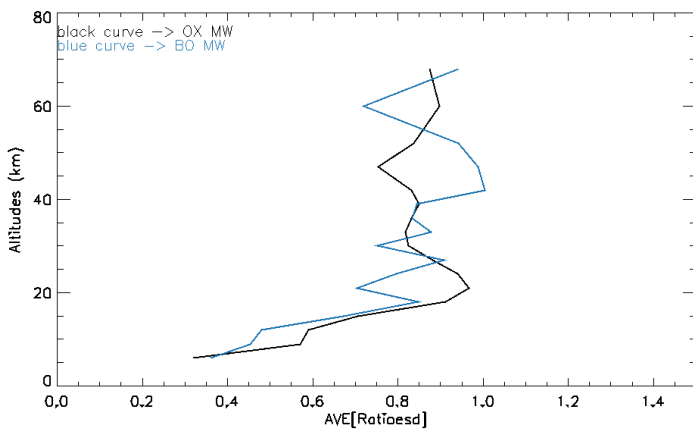
Figure 5.21: AVE[Ratio$_{esd}$] values for Ozone as a function of altitude for the two sets of MWs
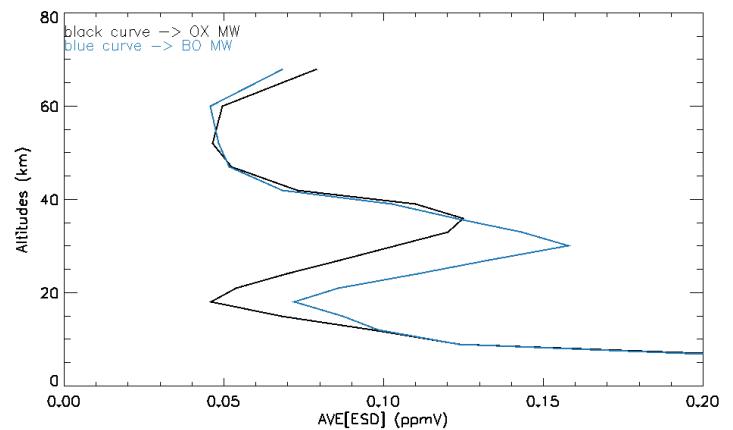
Figure 5.22: AVE[ESD] values for Ozone as a function of altitude for the two sets of MWs

5.2.6 Nitric Acid

Figs. 5.23, 5.24, 5.25 and 5.26 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Nitric Acid as a function of altitude. In Fig. 5.23 and 5.24, I observe for the OX set of MWs slightly worse performance from 35 to 40 km. Fig. 5.25 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.8) and OX (0.8) sets of MWs which indicates that the consistency between the real and estimated error is almost the same for the two sets of MWs. In Fig. 5.26 I observe that both sets of MWs have almost the same AVE[ESD] values.

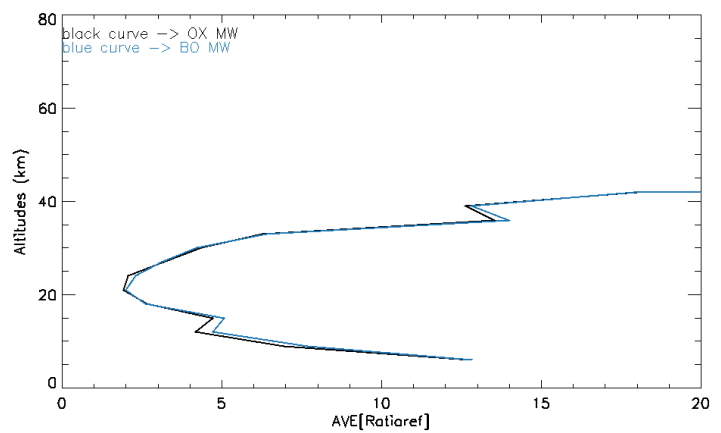Figure 5.23: AVE[DIFF] values for Nitric Acid as a function of altitude for the two sets of MWs

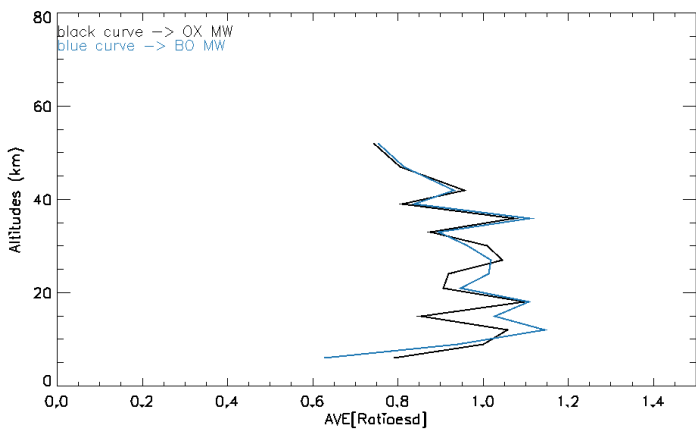Figure 5.24: AVE[Ratio$_{ref}$] values for Nitric Acid as a function of altitude for the two sets of MWs

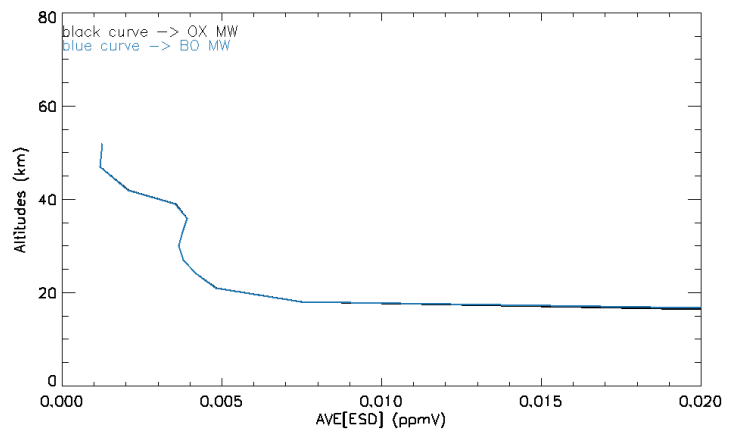Figure 5.25: AVE[Ratio$_{esd}$] values for Nitric Acid as a function of altitude for the two sets of MWs

Figure 5.26: AVE[ESD] values for Nitric Acid as a function of altitude for the two sets of MWs

41

5.2.7 Methane

Figs. 5.27, 5.28, 5.29 and 5.30 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Methane as a function of altitude. In Fig. 5.27 and 5.28, I observe for the OX set of MWs better performance from 16 to 68 km and worse performance from 6 to 16 km. Fig. 5.29 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.85) and OX (1.0) sets of MWs which indicates a better consistency between the real and estimated error for the OX set of MWs. In Fig. 5.30 I observe for the OX set of MWs lower AVE[ESD].
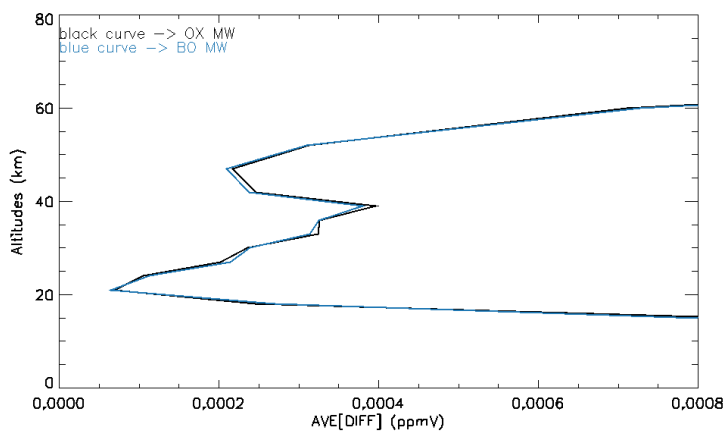

Figure 5.27: AVE[DIFF] values for Methane as a function of altitude for the two sets of MWs
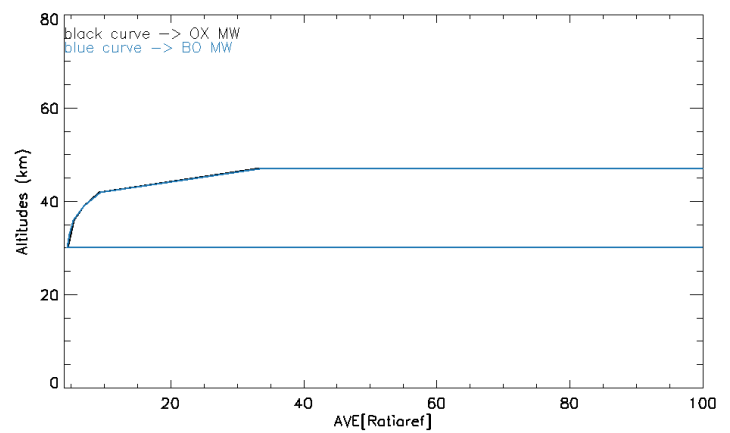

Figure 5.28: AVE[Ratio$_{ref}$] values for Methane as a function of altitude for the two sets of MWs

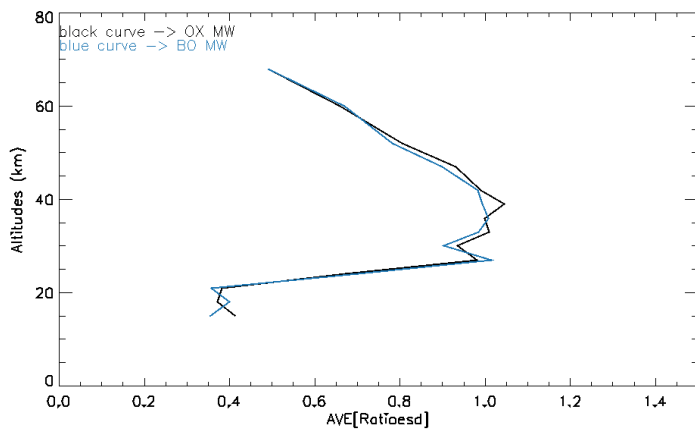Figure 5.29: AVE[Ratio$_{esd}$] values for Methane as a function of altitude for the two sets of MWs

Figure 5.30: AVE[ESD] values for Methane as a function of altitude for the two sets of MWs

5.2.8 Nitrous Oxide

Figs. 5.31, 5.32, 5.33 and 5.34 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Nitrous Oxide as a function of altitude. In Fig. 5.31 and 5.32, I observe for the OX set of MWs slightly better performance from 9 to 15 km,from 21 to 24 km and from 36 to 39 km. Fig. 5.33 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.9) and OX (0.9) sets of MWs which indicates that the consistency between the real and estimated error is almost the same for the two sets of MWs. In Fig. 5.26 I observe that both sets of MWs have almost the same AVE[ESD] values.

Figure 5.31: AVE[DIFF] values for Nitrous Oxide as a function of altitude for the two sets of MWs

Figure 5.32: AVE[Ratio$_{ref}$] values for Nitrous Oxide as a function of altitude for the two sets of MWs

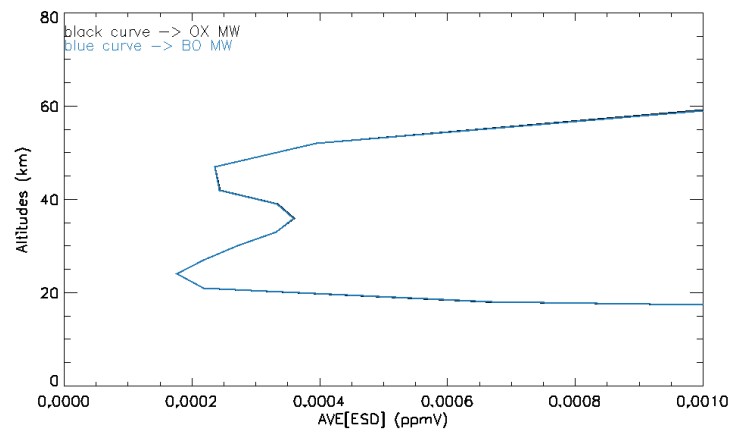Figure 5.33: AVE[Ratio$_{esd}$] values for Nitrous Oxide as a function of altitude for the two sets of MWs

Figure 5.34: AVE[ESD] values for Nitrous Oxide as a function of altitude for the two sets of MWs

5.2.9 Nitrogen Dioxide

Figs. 5.35, 5.36, 5.37 and 5.38 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Nitrogen Dioxide as a function of altitude. In Fig. 5.35 and 5.36, I observe for the OX set of MWs slightly better performance from 27 to 30 km and slightly worse performance from 33 to 48 km. Fig. 5.37 shows the consistency between ESD and AVE[DIFF] for the two sets of MWs, I also note the average values of the curves oscillations for BO (0.8) and OX (0.8) sets of MWs which indicates that the consistency between the real and estimated error is almost the same for the two sets of MWs. In Fig. 5.38 I observe that both sets of MWs have almost the same AVE[ESD] values.

Figure 5.35: AVE[DIFF] values for Nitrogen Dioxide as a function of altitude for the two sets of MWs

Figure 5.36: AVE[Ratio$_{ref}$] values for Nitrogen Dioxide as a function of altitude for the two sets of MWs

Figure 5.37: AVE[Ratio$_{esd}$] values for Nitrogen Dioxide as a function of altitude for the two sets of MWs

Figure 5.38: AVE[ESD] values for Nitrogen Dioxide as a function of altitude for the two sets of MWs

5.2.10 Summary

In table 5.1 all the comments regarding the comparison of the two sets of MWs are summarised as a function of altitude. In table 5.1 I remark that generally the OX set of MWs has a better performance and lower ESD values at low altitudes and on the contrary a worse performance and higher ESD values at higher altitudes. I also remark that for all targets the ESD values and the "real error" have a good consistency which indicates that when I will test these two sets of MWs on real observation (Chap. 7), I could rely on the retrieved ESD values. However, I note that the BO set of MWs has a higher tendency to overestimate the ESD values than the OX set of MWs.

Targets	Retrievals Performance of the two sets of MWs (Altitude in km)	Average values of Ratio$_{esd}$ oscillations		Retrieved ESD values of the two sets of MWs (Altitude in km)
		BO OM	OX OM	
Temperature	From 15 to 36 → B From 39 to 60 → W	0.7	0.85	From 18 to 33 → GL From 33 to 50 → SH
Pressure	From 9 to 33 → B From 40 to 47 → W	0.75	0.9	From 6 to 55 → SL From 55 to 68 → SH
Water	Around 21 → SB From 37 to 45 → B From 24 to 36 → W	0.75	0.8	From 21 to 24 → SL From 40 to 55 → L From 25 to 39 → H From 55 to 68 → H
Ozone	From 15 to 35 → GB From 40 to 55 → B From 9 to 12 → SW From 35 to 40 → SW From 55 to 68 → W	0.8	0.8	From 12 to 35 → GL
Nitric Acid	From 35 to 40 → SW	0.8	0.8	X
Methane	From 16 to 68 → B From 6 to 16 → SW	0.85	1.0	From 6 to 68 → L
Nitrous Oxide	From 9 to 15 → SB From 21 to 24 → SB From 36 to 39 → SB	0.9	0.9	X
Nitrogen Dioxide	X	0.8	0.8	X

Table 5.1 Results of the comparison of the two sets of MWs in function of the altitude

B → Better efficiency for the OX set of MWs
W → Worse efficiency for the OX set of MWs
L → Lower ESD values for the OX set of MWs
H → Higher ESD values for the OX set of MWs
S → Slightly
G → Greatly
X → No relevant divergence between the two sets of MWs.

6 Information Load analysis

The analysis of the Information Load (IL) (see Sect 4.2) will permit to compare the performance of the two sets of MWs and to evaluate the performance of each individual MW.

6.1 Comparison of the Information Load values for the two sets of MWs

I used the "GMTR_INFO" software (Sect. 2.4) to compute the weighted IL values (Eq. 4.3) of the two sets of MWs for each clove of the atmosphere for the four targets. In fig. 6.1 the obtained IL values for Temperature, Pressure, Ozone and Water are plotted as a function of altitudes and OC for the BO, OX and Comb (see section 6.2) set of MWs. For practical reasons I decided to introduce on the same page the maps concerning OX, BO and Combined (Comb) sets of MWs even if the discussion of the Comb set of MWs will take place in section 6.2. Thanks to these maps, I can observe the altitude coverage of the BO and OX sets of MWs. Thus I localise the differences in altitude coverage between the two sets of MWs. But in the case that both sets of MWs are covering the same altitude, it is difficult to argue which of them has the best efficiency.

I plotted the IL values as a function of OC at several samples altitudes (10, 20, 30, 40 and 50 km). Because of the high number of these plots I only show three of them as examples , Figs. 6.2, 6.3 and 6.4, which represent the IL values for Temperature as a function of OC at 20 km for BO, OX and Comb sets of MWs respectively. Based on these figures I can compare the IL intensity values between the two set of MWs. The set of MWs which has the higher intensity is the most efficient of the two set of MWs. On these figures, I can also compare the amplitude of the peak to peak oscillation between the two sets of MWs. Thanks to this I can compare the uniformity of the IL values over the full orbit between the two sets of MWs. The uniformity represent the homogeneity of the IL values over the full orbit, i.e. the higher the uniformity, the lower the position of the retrieved profiles over the full orbit is depending on the intensity of the IL values at these position. In absence of uniformity the retrieved profiles should be positioned where the intensity of the IL values are the highest.

In the next paragraphs I will show the maps and the plots described above to compare the efficiency of the two sets of MWs for Temperature, Pressure, Ozone and Water.

Temperature

In Fig. 6.1, I observe for the OX set of MWs a better coverage at low altitudes (below 10 km) and for the BO set of MWs a better coverage at higher altitudes (above 30 km). Thanks to the plots of the IL values for Temperature described above, I can refine my previous observations.

I observe, in the OX set of MWs generally higher IL values below 30 km (especially a large improvement around the equator) and lower IL values at higher altitudes. I also observe that the intensity of the peak to peak for the OX set of MWs is higher at low altitudes (lower than 30 km) and lower at higher altitudes.

I conclude that the OX set of MWs has a better efficiency at lower altitudes (below 30 km), and a worse efficiency at higher altitudes. I also note, in the OX set of MWs, a worse

uniformity at lower altitudes (below 30km) and a better uniformity at higher altitudes.

I will abridge the comparison of the other targets by describing directly the efficiency and the uniformity of the BO and OX sets of MWs. To proceed through this task, I decided not to use anymore the maps, since the comparison of the two sets of MWs is facilitated by the plots described above.

Pressure

The OX set of MWs is generally more efficient over all altitudes except around 30 km where it is a bit less. I note a large improvement around the equator for the OX set of MWs below 10 km. The uniformity is generally better for all altitudes for the BO set of MWs.

Ozone

The OX set of MWs is more efficient around 20, 30 and 50 km and less efficient around 10 and 40 km. The uniformity in most cases is correlated to the intensity of the IL values, i.e. the higher the intensity of the IL values, the lower the uniformity over the full orbit is.

Water

Water is mainly concentrated at low altitudes and around 50 km, so I will mainly describe the efficiency of the two sets of MWs at these altitudes. The OX set of MWs is more efficient at low altitudes (around 8 km) but less efficient in higher altitudes (around 40 km). The uniformity is generally lower for the OX set of MWs.

Summary

OX and BO sets of MWs show advantages and drawbacks, generally the OX set of MWs has better performance in low altitudes and the BO set of MWs has better performance in higher altitudes.
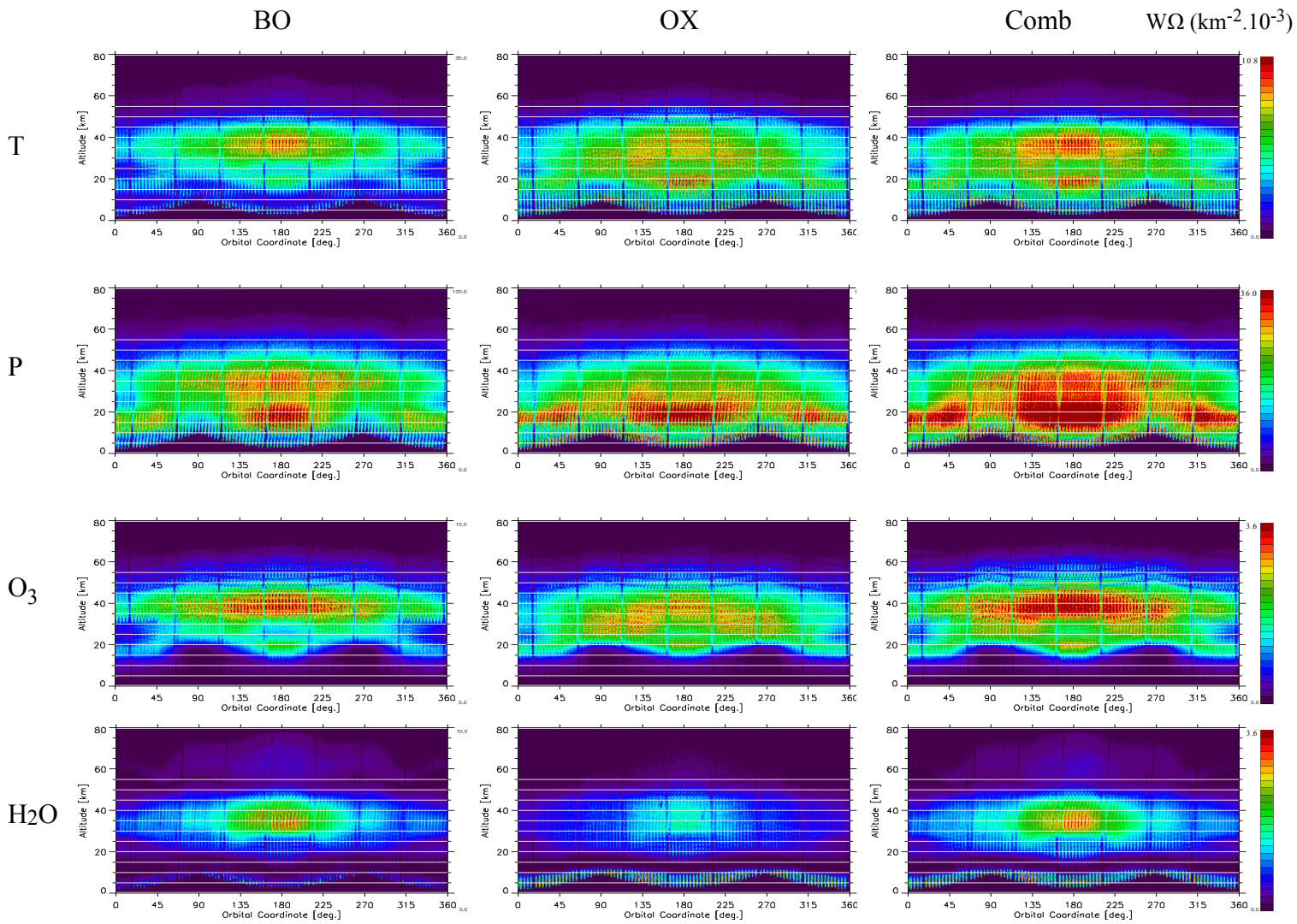
Figure 6.1: IL values for Temperature, Pressure, Ozone and Water as a function of altitude and OC for BO, OX and Comb set of MWs

The white horizontal lines are set every 5 km from 5 to 55 km, in order to help to localise, at a given altitude, an improvement or a worsening of the IL values
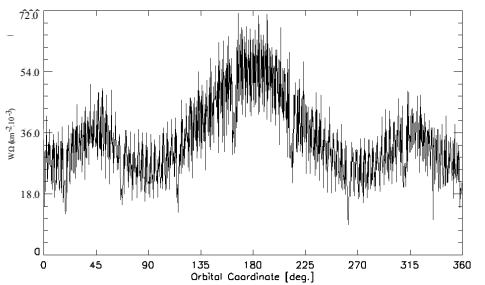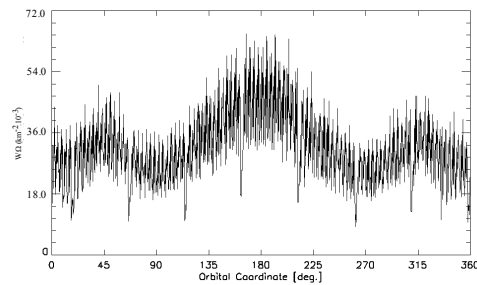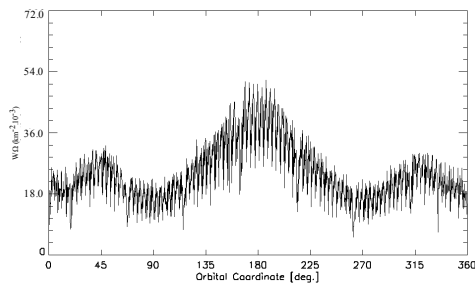
Figure 6.2: IL values for Temperature as a function of OC for the BO set of MWs at 20 km

Figure 6.3: IL values for Temperature as a function of OC for the OX set of MWs at 20 km

Figure 6.4: IL values for Temperature as a function of OC for the Comb set of MWs at 20 km

6.2 Selections of Micro-Windows for the "Combined" set

To improve the retrievals performance, I decided to merge these two sets of MWs. I will calculate the IL values and then describe the contributions of each individual MW. Thereafter, I will select the most efficient MW, to create a "Comb" set of MWs. The BO set and the OX set of MWs are composed of 8 (PTHO06..) and 10 (PTHO03..) MWs respectively. Table 6.1 gives their main characteristics.

Name	Frequency Interval (cm^{-1})	Altitudes Range of the MW (km)	Species present in the MW
PTHO0601	1036.313 / 1039.313	68 / 33	O_3
PTHO0602	1682.688 / 1685.688	60 / 27	H_2O
PTHO0603	790.625 / 793.625	47 / 27	CO_2
PTHO0604	1071.875 / 1074.875	39 / 6	CO_2 / O_3
PTHO0605	689.750 / 692.750	52 / 30	CO_2 / O_3
PTHO0607	686.688 / 689.688	68 / 60	CO_2 / H_2O / O_3
PTHO0609	731.188 / 734.188	36 / 6	CO_2 / O_3
PTHO0610	1651 / 1654	36 / 6	H_2O
PTHO0301	1074.188 / 1077.188	60 / 6	CO_2 / O_3
PTHO0302	1575.125 / 1578.125	60 / 15	H_2O
PTHO0303	719.375 / 722.375	68 / 21	CO_2 / O_3
PTHO0304	744.875 / 747.875	47 / 9	CO_2 / O_3
PTHO0305	825.125 / 828.125	18 / 6	CO_2 / H_2O / O_3
PTHO0306	1032.938 / 1035.938	68 / 33	O_3
PTHO0307	735 / 738	39 / 15	CO_2 / O_3
PTHO0308	1151.063 / 1154.063	18 / 6	H_2O / O_3
PTHO0309	740.875 / 743.875	68 / 15	CO_2 / O_3
PTHO0310	730 / 733	42 / 15	CO_2 / O_3

Table 6.1: MWs description (Frequency Interval, Altitudes Range and present Species)

I used the "GMTR_INFO" to compute the IL values of each individual MW. Then I mapped the obtained IL values as a function of altitude and OC. Since I had 18 MWs and four targets, I obtained 72 maps. I decided to show only as examples Figs. 6.5, 6.6 and 6.7 which represent the IL values for Temperature as a function of altitude and OC for MWs PTHO0305, PTHO0306 and PTHO0302 respectively. In these examples I observed intermediate IL values around 7.5 km on fig. 6.5, large IL values from 15 to 25 km around the South Pole (SP) on fig. 6.6 and very small IL values from 30 to 45 km on fig. 6.7. I noted that each of these MWs has a very specific contribution. In order to facilitate the MWs selection I reported the IL values intensity and the altitude of each individual MW in tables 6.2 and 6.3.
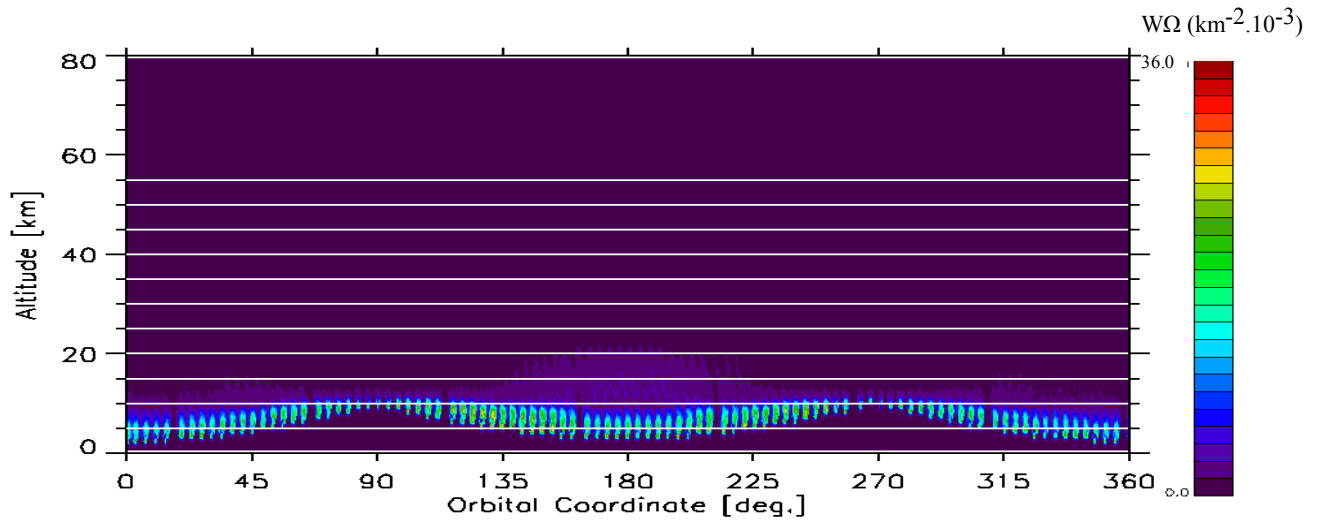
Figure 6.6: IL values for Temperature plotted as a function of altitude and OC for PTHO0305
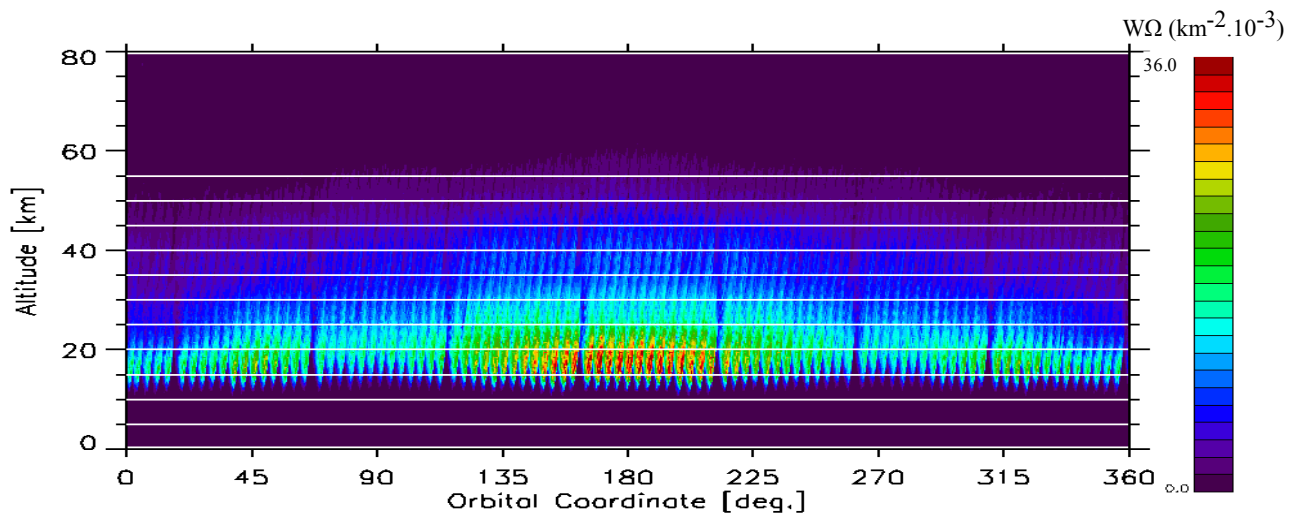
Figure 6.7: IL values for Temperature plotted as a function of altitude and OC for PTHO0302
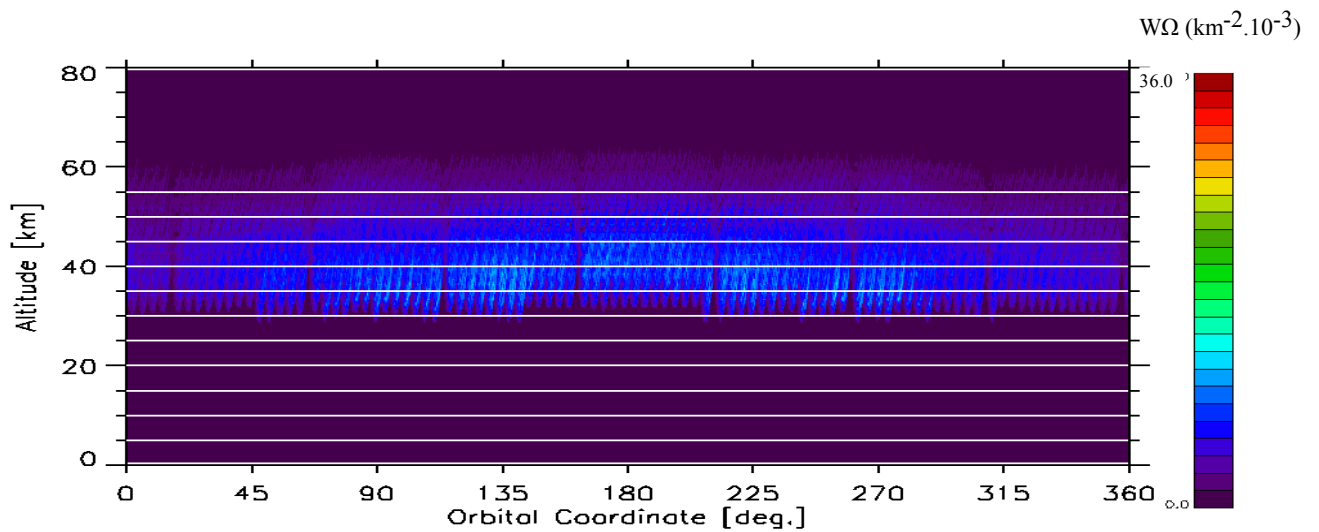
Figure 6.8: IL values for Temperature plotted as a function of altitude and OC for PTHO0306

MWs\Species	Temperature	Pressure
PTHO0301	Around 7.5 km ➜ S From 10 to 50 km ➜ VS	From 5 to 50 km ➜ VS
PTHO0302	From 15 to 30 km ➜ L From 30 to 45 km ➜ S From 45 to 55 km ➜ VS	From 30 to 45 km (SP) ➜ S From 15 to 55 km ➜ VS
PTHO0303	From 20 to 55 km ➜ VS	From 35 to 50 km (SP) ➜ S From 20 to 55 km ➜ VS
PTHO0304	From 5 to 25 km ➜ I From 25 to 35 km ➜ S From 35 to 50 km ➜ VS	From 5 to 20 km (SP) ➜ I From 20 to 50 km ➜ VS
PTHO0305	Around 7.5 km → I	Around 7.5 km ➜ I
PTHO0306	From 30 to 55 km ➜ VS	From 30 to 55 km ➜ VS
PTHO0307	From 10 to 30 km ➜ I From 30 to 45 km ➜ S From 45 to 60 km ➜ VS	From 15 to 40 km (SP) ➜ I
PTHO0308	Around 7.5 km ➜ I From 10 to 40 km ➜ VS	Around 7.5 km ➜ S From 10 to 20 km ➜ VS
PTHO0309	From 10 to 55 km ➜ VS	From 15 to 55 km ➜ VS
PTHO0310	From 10 to 35 km ➜ S From 35 to 50 km ➜ VS	From 10 to 45 km ➜ VS
PTHO0607	X	Around 60 km ➜ VS
PTHO0605	From 25 to 40 km ➜ S From 40 to 55 km ➜ VS	From 25 to 45 km ➜ S From 45 to 55 km ➜ VS
PTHO0609	From 5 to 20 km ➜ I From 20 to 35 km ➜ S From 35 to 50 km ➜ VS	From 10 to 25 km ➜ S From 25 to 50 km ➜ VS
PTHO0603	From 25 to 55 km ➜ VS	From 25 to 45 km ➜ VS
PTHO0601	From 30 to 50 km ➜ I From 50 to 60 km ➜ VS	From 30 to 45 km (SP) ➜ I From 30 to 50 km ➜ VS
PTHO0604	Around 7.5 km → I From 10 to 40 km → S	Around 7.5 km → I From 10 to 45 km → VS
PTHO0610	From 15 to 25 km (SP) → VL From 10 to 40 km → I From 40 to 55 km → S	From 30 to 45 km (SP) → I From 15 to 50 km → VS
PTHO0602	From 25 to 55 km → S	From 25 to 50 km → VS

Table 6.2: Description of the MWs contribution for Temperature and Pressure

MWs\Species	Ozone	Water
PTHO0301	From 15 to 50 km → VS From 25 to 40 km (SP) → S	Around 7.5 km → S
PTHO0302	X	From 15 to 55 km → VS From 20 to 45 km (SP) → S
PTHO0303	From 20 to 55 km → VS	X
PTHO0304	From 15 to 45 km → VS	X
PTHO0305	X	Around 7.5 km → L
PTHO0306	From 30 to 50 km → I From 50 to 55 km → S	X
PTHO0307	From 20 to 35 km → I From 15 to 20 km → VS From 35 to 45 km → VS	X
PTHO0308	From 15 to 45 km → VS	Around 7.5 km → S
PTHO0309	From 15 to 45 km → VS	X
PTHO0310	From 15 to 45 km → VS	X
PTHO0607	X	X
PTHO0605	X	X
PTHO0609	From 15 to 45 km → VS	X
PTHO0603	From 20 to 45 km → VS	X
PTHO0601	From 30 to 50 km → VL From 50 to 60 km → VS	X
PTHO0604	From 15 to 25 km (SP) → I From 15 to 40 km → S	Around 7.5 km → S
PTHO0610	X	From 30 to 45 km (SP) → VL From 20 to 50 km → S
PTHO0602	X	From 25 to 55 km → VS

Table 6.3: Description of the MWs contribution for Ozone and Water

VL → Very Large IL values
L → Large IL values
I → Intermediate IL values
S → Small IL values
VS → Very Small IL values
SP → indicates that the IL values are localised around the South Pole
X → IL values are nil or insignificant

I marked in blue the MWs which have a good efficiency, i.e. the MWs which have Large or Very Large IL values or intermediate IL values but for at least two targets. I marked in red the MWs which have the worse efficiency, i.e. the MWs which have only Small or Very Small IL values. Then finally I marked in green the MWs which have a intermediate efficiency, i.e. the MWs which have intermediate IL values but only for one target.

Selection and test of an optimised Occupation Matrix (OM):

Looking at tables 6.2 and 6.3, I selected the MWs with the best efficiency, which are PTHO0302, PTHO0304, PTHO0305, PTHO0307, PTHO0308, PTHO0601, PTHO0604 and PTHO0610. Then for the four targets, I checked that using this set of MWs the IL will not suffer from a lack of coverage at any altitude. Looking at the IL values resulting from this Comb set of MWs I decided to add also PTHO0306 and PTHO0609 (the intermediate efficiency MWs) which are efficient for Ozone and Temperature respectively. Thus I composed the optimised OM of 10 MWs.

Fig. 6.1 show us the plotted maps of the IL values for the four targets as a function of altitude and OC for the BO, OX and Comb sets of MWs. Looking at these maps I appreciate the better efficiency of the Comb sets of MWs by comparing the BO and OX sets of MWs. By looking at the IL values plotted as a function of OC for several sample altitudes (10, 20, 30, 40 and 50 km), as for Figs. 6.2, 6.3 and 6.4, I checked my previous observations. Indeed in most cases the intensity and the uniformity of the IL values are higher for the Comb set of MWs, except in the case of Pressure around 50 km, where IL values are slightly lower.

Since the retrieval of Pressure is of minor importance w.r.t. the retrieval of Temperature, I can state that the Comb set of MWs show much larger benefits than drawbacks. In the next chapter I test the Comb set of MWs on simulated and real retrievals (Chap. 7).

7 Performance of the optimised Occupation Matrix

7.1 Comparison of the sets of MWs using retrievals on simulated observations

As for Chap. 5, I decided not to use the maps but instead the plots of AVE[DIFF], AVE[Ratio$_{ref}$], AVE[Ratio$_{esd}$] and AVE[ESD] as a function of altitudes. I added in green the quantifier values of the Comb set of MWs to the previous plots of BO and OX sets of MWs (from Sub-Sect. 5.2.2 to 5.2.9).

7.1.1 Temperature

Figs. 7.1, 7.2, 7.3 and 7.4 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Temperature as a function of altitude. In Figs. 7.1 and 7.2, I observe for the Comb set of MWs better performance from 6 to 15 km, from 18 to 30 km and from 33 to 50 km and worse performance above 50 km. Fig. 7.3 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.85) sets of MWs. In Fig. 7.4 I observe for the Comb set of MWs lower AVE[ESD] values from 15 to 55 km.
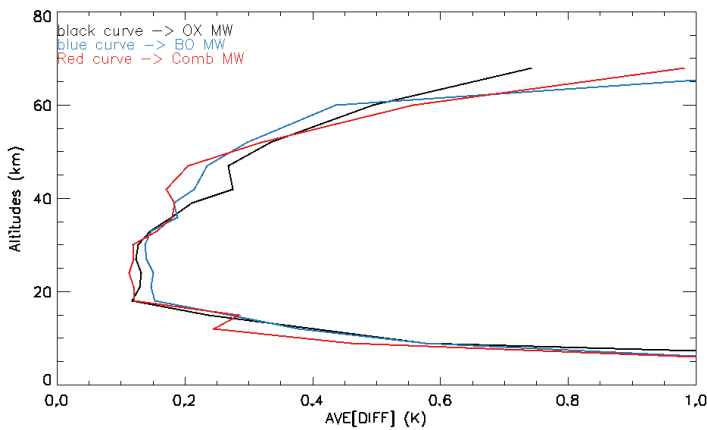
Figure 7.1: AVE[DIFF] values for Temperature as a function of altitude for the three sets of MWs
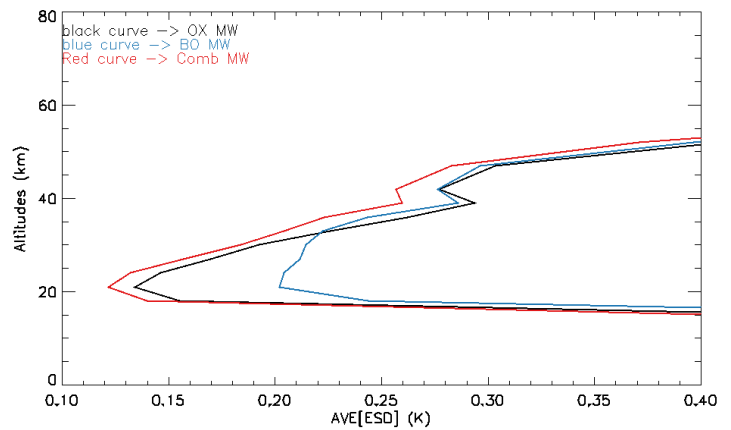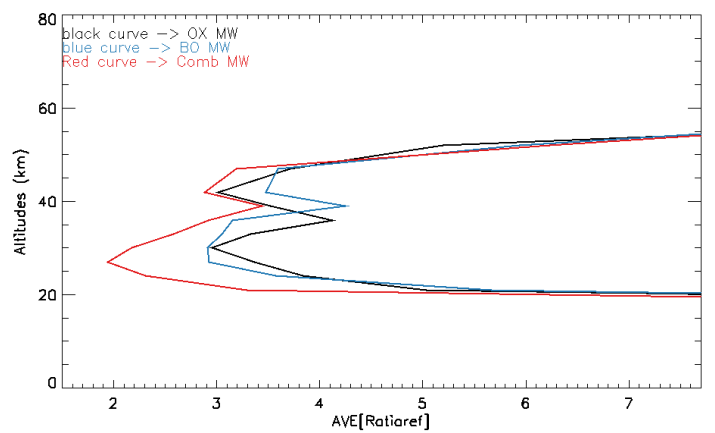

Figure 7.4: AVE[ESD] values for Temperature as a function of altitude for the three sets of MWs
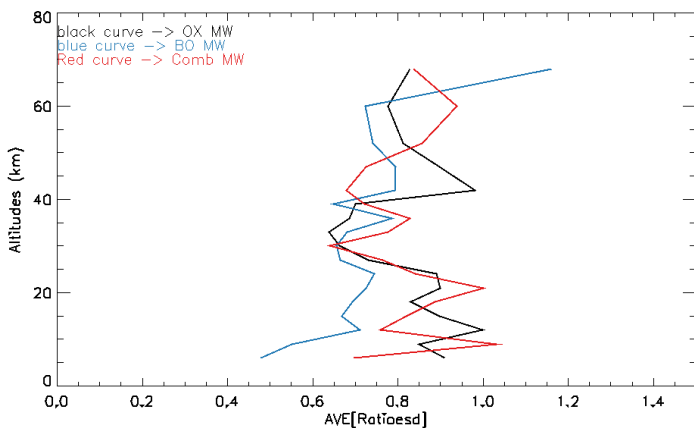
Figure 7.3: AVE[Ratio$_{esd}$] values for Temperature as a function of altitude for the three sets of MWs
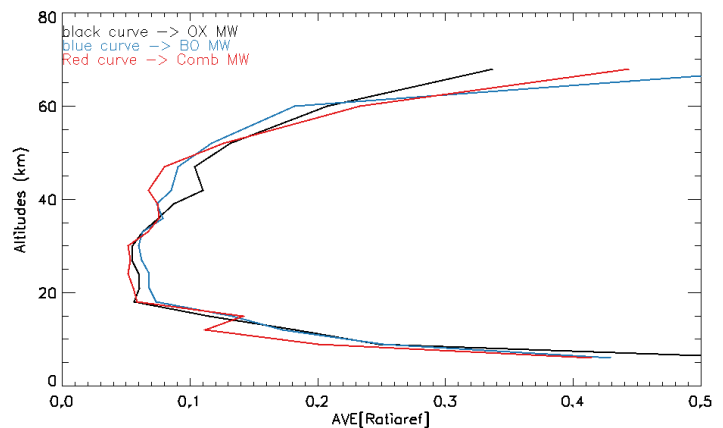
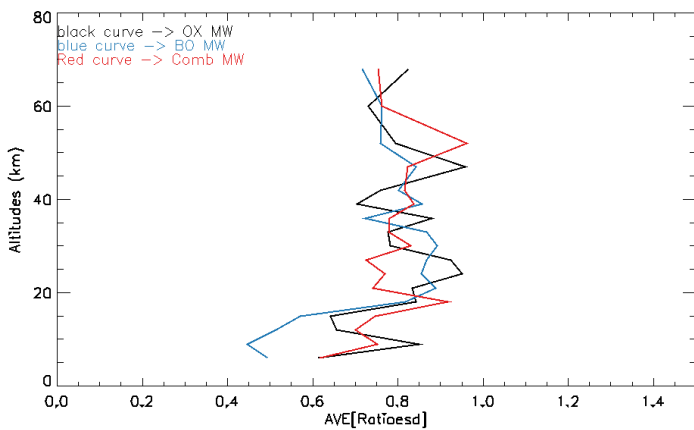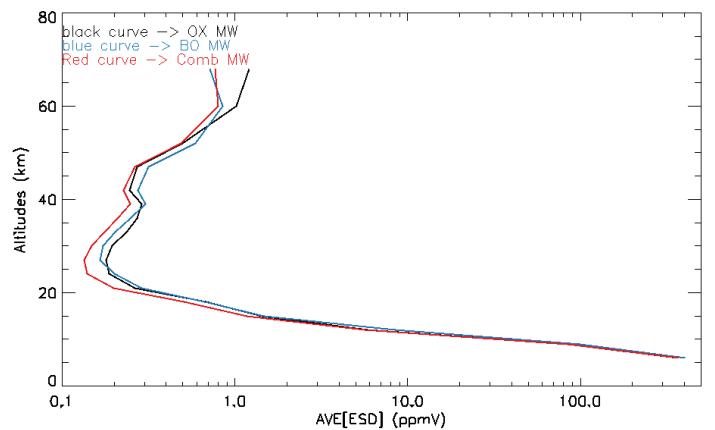Figure 7.2: AVE[Ratio$_{ref}$] values for Temperature as a function of altitude for the three sets of MWs

7.1.2 Pressure

Figs. 7.5, 7.6, 7.7 and 7.8 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Pressure as a function of altitude. In Figs. 7.5 and 7.6, I observe for the Comb set of MWs significantly worse performance from 30 to 68 km. Fig. 7.7 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.9) sets of MWs. In Fig. 7.8 I observe for the Comb set of MWs higher AVE[ESD] values above 45 km.
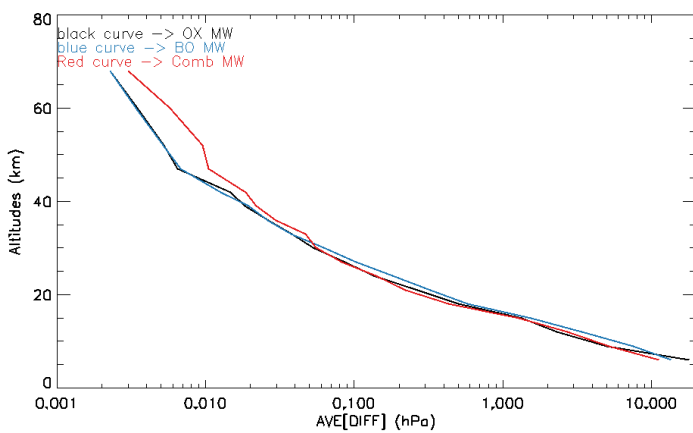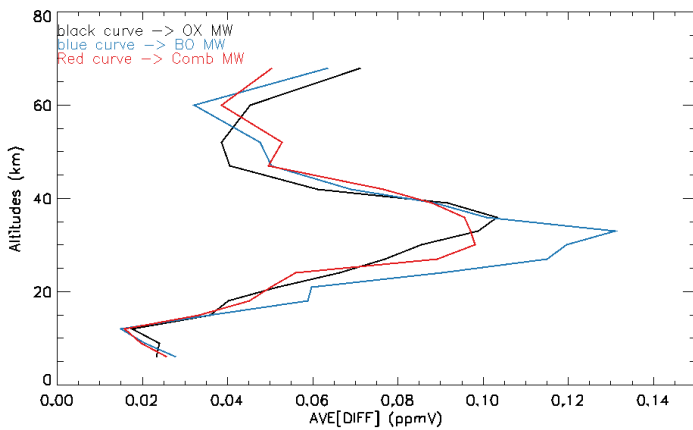
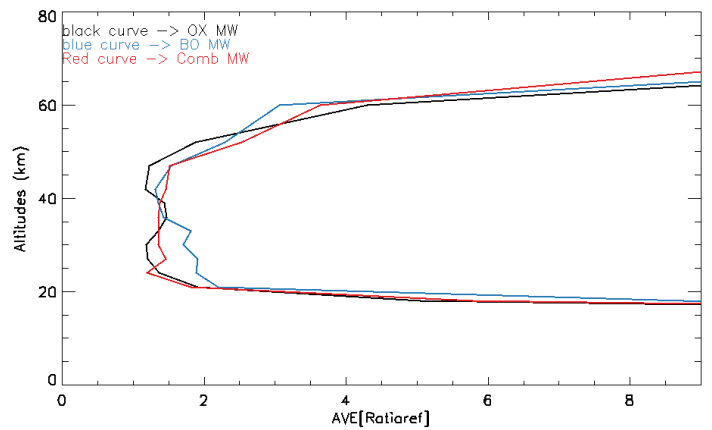Figure 7.5: AVE[DIFF] values for Pressure as a function of altitude for the three sets of MWs
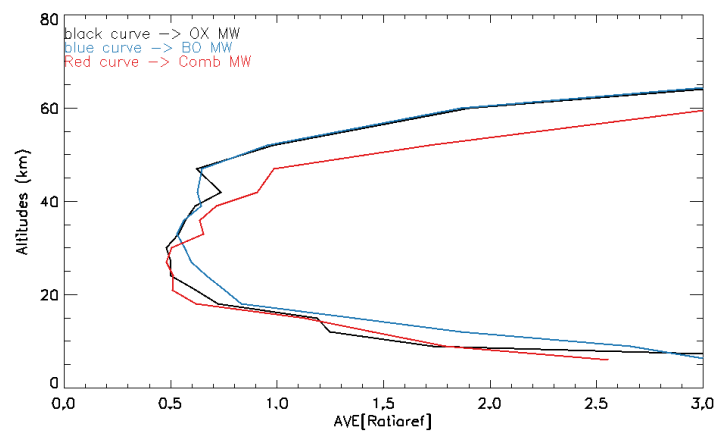
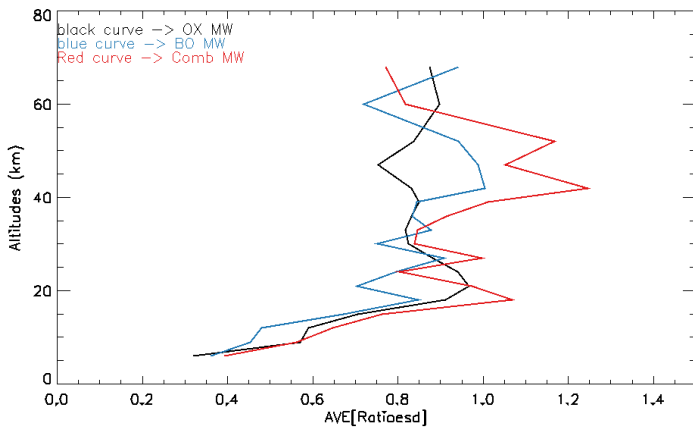Figure 7.6: AVE[Ratio$_{ref}$] values for Pressure as a function of altitude for the three sets of MWs
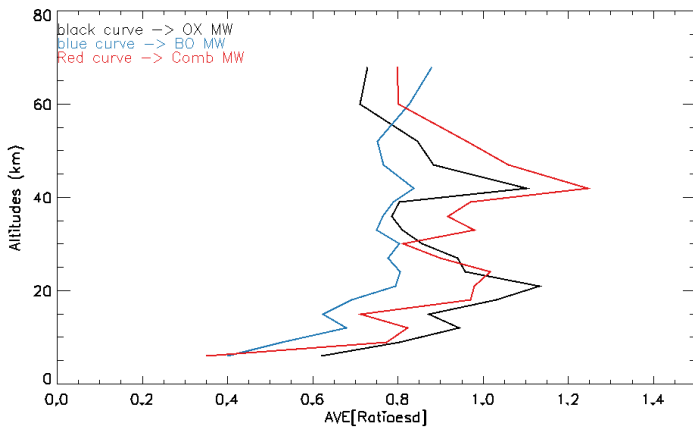
Figure 7.7: AVE[Ratio$_{esd}$] values for Pressure as a function of altitude for the three sets of MWs
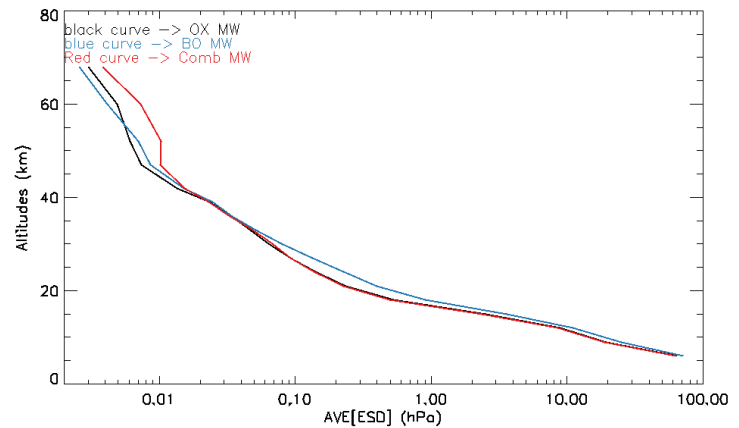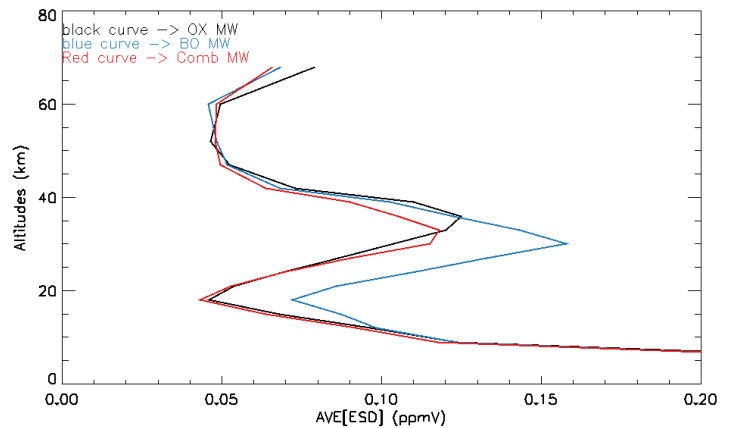
Figure 7.8: AVE[ESD])values for Pressure as a function of altitude for the three sets of MWs

55

7.1.3 Water

Figs. 7.9, 7.10, 7.11 and 7.12 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Ozone as a function of altitude. In Figs. 7.9 and 7.10, I observe for the Comb set of MWs better performance from 15 to 50 km and worse performance from 50 to 55 km. Fig. 7.11 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.8) sets of MWs. In Fig. 7.12 I observe for the Comb set of MWs lower AVE[ESD] values from 15 to 60 km.
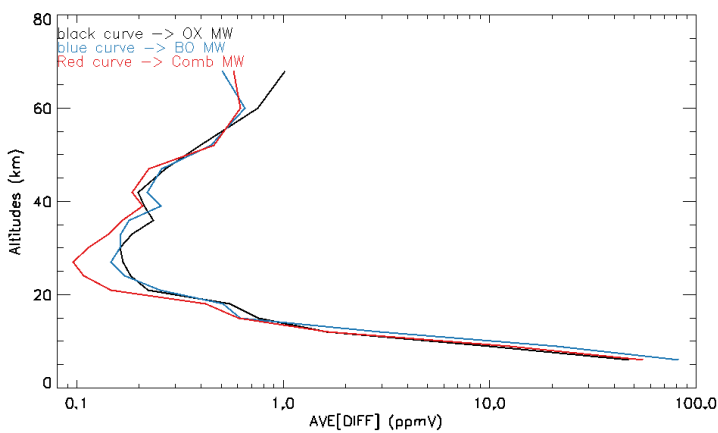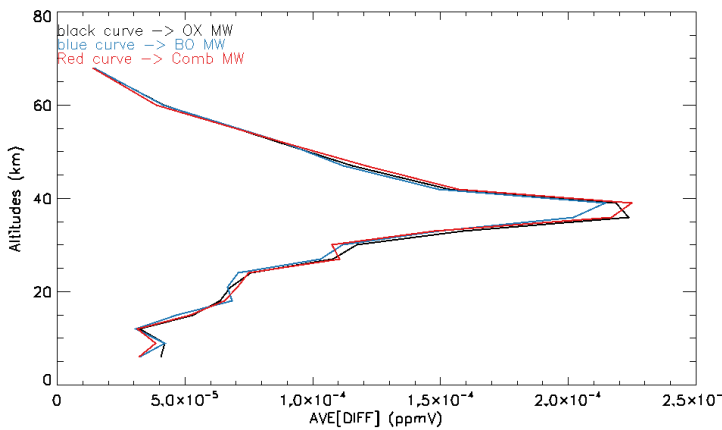
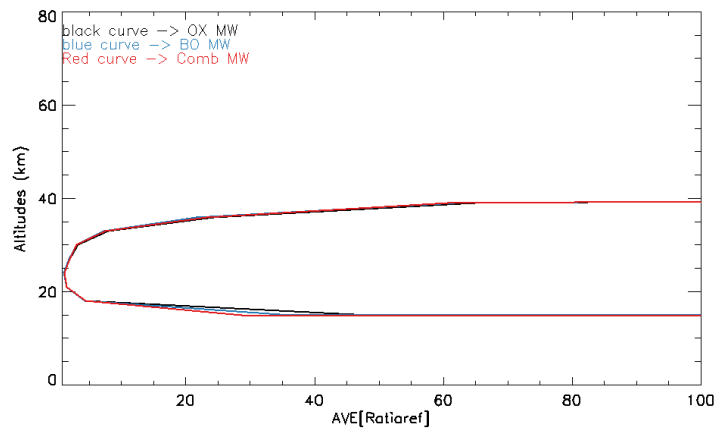Figure 7.9: AVE[DIFF] values for Water as a function of altitude for the three sets of MWs

Figure 7.10: AVE[Ratio$_{ref}$] values for Water as a function of altitude for the three sets of MWs

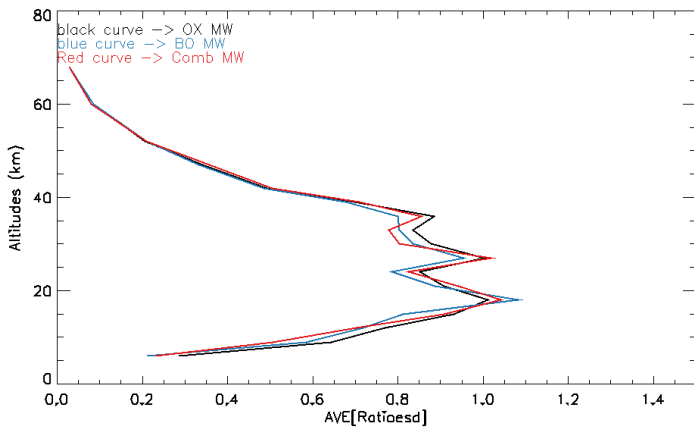Figure 7.11: AVE[Ratio$_{esd}$] values for Water as a function of altitude for the three sets of MWs

Figure 7.12: AVE[esd] values for Water as a function of altitude for the three sets of MWs

7.1.4 Ozone

Figs. 7.13, 7.14, 7.15 and 7.16 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Ozone as a function of altitude. In Figs. 7.13 and 7.14, I observe for the Comb set of MWs better performance from 21 to 24 km, from 36 to 39 km and above 60 km and worse performance from 25 to 33 km and from 39 to 60 km. Fig. 7.15 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.9) sets of MWs. In Fig. 7.16 I observe for the Comb set of MWs lower AVE[ESD] values from 6 to 21 km and from 33 to 50 km and higher AVE[ESD] values from 24 to 33 km and above 50 km.

Figure 7.13: AVE[DIFF] values for Ozone as a function of altitude for the three sets of MWs

Figure 7.14: AVE[Ratio$_{ref}$] values for Ozone as a function of altitude for the three sets of MWs

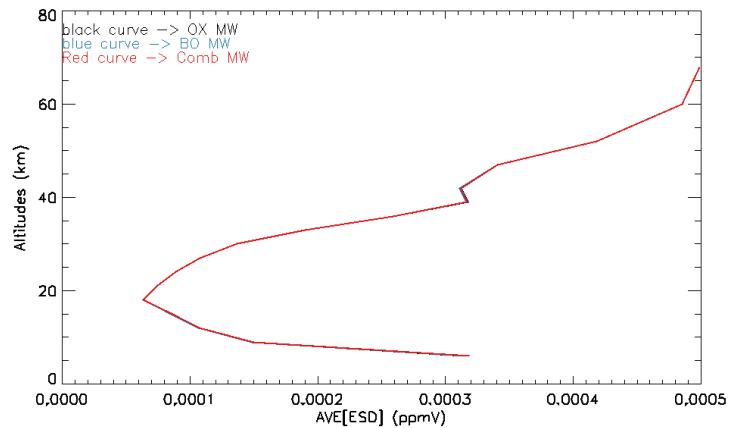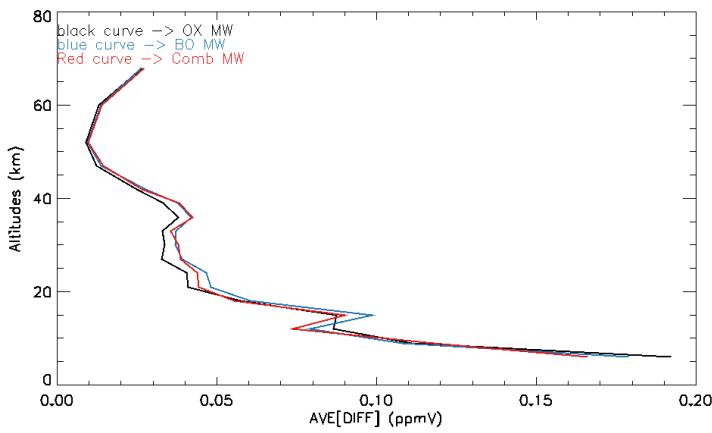Figure 7.15: AVE[Ratio$_{esd}$] values for Ozone as a function of altitude for the three sets of MWs

Figure 7.16: AVE[ESD] values for Ozone as a function of altitude for the three sets of MWs

7.1.5 Nitric Acid

Figs. 7.17, 7.18, 7.19 and 7.20 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Nitric Acid as a function of altitude. In Fig. 7.17 and 7.18, I observe for the Comb set of MWs slightly worse performance from 35 to 40 km. Fig. 7.19 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.8) sets of MWs. In Fig. 7.20 I observe that the three sets of MWs have almost the same AVE[ESD] values.

Figure 7.17: AVE[DIFF] values for Nitric Acid as a function of altitude for the three sets of MWs
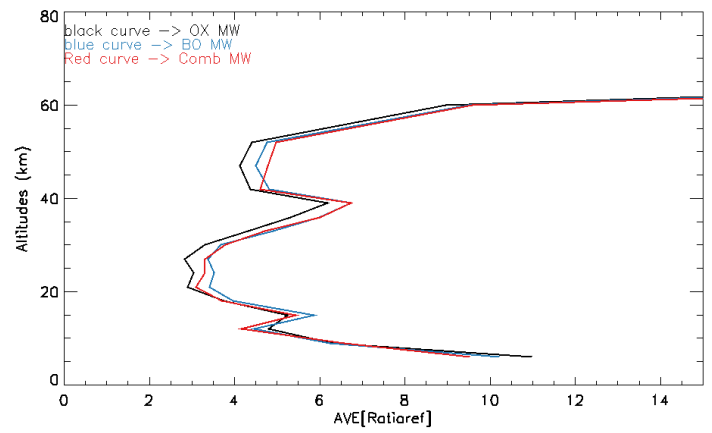

Figure 7.18: AVE[Ratio$_{ref}$] values for Nitric Acid as a function of altitude for the three sets of MWs
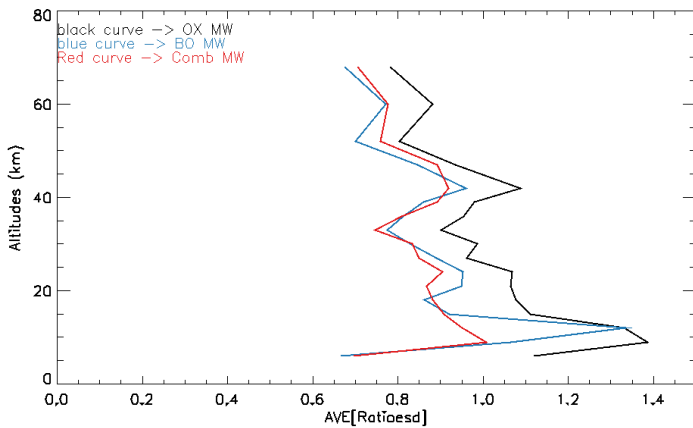

Figure 7.19: AVE[Ratio$_{esd}$] values for Nitric Acid as a function of altitude for the three sets of MWs
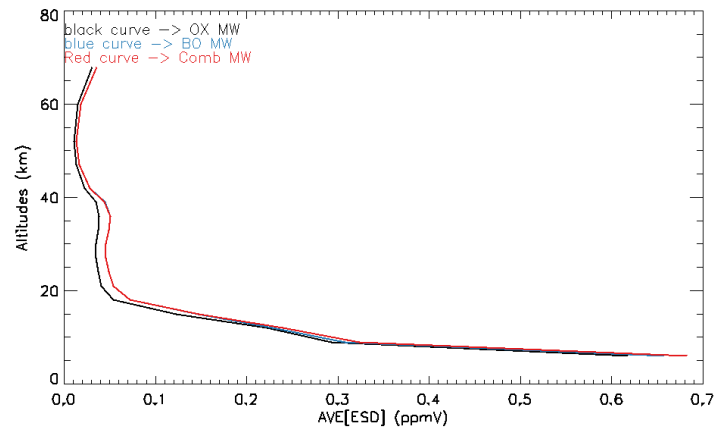

Figure 7.20: AVE[ESD] values for Nitric Acid as a function of altitude for the three sets of MWs

7.1.6 Methane

Figs. 7.21, 7.22, 7.23 and 7.24 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Methane as a function of altitude. In Figs. 7.21 and 7.22, I observe for the Comb set of MWs better performance around 12 km and worse performance from 18 to 68 km. Fig. 7.23 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.8) sets of MWs. In Fig. 7.24 I observe for the Comb set of MWs higher AVE[ESD] values from 6 to 68 km.
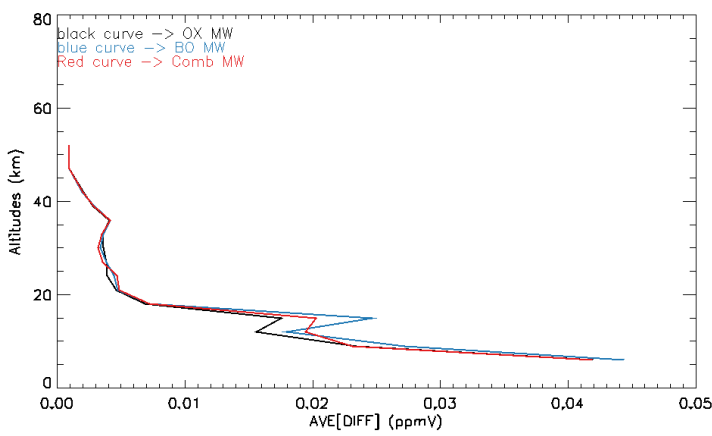

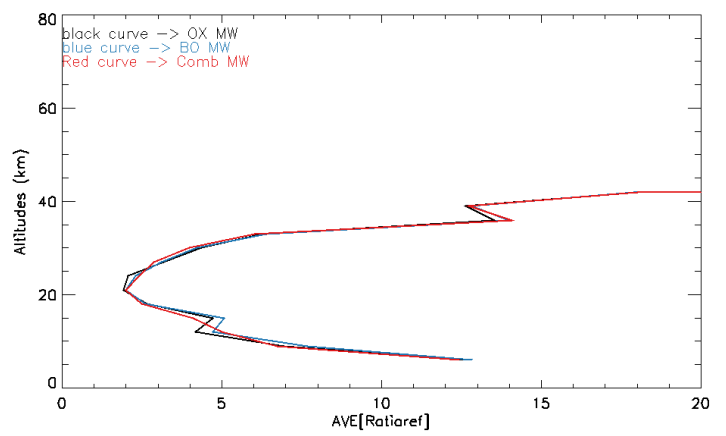Figure 7.21: AVE[DIFF] values for Methane as a function of altitude for the three sets of MWs

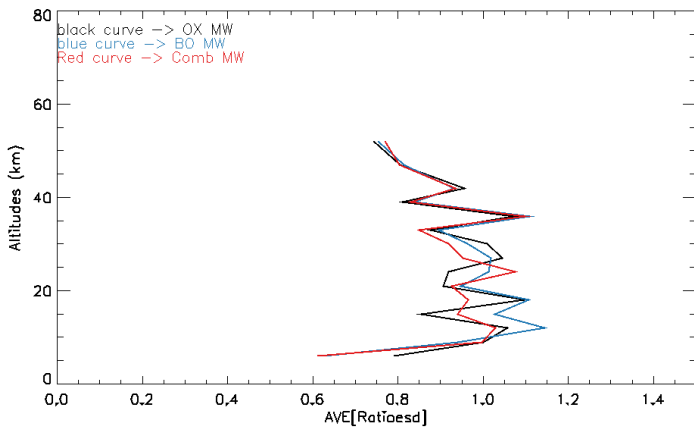Figure 7.22:AVE[Ratio$_{ref}$] values for Methane as a function of altitude for the three sets of MWs

Figure 7.23: AVE[Ratio$_{esd}$] values for Methane as a function of altitude for the three sets of MWs

Figure 7.24: AVE[ESD] values for Methane as a function of altitude for the three sets of MWs

7.1.7 Nitrous Oxide

Figs. 7.25, 7.26, 7.27 and 7.28 report respectively AVE[DIFF], AVE[Ratio$_{ref}$] AVE[Ratio$_{esd}$] and AVE[ESD] values for Nitrous Oxide as a function of altitude. In Figs. 7.25 and 7.26, I observe for the Comb set of MWs slightly better performance from 27 to 35 km and worse performance from 9 to 15 km, around 24 km and from 35 to 40 km. Fig. 7.27 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.9) sets of MWs. In Fig. 7.28 I observe that the three sets of MWs have almost the same AVE[ESD] values.

Figure 7.25: AVE[DIFF] values for Nitrous Oxide as a function of altitude for the three sets of MWs

Figure 7.26: AVE[Ratio$_{ref}$] values for Nitrous Oxide as a function of altitude for the three sets of MWs

Figure 7.27: AVE[Ratio$_{esd}$] values for Nitrous Oxide as a function of altitude for the three sets of MWs
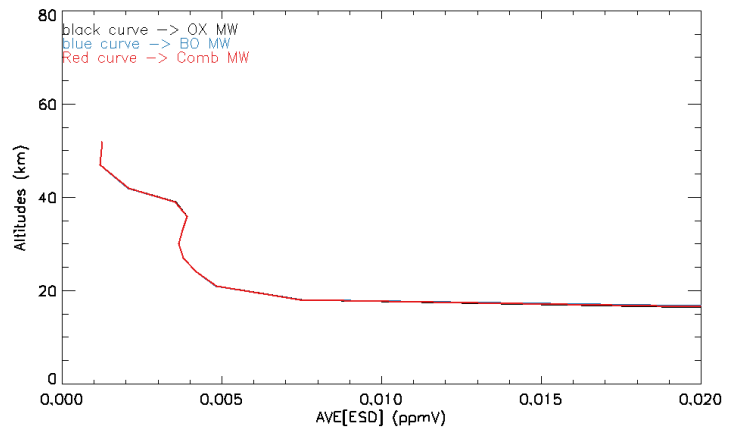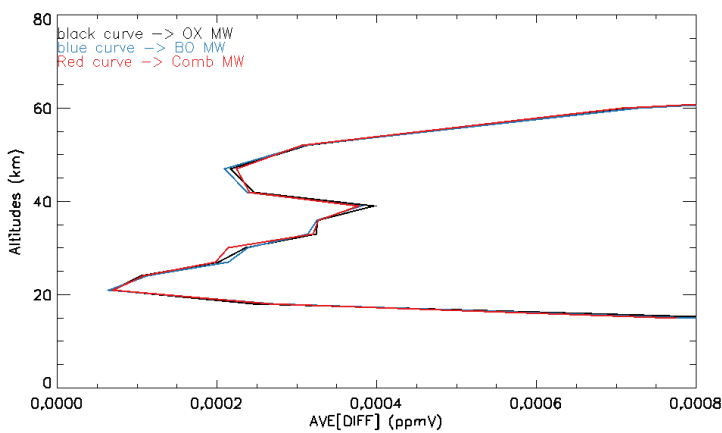

Figure 7.28: AVE[ESD] values for Nitrous Oxide as a function of altitude for the three sets of MWs

7.1.8 Nitrogen Dioxide

Figs. 7.29, 7.30, 7.31 and 7.32 report respectively AVE[DIFF], AVE[Ratio_ref] AVE[Ratio_esd] and AVE[ESD] values for Nitrogen Dioxide as a function of altitude. In Figs. 7.29 and 7.30, I observe for the Comb set of MWs slightly better performance from 27 to 30 km. Fig. 7.31 shows the consistency between ESD and AVE[DIFF] for the three sets of MWs, I also note the average values of the curves oscillations for the Comb (0.8) sets of MWs. In Fig. 7.32 I observe that the three sets of MWs have almost the same AVE[ESD] values.

Figure 7.29: AVE[DIFF] values for Nitrogen Dioxide as a function of altitude for the three sets of MWs
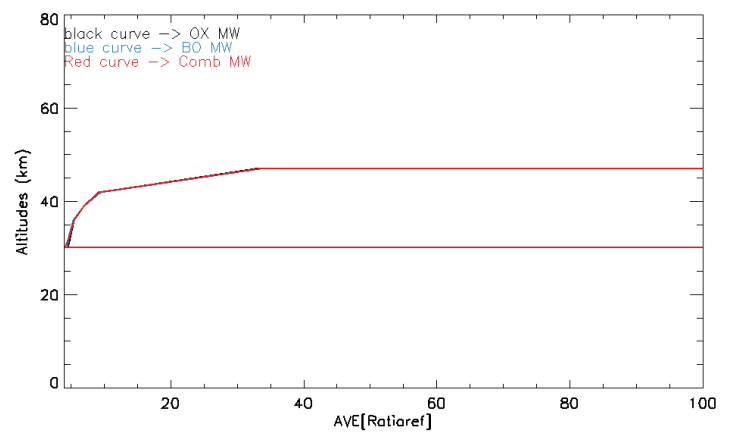

Figure 7.30: AVE[Ratio_ref] values for Nitrogen Dioxide as a function of altitude for the three sets of MWs
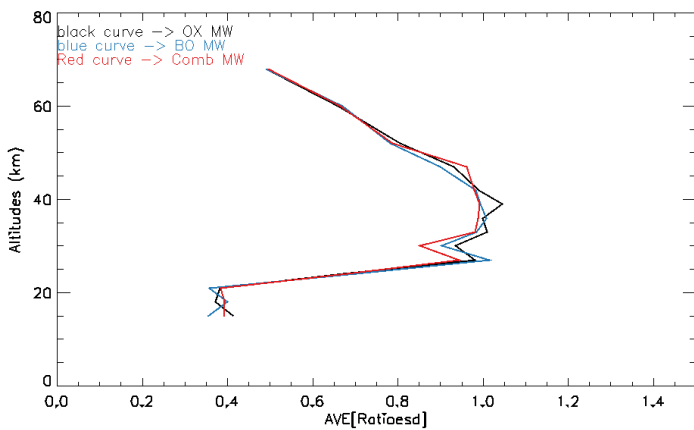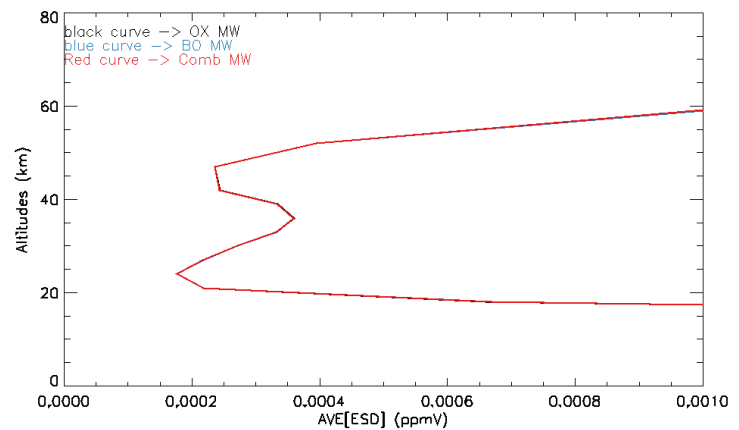

Figure 7.31: AVE[Ratio_esd] values for Nitrogen Dioxide as a function of altitude for the three sets sets of MWs

Figure 7.32: AVE[ESD] values for Nitrogen Dioxide as a function of altitude for the three sets of MWs

61

7.1.9 Summary

In table 7.1 all the comments regarding the comparison of the three sets of MWs are summarised as a function of altitude. In table 7.1 I observe that the efficiency and the ESD values of the Comb set of MWs are better for Temperature and Water, but worse at higher altitudes (Above 30 km) for Pressure. For the Ozone I note that the curves of OX and Comb sets of MWs oscillate around each other which indicates that the efficiency between the two sets is more or less equal, except from 45 to 55 km where the efficiency is definitely worse for the Comb set of MWs.

I also note for the three sets of MWs that the ESD values are overestimated in comparison with the "real error" for all targets.

Targets	Retrievals Performance of the three sets of MWs (Altitude in km)	Average values of Ratio$_{esd}$ oscillations			Retrieved ESD values of the three sets of MWs (Altitude in km)
		BO OM	OX OM	Comb OM	
Temperature	From 6 to 15 → B From 18 to 30 → SB From 33 to 50 → B From 50 to 68 → W	0.7	0.85	0.85	From 15 to 55 → L
Pressure	From 30 to 68 → GW	0.75	0.9	0.9	From 45 to 68 → GH
Water	From 15 to 50 → GB From 50 to 55 → W	0.75	0.8	0.8	From 15 to 60 → L
Ozone	From 21 to 24 → SB From 36 to 39 → SB From 60 to 68 → SB From 25 to 33 → W From 39 to 60 → W	0.8	0.8	0.9	From 6 to 21 → SL From 33 to 50 → L From 24 to 33 → SH From 50 to 68 → SH
Nitric Acid	From 35 to 40 → SW	0.8	0.8	0.8	X
Methane	Around 12 → B From 18 to 68 → W	0.85	1.0	0.8	From 6 to 68 → H
Nitrous Oxide	From 27 to 35 → SB From 9 to 15 → SW Around 24 → SW From 35 to 40 → SW	0.9	0.9	0.9	X
Nitrogen Dioxide	From 27 to 30 → SB	0.8	0.8	0.8	X

Table 7.1 Results of the comparison of the three sets of MWs in function of the altitude

B → Better efficiency for the Comb set of MWs   S → Slightly
W → worse efficiency for the Comb set of MWs   G → Greatly
L → Lower ESD values for the Comb set of MWs
H → Higher ESD values for the Comb set of MWs
X → No relevant divergence between the three sets of MWs.

7.2 Comparison of the sets of MWs using retrievals on real observations

On real observation, the "real" atmosphere isn't known, therefore to compare the three sets of MWs I only can rely on the ESDs. To have a good coverage of the Earth and a reasonable amount of data I selected 15 orbits which represent all spectra recorded by MIPAS during the 20th of June 2008. Two consecutive orbits are merged together as explained in Sub-Sect. 3.3.2. Then I retrieved the profiles from these composed orbits with the BO, OX and Comb sets of MWs. The computing time of one orbit is for the BO set of MWs between 3 and 5 hours (3283 spectral points[1]),for the OX set of MWs between 13 and 15 hours (5145 spectral points[1]) and for the Comb set of MWs between 9 and 11 hours (4655 spectral points[1]).

In Figs 7.33, 7.34, 7.35 and 7.36 I plotted the averaged ESD values of all the retrieved profiles for the three sets of MWs as a function of altitude for Temperature, Pressure, Water and Ozone respectively.
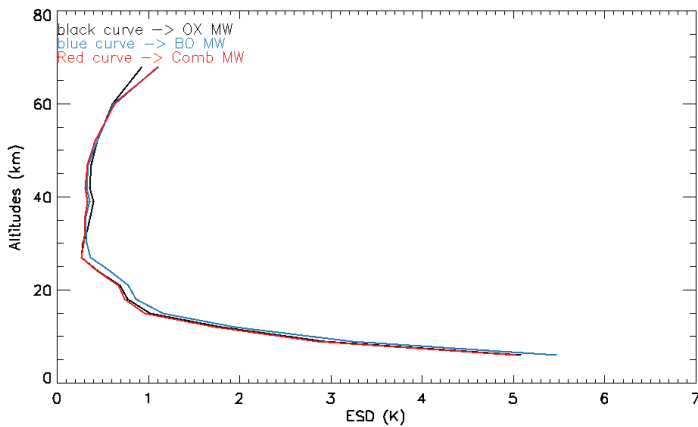

Figure 7.33: ESD values for Temperature as a function of altitude for the three sets sets of MWs
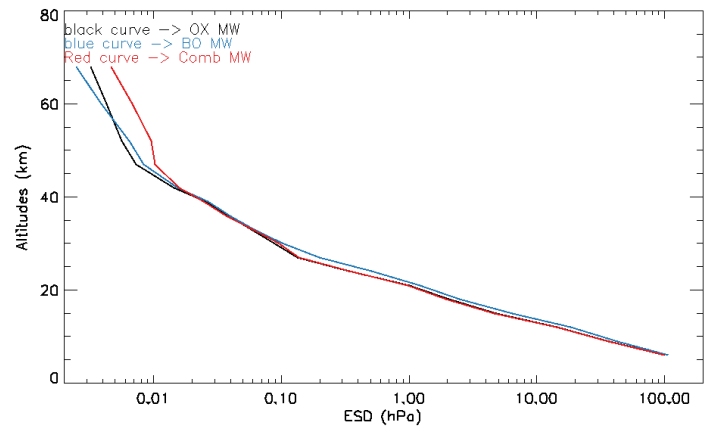

Figure 7.34: ESD values Pressure as a function of altitude for the three sets sets of MWs
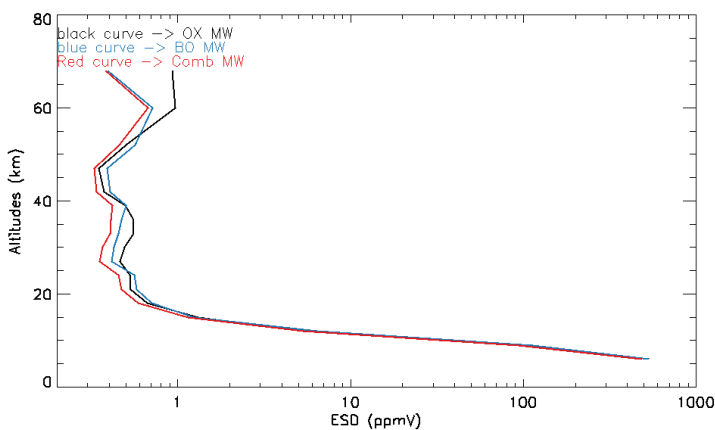

Figure 7.35: ESD values for Water scale as a function of altitude for the three sets sets of MWs
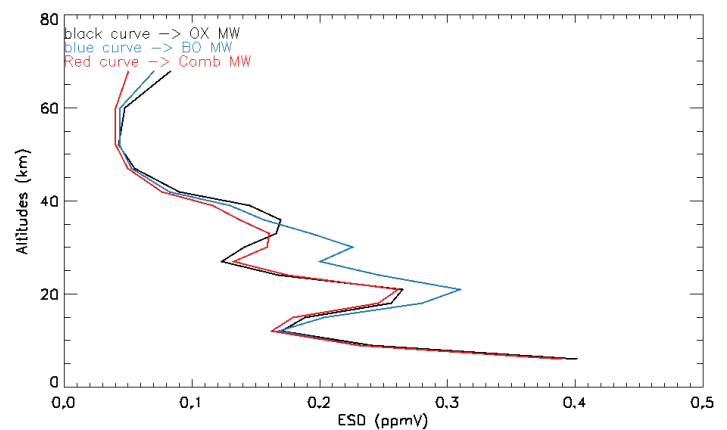

Figure 7.36: ESD values for Ozone as a function of altitude for the three sets sets of MWs

[1] The spectral points represent the number of point in the spectra which are computed by the retrieval process, the computing time increases exponentially with the number of spectral points.

In Figs. 7.33 and 7.34 I observe for the Comb set of MWs slightly lower or equal ESD values than for the two other sets of MWs, except at higher altitudes from 60 to 68 km for Temperature and from 40 to 68 km for Pressure. In Fig. 7.35 I observe for the Comb set of MWs lower ESD values. In Fig. 7.36 I observe for the Comb set of MWs slightly lower or equal ESD values, except from 25 to 33 km. The observation in Fig. 7.36 has to be nuanced since I observed in Fig. 7.15 (Sub-Sect. 7.1.4) that the ESD values were more overestimated for the BO and OX set of MWs than for the Comb set of MWs, especially from 40 to 60 km.

In Figs 7.37, 7.38, 7.39 and 7.40 I plotted the averaged ESD values of all retrieved profiles of one orbit (the one which seemed to have the lower ESD values for the Comb set of MWs) for the three sets of MWs as a function of altitude for Methane, Nitric Acid, Nitrous Oxide and Nitrogen Dioxide respectively. In Fig. 7.37 I observe for the Comb set of MWs slightly higher ESD values from 25 to 35 km and slightly lower ESD values from 50 to 68 km. In fig. 7.38 I observe for the three sets of MWs almost the same ESD values. In Fig. 7.39 I observe for the Comb set of MWs slightly higher ESD values from 25 to35 km and slightly lower ESD values from 40 to 52 km. In Fig. 7.40 I observe for the Comb set of MWs slightly higher ESD values from 25 to30 km and slightly lower ESD values from 45 to 60 km.
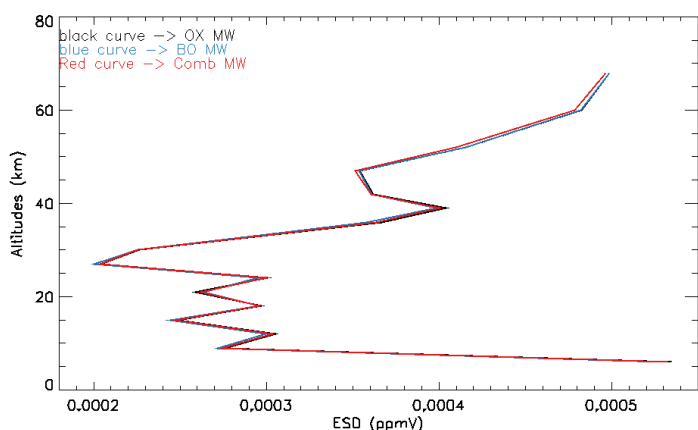

Figure 7.38: ESD values for Nitric Acid as a function of altitude for the three sets sets of MWs
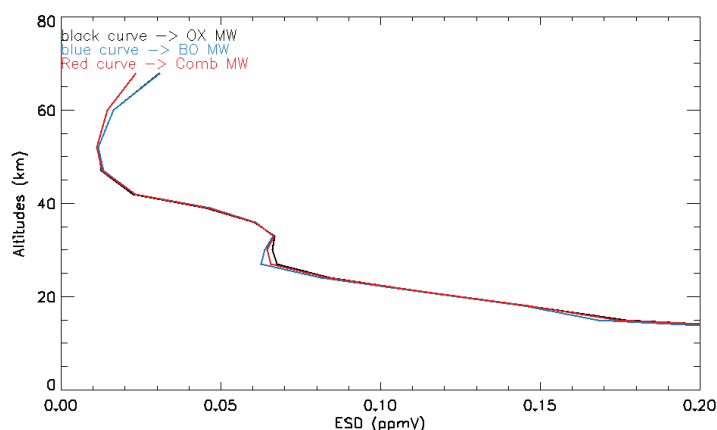

Figure 7.37: ESD values for Methane as a function of altitude for the three sets sets of MWs
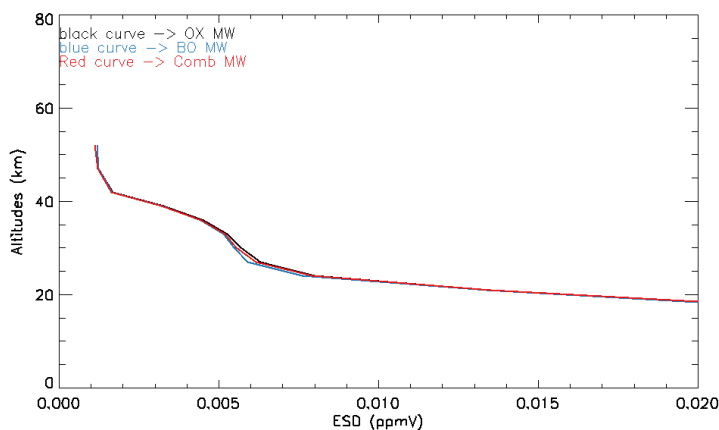

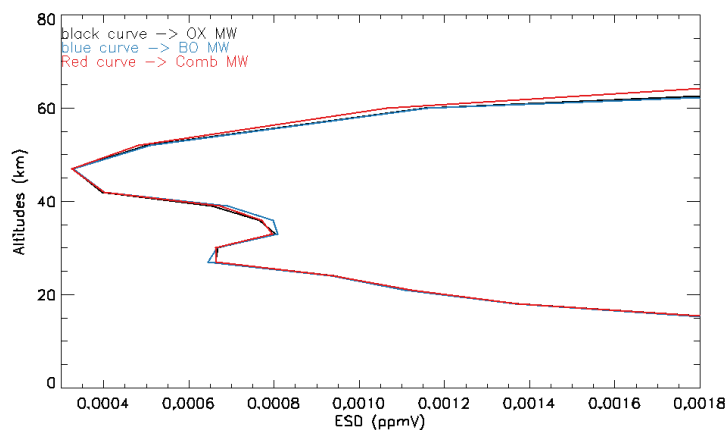Figure 7.39: ESD values for Nitrous Oxide as a function of altitude for the three sets sets of MWs

Figure 7.40: ESD values for Nitrogen Dioxide as a function of altitude for the three sets sets of MWs

7.2.1 Systematic and Total Error

In order to compare the systematic error of the three sets of MWs, in figs. 7.41, 7.42, 7.43 and 7.44 are plotted the computed systematic error values for the three sets of MWs as a function of altitude for Temperature, Pressure, Water and Ozone respectively.

In Fig. 7.41 I observe for the Comb set of MWs slightly lower systematic error values from 6 to15 km, lower from 18 to 30 km and higher from 20 to 60 km. In Fig. 7.42 I observe for the Comb set of MWs slightly lower systematic error values from 6 to 30 km and higher from 20 to 60 km. In Fig. 7.43 I observe for the Comb set of MWs lower systematic error values from 23 to 30 km and from 57 to 68 km, slightly higher from 6 to 23 km and higher from 30 to 57 km. In Fig. 7.44 I observe for the Comb set of MWs lower systematic error values from 12 to 15 km and from 27 to 30 km, slightly higher from 6 to 12 km and from 22 to 27 km and higher from 30 to 65 km.
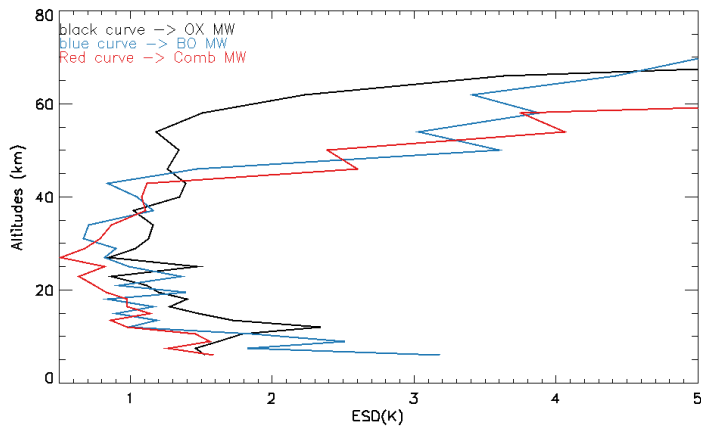

Figure 7.41: Systematic error values for Temperature as a function of altitude for the three sets sets of MWs
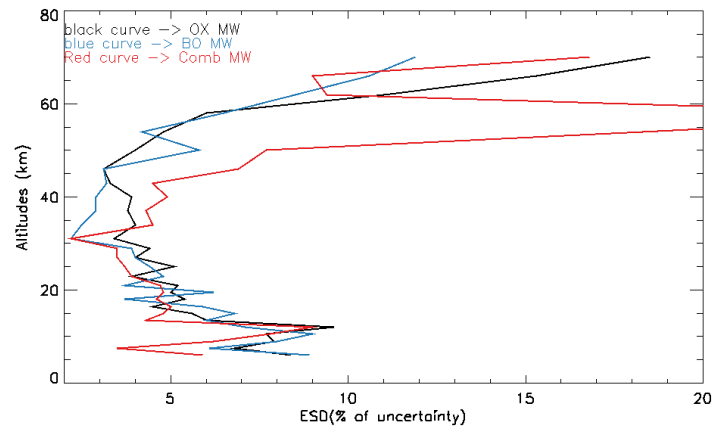

Figure 7.42: Systematic error values for Pressure as a function of altitude for the three sets sets of MWs
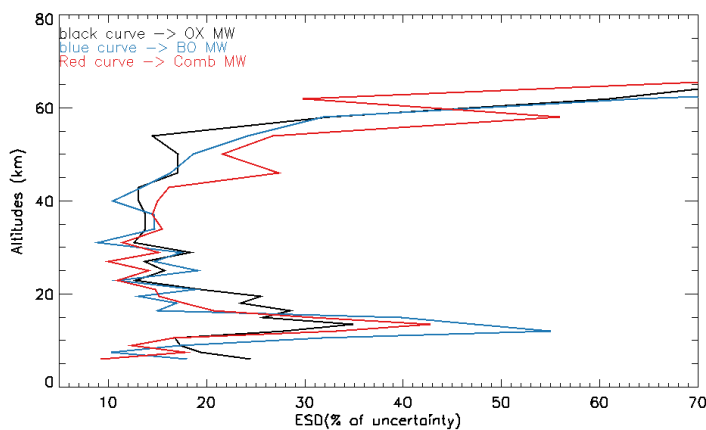

Figure 7.43: Systematic error values for Water as a function of altitude for the three sets sets of MWs
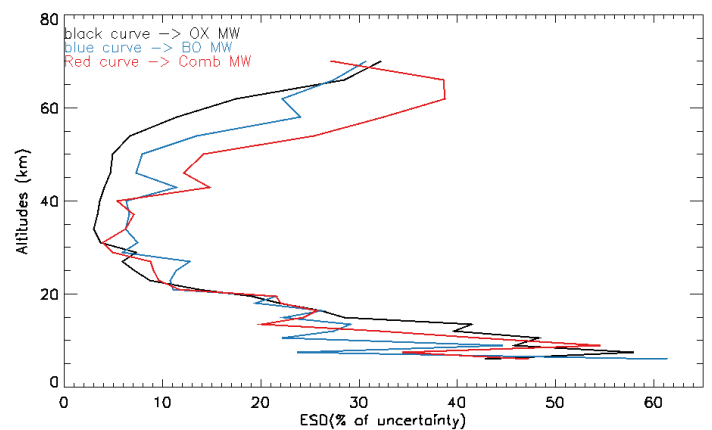

Figure 7.44: Systematic error values for Ozone as a function of altitude for the three sets sets of MWs

During the retrieval process, an estimated value of the random error is calculated, i.e. the ESD. In order to provide an example the total error I selected one profile at mid-latitude (orbit 39969+39970, Sequence 13, latitude 43.98°, longitude 34.04° of the 20 June 2008 at 7:53) and computed the root summation square between the systematic error and the ESD values of this profile.

In Figs 7.45, 7.46, 7.47 and 7.48 the computed total error values for the three sets of MWs are plotted as a function of altitude for Temperature, Pressure, Water and Ozone respectively. In Fig 7.45 I observe for the Comb set of MWs at slightly lower total error values from 6 to12 km and lower from 18 to 30 km and slightly higher from 30 to 42 km and higher from 12 to 18 and from 42 to 68. In Fig 7.46 I observe for the Comb set of MWs ,slightly lower total error values from 6 to 18 km and from 22 to 30 km and slightly higher from 18 to 22 km and higher from 30 to 68. In Fig 7.47 I observe for the Comb set of MWs at slightly lower total error values from 6 to 30 km and from 57 to 68 km and higher from 30 to 57 km. In Fig 7.48 I observe for the Comb set of MWs slightly higher total error values from 6 to 30 km and higher from 30 to 68 km

In Both cases, i.e. the systematic error and total error, I observe generally lower values for the Comb set of MWs at low altitudes and higher values at high altitudes.
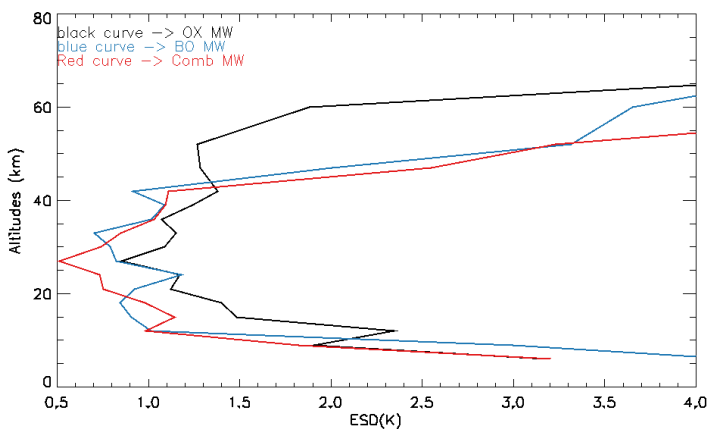

Figure 7.45: Total error values for Temperature as a function of altitude for the three sets sets of MWs
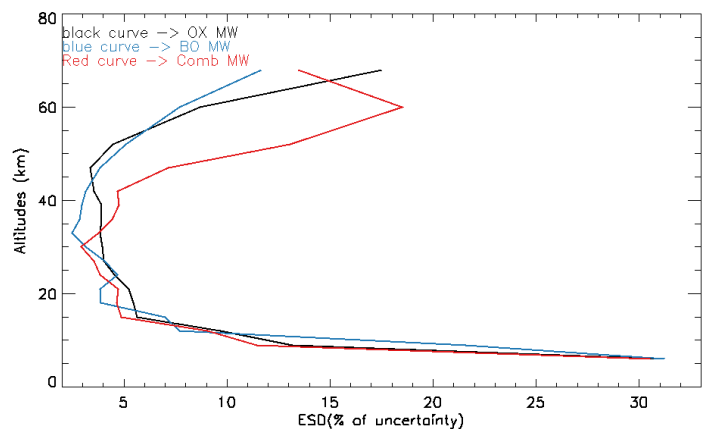
Figure 7.46: Total error values for Pressure as a function of altitude for the three sets sets of MWs
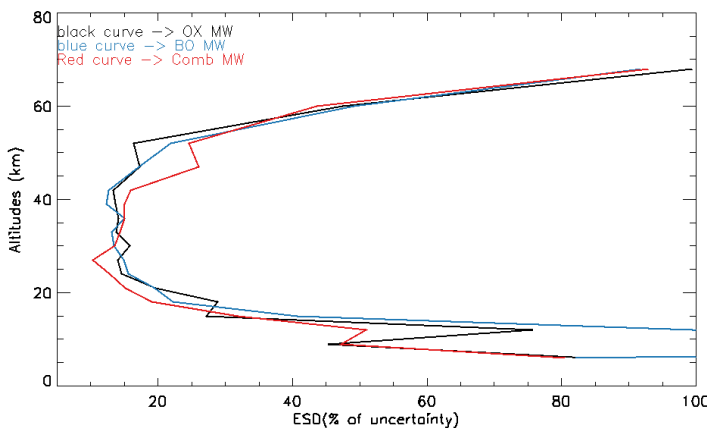
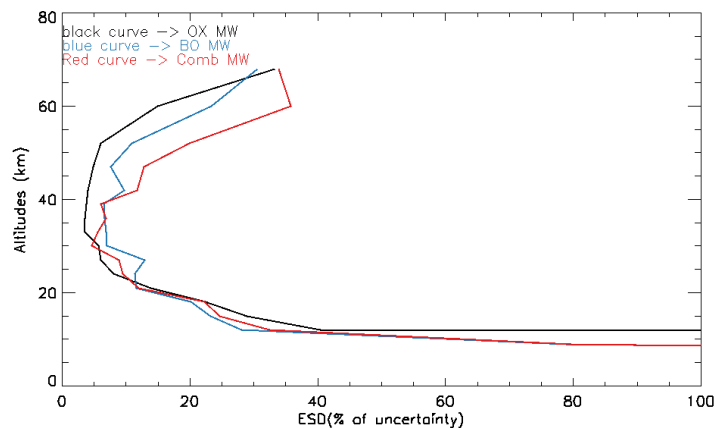Figure 7.47: Total error values for Water as a function of altitude for the three sets sets of MWs

Figure 7.48: Total error values for Ozone as a function of altitude for the three sets sets of MWs

Conclusion:

The first objective of this thesis was to compare the Bologna (BO) and Oxford (OX) sets of Micro-Windows (MWs). For this purpose I first compared the retrieved altitude profiles on simulated observations and then the Information Load (IL) of the two sets of MWs. The second objective was to combine the MWs of the two sets to create a more efficient set. Therefore I calculated the IL of each individual MW and I selected the MWs with the highest IL values. The third objective was to compare the combined set of MW w.r.t. the BO and OX sets. For this purpose I compared the retrieved altitude profiles of the three sets of MWs on simulated and real observations along with their systematic and total errors.

The comparison of the BO and OX sets of MWs on simulated retrieval (Chap. 5) showed for the studied targets (Temperature, Pressure, Water, Ozone, Nitric Acid, Methane, Nitrous Oxide and Nitrogen Dioxide) that the OX set has generally better performance and lower ESD values at low altitudes and on the contrary that the BO set has generally better performance and lower ESD values at high altitudes.

Comparing the IL values of the two sets of MWs (Chap. 6), I observed that the OX set has higher IL values at low altitudes and lower IL values at high altitude. The IL values of the combined set are higher for all the targets at all altitudes, except for Pressure around 50 km.

The comparison of the combined set w.r.t. the BO and OX sets of MWs on simulated retrieval (Sect. 7.1) show that the combined set has better performance and lower ESD values for Temperature and Water and similar performance and ESD values for Pressure and Ozone, except for Pressure above 40 km where the combined set has worse performance and higher ESD values. I note for Temperature a gain in efficiency from 10 to 45 km in the retrieval using the combined set. This is an important achievement since at these altitudes the Temperature values are quite low.

The retrieved ESD values on real observation are lower for the combined set of MWs, except for Ozone from 25 to 30 km and for Pressure above 40 km. The computed total errors for the Combined set are generally lower or equal at low altitudes and higher at high altitudes.

An important point to highlight is the computing time which is required for the retrieval of one full orbit with the BO, OX and combined sets: it is around 4, 14 and 10 hours respectively, which correspond to 3285, 5145 and 4655 computed spectral points.

References:

1. "ENVISAT information website,"
   http://www.esa.int/esaEO/SEMWYN2VQUD_index_0_m.html.

2. "MIPAS information website," http://envisat.esa.int/handbooks/mipas/toc.htm.

3. E. Papandrea, "PhD Thesis Tomographic Inversion of MIPAS Measurements ; Optimization of Retrieval Grids and Generation of Innovative Products," Universita di Bologna (2007).

4. H. Oelhaf and F. Draft, *Mipas Mission Plan* (2008).

5. "MIPAS current status," http://www-atm.physics.ox.ac.uk/group/mipas/.

6. C. D. Rodgers., "Inverse Methods for Atmospheric Sounding, Theory and Practice," World Scientific Publishing Co. Pte. Ltd. (2000).

7. D. S. Sivia, *Data Analysis: A Bayesian Tutorial* (Clarendon Press, 1998).

8. K. Levenberg., "A method for the solution of certain non-linear problems in least squares.," Quart. Appl. Math. (1944).

9. D. W. Marquardt., "An Algorithm for the Least-Squares Estimation of NonLinear Parameters.," J. Soc. Appl. Math. (1963).

10. Press and Al., "Numerical recipes: the art of scientific computing," Cambridge University Press (1995).

11. M. Carlotti, G. Brizzi, E. Papandrea, M. Prevedelli, M. Ridolfi, B. M. Dinelli, and L. Magnani, "GMTR: two-dimensional geo-fit multitarget retrieval model for michelson interferometer for passive atmospheric sounding/environmental satellite observations.," Applied optics **45**, 716-27 (2006).

12. "http://www.ifac.cnr.it/retrieval/orm.html," .

13. A. Dudhia, V. L. Jay, and C. D. Rodgers, "Microwindow selection for high-spectral-resolution sounders.," Applied optics **41**, 3665-73 (2002).

14. C. D. Rodgers., "Inverse Methods for Atmospheric Sounding: Theory and Practice, Vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics," World Scientific Publishing Co. Pte. Ltd. (2000).

15. M. Ridolfi, L. Magnani, M. Carlotti, and B. M. Dinelli, "MIPAS-ENVISAT limb-sounding measurements: trade-off study for improvement of horizontal resolution.," Applied optics **43**, 5814-24 (2004).

16. M. Carlotti, B. M. Dinelli, P. Raspollini, and M. Ridolfi, "Geo-fit Approach to the Analysis of Limb-Scanning Satellite Measurements," Applied Optics **40**, 1872 (2001).

17. 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," Journal of Quantitative Spectroscopy and Radiative Transfer **84**, 141-157 (2004).

18. B. M. Dinelli, "Algorithm Theoretical Baseline Document," 1-131 (2004).

19. M. Carlotti, E. Arnone, E. Castelli, B. M. Dinelli, and E. Papandrea, "POSITION ERROR IN VMR PROFILES RETRIEVED FROM MIPAS OBSERVATIONS WITH A 1-D ALGORITHM," 1-8 (n.d.).

20. M. Carlotti and L. Magnani, "Two-dimensional sensitivity analysis of MIPAS observations.," Optics express **17**, 5340-57 (2009).

21. "http://www.atm.ox.ac.uk/group/mipas/err/," .