Technical Note on:

Averaging Kernels for MIPAS off-line level 2 retrievals

Issue 1
Revision 0

9 January 2004

Delivery of the study:

"Support to MIPAS Level 2 Product Validation"
Contract No: 17580/03/I-OL
WP 6320

Prepared by:

Name	Institute
S. Ceccherini	IFAC-CNR

Approved by:

Name	Institute
B. Carli	IFAC-CNR

CONTENTS

1. Background

MIPAS (Michelson Interferometer for Passive Atmospheric Sounding) is an ESA developed instrument operating on board ENVISAT-1 as part of the first Polar Orbit Earth Observation Mission program (POEM-1). MIPAS performs limb sounding observations of the atmospheric emission spectrum in the middle infrared region. Concentration profiles of numerous trace gases can be derived from MIPAS observed spectra.
According to the current baseline, from MIPAS measurements altitude profiles of atmospheric pressure and temperature (p,T), and of volume mixing ratio (VMR) of six high priority species (H_2O, O_3, HNO_3, CH_4, N_2O and NO_2) are routinely retrieved in near real time (NRT). The retrieval of these parameters from calibrated spectra (level 1b) is indicated as NRT Level 2 processing.
The scientific code for NRT level 2 analysis optimised for the requirements of speed and accuracy is called the Optimised Retrieval Model (ORM) [RD1]. The ORM has been used for the development of the NRT level 2 processor of MIPAS.
The speed requirement of NRT Level 2 processing has imposed a compromise in accuracy. In order to overcome this accuracy limitation ESA has decided to produce, in addition to the NRT Level 2 products, a subset of improved data using the ORM code and a different set of auxiliary data files. Improvements include a reduction of the retrieval error because of more stringent convergence criteria and a reduction of the extrapolation error because of an extended retrieval range. The products of this processing are distributed off-line (OL) by ESA [RD2].
Each product is characterised by some auxiliary parameters among which are the averaging kernels. The averaging kernels for the NRT Level 2 products are available and they are described in [RD3].
The averaging kernels for the OL MIPAS products have been calculated and are described in this technical note.

2. Reference documents

[RD1] M. Ridolfi, B. Carli, M. Carlotti, T. von Clarmann, B. M. Dinelli, A. Dudhia, J-M Flaud, M. Hopfner, P. E. Morris, P. Raspollini, G. Stiller and R. J. Wellis, “Optimized forward model and retrieval scheme for MIPAS near-real-time data processing”, Appl. Opt. **39**, 1323-1340 (2000).

[RD2] MIPAS Quality Working Group, “ENVISAT MIPAS monthly report: October 2003”, Technical Note ENVI-SPPA-EOPG-TN-03-0026, Issue 1.0 (20 November 2003).

[RD3] S. Ceccherini and M. Ridolfi, “Averaging Kernels for MIPAS near real time level 2 retrievals”, Technical Note TN-IFAC-OST0201, Issue 1 (14 June 2002) of the ESA study “Development of an Optimised Algorithm for Routine p, T and VMR Retrieval from MIPAS Limb Emission spectra”, Contract No. 11717/95/NL/CN-CCN5.

[RD4] C. D. Rodgers, Inverse Methods for Atmospheric Sounding – Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics – Vol. 2, World Scientific, (2000).

[RD5] S. Ceccherini, B. Carli, E. Pascale, M. Prosperi, P. Raspollini and B. M. Dinelli, “Comparison of measurements made with two different instruments of the same atmospheric vertical profile”, Appl. Opt. **42** No. 32, 6465-6473 (2003).

[RD6] A.Dudhia and V.Payne, “AMIL2DA error analysis”, delivery D40 of the EC study “Advanced MIPAS Level 2 Data Analysis (AMIL2DA)”, contract n. EVG1-CT-1999-00015, (April 2002)

3. Rationale of the averaging kernels

In order to determine the accuracy of MIPAS level 2 data for all relevant atmospheric conditions a validation of MIPAS products must be performed. This will be achieved by comparing the MIPAS data products with temporally and spatially coincident measurements performed with independent instrumentation.
When the reliability of MIPAS products will be proved the level 2 profiles will be assimilated in atmospheric composition models. The assimilation involves the combination of the measurements with the model forecast.
The comparison/combination of measurements of the same quantity performed with different instruments has to take into account the different characteristics of the observing systems [RD4, RD5]. The different measurements may have, not only different measurement errors, but also a different representations of the retrieved profiles due to:
- different altitude grids
- different vertical resolutions
- different a priori information (if used).

As a consequence the variance covariance matrix (VCM) of the retrieved profiles does not provide a sufficient information to perform a correct comparison between independent measurements. In addition to error information it is necessary to provide a description of how the observing system modifies the true state of the atmosphere. This information is included in the averaging kernel matrix.
The mathematical formalism of the averaging kernels is already described in [RD3] but for the self consistency of this document it is here repeated in sections 4 and 5.

4. Averaging kernel matrix

In the ideal case of absence of both random and systematic errors in the measured signal and in the instrument's forward model for each true state $\mathbf{x}$ of the atmosphere the observing system provides a retrieved profile $\hat{\mathbf{x}}_v$ which is a function of $\mathbf{x}$.
Expanding $\hat{\mathbf{x}}_v$ up to the first order about a generic atmospheric state $\mathbf{x}_0$ we obtain:

(1) $$\hat{\mathbf{x}}_v - \hat{\mathbf{x}}_{v\mathbf{0}} = \left.\frac{\partial \hat{\mathbf{x}}_v}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_\mathbf{0})$$

The quantity:

(2) $$\mathbf{A} \equiv \left.\frac{\partial \hat{\mathbf{x}}_v}{\partial \mathbf{x}}\right|_{\mathbf{x}_0}$$

is called averaging kernel matrix (AKM) [RD4,RD6] and it is a function of the state $\mathbf{x}_0$.
In the real case when both random and systematic errors in the measured signal and in the instrument's forward model are present for each true state $\mathbf{x}$ of the atmosphere the observing system provides a retrieved profile

(3) $$\hat{\mathbf{x}} = \hat{\mathbf{x}}_v + \boldsymbol{\varepsilon}_\mathbf{x}$$

From (1-3) it follows the relation connecting the retrieved profile with the true state of the atmosphere:

(4) $$\hat{\mathbf{x}} - \hat{\mathbf{x}}_{v\mathbf{0}} = \mathbf{A}(\mathbf{x} - \mathbf{x}_\mathbf{0}) + \boldsymbol{\varepsilon}_\mathbf{x}$$

or equivalently:

(5) $$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x} + \hat{\mathbf{x}}_{v0} - \mathbf{A}\mathbf{x}_\mathbf{0} + \boldsymbol{\varepsilon}_\mathbf{x}$$

From (5) it is clear that in the case $\hat{\mathbf{x}}_{v0} \approx \mathbf{A}\mathbf{x}_0$ the AKM directly provides the description of how the observing system modifies the true state of the atmosphere.
The rows and the columns of $\mathbf{A}$ matrix are respectively the averaging kernels and the delta-function responses.
The averaging kernel $\frac{\partial \hat{x}_{vi}}{\partial \mathbf{x}}$ indicates where the information at each retrieved profile level i originates from, while the delta-function response $\frac{\partial \hat{\mathbf{x}}_v}{x_j}$ shows how the retrieval responds to a perturbation of the true profile at level j.
In a ideal inverse method $\mathbf{A}$ would be a unit matrix. In reality the rows of $\mathbf{A}$ are in general functions peaking at the appropriate level and with a half-width which is a measure of the spatial resolution of the observing system, thus providing a simple characterisation of the relationship between the retrieval and the true state.

5. Methods for the calculation of the AKM

There are two methods for the calculation of $\mathbf{A}$ [RD6]: the perturbation method and the analytical method.
The perturbation method is based on the fact that the columns of $\mathbf{A}$ give the response of the retrieval to a delta-function perturbation in the state vector. The delta-function response can be numerically calculated by finding the change in the retrieval which results when each element of the state vector is perturbed by some suitably small amount. The perturbation should be small enough so that the response is linear in the size of the perturbation, but large enough to make negligible rounding errors. The response is placed in the appropriate column of the $\mathbf{A}$ matrix.
The analytical method is based on the following expression for $\mathbf{A}$ [RD4]:

(6) $$\mathbf{A} = \mathbf{G}\mathbf{K}$$

where $\mathbf{G}$ is the gain matrix representing the mapping of the measurements $\mathbf{y}$ into the retrieval and $\mathbf{K}$ is the jacobian matrix calculated in the vertical grid of the true profile $\mathbf{x}$.
Recalling the expression, expanded up to the first order, of the measurements as a function of the state of the atmosphere:

(7) $$\mathbf{y} - \mathbf{y}_0 = \mathbf{K}(\mathbf{x} - \mathbf{x}_0) + \boldsymbol{\varepsilon}_\mathbf{y}$$

the expression, expanded up to the first order, of the retrieved atmospheric state as a function of the measurements:

(8) $$\hat{\mathbf{x}} - \hat{\mathbf{x}}_{v0} = \mathbf{G}(\mathbf{y} - \mathbf{y}_0)$$

and combining the two expressions we obtain:

(9) $$\hat{\mathbf{x}} - \hat{\mathbf{x}}_{v0} = \mathbf{G}\mathbf{K}(\mathbf{x} - \mathbf{x}_0) + \mathbf{G}\boldsymbol{\varepsilon}_\mathbf{y}$$

From (4) it follows (6).

6. Calculation of AKM for OL processor

The AKM for ORM was calculated using the perturbation method.
The procedure used for the calculation of the matrix $\mathbf{A}$ can be summarised in the followings steps:

1) simulation of the retrieval in the case of a profile $\mathbf{x}_0$, by performing the calculation of the forward model of the measurements with added a random error of magnitude comparable with the real MIPAS random error and retrieval of the so-called unperturbed profile,
2) perturbation of the profile $\mathbf{x}_0$ at a single altitude
3) calculation with the forward model of the measurements that are obtained in the case of perturbed profile with added the same random error used in step 1.
4) use of this simulated measurements for a retrieval with the ORM
5) differentiation of the retrieved profile with respect to the unperturbed profile of step 1
6) ratio of retrieved profile perturbation divided by input profile perturbation

The result of this procedure is a column of $\mathbf{A}$. By operating the perturbation on all the levels of the input profile the overall AKM is calculated.
The true profile $\mathbf{x}$ was defined on a grid of 1 km step from 0 to 120 km.
The perturbation of the input profile was fixed at 5 K for temperature, 1 ppmv for H_2O and 5% of the maximum value for the others VMRs. From several trials with different sizes of the perturbations these values resulted to be the best compromise between the requirements of small values in order to be in the linear response region, and large values in order to avoid numerical errors.
The AKMs for ORM have been calculated for the nominal occupation matrices corresponding to version 3.5 of auxiliary files (see Appendix A).
The procedure for the calculation of the AKM described above implies the use of the forward model to simulate the atmospheric spectra. Because the simulation of the spectra should approximate the atmospheric behaviour as well as possible we have decided to improve the calculation of the Field of View (FOV) convolution and to adopt a close spacing of 1 km for the grid of tangent altitudes used for the modelling of the line of sights and a linear interpolation between consecutive tangent altitudes for the determination of the function that is convoluted with the FOV. On the other hand we recall that the AKM for the NRT products [RD3] were obtained using in the forward model the same FOV convolution method as the one used in ORM retrieval (which was optimised for reducing the calculation time and performs spectra simulations on a coarse tangent altitude grid and a polynomial interpolation between consecutive tangent altitudes). So from this point of view the calculation of the AKM is more accurate for the OL products than for NRT products.
Since the AKM depend on the profile $\mathbf{x}_0$ the AKM should be calculated for a set of typical atmospheric conditions. For instance the AKM of the NRT retrieval were calculated for the four seasons (January, April, July and October) and for six latitude bands. However from the analysis of these AKMs we have seen [RD5] that they are practically coincident. On the basis of this consideration we have decided to calculate the AKM only for a typical profile of April at 45 degrees of latitude.
The engineering information on the pointing was used as a priori information. No other a priori information was used. Regularisation was not used.
As pointed out in the first section, in OL MIPAS processor the retrieval range was extended with respect to the retrieval range of the NRT processor in order to reduce the error induced by the assumed profile above the highest and below the lowest retrieved tangent altitudes. A comparison of the NRT retrieval range with the OL retrieval range is shown in Appendix B. This implies that the retrieval is improved at the boundaries of the NRT retrieval range, but it is bound to have problems at the new boundaries of the OL retrieval range.
This occurrence is highlighted by some tests performed with the AK. When the AKM is calculated using different "seeds" for the generation of the random errors added to the spectra the same AKMs must be obtained. If different AKMs are obtained it implies that the measurement error can change the retrieved values of an amount large enough to move the retrieval well outside its linear variation. The variability of the AK with the noise is therefore a test that identifies the points of the

extended retrieval range for which an insufficient information is present in the measurements. Fig. 1 and 2 show examples of the rows of the AKMs that display this variability as a function of the seed used for generating the random errors. In these cases the averaging kernels cannot be defined and are not provided. In appendix B the altitude points for which the AK cannot be defined are indicated with a shaded background. The values retrieved at these altitude points are reported as part of the MIPAS products for their mathematical value (they are the values used by the retrieval for the modelling of the profile), but are not expected to have a physical meaning.

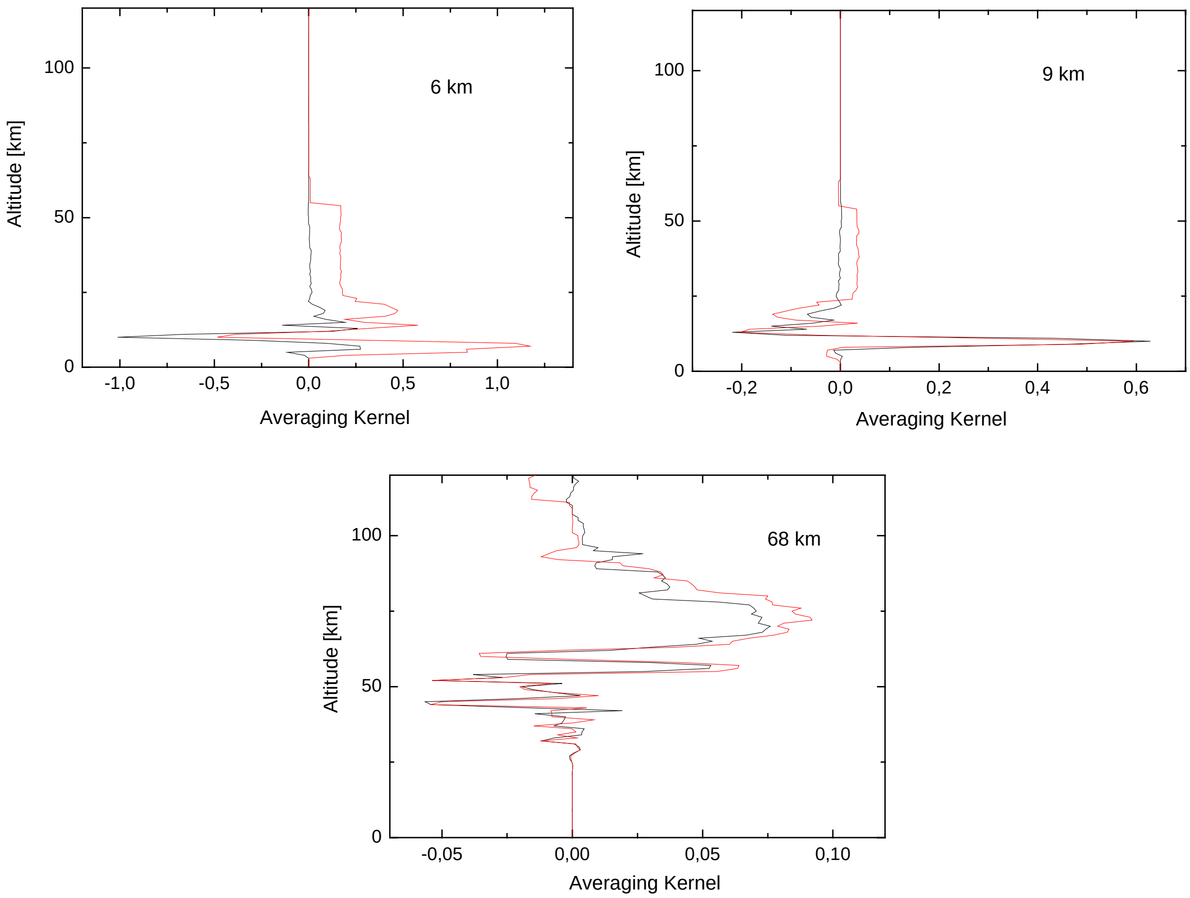

Fig. 1 Averaging kernels of H_2O retrieval corresponding to the altitudes 6 km, 9 km and 68 km calculated with two different seeds for generating the random errors of the spectra.

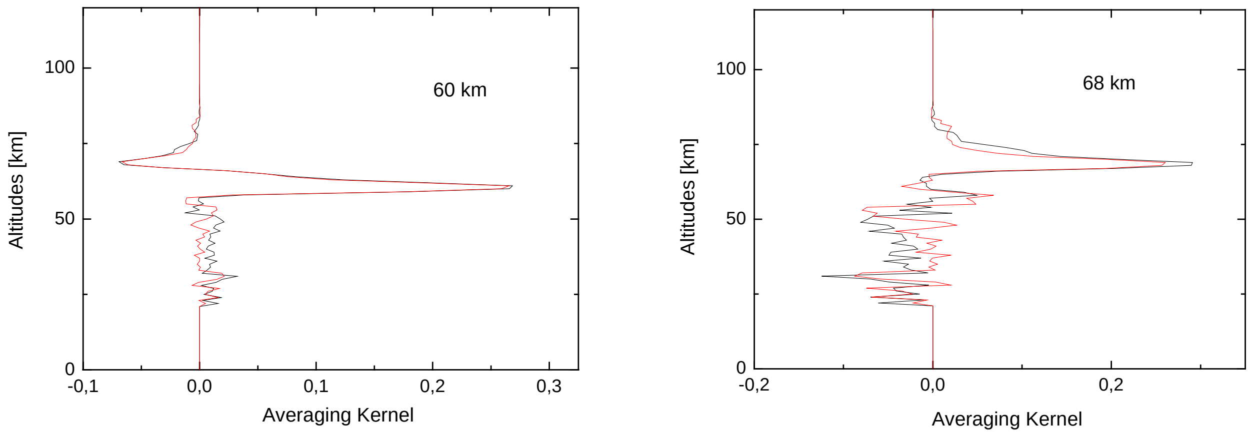

Fig. 2 Averaging kernels of NO_2 retrieval corresponding to the altitudes 60 km and 68 km calculated with two different seeds for generating the random errors of the spectra.

7. Results

The files containing the AKMs are:

AK_T.dat	= AKM of retrieved temperature profile
AK_H2O.dat	= AKM of retrieved H_2O VMR profile
AK_O3.dat	= AKM of retrieved O_3 VMR profile
AK_HNO3.dat	= AKM of retrieved HNO_3 VMR profile
AK_CH4.dat	= AKM of retrieved CH_4 VMR profile
AK_N2O.dat	= AKM of retrieved N_2O VMR profile
AK_NO2.dat	= AKM of retrieved NO_2 VMR profile

AKMs files are in ASCII format. Each of these files contains a matrix of 121 columns and a number of rows dependent on the species. Each of the 121 columns refers to one altitude level adopted for the "fine" discretisation of the atmosphere [RD6] (from 0 to 120 km in 1 km steps, column #1 refers to 120 km, column #121 refers to 0 km). The rows of the AKMs refer to a subset of altitudes at which profiles are retrieved. The information on the retrieval altitudes corresponding to the rows of AKMs is included in the Appendix C. In any case, the first row of the AKMs relates to the highest altitude while the last row refers to the lowest altitude.
AKM files can be read using the following FORTRAN lines:

```
open (unit = 1, file = 'AK_xxx.dat', status = 'old')
        do j=1,ipar
                read ( 1, '(121(1pe15.5))' ) (AK(j,jret),jret=1,121)
        end do
close (1)
```

where 'ipar' is the number of rows of AKM for the profile 'xxx'.
The figures 3-9 show curves representing both rows (left panels) and a subset of columns (right panels), corresponding to a subset of the retrieved profile altitudes, of AKMs for temperature and the six retrieved VMRs.

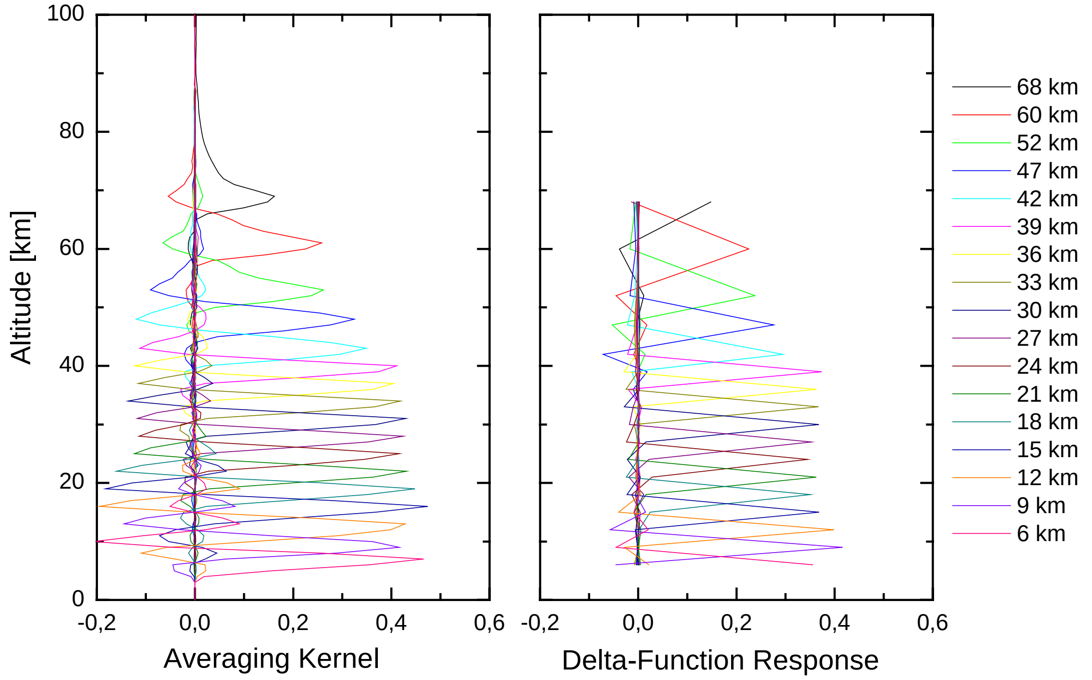

Fig. 3 Temperature AKM

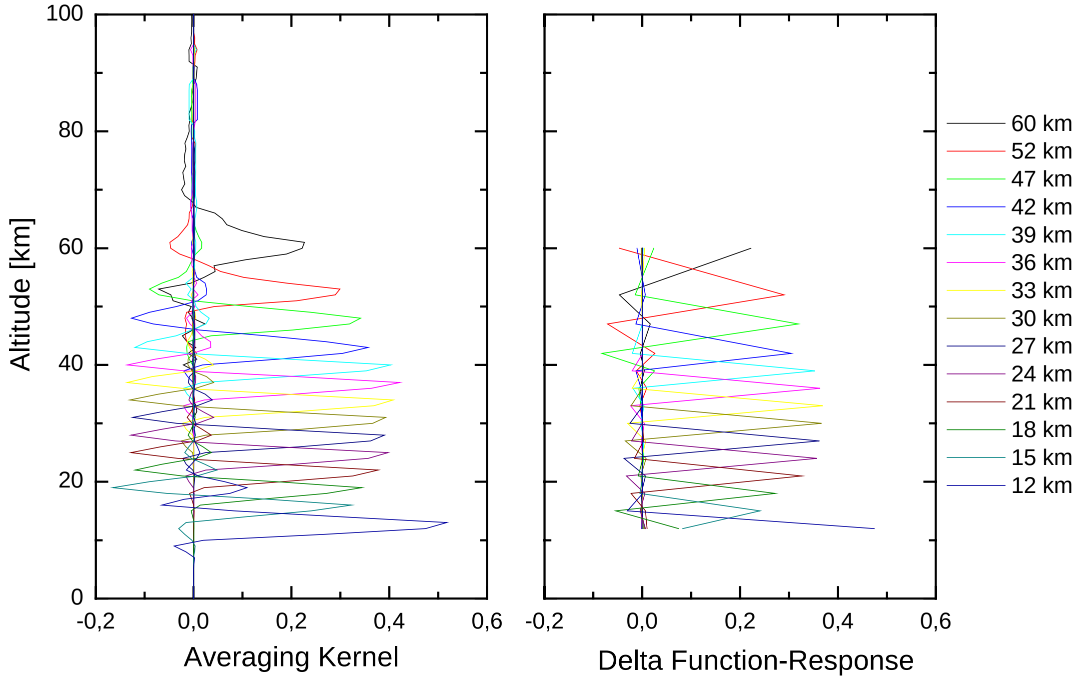

Fig. 4 H_2O AKM

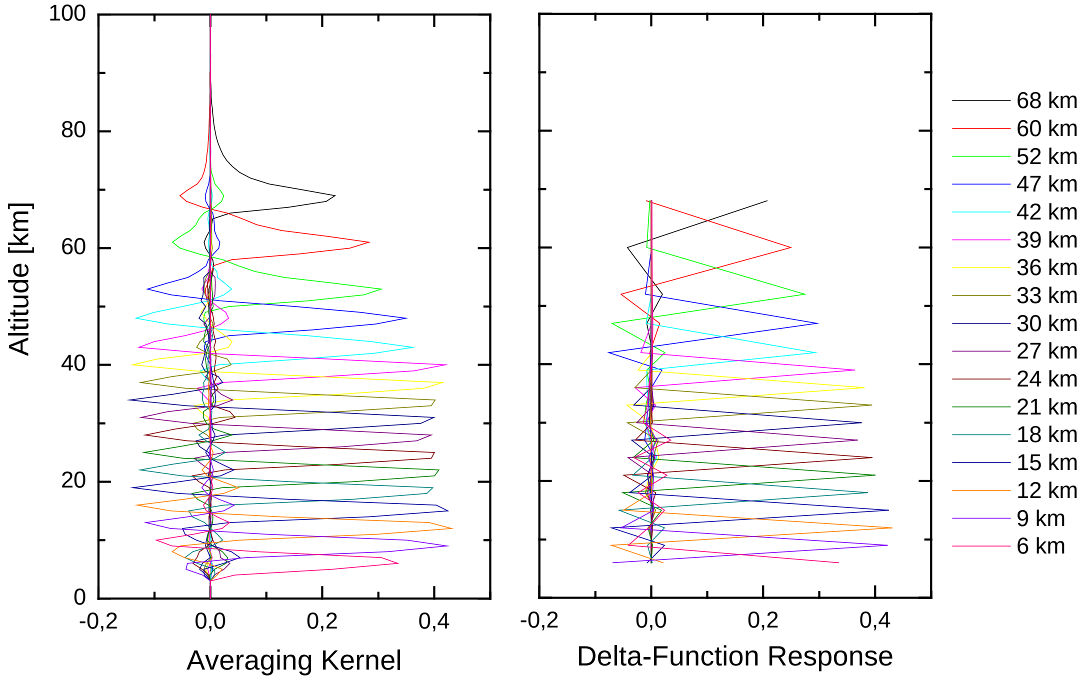

Fig. 5 O_3 AKM

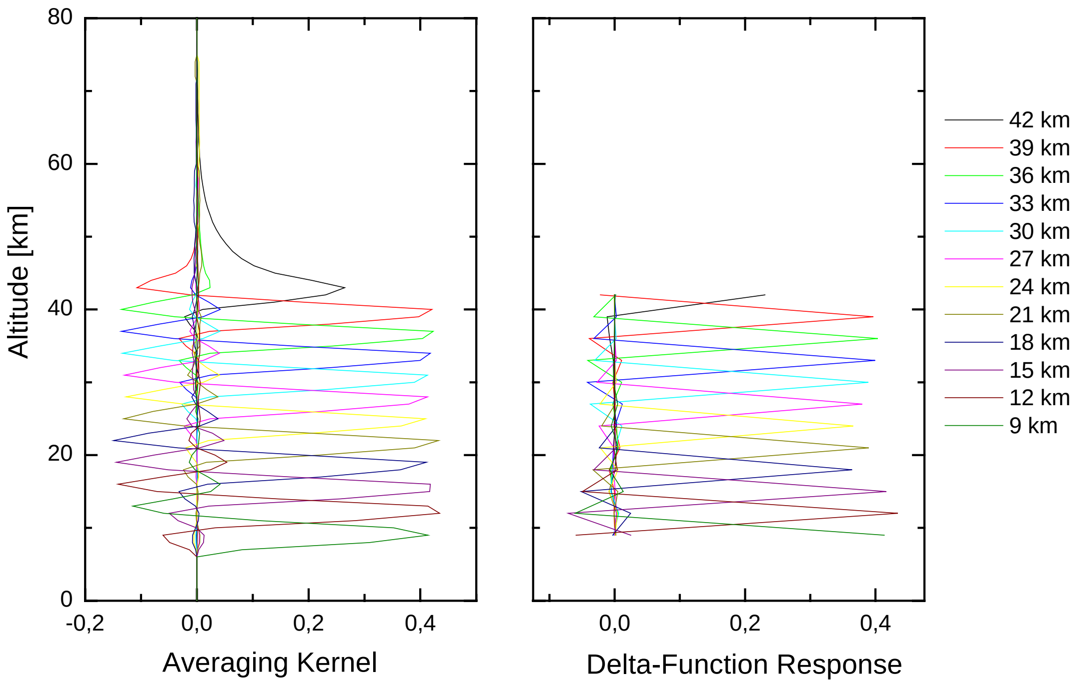

Fig. 6 HNO_3 AKM

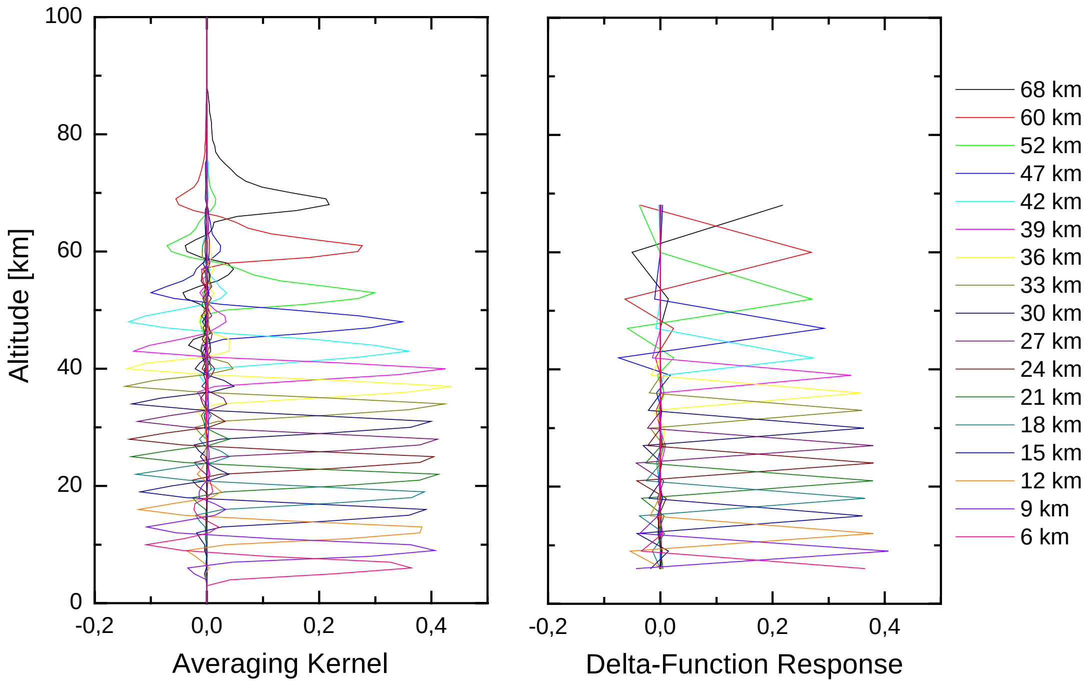

Fig. 7 CH_4 AKM

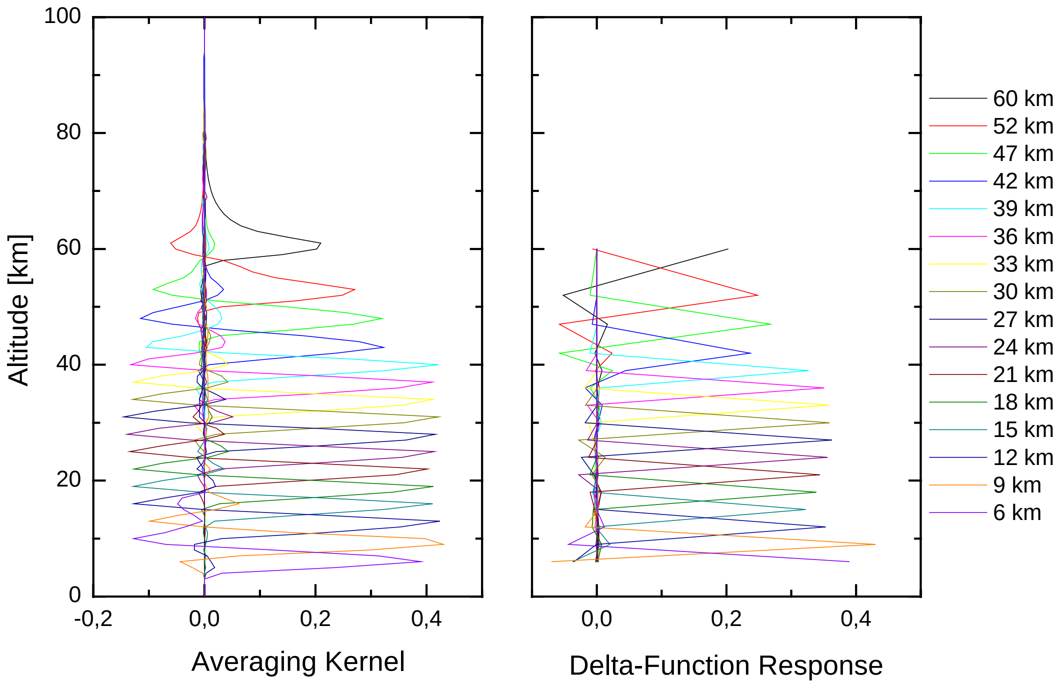

Fig. 8 N_2O AKM

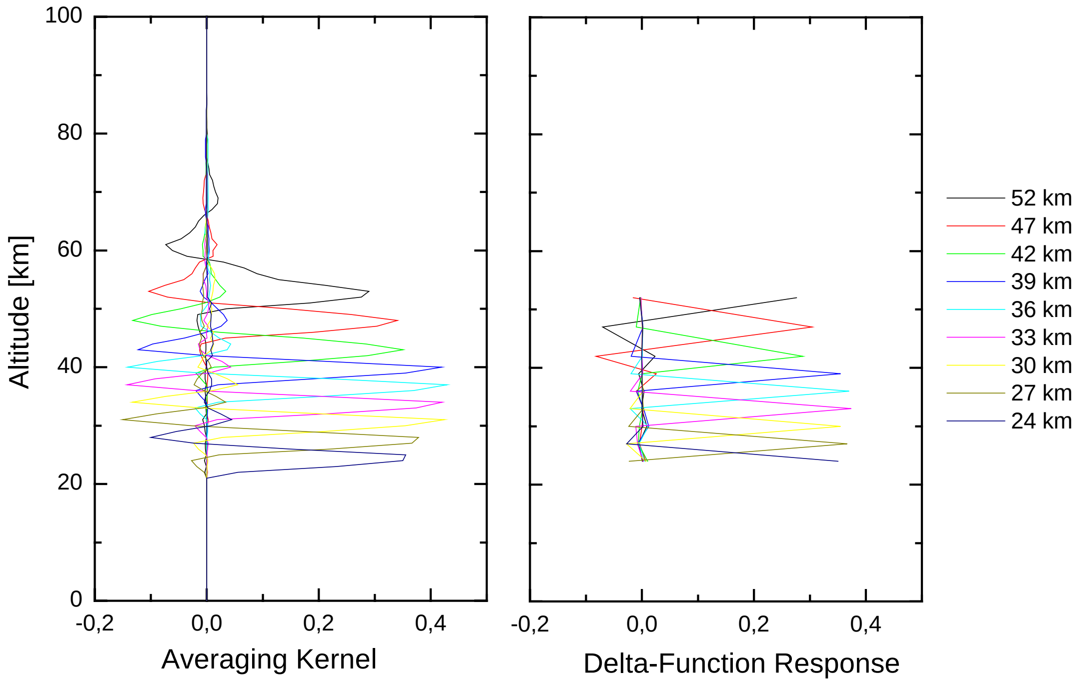
100
80
60
40
20
0
Altitude [km]
-0,2
0,0
0,2
0,4
Averaging Kernel
Delta-Function Response
52 km
47 km
42 km
39 km
36 km
33 km
30 km
27 km
24 km

Fig. 9 NO_2 AKM

Appendix A: spectral and altitude ranges of microwindows used for the calculation of AKMs

PT retrieval			
1	PT___oxf_039	685.7000 - 685.8250 cm^{-1}	33.0 – 47.0 km
2	PT___oxf_001	686.4000 - 689.4000 cm^{-1}	30.0 – 68.0 km
4	PT___oxf_037	694.8000 - 695.1000 cm^{-1}	27.0 – 36.0 km
5	PT___oxf_038	700.4750 - 701.0000 cm^{-1}	21.0 – 30.0 km
6	PT___oxf_004	728.3000 - 729.1250 cm^{-1}	15.0 – 27.0 km
7	PT___oxf_006	741.9750 – 742.2500 cm^{-1}	15.0 – 24.0 km
8	PT___oxf_002	791.3750 – 792.8750 cm^{-1}	6.0 – 33.0 km
H_2O retrieval			
1	H2O__oxf_002	807.8500 - 808.4500 cm^{-1}	12.0 – 18.0 km
2	H2O__oxf_022	946.6500 – 947.7000 cm^{-1}	6.0 – 18.0 km
3	H2O__oxf_007	1645.5250 – 1646.2000 cm^{-1}	27.0 – 60.0 km
4	H2O__oxf_001	1650.0250 – 1653.0250 cm-1	15.0 – 68.0 km
O_3 retrieval			
1	O3___oxf_021	763.3750 - 766.3750 cm^{-1}	6.0 – 68.0 km
2	O3___oxf_013	1039.3750 – 1040.3250 cm-1	52.0 – 68.0 km
3	O3___oxf_001	1122.8000 - 1125.8000 cm^{-1}	6.0 – 68.0 km
HNO_3 retrieval			
1	HNO3_oxf_001	876.3750 - 879.3750 cm^{-1}	6.0 – 68.0 km
2	HNO3_oxf_006	885.1000 - 888.1000 cm^{-1}	6.0 – 42.0 km
CH_4 retrieval			
1	CH4__oxf_012	1227.1750 - 1230.1750 cm^{-1}	6.0 - 60.0 km
4	CH4__oxf_001	1350.8750 - 1353.8750 cm^{-1}	12.0 - 68.0 km
N_2O retrieval			
2	N2O__oxf_012	1233.2750 - 1236.2750 cm^{-1}	6.0 – 27.0 km
6	N2O__oxf_001	1272.0500 - 1275.0500 cm^{-1}	12.0 – 60.0 km
NO_2 retrieval			
1	NO2__oxf_001	1607.2750 - 1610.2750 cm^{-1}	15.0 – 68.0 km
2	NO2__oxf_003	1613.7250 - 1616.6000 cm^{-1}	15.0 – 68.0 km
4	NO2__oxf_013	1622.5500 - 1623.4750 cm^{-1}	6.0 - 30.0 km

Appendix B: comparison between the NRT retrieval grid (km) and the OL retrieval grid (km).

PT		H_2O		O_3		HNO_3		CH_4		N_2O		NO_2	
NRT	OL	NRT	OL	NRT	OL	NRT	OL	NRT	OL	NRT	OL	NRT	OL
68	68		68		68				68				68
60	60	60	60	60	60			60	60		60		60
52	52	52	52	52	52			52	52		52		52
47	47	47	47	47	47			47	47	47	47	47	47
42	42	42	42	42	42	42	42	42	42	42	42	42	42
39	39	39	39	39	39	39	39	39	39	39	39	39	39
36	36	36	36	36	36	36	36	36	36	36	36	36	36
33	33	33	33	33	33	33	33	33	33	33	33	33	33
30	30	30	30	30	30	30	30	30	30	30	30	30	30
27	27	27	27	27	27	27	27	27	27	27	27	27	27
24	24	24	24	24	24	24	24	24	24	24	24	24	24
21	21	21	21	21	21	21	21	21	21	21	21		
18	18	18	18	18	18	18	18	18	18	18	18		
15	15	15	15	15	15	15	15	15	15	15	15		
12	12	12	12	12	12	12	12	12	12	12	12		
	9		9		9		9		9		9		
	6		6		6				6		6		

Appendix C: retrieval altitudes (km) corresponding to the rows of AKMs

PT	H_2O	O_3	HNO_3	CH_4	N_2O	NO_2
68	60	68	42	68	60	52
60	52	60	39	60	52	47
52	47	52	36	52	47	42
47	42	47	33	47	42	39
42	39	42	30	42	39	36
39	36	39	27	39	36	33
36	33	36	24	36	33	30
33	30	33	21	33	30	27
30	27	30	18	30	27	24
27	24	27	15	27	24	
24	21	24	12	24	21	
21	18	21	9	21	18	
18	15	18		18	15	
15	12	15		15	12	
12		12		12	9	
9		9		9	6	
6		6		6		