

## A blind test retrieval experiment for infrared limb emission spectrometry

T. von Clarmann,<sup>1</sup> S. Ceccherini,<sup>2</sup> A. Doicu,<sup>3</sup> A. Dudhia,<sup>4</sup> B. Funke,<sup>5</sup> U. Grabowski,<sup>1</sup> S. Hilgers,<sup>3</sup> V. Jay,<sup>6</sup> A. Linden,<sup>1</sup> M. López-Puertas,<sup>5</sup> F.-J. Martín-Torres,<sup>1,7</sup> V. Payne,<sup>4</sup> J. Reburn,<sup>6</sup> M. Ridolfi,<sup>8</sup> F. Schreier,<sup>3</sup> G. Schwarz,<sup>3</sup> R. Siddans,<sup>6</sup> and T. Steck<sup>9</sup>

Received 5 June 2003; revised 24 September 2003; accepted 2 October 2003; published 13 December 2003.

[1] The functionality and characteristics of six different data processors (i.e., retrieval codes in their actual software and hardware environment) for analysis of high-resolution limb emission infrared spectra recorded by the space-borne Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) have been validated by means of a blind test retrieval experiment based on synthetic spectra. For this purpose a self-consistent set of atmospheric state parameters, including pressure, temperature, vibrational temperatures, and abundances of trace gases and aerosols, has been generated and used as input for radiative transfer calculations for MIPAS measurement geometry and configuration. These spectra were convolved with the MIPAS field of view, spectrally degraded by the MIPAS instrument line shape, and, finally, superimposed with synthetic measurement noise. These synthetic MIPAS measurements were distributed among the participants of the project “Advanced MIPAS level-2 data analysis” (AMIL2DA), who performed temperature and species abundance profile retrievals by inverse radiative transfer calculations. While the retrieved profiles of atmospheric state parameters reflect some characteristics of the individual data processors, it was shown that all the data processors under investigation are capable of producing reliable results in the sense that deviations of retrieved results from the reference profiles are within the margin that is consistent with analytical error estimation. *INDEX TERMS*: 0340 Atmospheric Composition and Structure: Middle atmosphere—composition and chemistry; 1610 Global Change: Atmosphere (0315, 0325); 1640 Global Change: Remote sensing; *KEYWORDS*: MIPAS, retrieval, remote sensing

**Citation:** von Clarmann, T., et al., A blind test retrieval experiment for infrared limb emission spectrometry, *J. Geophys. Res.*, 108(D23), 4746, doi:10.1029/2003JD003835, 2003.

### 1. Introduction

[2] Quantitative knowledge of atmospheric composition plays a key role in the assessment of global change with its two interlinked major subproblems ozone destruction and global warming. The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) has been designed to

measure the abundances of a large number of trace species of the Earth’s atmosphere [Endemann and Fischer, 1993; Fischer and Oelhaf, 1998]. MIPAS is part of the core payload of ESA’s Environmental Satellite (Envisat), which was launched successfully into its Sun-synchronous polar orbit on 1 March 2002.

[3] Species measured by MIPAS are the nitrogen family (N<sub>2</sub>O, HNO<sub>3</sub>, NO, NO<sub>2</sub>, ClONO<sub>2</sub>, N<sub>2</sub>O<sub>5</sub> and others), chlorine reservoirs (again ClONO<sub>2</sub>, HOCl), chlorine source species (CFC-11, CFC-12, CFC-22, CCl<sub>4</sub>; possibly also CFC-113, HCFC-123, HCFC-141b, and HCFC-142b), the chlorine radical ClO, dynamic tracers (H<sub>2</sub>O, N<sub>2</sub>O, CH<sub>4</sub>, CF<sub>4</sub>, SF<sub>6</sub>, CO) and many more, e.g., CO<sub>2</sub>, OCS, NH<sub>3</sub>, HCN, C<sub>2</sub>H<sub>2</sub>, or C<sub>2</sub>H<sub>6</sub> [Stiller et al., 2001]. While only six species, the so-called key species, (H<sub>2</sub>O, O<sub>3</sub>, HNO<sub>3</sub>, CH<sub>4</sub>, N<sub>2</sub>O, NO<sub>2</sub>, along with temperature and pressure) are analyzed under ESA responsibility in an operational mode [Nett et al., 1999], analysis of many more scientifically interesting species is left to research groups. In order to fill this gap in the list of data products, various research groups have developed different innovative data analysis strategies.

[4] Different scientific goals of different research groups lead to different specifications of data analysis tools. The consequence is that most research groups use independent

<sup>1</sup>Institut für Meteorologie und Klimaforschung, Forschungszentrum Karlsruhe, Karlsruhe, Germany.

<sup>2</sup>Istituto di Fisica Applicata “Nello Carrara,” Florence, Italy.

<sup>3</sup>Deutsche Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, Wessling, Germany.

<sup>4</sup>Atmospheric, Oceanic and Planetary Physics, Oxford University, Oxford, UK.

<sup>5</sup>Instituto de Astrofísica de Andalucía, Consejo Superior de Investigaciones Científicas, Granada, Spain.

<sup>6</sup>Rutherford Appleton Laboratory, Chilton, Didcot, UK.

<sup>7</sup>Now at Analytical Services and Materials, NASA Langley Research Center, Hampton, Virginia, USA.

<sup>8</sup>Department of Physical and Inorganic Chemistry, Bologna University, Bologna, Italy.

<sup>9</sup>Institut für Meteorologie und Klimaforschung, Universität Karlsruhe, Karlsruhe, Germany.

and distinct retrieval software. While this situation has the advantage that each institute has a solution tailored to the problems to be solved in their predominant research work, this situation implies the risk that inconsistent data sets could be published as a result of diversity of data processors. This clearly would add difficulty to the scientific exploitation and would undermine the confidence of the scientific community in the data produced by the research groups. For this reason, in order to ensure quality and intercomparability of these independently generated value-added data products, a blind test retrieval experiment has been performed within the framework of the ‘‘Advanced MIPAS level-2 data analysis’’ (AMIL2DA) project, as the most realistic preflight data processor validation possible. Data processor in this context means the entity of computer hardware and software, including underlying analysis strategies and scientific algorithms, with a strong emphasis on the latter ones.

## 2. Retrieval Processors

[5] Five institutions contribute to AMIL2DA with their MIPAS data analysis processors, which will be used to retrieve atmospheric state parameters from calibrated geolocated spectra (level-2 processing). These institutions are the Institut für Meteorologie und Klimaforschung (IMK) with the Retrieval Control Program (IMK-RCP), which has been designed for scientific offline processing of MIPAS spectra, in particular with respect to species beyond ESA key species; Oxford University (OU) with the Oxford Processor To Invert MIPAS Observations (OPTIMO), which has been generated for the same purpose; Deutsches Zentrum für Luft- und Raumfahrt (DLR), operating two distinct processors (DLR-a, which is a workbench code for analysis of data from different instruments and measurement geometries, and which supports different retrieval approaches; and the D-PAC (German Processing and Archiving Center) processor, which has been installed for MIPAS offline data analysis); Istituto di Fisica Applicata ‘‘Nello Carrara’’ (IFAC) together with Bologna University with the Optimized Retrieval Algorithm (ORM), which is the prototype of the operational ESA MIPAS processing software; and Rutherford Appleton Laboratory (RAL) with RET2D, a general purpose processor for 1-D (single profile) or 2-D retrievals from mm wave and infrared instruments.

[6] To infer information on atmospheric state parameters from spectral radiance measurements requires the solution of the inverse radiative transfer problem. All retrieval processors discussed in this paper are based on minimization of a cost function:

$$\chi^2 = (\mathbf{y} - F(\hat{\mathbf{x}}))^T \mathbf{S}_y^{-1} (\mathbf{y} - F(\hat{\mathbf{x}})) + (\hat{\mathbf{x}} - \mathbf{x}_a)^T \mathbf{R} (\hat{\mathbf{x}} - \mathbf{x}_a), \quad (1)$$

where  $\hat{\mathbf{x}}$  is an estimate of the  $n$ -dimensional vector representing the atmospheric state parameters;  $\mathbf{x}_a$  is the a priori information of the atmospheric state;  $\mathbf{y}$  is the  $m$ -dimensional vector of real measurements, while  $F(\hat{\mathbf{x}})$  is the vector of measurements simulated by solution of the radiative transfer forward model  $F$  using  $\hat{\mathbf{x}}$  as input; the superscript  $T$  denotes transposed matrices.  $\mathbf{S}_y$  is the  $m \times m$  covariance matrix of the measurement, and  $\mathbf{R}$  is a user-defined regularization matrix of dimension  $n \times n$ . All

retrieval processors participating in this study minimize the cost function  $\chi^2$  in the context of a Newtonian iteration, where at the  $i$ th step the linear estimate of the parameter update is calculated by a variant of the following constrained least squares equation [see, e.g., Rodgers, 2000]:

$$\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_i + \left( \mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \mathbf{R} + \lambda \mathbf{I} \right)^{-1} \cdot \left( \mathbf{K}^T \mathbf{S}_y^{-1} (\mathbf{y} - F(\hat{\mathbf{x}}_i)) - \mathbf{R} (\hat{\mathbf{x}}_i - \mathbf{x}_a) \right). \quad (2)$$

$\mathbf{K}$  is the  $m \times n$  Jacobian matrix, containing the partial derivatives  $\frac{\partial F}{\partial x}$ ; typical settings for the regularization matrix  $\mathbf{R}$  are the inverse a priori covariance matrix ( $\mathbf{R} = \mathbf{S}_a^{-1}$ ) in the case of optimal estimation (IMK and D-PAC for tangent altitude retrievals (ZTAN); OPTIMO, RET-2D for all applications), or a Tikhonov-type smoothing term of the type  $\gamma \mathbf{B} \mathbf{L}_j \mathbf{L}_j^T \mathbf{B}^T$ , where  $\mathbf{L}_j$  is an  $n \times (n - j)$  matrix representation of the  $j$ th-order difference operator (IMK, DLR-a and D-PAC for volume mixing ratios,  $VMR$ ). Some schemes use weighted combinations of difference operators of different orders  $\mathbf{L}_0 \dots \mathbf{L}_j$  (DLR-a for  $O_3$ ).  $\mathbf{B}$  is a diagonal weighting matrix which can be used for altitude-dependent regularization. The strength of regularization is adjusted by a scaling factor  $\gamma$ .  $\lambda \mathbf{I}$  is a damping term as proposed by Levenberg [1944] and Marquardt [1963], where  $\lambda$  is typically a scalar, and  $\mathbf{I}$  is the  $n \times n$  unity matrix. It is hoped that after convergence, the estimated state of the atmosphere  $\hat{\mathbf{x}}$  is close to the discrete representation  $\mathbf{x}$  of the true state.

[7] While equation (1) is an appropriate tool for most retrieval applications, differences in implementation details may cause quite different results. Differences include, but are not limited to, aspects such as definition and discrete representation of the retrieval vector  $\mathbf{x}$ , choice of regularization term  $\mathbf{R}$ , details of setting and updating of the damping term  $\lambda \mathbf{I}$ , selection of an appropriate subset of measurements  $\mathbf{y}$  (usually small spectral regions containing major information on the target state parameter, so-called microwindows), a priori information  $\mathbf{x}_a$ , approaches to decompose the retrieval problem into subproblems, involving only subspaces of the measurement space and/or parameter space, numerical thresholds, e.g., for convergence, as well as the radiative transfer forward model  $F(\hat{\mathbf{x}})$ . The forward models used within the AMIL2DA consortium were cross validated in a separate study [von Clarmann et al., 2003a], hence this paper focuses on the different approaches to tackle the inverse problem. In the case of application to MIPAS, the retrieval vector  $\hat{\mathbf{x}}$  includes components of the atmospheric state vector (discrete vertical profiles of temperature  $T$ ,  $VMR$  of species, pressure  $p$ ), a frequency-dependent so-called ‘‘offset,’’ which is an altitude-independent radiometric zero-level calibration correction, and an altitude-dependent and frequency-dependent empirical continuum, which compensates in the retrieval for deficiencies in the forward modelling of aerosol emission, line-wing-induced continua etc. This continuum typically is characterized by one value per microwindow per tangent altitude but there are exceptions: The IMK code assigns one value per altitude to all microwindows within a  $5\text{-cm}^{-1}$  interval (‘‘clustering’’). The ORM constrains continuum values within a  $10\text{ cm}^{-1}$  region to be a linear function. Furthermore the IMK code allows

**Table 1.** Processor Setup, General

Characteristics	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM	RET2D
Processor reference	von Clarmann et al. [2001, 2003b]	RFM http://www.atm.ox.ac.uk/RFM	Doicu et al. [2002, 2003, 2004] MIRART	KOPRA Stiller et al. [2001]	Ridolfi et al. [2000] OFM Ridolfi et al. [2000]	FM2D Reburn et al. [2000]
Forward model	KOPRA					
Forward model reference	Stiller [2000]	http://www.atm.ox.ac.uk/RFM	Schreier and Schimpf [2001]	Stiller et al. [2001]	Ridolfi et al. [2000]	Reburn et al. [2000]
Retrieval grid	fixed	tangent altitude-dependent defined by tangent altitudes of measurements	tangent altitude-dependent defined by tangent altitudes of measurements	tangent altitude-dependent defined by tangent altitudes of measurements	tangent altitude-dependent defined by tangent altitudes of measurements	fixed
Retrieval grid width, km	<44: 1; 44–70: 2; >70: 5					<42: 3
Microwindows reference	Echle et al. [2000]	Dudhia et al. [2002]	Ridolfi et al. [2000]	Streck et al. [2000]	Dudhia et al. [2002]	
Spectral masks	yes	yes	no	yes	yes	no
Altitude-dependent microwindow selection	yes	yes	no	yes	yes	no
Continuum set to zero above, km	30	30	variable <sup>a</sup>	N/A	36	N/A
Continuum retrieval	clustering <sup>b</sup>	standard <sup>c</sup>	standard <sup>c</sup>	standard <sup>c</sup>	linear <sup>d</sup>	standard <sup>c</sup>
Zero-offset retrieval	standard <sup>c</sup>	standard <sup>c</sup>	standard <sup>c</sup>	standard <sup>c</sup>	standard <sup>c</sup>	standard <sup>c</sup>
Sequence of retrievals	1. T + ZTAN 2. H <sub>2</sub> O 3. O <sub>3</sub> 4. HNO <sub>3</sub> 5. CH <sub>4</sub> + N <sub>2</sub> O 6. NO <sub>2</sub>	1. T + log(p) 2. species in parallel ln(V/MR)	1. O <sub>3</sub> 2. CH <sub>4</sub> 3. HNO <sub>3</sub> 4. N <sub>2</sub> O 5. NO <sub>2</sub>	1. T + ZTAN 2. O <sub>3</sub> 3. H <sub>2</sub> O 4. CH <sub>4</sub> 5. N <sub>2</sub> O 6. HNO <sub>3</sub> 7. NO <sub>2</sub>	1. T + p 2. H <sub>2</sub> O 3. O <sub>3</sub> 4. HNO <sub>3</sub> 5. CH <sub>4</sub> 6. N <sub>2</sub> O 7. NO <sub>2</sub>	H <sub>2</sub> O, O <sub>3</sub> , CO <sub>2</sub> , N <sub>2</sub> O, CH <sub>4</sub> , NO <sub>2</sub> , and HNO <sub>3</sub> in parallel
Processing in tangent altitude domain	simultaneous	simultaneous	simultaneous	simultaneous	simultaneous	simultaneous
Processing in microwave domain	simultaneous	sequential	N/A	simultaneous	simultaneous	simultaneous

<sup>a</sup>Depends on target species.

<sup>b</sup>Microwindows within a 5 cm<sup>-1</sup> region are combined to a cluster. One continuum value per cluster and per altitude is retrieved.

<sup>c</sup>One continuum value per microwindow and per altitude is retrieved.

<sup>d</sup>One continuum value per microwindow and per altitude is retrieved, except that for all microwindows within a 10 cm<sup>-1</sup> region the continuum is assumed to be a linear function of wave number.

<sup>e</sup>One zero-offset value per microwindow is retrieved and applied to all altitudes of the limb scan.

**Table 2.** Processor Setup, Temperature, and Pressure or Pointing Retrieval

Characteristics	IMK-RCP	OPTIMO	D-PAC	ORM
Number of microwindows	22	10	17	10
Spectral masks	yes	yes	no	yes
Regularization $T$	$L_1$ -based	$S_a^{-1}$	$L_1$ -based	none
Regularization $ZTAN$ (relative)	$S_a^{-1}$ (engineering information)	$S_a^{-1}$ (engineering information)	$S_a^{-1}$ (engineering information)	hard hydrostatic constraint
Regularization $ZTAN$ (absolute)	$S_a^{-1}$ (engineering information)	hard constraint (lowest tangent altitude)	$S_a^{-1}$ (engineering information)	hard constraint (lowest tangent altitude)
Regularization $p$	hard hydrostatic constraint	hard hydrostatic constraint	hard hydrostatic constraint	none

smoothing of the continuum not only in the altitude domain but also in the wave number domain. All codes under assessment solve the multitangent altitude aspect of the inverse radiative transfer problem in a global fit sense [Carlotti, 1988], where measurements of all tangent altitudes of a limb scan are analyzed simultaneously. However, OPTIMO solves the inversion problem sequentially in the frequency domain, i.e., microwindow by microwindow. Some processors apply altitude-dependent spectral masks to their analysis microwindows; that is, they do not use all spectral grid points within a microwindow at each tangent altitude but ignore some of them which contain uncertain measurement information. Since neither the tangent altitudes nor the spacing between adjacent tangent altitudes (henceforth called “relative tangent altitudes”) are perfectly known for the MIPAS instrument, some of the processors (IMK-RCP, OPTIMO, D-PAC, ORM) are designed to also retrieve components of this pointing information from the measured spectra.

[8] Components of the atmospheric state vector which are not part of the retrieval vector, i.e., mixing ratios of interfering species (so-called contaminants) or, e.g., vibrational temperatures, contribute to the so-called systematic error budget (which, however, is not necessarily systematic in its strict meaning, since variation of these quantities may have random components in the temporal domain).

[9] The processor settings used for this study are compiled in Tables 1–8. In this paper, only the distinctive features and parameter settings as used in the blind test retrieval experiment are discussed. The main characteristics of the processors are summarized in Table 1. Most of the processors support applications beyond those assessed here. These, however, are discussed in this paper only if relevant to this study.

### 3. Blind Test Experiment

[10] The purpose of the blind test retrieval experiment was to prove that the participants’ codes can successfully retrieve temperature and volume mixing ratios of species from spectra as expected from the MIPAS measurement. This test was intended as a test under realistic conditions

rather than a mere sensitivity study. For this reason, the participants were provided with limited information, which was: time and location of the measurement, nominal measurement geometry, and MIPAS instrument specification. All other data, such as pressure-temperature model, abundances of species, vibrational temperatures, and exact measurement geometry were not known to the participants. Simulated measurements were generated by the Instituto de Astrofísica de Andalucía, who did not participate in the retrieval study with a retrieval processor of their own. This rigorous strategy was decided to guarantee realistic conditions. However, the following sources of error have been neglected throughout the blind test retrieval experiment: Instrument line shape errors, horizontal gradients of state parameters, scattering.

[11] A separate test case has been defined where pressure and temperature profiles as well as tangent altitude information were provided to the participants. This was to make participation possible for groups which plan not to retrieve these quantities but to rely on the operational ESA data product, and to isolate pressure/temperature propagation errors in the constituent retrieval. While the final goal of some participants is to retrieve data beyond the operational ESA data product, this retrieval experiment focuses on the six key species, since key species retrievals are a good example to demonstrate the capability of a retrieval processor. Key species can be treated by all groups, while handling of further species often needs more sophisticated retrieval approaches or forward models (e.g., consideration of non-local thermodynamic equilibrium). In this sense retrievals of key species are more readily intercomparable among the groups. Also, to improve the data quality of key species beyond the quality of the operational ESA data products is a further goal in itself.

#### 3.1. Simulated Measurements

[12] Pressure, temperature, and volume mixing ratio profiles were compiled for a polar summer scenario (80°N latitude, 12.00 h local time, 70° solar zenith angle) from the ground up to 200 km altitude. In order to provide a self-consistent data set, in particular for chemically active species, profiles of atmospheric state parameters were

**Table 3.** Processor Setup, H<sub>2</sub>O Retrieval

Characteristics	IMK-RCP	OPTIMO	D-PAC	ORM	RET2D
Number of microwindows	23	5	16	6	10 <sup>a</sup>
Regularization	$L_1$ -based; $B \propto 1/VMR$ [Steck, 2002]	$S_a^{-1}$	$L_1$ -based	none	$S_a^{-1}$

<sup>a</sup>Joint retrieval of H<sub>2</sub>O and O<sub>3</sub>.

**Table 4.** Processor Setup, O<sub>3</sub> Retrieval

Characteristics	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM	RET2D
Number of microwindows	16	4	1 <sup>a</sup>	12	3	10 <sup>b</sup>
Regularization	L <sub>1</sub> -based; $\mathbf{B} \propto 1/VMR$	S <sub>a</sub> <sup>-1</sup>	$\gamma(L_0) = 0.25$ ; $\gamma(L_1) = 0.0$ ; $\gamma(L_2) = 0.75$	L <sub>1</sub> -based	none	S <sub>a</sub> <sup>-1</sup>

<sup>a</sup>Microwindow 715.00–717.00 cm<sup>-1</sup>.

<sup>b</sup>Joint retrieval of H<sub>2</sub>O and O<sub>3</sub>.

generated by means of chemical transport models (CTMs) (Table 9). These profiles were sampled on a 1-km altitude grid, which is finer than the retrieval resolution, in order to simulate realistic conditions. A wave structure was superimposed to the modeled temperature profile below 100 km in order to generate some small-scale structures, and pressures were readjusted by means of the hydrostatic approximation. For altitudes where the available models did not provide enough information, volume mixing ratio profiles were extrapolated to reasonable values.

[13] Two different sets of spectral radiances covering all MIPAS channels were generated. The first set (henceforth Blind Test 1, BT1) was calculated for the tangent height of the nominal observation scenario, while for the second set (Blind Test 2, BT2) these tangent heights were perturbed within the expected pointing error (absolute error: –663 m, 1 $\sigma$  relative error between tangent altitudes: 136 m) of MIPAS. The rationale of BT2 was to investigate the retrievability of pointing, pressure and temperature, and to assess the propagation of related errors to species retrievals. The need to retrieve tangent altitude information directly from the spectra is discussed by *von Clarmann et al.* [2003b]. Tangent altitudes used for BT1 and BT2 are summarized in Table 10.

[14] All spectra were calculated with the KOPRA forward model [*Stiller et al.*, 2000]. The following gases have been included for line-by-line calculation, H<sub>2</sub>O, CO<sub>2</sub>, O<sub>3</sub>, N<sub>2</sub>O, CO, CH<sub>4</sub>, O<sub>2</sub>, NO, SO<sub>2</sub>, NO<sub>2</sub>, NH<sub>3</sub>, HNO<sub>3</sub>, OH, HF, HCl, HBr, ClO, OCS, H<sub>2</sub>CO, HOCl, N<sub>2</sub>, HCN, CH<sub>3</sub>Cl, H<sub>2</sub>O<sub>2</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>6</sub>, COF<sub>2</sub>, HO<sub>2</sub>. The contributions of the following species were calculated on the basis of empirical cross section spectra: CFC-11 (CCl<sub>3</sub>F), CFC-12 (CCl<sub>2</sub>F<sub>2</sub>), CFC-22 (CHClF<sub>2</sub>), SF<sub>6</sub>, CFC-14 (CF<sub>4</sub>), CCl<sub>4</sub>, CFC-13 (CCl<sub>2</sub>FCCl<sub>2</sub>), CFC-114 (CClF<sub>2</sub>CClF<sub>2</sub>), N<sub>2</sub>O<sub>5</sub>, ClONO<sub>2</sub>.

[15] The choice of KOPRA for generation of simulated measurements may be suspected to bias the blind test, since two of the retrieval processors under assessment are also based on KOPRA. However, it has been shown that spectra generated with KOPRA depend largely on the actual parameter settings such as discretization, truncation of far wings of spectral lines, or choice of submodels supported, just to name a few, while results of different forward models for consistent parameter settings are quite similar [*von Clarmann et al.*, 2003a]. Since these parameter settings were not known to the participants, and indeed differed largely between generation of reference spectra and retrieval applications, such a bias can be excluded.

[16] HITRAN2000 (<http://www.hitran.com>) spectroscopic data were used for all species, with the following exceptions: Updates from this HITRAN Web page were considered for CH<sub>4</sub>, O<sub>2</sub>, and C<sub>2</sub>H<sub>2</sub>. H<sub>2</sub>O, O<sub>3</sub>, HNO<sub>3</sub>, and HOCl line data were taken from the IROE HITRAN96

update [*Flaud et al.*, 1998, 2003; *Vander Auwera et al.*, 2000]. An updated version of NO<sub>2</sub> lines has been used.

[17] CO<sub>2</sub> Q branch line mixing was calculated with the direct diagonalization method using line mixing data from *Hartmann et al.* [1995]. Gas continua for N<sub>2</sub>, O<sub>2</sub>, CO<sub>2</sub> and H<sub>2</sub>O were included [*von Clarmann et al.*, 2003a, and references therein]. Volcanic aerosol continuum was calculated with the KOPRA-internal Mie model using a Pinatubo aerosol size distribution [*Deshler*, 1994] of which the number density was reduced by a factor of two in order to account for polar conditions, and H<sub>2</sub>SO<sub>4</sub> refraction indices from *Tisdale et al.* [1998]. Vibrational temperatures of CO<sub>2</sub>, H<sub>2</sub>O, O<sub>3</sub>, NO, NO<sub>2</sub>, N<sub>2</sub>O, HNO<sub>3</sub>, CO, and CH<sub>4</sub> were calculated for the previously defined atmospheric conditions by means of most recent nonlocal thermodynamic equilibrium (non-LTE) models and updated kinetic rates [*López-Puertas et al.*, 2002]. Care was taken to feed the non-LTE models for all species with consistent chemical and hydrostatic atmospheric input data. To achieve best possible accuracy, the non-LTE-driving tropospheric upwelling fluxes were included by means of line-by-line calculations. Non-LTE was considered for CO<sub>2</sub> (112 bands), H<sub>2</sub>O (11 bands), O<sub>3</sub> (224 bands), NO (12 bands), NO<sub>2</sub> (6 bands), N<sub>2</sub>O (53 bands), HNO<sub>3</sub> (7 bands), CO (1 band), OH (63 bands), and CH<sub>4</sub> (13 bands).

[18] A trapezium shape field-of-view (1.22 mrad base, 0.854 mrad top) was applied using a discretization of 7 pencil beams. A pretabulated instrumental line shape was used [*von Clarmann et al.*, 2003a]. Synthetic noise was superimposed to the spectra according to the noise equivalent spectral radiance (NESR) values summarized in Table 11. Finally, an altitude-independent radiance zero offset, which varies smoothly with frequency, was added to the spectra (Table 11). While we now know from analysis of real measurement data that the actual MIPAS offset is negligibly small, the option to retrieve radiance offsets was tested in this preflight study in order to be prepared for the worst case. Offset values found for real MIPAS measurements typically are about, e.g., 4.5 nW/(cm<sup>2</sup> sr cm<sup>-1</sup>), retrieved by IMK-RCP in band A, near 791 cm<sup>-1</sup>, for all available spectra measured from 18 September 2002 to 13 October 2002. These values are a factor of 22 better than the specified values which were used for this study.

**Table 5.** Processor Setup, HNO<sub>3</sub> Retrieval

Characteristics	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM
Number of microwindows	16	4	1 <sup>a</sup>	14	5
Regularization	L <sub>1</sub> -based; $\mathbf{B} \propto 1/VMR$	S <sub>a</sub> <sup>-1</sup>	identity	L <sub>1</sub> -based	none

<sup>a</sup>Microwindow 870.65–871.35 cm<sup>-1</sup>.

**Table 6.** Processor Setup, N<sub>2</sub>O Retrieval

Characteristics	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM
Number of microwindows	12 <sup>a</sup>	5	1 <sup>b</sup>	18	5
Regularization	L <sub>1</sub> -based; B ∝ 1/VMR	S <sub>a</sub> <sup>-1</sup>	L <sub>0</sub> -based and bound constraints	L <sub>1</sub> -based	none

<sup>a</sup>N<sub>2</sub>O-CH<sub>4</sub> joint fit.

<sup>b</sup>Microwindow 1270.82–1272.03 cm<sup>-1</sup>.

Similar small values were found for the ORM applied to Envisat orbit 2081, measured on 24 July 2002 [Raspollini *et al.*, 2002].

### 3.2. Error Estimation

[19] Error estimates were calculated in order to judge whether the deviations of the retrieved profiles from the reference profiles  $\hat{\mathbf{x}} - \mathbf{x}$  are understood or need some further investigation. The advantage of the strategy of a blind test experiment based on synthetic measurements is, contrary to the approach of using real measurements, that the true profile  $\mathbf{x}$  is known.

[20] A retrieved profile  $\hat{\mathbf{x}}$  can generally be represented as some smoothing of the “true” profile  $\mathbf{x}$  plus error terms [Rodgers, 2000]:

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \mathbf{G}(\delta\mathbf{y} - \delta\mathbf{f}), \quad (3)$$

where  $\mathbf{A} = \partial\hat{\mathbf{x}}/\partial\mathbf{x}$  is the averaging kernel matrix for the retrieval grid,  $\mathbf{G} = \partial\hat{\mathbf{x}}/\partial\mathbf{y}$  is the gain matrix representing the mapping of the measurements  $\mathbf{y}$  into the retrieval,  $\delta\mathbf{y}$  is the vector representing the error in each measurement,  $\delta\mathbf{f}$  is the vector representing the error in each corresponding forward model calculation.

[21] Rearranging this gives the difference between the retrieval and the “true” profile  $\mathbf{x}$ :

$$\hat{\mathbf{x}} - \mathbf{x} = (\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}_a) + \mathbf{G}(\delta\mathbf{y} - \delta\mathbf{f}). \quad (4)$$

[22] It is convenient to define  $\delta\mathbf{y}$  to contain just the random noise component of measurement error, while including correlated or persistent effects such as calibration errors in  $\delta\mathbf{f}$  along with the forward model errors. The random noise is routinely propagated through each retrieval and the covariance of  $\mathbf{G}\delta\mathbf{y}$  gives the random errors reported with each profile.

[23] The covariance of  $\mathbf{G}\delta\mathbf{f}$  represents the error component due to a less than perfect forward model and uncertainties in forward model input parameters which are not part of the retrieval vector  $\hat{\mathbf{x}}$ . Since these errors often are, according to various time/spatial scales, correlated between measurements or forward model calculations, they commonly are called

“systematic errors.” They have been estimated by assigning typical uncertainties to the driving parameters. The following systematic errors have been considered and evaluated for midlatitude day-time conditions: climatological 1 $\sigma$  variabilities of 28 species, errors due to assumption of local thermodynamic equilibrium applying at all altitudes, and due to ignoring CO<sub>2</sub> line-mixing effects. While some of the forward models used support both line coupling and non-LTE, these features have not been activated in any of the codes for this retrieval study.

[24] The estimate of the smoothing characteristics of a retrieval is based on the averaging kernel matrix. The averaging kernel for the retrieval grid can be calculated as

$$\mathbf{A} = (\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} + \mathbf{R})^{-1} \mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K} = \mathbf{GK}. \quad (5)$$

[25] The averaging kernels have been used to calculate the smoothing error component by

$$\hat{\mathbf{x}} - \mathbf{x} = (\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}_a). \quad (6)$$

The analysis of the smoothing error is necessary to judge if any deviation of the retrieved profile from the true profile can be explained by the reduced altitude resolution of the retrieval or any other mapping of a priori information on the retrieval.

## 4. Results

[26] IMK-RCP, OPTIMO, DLR-a and ORM results presented here are the original blind test results, while RET2D and D-PAC results are based on upgraded retrieval schemes.

### 4.1. Pressure, Temperature, and Tangent Altitudes

[27] Only the IMK-RCP, OPTIMO, D-PAC and ORM processors participated in the pressure, temperature and tangent-altitude retrieval intercomparison. While all processors retrieve the similar quantities, i.e., temperature and tangent altitudes, either in geometrical or pressure coordinates, the specification of the details of tangent altitude retrievals are quite different for the IMK-RCP, Oxford (OPTIMO) and IFAC (ORM) codes. All these approaches

**Table 7.** Processor Setup, CH<sub>4</sub> Retrieval

Characteristics	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM
Number of microwindows	12 <sup>a</sup>	7	1 <sup>b</sup>	14	5
Regularization	L <sub>1</sub> -based; B ∝ 1/VMR	S <sub>a</sub> <sup>-1</sup>	L <sub>0</sub> -based	L <sub>1</sub> -based	none

<sup>a</sup>N<sub>2</sub>O-CH<sub>4</sub> joint fit.

<sup>b</sup>Microwindow 1355.15–1356.65 cm<sup>-1</sup>.

**Table 8.** Processor Setup, NO<sub>2</sub> Retrieval

Characteristics	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM
Number of microwindows	15	8	1 <sup>a</sup>	15	5
Regularization	L <sub>1</sub> -based; B ∝ 1/VMR	S <sub>a</sub> <sup>-1</sup>	L <sub>1</sub> -based and bound constraints	L <sub>1</sub> -based	none

<sup>a</sup>Microwindow 1600.50–1601.40 cm<sup>-1</sup>.

**Table 9.** Volume Mixing Ratio Sources Used for Simulated Measurements

Data Source	Month	Latitude, °N	Atmospheric Parameter
2D CTM; <i>Garcia et al.</i> [1992]; <i>Garcia and Solomon</i> [1994]	July	79	temperature and all gases available from model below 112 km
MSIS high solar activity (F10.7 = 200, AP = 10); <i>Hedin</i> [1991]	May	85	N, O <sub>2</sub> , and O >112 km
2D CTM; <i>Gérard and Roble</i> [1988]; solar cycle maximum	June	80	NO >120 km
MIPAS reference atmosphere compilation (J. J. Remedios, personal communication, 2000)	April	80	CO <sub>2</sub> < 120 km
U.S. Standard Atmosphere [1976]			some minor constituents

use some basic ideas of the technique of *Gille and House* [1971] (use of CO<sub>2</sub> as tracer of atmospheric density; use of different spectral regions to make the multiparameter retrieval problem solvable). However, due to the particular characteristics of the MIPAS instrument (uncertainties of relative tangent altitudes, need to retrieve continua and possibly zero offset, unavailability of opaque spectral region suggested in their paper), some dedicated tangent altitude retrieval approaches had to be developed. Some of these (IMK-RCP and OPTIMO) take advantage of independent a priori information in order to improve results beyond the information content of the pure spectral measurement.

[28] The IMK-RCP attempts to retrieve both absolute and relative tangent altitudes [*von Clarmann et al.*, 2003b] but depends on a priori knowledge of a reference point where altitude and pressure are known (typically from European Centre for Medium-Range Weather Forecasts (ECMWF) analysis; climatological values for this study). This point is used to build up a hydrostatic atmosphere. In spite of a bad temperature/pressure a priori profile in this study, the knowledge of the absolute tangent altitude information is improved by a factor of two with respect to the a priori (Figure 1). Relative tangent altitude errors are in the order of a few tens of meters above 30 km altitude but increase toward larger values below.

[29] For tangent altitude and temperature retrievals, the D-PAC processor follows an approach similar to the IMK processor, where a reference point at a given altitude is used to obtain absolute tangent point altitudes. Differences in the detailed implementation and parameter settings, however, lead to different results.

**Table 10.** Tangent Heights

BT1, km	BT2, km
6.00	5.09
9.00	8.13
12.00	11.19
15.00	14.26
18.00	17.32
21.00	20.39
24.00	23.46
27.00	26.52
30.00	29.56
33.00	32.60
36.00	35.51
39.00	38.37
42.00	41.34
47.00	46.31
52.00	51.25
60.00	59.29
68.00	67.32

[30] The Oxford (OPTIMO) and IFAC (ORM) codes retrieve temperature and pressure (on a logarithmic scale in the case of OPTIMO) at tangent points and treat the lowest tangent point as fixed at 6 km, which is the nominal value of the lowest tangent height in the MIPAS standard measurement scenario. Remaining tangent point altitudes are calculated using the hydrostatic equation. In the case of OPTIMO and ORM, pointing uncertainties are included in the pressure a priori covariance matrix. By this means the relative tangent altitudes (i.e., differences between adjacent tangent altitudes) are recovered at an accuracy of ~120 m for tangent altitudes between 10 and 50 km. In this study, IFAC underestimates relative tangent altitude steps below 39 km and overestimates them above.

[31] All codes successfully recover the reference temperature profile (Figure 2). The random component of the temperature retrieval error is slightly below 1 K, while the total estimated retrieval error is between 1 and 2 K for most altitudes.

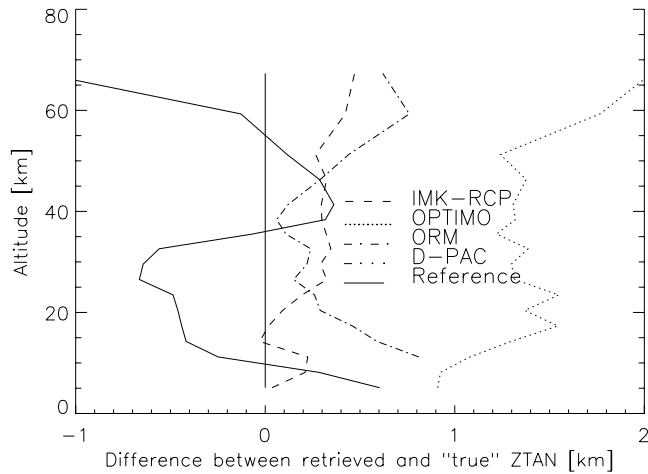
#### 4.2. Constituents, BT1

[32] Since not all participating groups use processors which support the retrieval of pressure and temperature, the discussion is focused on BT1, where retrievals were performed on the basis of known pressure, temperature and pointing information.

[33] Results for the six key species are shown in Figures 3–8. Error bars for random error due to measurement noise (solid line) and systematic errors, including all errors of the type Gδf (dotted line) are shown. The actual smoothing error was calculated for each data point according equation (6), which is possible in a blind test study, since the “true” state vector  $\mathbf{x}$  is known, contrary to real measurement applications where only a statistical assessment on the basis of assumed a priori covariances is possible. Since information on the sign of the smoothing error would be lost when adding this error component quadratically to the other parts of the error budget, we instead centered the error bars, which represent the noise-induced and systematic errors, around the smoothing-error-corrected retrieved data point. This helps to better

**Table 11.** Instrumental Noise (Unapodised) and Offset (1σ Deviation) for Each Channel

Channel	Wave Number Coverage, cm <sup>-1</sup>	NESR, nW/(cm <sup>2</sup> sr cm <sup>-1</sup> )	Offset, nW/(cm <sup>2</sup> sr cm <sup>-1</sup> )
A	685–970	30.0	100.0
AB	1020–1170	17.0	80.0
B	1215–1500	12.0	40.0
C	1570–1750	4.0	12.0
D	1820–2410	4.0	8.4



**Figure 1.** Results of blind test tangent altitude retrieval.

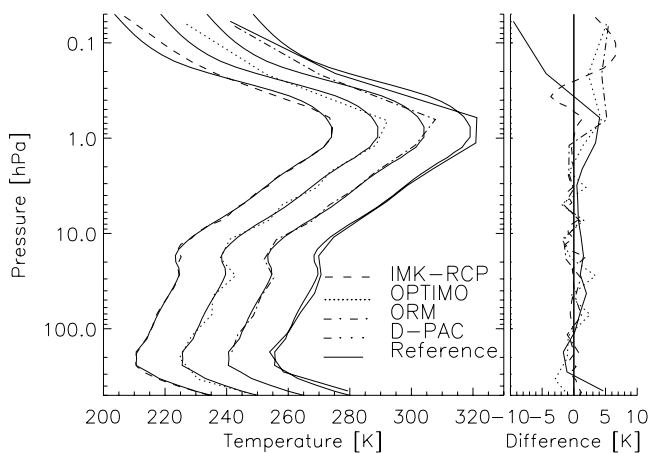
judge whether the difference between true data points and reference data points can be explained by one of the assessed error terms.

[34] As a proxy for the chi square of the retrieval, the normalized retrieval error

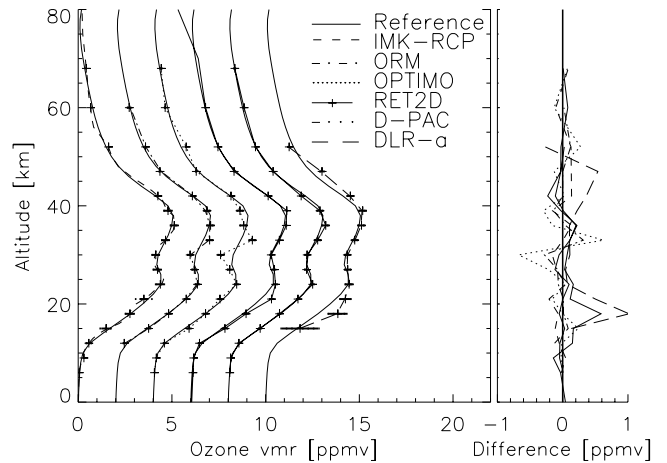
$$\sigma_{\text{norm}} = \left( \sum_{i=1}^n \frac{(\hat{x}_i - x_i)^2}{(\sigma_{\text{estimated}}(\hat{x}_i))^2} \right) / n, \quad (7)$$

where  $\hat{x}_i$  and  $x_i$  are the  $i$ th element of the vectors of the estimated (retrieved corrected by the smoothing error) and true target state parameters  $\hat{\mathbf{x}}$  and  $\mathbf{x}$ , respectively, and where  $\sigma_{\text{estimated}}(\hat{x}_i)$  is the related estimate of the total (random and systematic, excluding smoothing) retrieval error, was also computed (see Table 12). Unfortunately, since the off-diagonal elements of the retrieval error covariance matrices are not available for all data processors, the true chi square values could not be calculated.

[35] In most cases, the normalized retrieval errors  $\sigma_{\text{norm}}$  are in the order of 1, as expected. For some cases, in

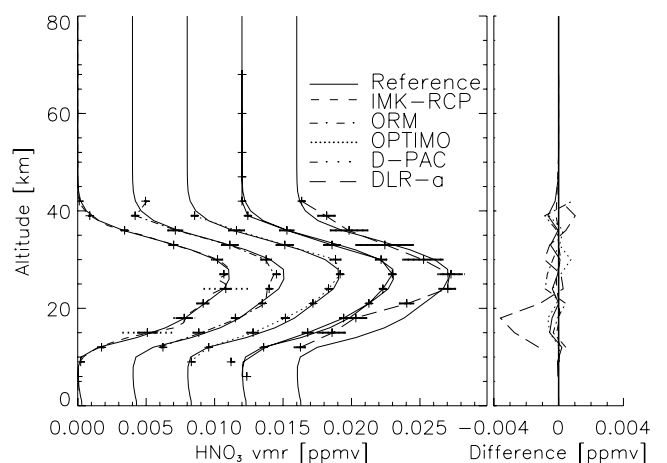


**Figure 2.** Results of blind test temperature retrieval. Results of different processors are shifted horizontally by 10 K each for a clearer representation. (left) The true temperature profile. (right) Differences between the retrieved profiles and the reference profile.

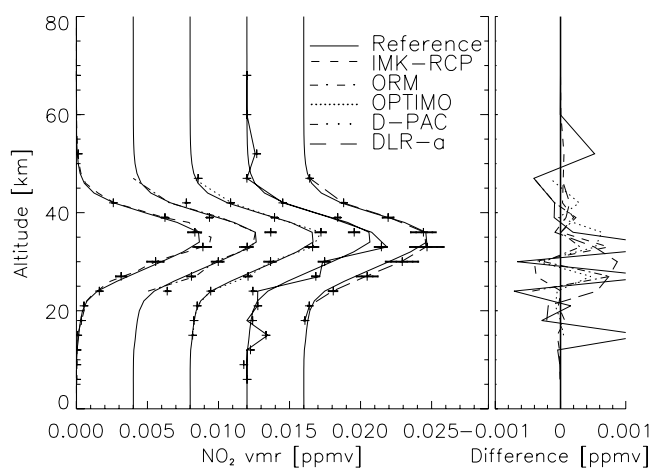


**Figure 3.** Retrieved ozone profiles, BT1. Error bars for random error (solid lines) and systematic errors (dotted lines) are included. The offset between the centers of the error bars and the retrieved data points represents the smoothing error. Results of different processors are shifted horizontally by 2 ppmv each for a clearer representation. (left) The true values. (right) Differences between the retrieved profiles and the reference profile.

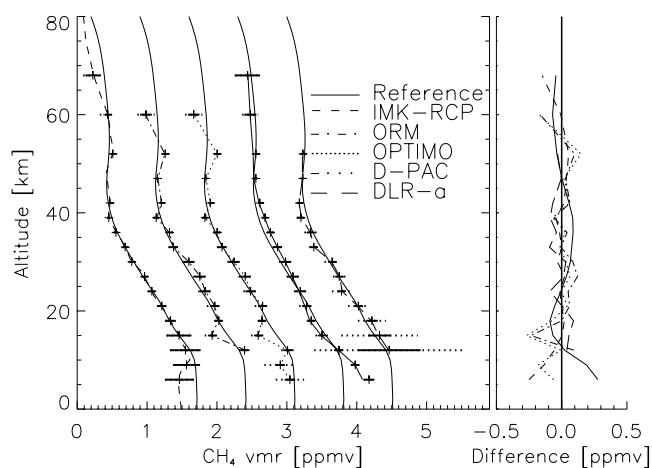
particular for IMK-RCP retrievals,  $\sigma_{\text{norm}}$  are below one, which is attributed to somewhat overpessimistic assumptions in the error estimation: (1) the standardized error estimation does not consider the individual regularization approach and thus sometimes overestimates the random error, and (2) it assumes parallel processing of all species where uncertainties of interfering species are determined by their climatological variability, while some of the processors retrieve species abundances in a predetermined sequence, where the uncertainty of interfering species retrieved in a preceding step is determined by the related retrieval error, which is smaller than climatological variability. Large values of  $\sigma_{\text{norm}}$  are often caused by large deviations between



**Figure 4.** Retrieved  $\text{HNO}_3$  profiles, BT1. Results of different processors are shifted horizontally by 0.004 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.



**Figure 5.** Retrieved  $\text{NO}_2$ , BT1. Results of different processors are shifted horizontally by 0.004 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.



**Figure 7.** Retrieved  $\text{CH}_4$ , BT1. Results of different processors are shifted horizontally by 0.7 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

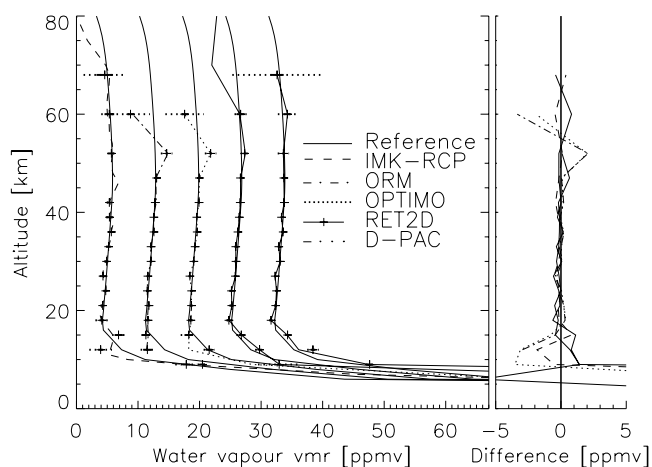
the retrieved and the “true” profile at very few single altitudes, which dominate the sum in equation (7), e.g., in the cases BT1,  $\text{H}_2\text{O}$ , lowermost altitude.

[36] Error bars (including random and systematic error) of ORM and OPTIMO retrievals are small because the assumptions made in their microwindow optimization were also used for the error estimation. Optimization of microwindows against other criteria naturally leads to increased estimated errors.

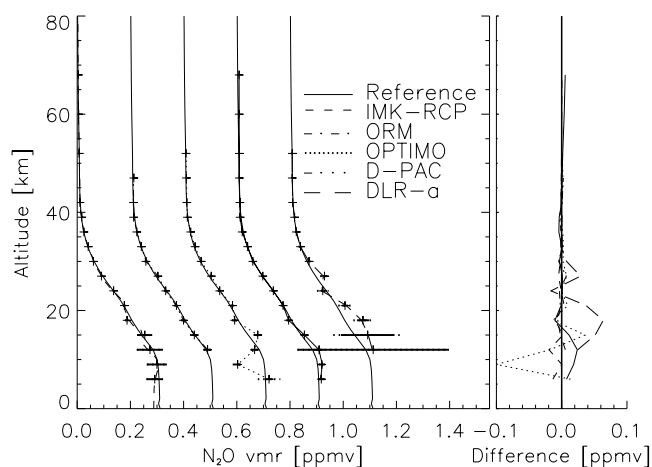
[37] The  $\text{O}_3$  and  $\text{HNO}_3$  profiles are well recovered by the IMK-RCP, ORM, and OPTIMO processors (Figures 3 and 4); this is true also for the RET2D and DLR-a  $\text{O}_3$  retrieval; however,  $\sigma_{\text{norm}}$  values larger than one for all processors suggest that retrieval errors may be somewhat underestimated, particularly for  $\text{HNO}_3$  (c.f. Table 12). Some overshooting at the double maximum of VMR of  $\text{O}_3$  is detected for the IMK-RCP, ORM and OPTIMO processors, and reaches an amplitude of about two stan-

dard deviations of the estimated total error for the latter one (Figure 3). The  $\text{HNO}_3$  maximum is somewhat underestimated by the ORM, while DLR-a underestimates  $\text{HNO}_3$  at low altitudes.

[38] For the third species with maximum VMR in the stratosphere,  $\text{NO}_2$ , the estimated errors explain the differences between retrieved and reference profiles well at least for IMK-RCP, OPTIMO and DLR-a (Figure 5);  $\sigma_{\text{norm}}$  for these three retrievals are around one (Table 12). The overestimated peak mixing ratio in the IMK-RCP profile could be explained by inappropriate microwindows in the MIPAS A band ( $757.925\text{--}762.725\text{ cm}^{-1}$ ,  $790.800\text{--}791.025\text{ cm}^{-1}$ , and  $825.750\text{--}825.950\text{ cm}^{-1}$ ), which carry information on  $\text{NO}_2$  in the lowermost altitudes only, while it is useless in the altitude range of interest. This microwindow will be avoided for future applications by not forcing the microwindow optimization program to minimize retrieval errors in the tropopause region.



**Figure 6.** Retrieved water vapor profiles, BT1. Results of different processors are shifted horizontally by 7.0 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.



**Figure 8.** Retrieved  $\text{N}_2\text{O}$ , BT1. Results of different processors are shifted horizontally by 0.2 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

**Table 12.** Normalized Retrieval Errors

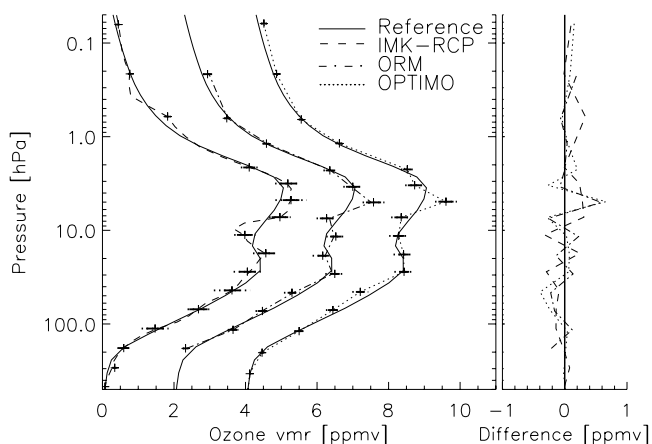
	IMK-RCP	OPTIMO	DLR-a	D-PAC	ORM	RET2D
O <sub>3</sub> (BT1)	1.42	4.52	2.83	0.76	1.11	4.20
HNO <sub>3</sub> (BT1)	2.21	4.90	3.72	215.4	11.2	
NO <sub>2</sub> (BT1)	0.98	1.27	1.01	52.6	18.1	
H <sub>2</sub> O (BT1) <sup>a</sup>	3.19	82.5		2469.		5.11
H <sub>2</sub> O (BT1) <sup>b</sup>	0.64	1.12		3.17	1.46	1.63
CH <sub>4</sub> (BT1)	0.74	2.01	0.70	3.06	1.30	
N <sub>2</sub> O (BT1)	2.11	3.09	1.04	1.25	2.46	
O <sub>3</sub> (BT2)	1.60	1.10			1.58	
HNO <sub>3</sub> (BT2)	8.10	31.9			2.80	
NO <sub>2</sub> (BT2)	1.29	4.77			8.24	
H <sub>2</sub> O (BT2) <sup>a</sup>	0.99	9.31				
H <sub>2</sub> O (BT1) <sup>b</sup>	0.87	1.65			1.80	
CH <sub>4</sub> (BT2)	0.65	1.97			1.08	
N <sub>2</sub> O (BT2)	0.56	1.43			1.15	

<sup>a</sup>With consideration of tropospheric values.

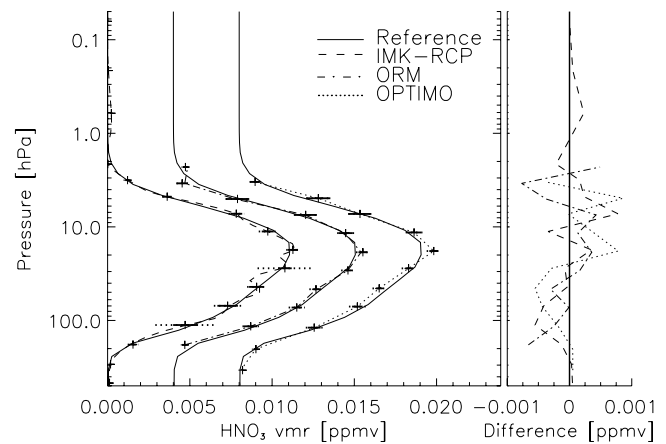
<sup>b</sup>Without consideration of tropospheric values.

[39] The different retrieval setups, in particular the different approaches to microwindow selection, evidently lead to differently characterized retrievals. Microwindow selection for the ORM and OPTIMO is based on a maximization of the information content and rigorous minimization of the computational effort [Rodgers, 1996; Dudhia et al., 2002]. These microwindows produce retrievals with tolerably small absolute errors. At altitudes where most of the information is in the measurement, i.e., in the region of the peaks of the VMR profiles, error bars are particularly small. The IMK microwindow selection used here aimed at minimization of the relative retrieval errors and leads to a good matching of retrieved profiles and reference profiles of species peaking in the stratosphere (e.g., O<sub>3</sub> and HNO<sub>3</sub>) at extreme (high and low) altitudes that is obtained at the cost of a larger number of microwindows. This leads to larger error bars of IMK retrievals near profile maxima and small error bars at the extreme ends. The choice of altitude-dependent regularization of IMK retrievals amplifies this effect.

[40] The water vapor profile is recovered well by IMK-RCP, ORM, OPTIMO, D-PAC and RET2D between 15 and 50 km altitude (Figure 6); all processors have, with the processor setup chosen, either problems in resolving the



**Figure 9.** Retrieved ozone profiles, BT2. Results of different processors are shifted horizontally by 2 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

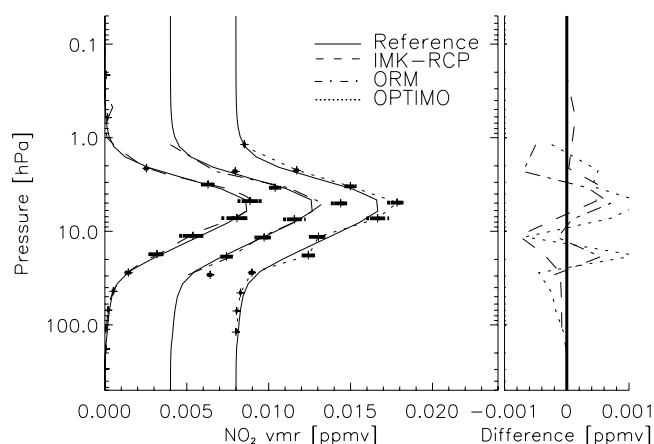


**Figure 10.** Retrieved HNO<sub>3</sub>, BT2. Results of different processors are shifted horizontally by 0.004 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

hyropause and/or in retrieving correct water vapor VMRs below. For altitudes above 50–60 km large differences between retrieved and reference profiles reflect non-LTE influences. These differences have been predicted by the analytical error estimation. Large errors for altitudes below 10 km have not been predicted, leading to large normalized errors (Table 12), which are dominated by the errors in the tropospheric part of the profile. Normalized errors calculated for altitudes above 10 km only are considerably smaller for all processors.

[41] The retrieved profiles of source gases CH<sub>4</sub> and N<sub>2</sub>O are very close to the reference profile for stratospheric altitudes (Figures 7 and 8). Above 50 km the reduced signal to noise of the spectra cause major deviations, while for altitudes around the tropopause the optically thick atmosphere, in particular due to strong aerosol loading of the reference atmosphere, causes large retrieval errors. Low sensitivities of spectral radiances with respect to the target parameters, i.e., low values of the Jacobian  $\mathbf{K}$ , destabilize the retrieval and lead to an increased random error component. The second mechanism by which the strong aerosol loading causes problems is that the empirical continuum, which is constrained in the frequency domain either explicitly by regularization or implicitly by linear parameterization or local flatness, does not perfectly compensate the continuum signal in the reference spectra, and thus leads to systematic errors, which are not part of the error budget reported here.

[42] For all species under investigation except H<sub>2</sub>O and all altitudes up to the middle atmosphere there is no evidence of any systematic non-LTE induced retrieval errors. Since for this study all retrievals were based on the assumption of LTE, while simulated measurements had been calculated for non-LTE, this indicates that non-LTE sensitive microwindows were successfully dropped by the microwindow selection schemes. Non-LTE induced errors were taken into account in the microwindow selection schemes on the basis of precomputed non-LTE populations [López-Puertas et al., 2002] and were also included in the error budget of this analysis using the second term in equation (3). This result proves for the species under consideration the robustness of the approach to avoid non-

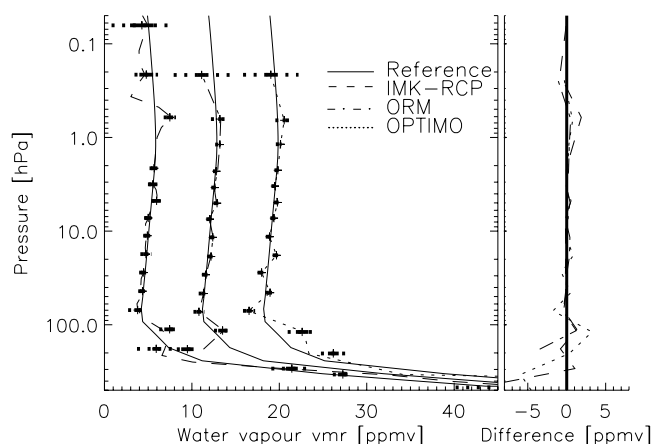


**Figure 11.** Retrieved  $\text{NO}_2$ , BT2. Results of different processors are shifted horizontally by 0.004 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

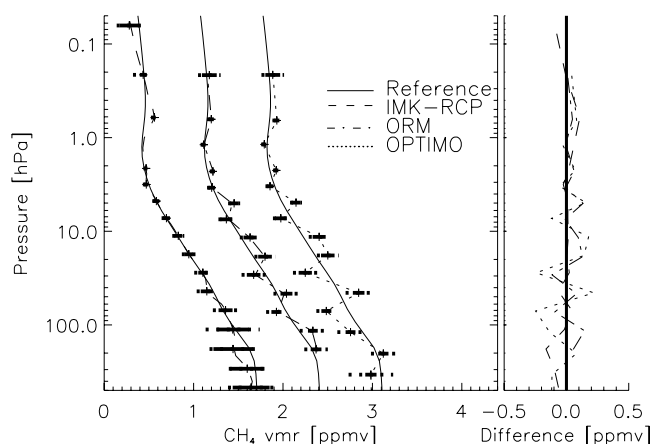
LTE in the retrieval rather than to consider it. Only for water vapor above 50 km is there some indication of non-LTE induced retrieval errors, which are the driving source of estimated so-called systematic retrieval errors.

#### 4.3. Constituents, BT2

[43] Error propagation from the temperature, pressure and/or tangent altitude retrieval to species retrievals amplifies some of the oscillations in retrieved profiles (Figures 9–14). This is true for species peaking in the stratosphere (e.g.,  $\text{O}_3$  and  $\text{HNO}_3$ ) as well as for source species such as  $\text{N}_2\text{O}$  and  $\text{CH}_4$ . The IMK-RCP BT2 water vapor retrieval (Figure 12) recovers the hygropause better than the one based on the reference information (BT1, Figure 6). The OPTIMO BT2 retrieval recovers the  $\text{HNO}_3$  peak better than the BT1 retrieval. Both these results are attributed to some cancellation of errors. In many cases the normalized errors in BT2 are smaller than the related ones in BT1. Also this result suggests that the use of the retrieved temperatures and instrument pointing information may be superior over the



**Figure 12.** Retrieved  $\text{H}_2\text{O}$ , BT2. Results of different processors are shifted horizontally by 7.0 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

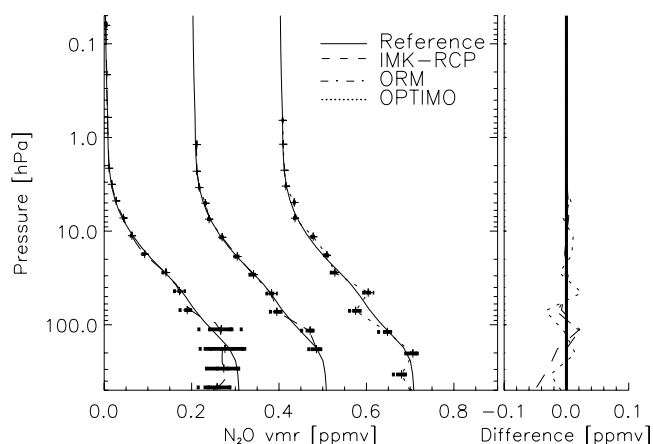


**Figure 13.** Retrieved  $\text{CH}_4$ , BT2. Results of different processors are shifted horizontally by 0.7 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

use of the true values due to error compensation. Generally, the BT2 retrievals are roughly consistent with the reference profiles within the estimated retrieval errors for most processors and species.

#### 4.4. Errors Not Considered in this Study

[44] Some uncertainties have not been included in the blind test experiment: Reference spectra were calculated for perfect instrument lineshape (ILS), perfect spectral and radiometric gain calibration, unperturbed spectroscopic data and horizontally homogeneous atmosphere. Many of the microwindows used by the different groups are particularly sensitive to horizontal structure of the atmosphere, and assumption of a horizontally homogeneous atmosphere was found to be the dominating component of systematic errors in some cases. IMK-RCP, however, is designed to reduce this effect to higher order by two dimensional radiative transfer forward calculations fed with information on 3-D fields of relevant state parameters which will be available from the operational MIPAS analysis and/or



**Figure 14.** Retrieved  $\text{N}_2\text{O}$ , BT2. Results of different processors are shifted horizontally by 0.2 ppmv each for a clearer representation. For a detailed explanation, see Figure 3.

meteorological analysis, and which are projected to the 2-D plane containing the line of sight. Also RET2D has this capability. In ORM and OPTIMO this error source is, in tendency, suppressed by taking this error source into account during microwindow optimization. In some cases, errors due to spectroscopic data uncertainties dominate the error budget (e.g., pressure retrievals with OPTIMO and ORM). ILS and spectral shift are the dominant error sources only in a few exceptional cases, e.g., for IMK-RCP temperature retrievals above 60 km. This processor, however, supports retrieval of ILS and spectral shift on a routine basis and thus it is not very dependent on reliable a priori information for these quantities.

## 5. Conclusions and Lessons Learned

[45] In summary, it can be stated that all processors under investigation are capable of retrieving profiles of atmospheric state parameters within the predicted retrieval errors relevant to the particular microwindow settings and regularization choices. This proves the functionality of all retrieval processors involved, even in a test which is assumed to be close to reality and which exceeds pure sensitivity studies by far. Effects of microwindow selection and regularization, which are the driving parameters of retrieval characteristics, are well understood. The participants were forced to work with initial guess and a priori information which often had very little to do with the true state of the atmosphere. However, all retrieval processors proved to be quite robust to the choice of this type of input data. In some cases, estimated error bars are quite large in comparison to the actual errors. In particular, errors of IMK-RCP profiles seem overestimated, which is confirmed for many cases by normalized errors, which often are below one and smaller than for the other processors. Large estimated random errors are explained by the fact that the standardized error estimation does not consider the regularization chosen, while large systematic errors are due to the fact that the climatological variability on which the error estimation is based is larger than the difference between the particular test case and climatology. The smoothing error components encountered in this study are negligibly small (below 1% of the squared sum of noise-induced and systematic errors in 94% of analyzed cases; driving error source, i.e., larger than the squared sum of noise-induced and systematic errors in only 0.7% of the analyzed cases, which refer to extreme high or low altitudes, where the measurements contains virtually no information. This means that none of the regularization or discrete sampling approaches chosen by the participating groups significantly distorts the retrieval as long as the truth is as smooth as the reference profiles used here, which, since calculated by chemical transport models for a particular situation rather than being taken from a climatology, are supposed to be fairly realistic. The most serious problem is the retrieval of the instrument pointing. While the absolute pointing is not in all cases necessary, knowledge of relative spacing of tangent altitudes is necessary for accurate species abundances retrievals. The strong aerosol loading of the reference atmosphere led to problems in retrieval of some source species ( $N_2O$  and  $CH_4$ ) at lowermost altitudes and gave the consortium a flavor of what the problems with real measure-

ments may be. In none of the cases under investigation was there any evidence of problems with respect to retrieval of zero offset or its propagation to the retrievals of target species.

[46] Lessons learned during this study are summarized as follows: In IMK-RCP some inappropriate microwindows have been removed by modification of the microwindow optimization specification such that less weight is put on extreme altitudes; furthermore, the dynamic range of matrix  $\mathbf{B}$  which controls the altitude dependence of regularization has been slightly decreased (These modifications have not been applied to results shown here, which are the original ones). Also, simultaneous retrieval of  $N_2O$  and  $CH_4$  on the basis of a dedicated joint fit microwindow selection turned out to be more accurate and efficient than sequential retrieval of both species. The additional continuum constraint in the wave number domain proved to stabilize the retrievals. The OPTIMO retrieval approach was refined with respect to atmospheric layering and the use of pointing a priori information. The RET2D results shown here are based on an updated retrieval code which preconditions the retrieval by scaling the retrieved products to allow for widely differing values of weighting functions. This is particularly relevant to retrieval of instrument parameters in conjunction with mixing ratio profiles where values may differ by many orders of magnitude. D-PAC have revised their processing strategy after the blind test experiment. Results shown here are the revised ones. Upgrades are related to the use of new microwindows qualified for retrievals down to 6 km tangent height with continuum contaminated spectra as well as several improvements with respect to the retrieval code, in particular a robust calculation of pointing corrections.

[47] This blind test experiment is regarded as a successful “readiness for real data test.” The advantage over a test with real measurements, which certainly would be even more realistic, is, that in the case of synthetic measurements the true state of the atmosphere is known.

[48] Certainly, one could think of even more realistic synthetic test cases: Measurements could be simulated for a horizontally structured atmosphere; a distorted instrument line shape could be used, as well as spectroscopic data statistically perturbed by their uncertainties. However, since the usefulness of such a blind test experiment is directly correlated with its timeliness, a line must be drawn somewhere, and not all possible improvements of the test setup justify further iterations, which would have become necessary to implement them.

[49] After having understood how MIPAS data processors behave when applied to synthetic but realistic spectra, the next step will be to compare results obtained with these processors when applied to real MIPAS measurement data. After geophysical validation of these data products, these will be made available to the scientific community via direct cooperation with the data-generating groups within the framework of approved ESA announcement of opportunity proposals.

[50] **Acknowledgments.** AMIL2DA is a Shared Cost Action within the RTD generic activities of the 5th FP EESD Programme of the European Commission, Project EVG1-CT-1999-00015. M.L.-P. has also been partially supported by PNE under contract PNE-017/2000-C. B.F. has been supported by a Marie Curie host fellowship. A.D. and F.S. have been partially supported

by HGF-Vernetzungsfonds "ENVISAT" under contract 01SF9954. The authors would like to thank Dr. Michel Schoupe for smooth and helpful interfacing with the European Commission.

## References

- Carlotti, M., Global-fit approach to the analysis of limb-scanning atmospheric measurements, *Appl. Opt.*, 27(15), 3250–3254, 1988.
- Deshler, T., In situ measurements of Pinatubo aerosol over Kiruna on four days between 18 January and 13 February 1992, *Geophys. Res. Lett.*, 21(13), 1323–1326, 1994.
- Doicu, A., F. Schreier, and M. Hess, Iteratively regularized Gauss-Newton method for atmospheric remote sensing, *Comput. Phys. Commun.*, 148, 214–226, 2002.
- Doicu, A., F. Schreier, and M. Hess, Iteratively regularized Gauss-Newton method for bound-constraint problems in atmospheric remote sensing, *Comput. Phys. Commun.*, 153, 59–65, 2003.
- Doicu, A., F. Schreier, and M. Hess, Iterative regularization methods for atmospheric remote sensing, *J. Quant. Spectrosc. Radiat. Transfer*, 83, 47–61, doi:10.1016/S0022-4073(02)00292-3, 2004.
- Dudhia, A., V. L. Jay, and C. D. Rodgers, Microwave selection for high-spectral-resolution sounders, *Appl. Opt.*, 41(18), 3665–3673, 2002.
- Echle, G., T. von Clarmann, A. Dudhia, J.-M. Flaud, B. Funke, N. Glatthor, B. Kerridge, M. López-Puertas, F. J. Martín-Torres, and G. P. Stiller, Optimized spectral microwindows for data analysis of the Michelson Interferometer for Passive Atmospheric Sounding on the Environmental Satellite, *Appl. Opt.*, 39(30), 5531–5540, 2000.
- Endemann, M., and H. Fischer, Envisat's high-resolution limb sounder: MIPAS, *ESA Bull.*, 76, 47–52, 1993.
- Fischer, H., and H. Oelhaf, Remote sensing of vertical profiles of atmospheric trace constituents with MIPAS limb emission spectrometers, in *Optical Remote Sensing of the Atmosphere and Clouds, Beijing, China, 15–17 September 1998*, vol. 3501, edited by J. Wang et al., pp. 42–46, Soc. of Photo-Opt. Instrum. Eng., Bellingham, Wash., 1998.
- Flaud, J.-M., M. Birk, G. Wagner, J. Orphal, S. Klee, and W. J. Lafferty, The far infrared spectrum of HOCl: Line positions and intensities, *J. Mol. Spectrosc.*, 191(2), 362–367, 1998.
- Flaud, J.-M., C. Piccolo, B. Carli, A. Perrin, L. H. Coudert, J.-L. Teffo, and L. R. Brown, Molecular line parameters for the MIPAS (Michelson Interferometer for Passive Atmospheric Sounding) experiment, *Atmos. Oceanic Opt.*, 16, 172–182, 2003.
- Garcia, R. R., and S. Solomon, A new numerical model of the middle atmosphere: 2. Ozone and related species, *J. Geophys. Res.*, 99, 12,937–12,951, 1994.
- Garcia, R. R., F. Stordal, S. Solomon, and J. T. Kiehl, A new numerical model of the middle atmosphere: 1. Dynamics and transport of tropospheric source gases, *J. Geophys. Res.*, 97, 12,967–12,991, 1992.
- Gérard, J.-C., and R. G. Roble, The role of nitric oxide on the zonally averaged structure of the thermosphere: Solstice conditions for solar cycle maximum, *Planet. Space Sci.*, 36, 271–279, 1988.
- Gille, J. C., and F. B. House, On the inversion of limb radiance measurements I: Temperature and thickness, *J. Atmos. Sci.*, 28, 1427–1442, 1971.
- Hartmann, J.-M., C. Boulet, M. Margottin-Maclou, F. Rachet, B. Khalil, F. Thibault, and J. Boissoles, Simple modelling of Q-branch absorption, I. Theoretical model and application to CO<sub>2</sub> and N<sub>2</sub>O, *J. Quant. Spectrosc. Radiat. Transfer*, 54(4), 705–722, 1995.
- Hedin, A. E., Extension of the MSIS thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, 96, 1159–1172, 1991.
- Levenberg, K., A method for the solution of certain non-linear problems in least squares, *Quart. Appl. Math.*, 2, 164–168, 1944.
- López-Puertas, M., B. Funke, M. A. López-Valverde, F. J. Martín-Torres, T. von Clarmann, G. P. Stiller, H. Oelhaf, H. Fischer, and J. M. Flaud, Non-LTE studies for the analysis of MIPAS-ENVISAT data, in *Remote Sensing of Clouds and the Atmosphere VI*, vol. 4539, edited by K. Schäfer et al., pp. 381–395, Soc. of Photo-Opt. Instrum. Eng., Bellingham, Wash., 2002.
- Marquardt, D. W., An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Indust. Appl. Math.*, 11(2), 431–441, 1963.
- Nett, H., B. Carli, M. Carlotti, A. Dudhia, H. Fischer, J.-M. Flaud, G. Perron, P. Raspollini, and M. Ridolfi, MIPAS ground processor and data products, in *Proceedings of the IEEE 1999 International Geoscience and Remote Sensing Symposium, 28 June–2 July 1999, Hamburg, Germany*, pp. 1692–1696, Inst. of Electr. and Electron. Eng., Piscataway, N. J., 1999.
- Raspollini, P., D. Alpaslan, B. Carli, M. Carlotti, E. Castelli, S. Ceccherini, B. M. Dinelli, L. Magnani, M. Prosperi, and M. Ridolfi, MIPAS instrument and level 1 verifications using level 2 retrieval code, in *Proceedings of the ENVISAT Validation Workshop ESRIN, 9–13 November 2002* [CD-ROM], Rep. SP-531, Eur. Space Agency, Noordwijk, Netherlands, 2002.
- Reburn, W. J., R. Siddans, B. J. Kerridge, S. Bühler, A. von Engeln, P. Eriksson, T. Kuhn-Sander, K. Künzi, and C. Verdes, Critical assessments in millimetre-wave atmospheric limb sounding, *Final Rep. 13348/98/NL/GD*, part V, Eur. Space Agency, Noordwijk, Netherlands, 2000.
- Ridolfi, M., et al., Optimized forward and retrieval scheme for MIPAS near-real-time data processing, *Appl. Opt.*, 39(8), 1323–1340, 2000.
- Rodgers, C. D., Information content and optimisation of high spectral resolution measurements, in *Optical and Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II*, vol. 2830, edited by P. B. Hays and J. Wang, pp. 136–147, Soc. of Photo-Opt. Instrum. Eng., Bellingham, Wash., 1996.
- Rodgers, C. D., *Inverse Methods for Atmospheric Sounding: Theory and Practice*, Ser. Atmos. Oceanic Planet. Phys., vol. 2, edited by F. W. Taylor, World Sci., River Edge, N. J., 2000.
- Schreier, F., and B. Schimpf, A new efficient line-by-line code for high resolution atmospheric radiation computations incl. derivatives, in *IRS 2000: Current Problems in Atmospheric Radiation*, edited by W. L. Smith and Y. M. Timofeyev, pp. 381–384, A. Deepak Publ., Hampton, Va., 2001.
- Steck, T., Bestimmung von Vertikalprofilen von Spurengasen aus MIPAS—Messungen unter Hinzunahme von a priori Wissen, Ph.D. thesis, Inst. für Meteor. und Klimaforsch., Univ. Karlsruhe, Kernforschungszent. Karlsruhe, Germany, 2000.
- Steck, T., Methods for determining regularization for atmospheric retrieval problems, *Appl. Opt.*, 41(9), 1788–1797, 2002.
- Stiller, G. P., (Ed.), *The Karlsruhe Optimized and Precise Radiative Transfer Algorithm (KOPRA)*, Wissenschaft. Ber., vol. FZKA 6487, Forschungszent. Karlsruhe, Germany, 2000.
- Stiller, G. P., N. Glatthor, S. Kellmann, E. Kimmich, A. Linden, M. Milz, and H. Fischer, MIPAS on ENVISAT as an UT/LS sounder, in *IRS 2000: Current Problems in Atmospheric Radiation*, edited by W. L. Smith and Y. M. Timofeyev, pp. 839–842, A. Deepak Publ., Hampton, Va., 2001.
- Tisdale, R. T., D. L. Glandorf, M. A. Tolbert, and O. B. Toon, Infrared optical constants of low-temperature H<sub>2</sub>SO<sub>4</sub> solutions representative of stratospheric sulfate aerosols, *J. Geophys. Res.*, 103, 25,353–25,370, 1998.
- Vander Auwera, J., J. Kleffmann, J.-M. Flaud, G. Pawelek, H. Bürger, D. Hurtmans, and R. Pétrisse, Absolute v<sub>2</sub> line intensities of HOCl by simultaneous measurements in the infrared with a tunable diode laser and far-infrared region using a Fourier transform spectrometer, *J. Mol. Spectrosc.*, 204(1), 36–47, 2000.
- von Clarmann, T., H. Fischer, B. Funke, N. Glatthor, U. Grabowski, M. Höpfner, M. Kiefer, F. J. Martín-Torres, M. Milz, and G. P. Stiller, MIPAS interactive semi-operational level-2 data processing, in *IRS 2000: Current Problems in Atmospheric Radiation*, edited by W. L. Smith and Y. M. Timofeyev, pp. 785–788, A. Deepak Publ., Hampton, Va., 2001.
- von Clarmann, T., et al., Modelling of atmospheric mid-infrared radiative transfer: The AMIL2DA algorithm intercomparison experiment, *J. Quant. Spectrosc. Radiat. Transfer*, 78(3–4), 381–407, doi:10.1016/S0022-4073(02)00262-5, 2003a.
- von Clarmann, T., et al., Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), *J. Geophys. Res.*, doi:10.1029/2003JD003602, in press, 2003b.
- S. Ceccherini, Istituto di Fisica Applicata "Nello Carrara," Via Panciatichi 64, I-50127 Firenze, Italy. (s.ceccherini@ifac.cnr.it)
- A. Doicu, S. Hilgers, F. Schreier, and G. Schwarz, Deutsche Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, D-82230 Wessling, Germany. (adrian.doicu@dlr.de; siegfried.hilgers@dlr.de; franz.schreier@dlr.de; gottfried.schwarz@dlr.de)
- A. Dudhia and V. Payne, Atmospheric, Oceanic and Planetary Physics, Clarendon Laboratory, Oxford University, Parks Road, Oxford OX1 3PU, UK. (dudhia@atm.ox.ac.uk; payne@atm.ox.ac.uk)
- B. Funke and M. López-Puertas, Instituto de Astrofísica de Andalucía, Consejo Superior de Investigaciones Científicas, Apartado Postal 3004, E-18080 Granada, Spain. (bernd@iaa.es; puertas@iaa.es)
- V. Jay, J. Reburn, and R. Siddans, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK. (v.l.jay@rl.ac.uk; w.j.reburn@rl.ac.uk; r.siddans@rl.ac.uk)
- F.-J. Martín-Torres, Analytical Services and Materials, MS 936, NASA Langley Research Center, Hampton, VA 23681-2199, USA. (f.martin-torres@larc.nasa.gov)
- M. Ridolfi, Department of Physical and Inorganic Chemistry, Bologna University, Via Risorgimento 4, I-40136 Bologna, Italy. (ridolfi@ms.fci.unibo.it)
- T. Steck, Universität Karlsruhe, Institut für Meteorologie und Klimaforschung, Postfach 3640, D-76021 Karlsruhe, Germany. (tilman.steck@imk.fzk.de)
- T. von Clarmann and A. Linden, Forschungszentrum Karlsruhe, Institut für Meteorologie und Klimaforschung, Postfach 3640, D-76021 Karlsruhe, Germany. (thomas.clarmann@imk.fzk.de; andrea.linden@imk.fzk.de)