

Assimilation of MIPAS observations using a three-dimensional global chemistry-transport model

By F. BAIER^{1*}, T. ERBERTSEDER¹, O. MORGENSTERN², M. BITTNER¹ and G. BRASSEUR³

¹*Deutsches Zentrum für Luft- und Raumfahrt, Deutsches Fernerkundungsdatenzentrum, Germany*

²*Centre for Atmospheric Chemistry, Cambridge University, UK*

³*Max-Planck-Institut für Meteorologie, Hamburg, Germany*

(Received 23 May 2005; revised 31 October 2005)

SUMMARY

MIPAS observations are assimilated using a modified version of the chemistry-transport model ROSE to derive consistent global chemical analyses of the stratosphere. Due to different retrieval schemes applied, available MIPAS datasets are expected to differ in quality and coverage. In this study we investigate the sensitivity of the data assimilation scheme to two different sample datasets. ENVISAT/MIPAS baseline observations of H₂O, O₃, HNO₃, CH₄, N₂O and NO₂, covering October–November 2003, are considered. Sequential assimilation is performed using an optimum interpolation scheme with error propagation. It is shown that all assimilated model species benefit significantly from observations. Results are analysed using observation minus first-guess error statistics and are additionally compared to UARS/HALOE data. Optimized assimilation parameters are derived using χ^2 diagnostics. Two different MIPAS data products are examined: the European Space Agency operational product and the Institute for Meteorology Karlsruhe (IMK) scientific product. Assimilation results show some significant differences with respect to the dataset type used. For example, regions with increased stratospheric H₂O concentrations near the tropical tropopause are only present when IMK data are applied. Both datasets are found to be well suited for global assimilation experiments to study the chemistry and dynamics of the middle atmosphere.

KEYWORDS: Chemical data assimilation Error analysis Satellite observations Stratospheric chemistry

1. INTRODUCTION

In order to derive consistent global three-dimensional (3D) chemical analyses from asynchronous and inhomogeneously distributed remote-sensing observations, sequential assimilation into chemistry-transport models (CTMs) has been successfully demonstrated by several former studies (e.g. Khattatov *et al.* 2000; Chipperfield *et al.* 2002; El Amraoui *et al.* 2004). Giving global coverage within several days, instruments on polar orbiting satellites are in general well suited for providing substantial information on the variability of chemical species within the stratosphere. To further analyse atmospheric processes (e.g. ozone depletion), optimal combination of models and observations is essential.

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS; Fischer and Oelhaf 1996) was launched aboard the European Environmental Satellite (ENVISAT) in March 2002. It performs global limb measurements in the near- to mid-infrared and allows retrieval of temperature, pressure and trace-gas profiles of the middle atmosphere and upper troposphere (<http://envisat.esa.int/instruments/mipas>). Besides the operational European Space Agency (ESA) standard products (Carli *et al.* 2004) that focus on rapid processing of all observations, scientific off-line products are available. Depending on the trace-gas retrieval process applied, they show significant differences in quality and coverage. For this study we use data from the Institute for Meteorology Karlsruhe (IMK; Glatthor *et al.* 2005) for comparison with the standard ESA product.

The quality of data assimilation results strongly depends on the chemical characteristics of species observed (e.g. chemical time constants). The final analysis will

* Corresponding author: German Aerospace Centre, D-82234 Wessling, PO Box 1116, Germany.
e-mail: Frank.Baier@dlr.de

not be consistent if inadequate information is assimilated, e.g. if key species or error parameters are missing. In particular, information on the errors of remote-sensing data is often not sufficient for optimal assimilation. Therefore empirical adjustments and parametrizations are prerequisites (e.g. Khattatov *et al.* 2000).

Unlike variational assimilation schemes (e.g. Errera and Fonteyn 2001), the sequential method corrects the model first guess whenever observations are available. For a rigorous error treatment, giving non-biased statistically optimum results, Kalman-filter-based methods are principally well suited (Ménard *et al.* 2000). However, to reduce computational costs, approximations to the original formulation are necessary (Khattatov *et al.* 2000). Based on the CTM ROSE/DLR, we use a sequential assimilation scheme with optimum interpolation of first guess minus observation residuals and propagation of model errors (i.e. only the diagonal elements of the fully-fledged background covariance matrix are processed). Isotropic background correlations are parametrized using constant horizontal scale lengths (e.g. Riishøjgaard 1998). Vertical covariances are calculated from the averaging kernels used for the retrieval process and delivered with the MIPAS datasets.

This paper is structured as follows: we will first give a brief description of the different MIPAS datasets used in this study. The current version of the CTM ROSE/DLR is outlined in section 3. The sequential assimilation scheme is described in section 4 with a focus on the set-up of assimilation parameters. (More details are given in the appendix.) Section 5 follows with a discussion of results using χ^2 and observation minus first guess (OMF) error diagnostics, a comparison of assimilation results with Halogen Occultation Experiment (HALOE) observations, and our findings on typical zonal mean distributions of H₂O, O₃ and CH₄ for November 2003 are discussed. We close with a brief discussion (section 6).

2. DATA

The MIPAS instrument is a high-resolution atmospheric limb sounder aboard ESA's ENVISAT launched in March 2002 and operating in a sun-synchronous orbit. It aims at global and simultaneous measurements of the chemical composition of the middle atmosphere and upper troposphere. By means of sounding the earth's limb in the mid-infrared, it gathers emission spectra during night and daytime conditions. Nearly global coverage is achieved within three days. Based on the analysis of emission features, profiles of temperature, pressure and a multitude of trace gases can be obtained with a maximum vertical resolution of 3 km and a horizontal averaging along the line of sight of approximately 400 km (Stiller *et al.* 2002). For a detailed description of the MIPAS trace-gas retrieval as applied by ESA and IMK, the reader is referred to Ridolfi *et al.* (2000), Steck and von Clarmann (2001), Carli *et al.* (2004), and references therein. Since the operational start of MIPAS, extensive calibration and validation campaigns have been carried out on emission spectra and trace-gas profiles (Snoeij *et al.* 2004 and references therein). For example, comparisons of MIPAS observations to collocated HALOE and Stratospheric Aerosol and Gas Experiment (SAGE II) measurements (Bracher *et al.* 2004) show in general good correspondence for the middle stratosphere, with higher deviations below 20 km altitude. Typical r.m.s. errors for O₃ are found to be within 5 and 15%. Corresponding errors for H₂O, CH₄ and NO₂ are considerably higher with maximum values between 25% and 35%. Significant biases are found for all standard species, mostly positive about 10%. For the ESA operational products, a significant scatter is present in the O₃, CH₄ and N₂O profiles. MIPAS profiles seem to degrade rapidly below 15 km altitude.

For this study we use the IMK dataset MIPAS-E v2 (hereafter MIMK) for October and November 2003 (see Glatthor *et al.* 2005 for O₃, CH₄, N₂O; Funke *et al.* 2005 for NO₂; Milz personal communication for H₂O; Mengistu Tsidu *et al.* 2005 for HNO₃) and the corresponding ESA operational level 2 dataset v4.61 (hereafter MESA; Ceccherini and Ridolfi 2002; Carli *et al.* 2004; Ceccherini 2004). For comparability reasons, we use only the baseline observations of O₃, H₂O, CH₄, N₂O, NO₂ and HNO₃. It should be mentioned that profiles of NO, ClONO₂ and N₂O₅ are also available from IMK. For this study only stratospheric data are used. The stratosphere is therefore defined by potential vorticity values greater than 2 PVU (1 PVU = 10⁻⁶ K m²kg⁻¹s⁻¹) or potential temperature above 380 K.

While MIMK observations are more homogeneously distributed with respect to latitude and altitude, corresponding MESA observations show a better coverage of the upper stratosphere and lower mesosphere. The datasets also differ in temporal coverage. MESA observations cover the whole time period considered, with some days missing mainly during October 2003. For MIMK data there are two short gaps in October and November. There are no MIMK data available at all after 12 November 2003.

Both datasets give a good coverage of the middle atmosphere between 10 and 60 km altitude, which is the focus of this paper. MIMK-retrieved trace-gas profiles use a fixed vertical grid with 1 km step size. Each species' profile is provided with its own height-dependent vertical resolution. MESA retrievals are characterized by a variable vertical step size between 3 and 6 km. The vertical resolution of MESA profiles can be derived from the averaging kernel matrix accompanying the data, which depends on species, latitude, height and season. Depending on the species considered, the vertical resolution of MESA data shows in general a weaker variability than the corresponding MIMK resolution. MIMK CH₄ and N₂O resolutions are significantly reduced in the lower stratosphere, while the MESA N₂O vertical resolution is reduced in the upper stratosphere and lower mesosphere. The MIMK H₂O vertical resolution decreases particularly within the lower tropical stratosphere, where the corresponding MESA H₂O resolution shows also a reduction, but to a lesser extent.

3. CHEMISTRY-TRANSPORT MODEL ROSE/DLR

For the sequential assimilation of MIPAS observations, as described in section 4, the DLR (German Aerospace Centre) version of the 3D global CTM NCAR-ROSE is applied (hereafter ROSE/DLR). The original model is described in detail in Rose and Brasseur (1989) and Granier and Brasseur (1991). ROSE/DLR is based on a modified and improved version of Marsh *et al.* (2001) with focus on stratospheric chemistry. The model covers the relevant stratospheric gas-phase chemical processes and heterogeneous processes on sulphuric acid aerosols. Heterogeneous processes on polar stratospheric clouds (PSCs) are calculated using an Ice-NAT* scheme as in Chipperfield (1999).

The chemical rate constants and cross-sections are taken from Sander *et al.* (2003). Photolysis rates are derived from a look-up table depending on zenith angle, the ozone column above and altitude. The chemical rate equations are solved considering the reactivity of stratospheric species. The rather long-lived ones, e.g. HNO₃ and N₂O₅, are treated with a semi-implicit Gauss–Seidel solver. The fast-reacting species within the HO_x, NO_x, ClO_x, O_x and BrO_x families are solved fully implicitly using Newton–Raphson iteration. O(1D) is determined by considering the photochemical equilibrium state for the very short-lived species (e.g. ClO, NO, HO, BrO). The basic chemical time step is one hour. It is reduced adaptively to fulfil convergence.

* Nitric acid trihydrate.

All species are transported using the Lin and Rood finite-volume advection scheme (Lin and Rood 1996). Wind and temperature fields are derived from 24-hour UK Met Office analyses which are available from ground up to 0.1 hPa (Swinbank and O'Neill 1994). This dataset defines a consistent synoptic state using satellite-based temperature soundings and radiosonde observations assimilated in a global circulation model.

For this study ROSE/DLR consists of a $3.75^\circ \times 2.5^\circ$ longitude–latitude spherical grid on 43 log-pressure levels between 0 and approximately 56 km altitude, resulting in a vertical step size of 1.3 km. Within the model's troposphere, ozone is relaxed to 2D SOCRATES data (Brasseur *et al.* 1995) with a time-scale of 10 days. The tropopause is therefore defined by a potential vorticity of 2 PVU or a potential temperature of 380 K for the lower latitudes. All other species are prescribed by SOCRATES values at the model's upper and lower boundaries. The CTM is initialized with data from the 2D SOCRATES model corresponding to October conditions.

4. DATA ASSIMILATION

The NCAR-ROSE CTM has been used for assimilation of satellite data in several previous studies. Levelt *et al.* (1998) studied the performance of a simple optimal interpolation (OI) scheme using Upper-atmosphere Research Satellite (UARS) Microwave Limb Sounder (MLS) ozone observations. Ménard *et al.* (2000) proposed several modifications to the extended Kalman filter to reduce numerical errors. They analysed different aspects of the covariance calculus and its influence on assimilation results. Khattatov *et al.* (2000) applied a similar scheme using parametrized covariances. It consists primarily of OI with error propagation of analysis variances. They used the same forward model for both tracers and variances. This set-up has been applied successfully for different CTMs and input data (e.g. Chipperfield *et al.* 2002; El Amraoui *et al.* 2004). We will use a related set-up for the sequential assimilation of MIPAS chemical data.

For the observational error we use the random errors (precision) as delivered with the datasets. It is well known that these errors are in general not sufficient for assimilation. As in Khattatov *et al.* (2000), a representativeness error \mathbf{R} is therefore added. \mathbf{R} consists only of the diagonal components (variances). The initial analysis error variances, σ_{ii}^2 , and the model error growth rates, ϵ , were tuned as described in section 5 depending on dataset type, latitude and species. The correlation part of background covariances \mathbf{B} is parametrized by Gaussian functions depending on the local Euclidian grid-point distances. The horizontal scale is set to 1000 km and the vertical scale to 1.3 km, i.e. the vertical model step size. We will use χ^2 diagnostics (Khattatov *et al.* 2000; Chipperfield *et al.* 2002; Fierli *et al.* 2002; El Amraoui *et al.* 2004) to derive improved parameter sets for σ_{ii}^2 , \mathbf{R} , and ϵ for both MIMK and MESA data (section 5).

Depending on the observing geometry, the sample rate and the retrieval scheme used, the effective resolution can vary considerably. This has to be considered for the interpolation from model grid points to observations. The horizontal interpolation uses a linear weighting with a fixed horizontal box size of 1000 km. For the vertical interpolation, the full width half maximum (FWHM) is used as a measure of the vertical resolution. The latter is derived from the averaging kernels as provided by MESA which depend on species, latitude and season. MIMK gives the respective FWHM values for each individual profile and species. See the appendix for further details on the assimilation scheme.

TABLE 1. MODEL ERROR GROWTH RATE PARAMETER, ϵ (hr^{-1}), FOR DIFFERENT ASSIMILATED SPECIES

H ₂ O	O ₃	HNO ₃	CH ₄	N ₂ O	NO ₂
0.007	0.007	0.020	0.007	0.014	0.020

5. RESULTS

(a) χ^2 and OMF diagnostics

Global mean χ^2 and OMF values were calculated to check the consistency of assimilation parameters and investigate the influence of assimilated observations on the system. Results were also used to tune model error growth rates and representativeness errors. In the following \mathbf{y} , \mathbf{x} and \mathbf{H} depict observations, model first-guess values and the (linear) interpolation operator, respectively:

$$\text{OMF} \equiv (\mathbf{y} - \mathbf{H}\mathbf{x}) \quad (1)$$

$$\chi^2 = \frac{1}{N} \sum \text{OMF}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O} + \mathbf{R})^{-1} \text{OMF}. \quad (2)$$

\mathbf{B} , \mathbf{R} and \mathbf{O} indicate the background-error covariance matrix and the corresponding representativeness and observational error variances. We will use the analysis error as a synonym for the variance component of the diagnosed background-error covariance (see appendix). The sum in Eq. (2) consists of at least all N observations within a profile and a latitude band. Ideally, within Gaussian bias-free statistics, χ^2 values are expected to be $\simeq 1$, i.e. OMF error residuals and a priori errors are comparable in size. Initially, systematic differences between model and observations will give exceptionally high OMF and χ^2 values. If representativeness and background errors are adequate to the problem, then χ^2 will converge to identity during the course of assimilation.

Initially, relative representativeness errors were all set to 10% of corresponding model values, while the model error growth rate, ϵ , was set to 0.014 hr^{-1} for all species as in Khattatov *et al.* (2000). It was found that in this case the resulting χ^2 values were much too high. As discussed by Ménard and Chang (2000), χ^2 is mainly controlled by \mathbf{R} and ϵ . Therefore, adjustments to \mathbf{R} and ϵ parameters were applied as follows: a first assimilation run was performed to derive latitudinal-dependent temporal mean χ^2 values; these were then used as a correction factor for the representativeness error and the initial background error. The first 10 days of the analysis were not considered, to allow the model to adjust to observations. Next, corrections were applied to ϵ only. By reducing ϵ values for H₂O and CH₄, corresponding χ^2 values were increased. In the case of HNO₃ and NO₂, χ^2 values were decreased by increasing ϵ .

Table 1 shows the final ϵ settings used for all the following assimilation experiments. Compared to other strategies (e.g. El Amraoui *et al.* 2004), our approach has the advantage of giving improved parameters using a single assimilation experiment only. In principle, the results can be improved further by recursive application of the χ^2 diagnostics.

By definition, the resulting corrected representativeness errors depend on latitude (see Fig. 1). Most error values derived in this way correspond to a relative error between 5 and 20%. Relatively high values are found for MESA and MIMK HNO₃ in the subtropics and high latitudes. Corresponding N₂O errors show maximum values in the tropics and high southern latitudes. For NO₂, a strong increase of representativeness errors is derived for high northern latitudes which is especially significant in the case of

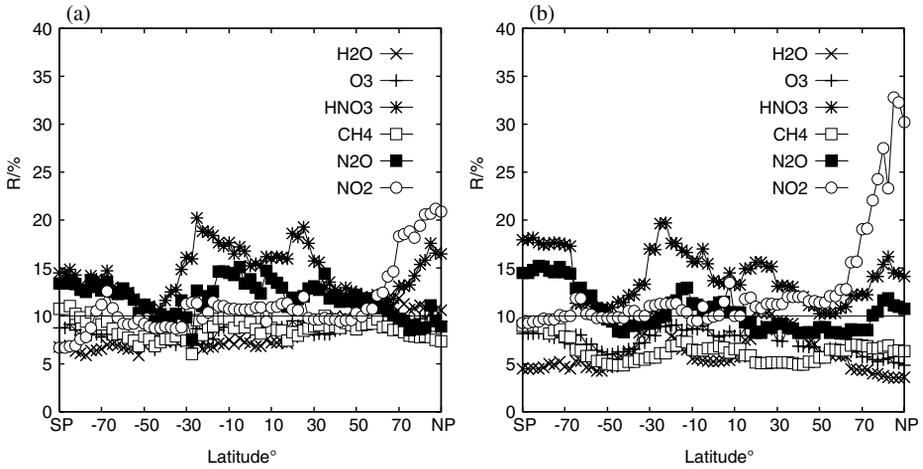


Figure 1. Latitudinal distribution of the mean representativeness errors (%) derived for (a) MESA and (b) MIMK datasets. The initial relative error of 10% is adjusted according to χ^2 results of the initial assimilation experiments covering 31 October to 20 November 2003 (see text).

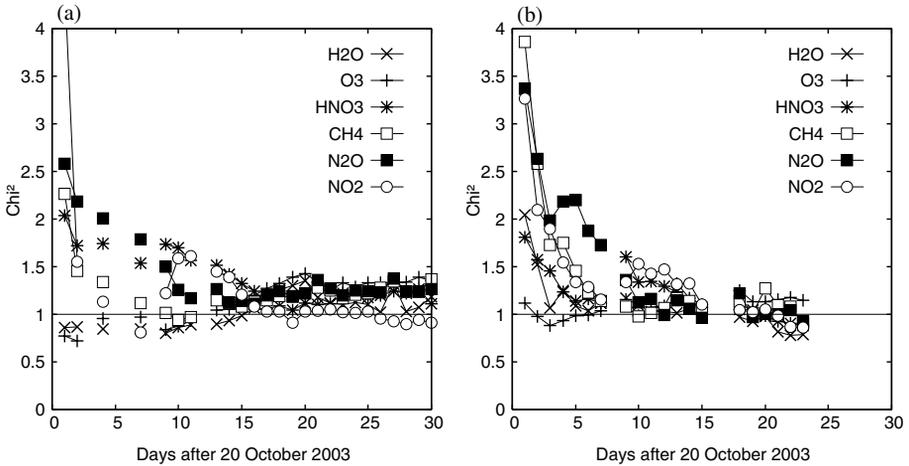


Figure 2. Time series of global mean χ^2 values for the final assimilation experiment using (a) MESA and (b) MIMK datasets covering 21 October to 20 November 2003.

MIMK data. This increase of NO_2 errors can be clearly related to the solar proton events in late October and early November 2003 (Degenstein *et al.* 2005).

Figure 2 shows χ^2 results for the six standard species of both MESA and MIMK data, during the course of assimilation. As can be seen from the plots, initial χ^2 values are in general too high but decrease considerably within the first two weeks. At the end of the assimilation period, values deviate from unity by no more than 50%. Some improvements could still be achieved by fine-tuning error growth rates for each dataset used. However, since the focus of this paper is on the comparison of different datasets, no further adjustments of assimilation parameters were applied.

In order to evaluate the performance of the assimilation system in more detail, relative OMF and analysis errors were analysed separately. In general, correspondence

TABLE 2. HALOE COMPARISON RESULTS

Species	MESA 2003		MIMK 2003		Reference	
	Bias	r.m.s.	Bias	r.m.s.	Bias	r.m.s.
O ₃	-0.54	11.89	-3.48	13.72	-0.89	16.12
H ₂ O	3.30	16.18	4.83	14.87	-13.09	23.75
NO _x	-3.48	36.73	4.93	33.21	-7.63	45.30
CH ₄	7.98	16.82	13.36	21.26	12.50	28.29
HCl	-6.37	23.83	-5.47	23.18	-6.08	23.50

Results of analysis minus HALOE comparisons showing the mean relative error values (%) for the assimilation experiments using MESA and MIMK data from 21 October to 20 November 2003. The reference columns show results without assimilation of MIPAS data. See text for more details.

between OMF and analysis errors is good. As can be anticipated from the χ^2 analyses, higher deviations are found for HNO₃, N₂O and NO₂. In the case of HNO₃ and N₂O, the derived representativeness errors are probably overestimated, resulting in analysis errors which are too high. As further investigations show, the respective N₂O χ^2 analysis is strongly influenced by OMF errors in the upper stratosphere. For HNO₃, OMF errors show maximum values within the tropical lower stratosphere. HNO₃ errors also correlate strongly with respective NO₂ errors in the upper northern stratosphere.

For NO₂, rather high OMF error values are expected due to its short photochemical time constant and the corresponding increase of representation errors, i.e. large times and distances between model and observations (Bracher *et al.* 2005). Differences are most pronounced when observed NO₂ values increase due to exceptional strong solar proton events (Degenstein *et al.* 2005). The model is intrinsically inadequate to simulate the observed NO₂ peak values. This can be attributed to the absence of any NO₂ source within the model domain that can mimic the effects of solar particle radiation. However, decreasing NO₂ OMF errors at the end of the assimilation period clearly indicates that the assimilation in general benefits from the NO₂ observations.

(b) Comparisons with HALOE observations

To give an independent estimate of the final analysis quality, comparisons against independent measurements are necessary. Therefore, we now discuss comparisons with HALOE observations. We use HALOE version 19 data available for October and November 2003 (<http://haloedata.larc.nasa.gov>) for sunrise and sunset conditions. Observations are limited to the middle northern latitudes and two latitude bands near 70°S and 30°S. Approximately two weeks are covered by observations in October and also in November 2003. Only data above the tropopause (defined by 2 PVU) are considered. Because of the strong scatter of HALOE NO_x observations above 1 hPa, such data were rejected.

Table 2 shows the mean analysis minus observation errors (biases) and the mean r.m.s. errors for the assimilation results using MESA and MIMK data. Corresponding results for the reference run without assimilation of MIPAS data are also given. Due to the general high zenith angle sensitivity of nitrous oxides, NO_x (= NO₂ + NO) is considered instead of single NO species. In all cases, biases are much smaller than the mean r.m.s. errors. Maximum r.m.s. errors are found for NO_x for both MIMK and MESA assimilation experiments. With the exception of HCl (not assimilated), relative improvements compared to the reference run reach 40%. Only MIMK O₃ and CH₄ mean biases increase slightly relative to HALOE.

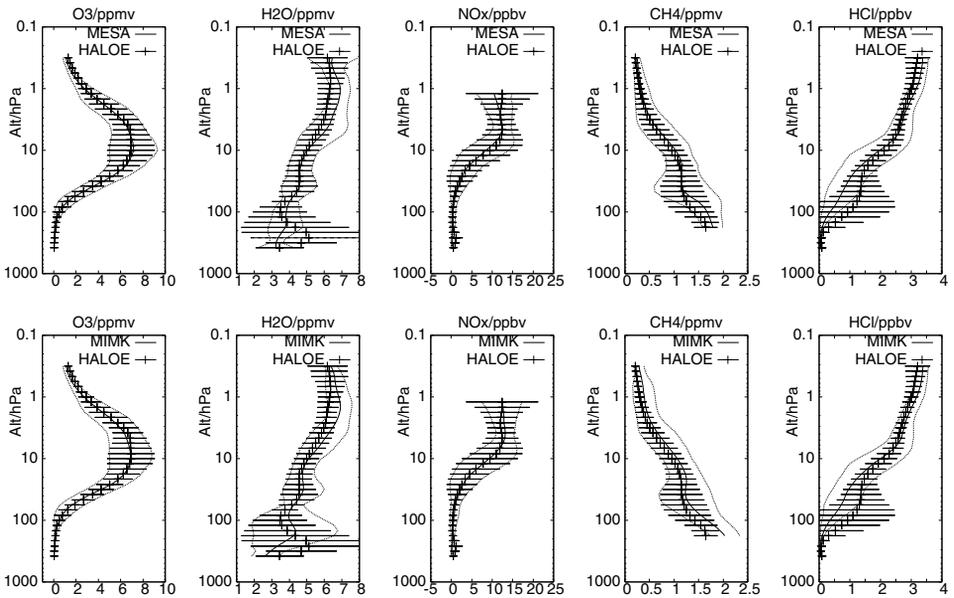


Figure 3. Mean collocated profiles of O_3 , H_2O , NO_x , CH_4 and HCl as observed by HALOE (vertical bars) and by (upper row) MESA and (lower row) MIMK analyses (both central solid curves) for 21 October to 20 November 2003. Horizontal bars (HALOE) and outer curves (MESA/MIMK) indicate the mean standard deviation at each model level.

Figure 3 shows the altitude dependency of the mean mixing ratio profiles as observed by HALOE compared with collocated assimilation results. Differences are within respective standard deviations (i.e. variability between single profiles). Only minor differences are found between MIMK and MESA assimilation results. In both cases, a positive bias for H_2O (upper stratosphere) and CH_4 (lower stratosphere) is visible. Above the ozone maximum, there is also a weak positive bias present especially in the MIMK results. In the case of the MIMK analysis, the CH_4 bias extends well into the upper stratosphere. The positive H_2O bias near 100 hPa is more pronounced in the MIMK results. Below 100 hPa, both analyses underestimate H_2O mixing ratios compared with HALOE observations. HCl mixing ratios are underestimated below 10 hPa and overestimated in the mesosphere. Compared to results without assimilated MIPAS observations, HCl shows no improvements, as expected, while all other species clearly improve their mean values and standard deviations.

To summarize, comparison results are in general within the known accuracy ranges of HALOE observations (O_3 , Brühl *et al.* 1996; H_2O , Harries *et al.* 1996; CH_4 , Park *et al.* 1996; HCl , Russell III *et al.* 1996; NO_2 , N_2O , Gordley *et al.* 1996). Our findings for H_2O and CH_4 , showing an overall positive bias compared to HALOE, are consistent with HALOE comparisons against ATLAS2/ATMOS* measurements. HALOE HCl observations are known to give values which are too low in the stratosphere. Because our results show significantly lower HCl values than HALOE below 10 hPa, this hints at a general model problem. As indicated by strong scatter, HALOE H_2O observations from below 100 hPa are suspicious (e.g. due to cloud contamination). Keeping this in mind, the variability of MIMK results compares better with HALOE observations than the corresponding MESA results.

* Atmospheric Trace Molecule Spectroscopy.

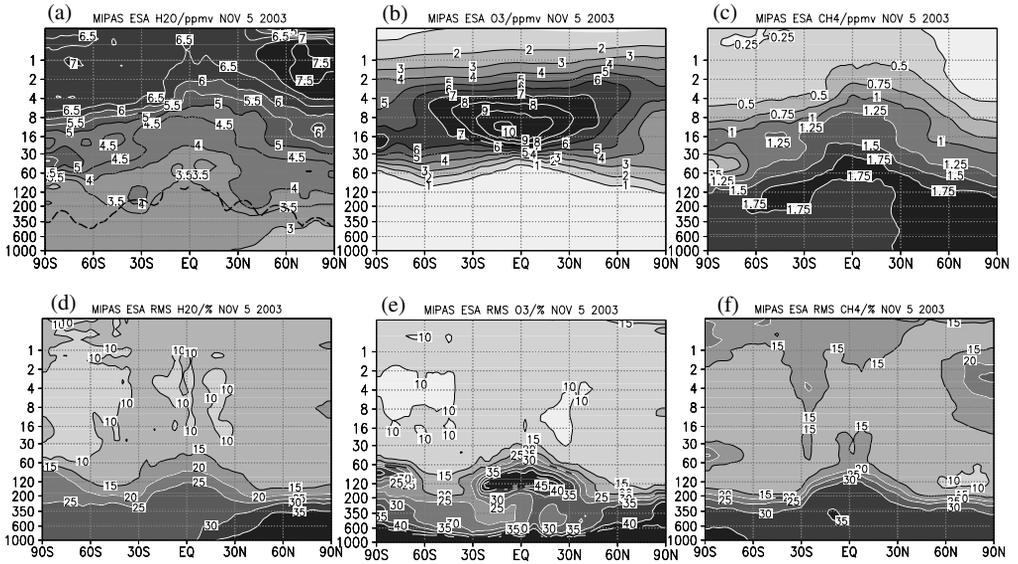


Figure 4. Zonal mean MESA analyses (ppmv) for (a) H_2O , (b) O_3 and (c) CH_4 for 5 November 2003, with corresponding r.m.s. analysis errors (%) at (d), (e) and (f). In (a), the 2 PVU contour (dashed) denotes the tropopause. The vertical axes show altitude as pressure levels in hPa.

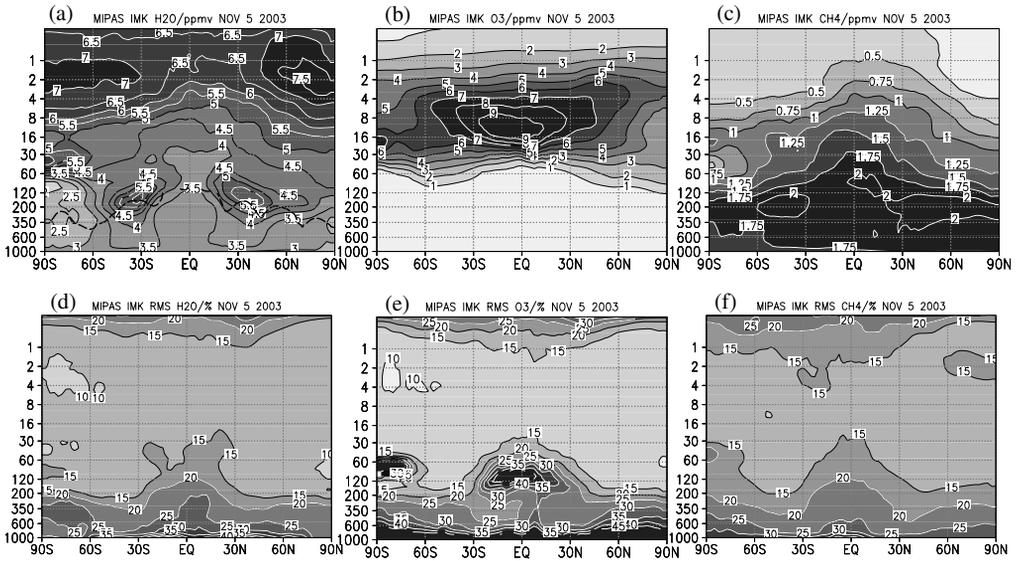


Figure 5. As Fig.4, but for MIMK analyses.

(c) Zonal mean results

In order to give an example of the general quality of the assimilation results, we finally examine the zonal mean analyses and zonal mean analysis errors for H_2O , O_3 and CH_4 . Figures 4 and 5 show the distributions for 5 November 2003 using MIMK and MESA data, respectively. This date is well covered by both MESA and MIMK observations and lies in the centre of the 30-day assimilation period (see Fig. 2).

Differences with respect to MESA and MIMK input data are small in general, except for two regions with increased H₂O and CH₄ mixing ratios near the tropical tropopause (indicated by the dashed line in H₂O plots). These are only visible when MIMK data are assimilated. MESA and MIMK analysis errors are in general very similar, but differ significantly in some regions mainly due to different data coverage. In both cases, minimum relative error values are found between 60 and 1 hPa. Errors increase strongly where no observations are available. For the species shown, this is mainly true for the troposphere and the tropopause region. For the same reason, HNO₃, N₂O and NO₂ analysis errors (not shown) increase considerably in the upper stratosphere and lower mesosphere. MIMK analyses of H₂O, O₃ and CH₄ also show a strong increase of analysis errors above 1 hPa.

6. DISCUSSION

We have developed a sequential data assimilation scheme to derive an improved analysis of the chemical state from different MIPAS observational datasets for October and November 2003. Using the CTM ROSE/DLR, results clearly improve in all cases considered when the MIPAS standard species H₂O, O₃, HNO₃, CH₄, N₂O and NO₂ are assimilated. However, significant differences due to different coverage and quality of input data (MESA/MIMK) are evident. MESA data are processed within an operational environment with good temporal coverage in general, while MIMK data are only available for single episodes with more homogeneous spatial coverage and less scatter.

In order to strengthen the consistency of the analyses, results of χ^2 diagnostics were employed to improve the initial background and representativeness errors. Improvements to model results without assimilation of MIPAS data were quantified by comparisons with HALOE observations. Results show an r.m.s. error reduction of up to 40% for the assimilated species. The positive influence of data assimilation is also found in the analysis error which significantly decreases where observations are available. At the end of October and again in early November 2003, observed NO_x concentrations increased rapidly in the mesosphere and upper stratosphere due to several solar proton events (Degenstein *et al.* 2005). This is evident as increased analysis errors in both MIMK and MESA NO₂ results. Nevertheless, assimilation of NO_x species proves valuable, as shown by comparisons with HALOE.

Regarding the different datasets used, mean assimilation results are comparable with some exceptions: assimilation of MIMK data leads in particular to higher H₂O mixing ratios in the lowermost stratosphere; two regions with increased H₂O values near the tropical tropopause, covered both by MESA and MIMK data, show only up in the results with MIMK data; and, compared with HALOE observations, MIMK CH₄ results show a strong positive bias throughout the stratosphere. In summary, both MIMK and MESA datasets are found to be well suited for global chemical data assimilation, when the special characteristics of each dataset are taken into account. The presented assimilation system contributes significantly in fulfilling the general objective to better monitor and study the chemical composition of the middle atmosphere.

ACKNOWLEDGEMENTS

We thank the reviewers for their constructive comments and helpful advice. ESA is acknowledged for providing the MIPAS Level 2 data within the ESA-AO project EVIVA. We are grateful to the Institute of Meteorology, Karlsruhe, for supply of MIPAS IMK data and for fruitful discussion. The Instituto de Astrofísica de Andalucía

(CSIC) provided crucial data for the IMK NO₂ retrieval. Further, we thank the German Ministry for Education and Research for funding HGF-ENVISAT (HGF/BMBF-Vernetzungsfonds) and the INVERT project (via grant 07ATF11). The UK Met Office and the HALOE team kindly provided the necessary meteorological and correlative data. We would also like to thank our colleagues from DLR D-PAC. We finally thank Anne Smith and Dan Marsh for providing the original NCAR-ROSE model code.

APPENDIX

Assimilation scheme

For this study an optimum interpolation scheme with propagation of analysis errors as in Khattatov *et al.* (2000) is used to correct the model first guess at each model time step. For a statistically viable weighting of model forecast and observations, background, \mathbf{B} , observational, \mathbf{O} , and representation, \mathbf{R} , error variances and covariances have to be specified. The statistically optimum analysis \mathbf{x}' can then be calculated from the model's first-guess state vector \mathbf{x} and the observation vector \mathbf{y} as follows (e.g. Daley 1991):

$$\mathbf{x}' = \mathbf{x} + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O} + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}). \quad (\text{A.1})$$

The observation vector \mathbf{y} holds all sampled observations within one profile. Interpolation from and to the neighbouring model grid points is performed using the forward interpolation operator \mathbf{H} . \mathbf{H} comprises the product of a vertical and horizontal linear interpolation using vertical log-pressure and local Euclidian coordinates, respectively. Details regarding the application of MIPAS observations are given in section 4. No error correlations are used (\mathbf{R} and \mathbf{O} matrixes are diagonal). Note that adding a representation error variance, \mathbf{R} , in fact increases the observational error. As discussed in section 4, this is necessary to make the errors consistent. The \mathbf{B} covariance matrix disperses the resulting corrections back to the model grid. It therefore must take the model's resolution and presumed uncertainty into account. With respect to the model's grid points, the matrix elements of \mathbf{B} are modelled by Gaussian functions (e.g. Riishøjgaard 1998):

$$B_{ij} = \sigma_{ii}\sigma_{jj} \exp\left(\frac{-d^2}{2D^2}\right) \exp\left(\frac{-r^2}{2R^2}\right), \quad (\text{A.2})$$

where d , r and D , R describe the horizontal and vertical distances between model grid points and scales respectively. D and R scales are set to 1000 km and 1.3 km (the model's vertical step size). σ_{ii}^2 and σ_{jj}^2 are the model's first-guess variances at grid points i and j . For the analysis (A.1), error covariances are calculated as follows:

$$\mathbf{E} = \mathbf{B} - \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O} + \mathbf{R})^{-1}\mathbf{H}\mathbf{B}. \quad (\text{A.3})$$

Only the diagonal components E_{ii} (variances) are retained for the subsequent temporal error propagation. Variances of all observed species are then transported as quasi-tracers. We add a chemical correction using the change of mixing ratios μ'_{ii}/μ_{ii} due to model chemistry. Finally an error increment is added to mimic the model error increase (Ménard *et al.* 2000). In summary, the propagated analysis variance after one time step takes the form:

$$\sigma_{ii}^2 \doteq T(E_{ii}) \cdot \frac{\mu'_{ii}}{\mu_{ii}} + (\epsilon\mu_{ii})^2, \quad (\text{A.4})$$

with T describing the advection step and ϵ the error increase rate per hour (see section 4). The adjustable parameter ϵ and the initial variances σ_{ii}^2 are defined as prescribed in section 5(a).

REFERENCES

- El Amraoui, L., Ricaud, P., Urban, J., Théodore, B., Hauchecorne, A., Lauti, J., De La Noë, N., Guirlet, M., Le Flocjoën, E., Murtagh, D., Dupuy, E., Frisk, U. and d'Andon, O. F. 2004 Assimilation of Odin/SMR O₃ and N₂O measurements in a three-dimensional chemistry transport model. *J. Geophys. Res.*, **109**, D22304, doi: [10.1029/2004JD004796](https://doi.org/10.1029/2004JD004796)
- Bracher, A., Sinnhuber, M., Rozanov, A. and Burrows, J. P. 2005 Using a photochemical model for the validation of NO₂ satellite measurements at different solar zenith angles. *Atmos. Chem. Phys.*, **5**, 393–408
- Bracher, A., Bramstedt, K., Sinnhuber, M., Weber, M. and Burrows, J. P. 2004 'Validation of MIPAS O₃, NO₂, H₂O and CH₄ profiles (v4.61) with collocated measurements of HALOE and SAGE II'. Proceedings of the Second Workshop on the atmospheric chemistry validation of ENVISAT; Frascati, Italy, 3–7 May 2004, ESA SP-562. <http://envisat.esa.int/workshop/acve2/papers/EPOMIABR.pdf>
- Brasseur, G., Hitchman, M. H., Walters, S., Dymek, M., Falise, E. and Pirre, M. 1995 An interactive chemical dynamical radiative two-dimensional model of the middle atmosphere. *J. Geophys. Res.*, **95**, 5639–5655
- Brühl, C., Drayson, S. R., Russell III, J. M., Crutzen, P. J., McInerney, J. M., Purcell, P. N., Claude, H., Gernandt, H., McGee, T. J., McDermid, I. S. and Gunson, M. R. 1996 Halogen occultation experiment ozone channel validation. *J. Geophys. Res.*, **101**, 10217–10240
- Carli, B., Alpaslan, D., Carlotti, M., Castelli, E., Ceccherini, S., Dinelli, B. M., Dudhia, A., Flaud, J. M., Höpfner, M., Jay, V., Magnani, L., Oelhaf, H., Payne, V., Piccolo, C., Prosperi, M., Raspollini, P., Ridolfi, M., Remedios, J. and Spang, R. 2004 First results from MIPAS/ENVISAT with operational Level 2 code. *Adv. Space Res.*, **33**(7), 1012–1019, doi: [10.1016/S0273-1177\(03\)00584-2](https://doi.org/10.1016/S0273-1177(03)00584-2)
- Ceccherini, S. 2004 'Averaging kernels for MIPAS off-line level 2 retrievals'. IFAC Technical Note, Italian National Research Council, Prog. Doc. No.: IFAC-GA-2004-01-SC
- Ceccherini, S. and Ridolfi, M. 2002 'Averaging kernels for MIPAS near-real-time level 2 retrievals'. IFAC Technical Note, Italian National Research Council, Prog. Doc. No.: TN-IFAC-OST0201
- Chipperfield, M. P. 1999 Multiannual simulations with a three-dimensional chemical transport model. *J. Geophys. Res.*, **104**(D1), 1781–1805
- Chipperfield, M. P., Khattatov, B. V. and Lary, D. J. 2002 Sequential assimilation of stratospheric chemical observations in a three-dimensional model. *J. Geophys. Res.*, **107**(D21), 4585, doi: [10.1029/2002JD002110](https://doi.org/10.1029/2002JD002110)
- Daley, R. 1991 *Atmospheric data analysis*. Cambridge University Press, Cambridge, UK
- Degenstein, D. A., Lloyd, N. D., Bourassa, A. E., Gattinger, R. L. and Llewellyn, E. J. 2005 Observations of mesospheric ozone depletion during the October 28, 2003 solar proton event by OSIRIS. *Geophys. Res. Lett.*, **32**, L03S11, doi: [10.1029/2004GL021521](https://doi.org/10.1029/2004GL021521)
- Errera, Q. and Fonteyn, D. 2001 Four-dimensional variational chemical data assimilation of CRISTA stratospheric measurements. *J. Geophys. Res.*, **106**, 12253–12265
- Fierli, F., Hauchecorne, A., Bekki, S., Théodore, B. and d'Andon, O. F. 2002 Data assimilation of stratospheric ozone using a high-resolution transport model. *Geophys. Res. Lett.*, **29**(10), doi: [10.1029/2001GL014272](https://doi.org/10.1029/2001GL014272)
- Fischer, H. and Oelhaf, H. 1996 Remote sensing of vertical profiles of atmospheric trace constituents with MIPAS limb-emission spectra. *Appl. Opt.*, **35**(16), 2787–2796

- Funke, B., Lopez-Puertas, M., von Clarmann, T., Stiller, G. P., Fischer, H., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Milz, M., Steck, T. and Wang, D.-Y. 2005 Retrieval of stratospheric NO_x from 5.3 and 6.2 μm non-LTE emissions measured by MIPAS on ENVISAT. *J. Geophys. Res.*, **110**(D9), D09302, doi: [10.1029/2004JD005225](https://doi.org/10.1029/2004JD005225)
- Glatthor, N., von Clarmann, T., Fischer, H., Funke, B., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Milz, M., Steck, S., Stiller, G. P., Mengistu Tsidu, G. and Wang, D.-Y. 2005 Mixing processes during the Antarctic vortex split in September/October 2002 as inferred from source gas and ozone distributions from MIPAS/ENVISAT. *J. Atmos. Sci.*, **62**(3), 787–800
- Gordley, L. L., Russell III, J. M., Mickley, L. J., Frederick, J. E., Park, J. H., Stone, K. A., Beaver, G. M., McInerney, J. M., Deaver, L. E., Toon, G. C., Murcray, F. J., Blatherwick, R. D., Gunson, M. R., Abbatt, J. P. D., Mauldin III, R. L., Mount, G. H., Sen, B. and Blavier, J.-F. 1996 Validation of nitric oxide and nitrogen dioxide measurements made by the Halogen Occultation Experiment for the UARS platform. *J. Geophys. Res.*, **101**(D6), 10241–10266
- Granier, C. and Brasseur, G. 1991 Ozone and other trace gases in the Arctic and Antarctic regions: Three-dimensional model simulations. *J. Geophys. Res.*, **96**, 2995–3011
- Harries, J. E., Russell III, J. M., Tuck, A. F., Gordley, L. L., Purcell, P., Stone, K., Bevilacqua, R. M., Gunson, M. R., Nedoluha, G. and Traub, W. A. 1996 A validation of measurements of water vapor from the Halogen Occultation Experiment, HALOE. *J. Geophys. Res.*, **101**(D6), 10205–10216
- Khattatov, B. V., Lamarque, J.-F., Lyjak, L. V., Ménard, R., Levelt, P., Tie, X., Brasseur, G. and Gille, J. C. 2000 Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. *J. Geophys. Res.*, **105**(D23), 29135–29144
- Levelt, P. F., Khattatov, B. V., Gille, J. C., Brasseur, G. P., Tie, X. and Waters, J. W. 1998 Assimilation of MLS ozone measurements in the global three-dimensional chemistry transport model ROSE. *Geophys. Res. Lett.*, **25**(24), 4493–4496
- Lin, S.-J. and Rood, R. B. 1996 Multidimensional flux form semi-Lagrangian transport schemes. *Mon. Weather Rev.*, **124**, 2046–2070
- Marsh, D., Smith, A., Brasseur, G., Kaufmann, M. and Grossman, K. 2001 The existence of a tertiary ozone maximum in the high-latitude middle mesosphere. *Geophys. Res. Lett.*, **28**, 4531–4534
- Ménard, R. and Chang, L. P. 2000 Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: χ^2 validated results and analysis of variance and correlation dynamics. *Mon. Weather Rev.*, **128**, 2672–2686
- Ménard, R., Cohn, S. E., Chang, L. P. and Lyster, P. M. 2000 Stratospheric assimilation of chemical tracer observations using a Kalman filter. Part I: Formulation. *Mon. Weather Rev.*, **128**, 2654–2671
- Mengistu Tsidu, G., Stiller, G. P., von Clarmann, T., Funke, B., Fischer, H., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Milz, M., Steck, T. and Wang, D. Y. 2005 NO_y from the Michelson Interferometer for Passive Atmospheric Sounding on the Environmental Satellite during the southern hemisphere polar vortex split in September/October 2002. *J. Geophys. Res.*, **110**(D11), D11301, doi: [10.1029/2004JD005322](https://doi.org/10.1029/2004JD005322)

- Park, J. H., Russell III, J. M., 1996
Gordley, L. L., Drayson, S. R.,
Benner, D. C.,
McInerney, J. M.,
Gunson, M. R., Sen, G. G. B.,
Blavier, J.-F., Webster, C. R.,
Zipf, E. C., Erdman, P.,
Schmidt, U. and Schiller, C.
- Ridolfi, M., Carli, B., Carlotti, M., 2000
von Clarmann, T.,
Dinelli, B. M., Dudhia, A.,
Flaud, J. M., Höpfner, M.,
Morris, P. E., Raspollini, P.,
Stiller, G. and Wells, R. J.
- Riishøgaard, L. P. 1998
- Rose, K. and Brasseur, G. 1989
- Russell III, J. M., Deaver, L. E., 1996
Mingzhao, L., Park, J. H.,
Gordley, L. L., Tuck, A. F.,
Toon, G. C., Gunson, M. R.,
Traub, W. A., Johnson, D. G.,
Jucks, K. W., Murcray, D. G.,
Zander, R., Nolt, I. G. and
Webster, C. R.
- Sander, S. P., Friedl, R. R., 2003
Ravishankara, A. R.,
Golden, D. M., Kolb, C. E.,
Kurylo, M. J., Huie, R. E.,
Orkin, V. L., Molina, M. J.,
Moortgat, G. K. and
Finlayson-Pitts, B. J.
- Snoeij, P., Koopman, R., 2004
Wursteisen, P., Zehner, C. and
Attema, E.
- Steck, T. and von Clarmann, T. 2001
- Stiller, G. P., von Clarmann, T., 2002
Funke, B., Glatthor, N.,
Hase, F., Höpfner, M. and
Linden, A.
- Swinbank, R. and O'Neill, A. 1994
- Validation of the Halogen Occultation Experiment CH₄ measurements from the UARS. *J. Geophys. Res.*, **101**(D6), 10183–10203
- Optimized forward model and retrieval scheme for MIPAS near-real-time data processing. *Appl. Opt.*, **39**, 1323–1340
- A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus*, **50A**, 42–57
- A three-dimensional model of chemically active trace species in the middle atmosphere during disturbed winter conditions. *J. Geophys. Res.*, **94**, 16387–16403
- Validation of hydrogen chloride measurements made by the Halogen Occultation Experiment from the UARS platform. *J. Geophys. Res.*, **101**(D6), 10241–10266
- ‘Chemical kinetics and photochemical data for use in atmospheric studies’. Evaluation No. 14, JPL Publication 02-25. California Institute of Technology, Pasadena, USA.
<http://jpldataeval.jpl.nasa.gov/download.html>
- ‘The ENVISAT Atmospheric Chemistry Instrument Validation Programme’. Proceedings of the Second Workshop on the atmospheric chemistry validation of ENVISAT, Frascati, Italy, 3–7 May 2004. ESA SP-562,
<http://envisat.esa.int/workshop/acve2/papers/EGE01PS.pdf>
- Constrained profile retrieval applied to the observation mode of the Michelson Interferometer for Passive Atmospheric Sounding. *Appl. Opt.*, **40**(21), 3559–3571
- Sensitivity of trace gas abundances retrievals from infrared limb emission spectra to simplifying approximations in radiative transfer modelling. *J. Quant. Spectrosc. Radiat. Transfer*, **72**, 249–280
- A stratosphere–troposphere data assimilation system. *Mon. Weather Rev.*, **122**, 686–702