The Assimilation of Envisat data (ASSET) project


1 Data Assimilation Research Centre, University of Reading, Reading, UK
2 Service Aeronomie, Universite Pierre et Marie Curie, Paris, France
3 ECMWF, Reading, UK
4 CNR-IFAC, Florence, Italy
5 University of Köln, Köln, Germany
6 BIRA-IASB, Brussels, Belgium
7 KNMI, De Bilt, The Netherlands
8 Met Office, Exeter, UK
9 Fusion Numerics, Boulder, Colorado, USA
10 CERFACS, Toulouse, France
11 Météo-France, Toulouse, France
12 University of Bologna, Bologna, Italy
13 LMD, Paris, France
The ASSET project
W. A. Lahoz et al.

Received: 4 October 2006 – Accepted: 29 November 2006 – Published: 8 December 2006
Correspondence to: W. A. Lahoz (swslahoz@rdg.ac.uk)
Abstract

This paper discusses the highlights of the EU-funded “Assimilation of Envisat data” (ASSET) project, which has involved assimilation of Envisat atmospheric constituent and temperature data into systems based on Numerical Weather Prediction (NWP) models and chemical transport models (CTMs). Envisat was launched in 2002 and is one of the largest Earth Observation (EO) satellites ever built. It carries several sophisticated EO instruments providing insights into chemistry and dynamics of the atmosphere. The overarching theme of the ASSET project has been to bring together experts from all aspects of the data assimilation problem. This has allowed ASSET to address several themes comprehensively: enhancement of NWP analyses by assimilation of research satellite data; studies of the distribution of stratospheric chemical species by assimilation of research satellite data into CTM systems; objective assessment of the quality of ozone analyses; studies of the spatial and temporal evolution of tropospheric pollutants; enhanced retrievals of Envisat data; and data archival and dissemination.

Among the results from the ASSET project, many of which are firsts in their field, we can mention: a positive impact on NWP analyses from assimilation of height-resolved stratospheric humidity and temperature data, and assimilation of limb radiances; the extraction of temperature information from the assimilation of chemical species into CTMs; a first intercomparison between ozone assimilation systems; the extraction of information on tropospheric pollution from assimilation of Envisat data; and the large potential of the Envisat MIPAS dataset. This paper discusses these, often novel, developments and results. Finally, achievements of, and recommendations from, the ASSET project are presented.
1 Introduction

The EU-funded “Assimilation of Envisat data” (ASSET) project (http://darc.nerc.ac.uk/asset) is a major European initiative in Earth Observation (EO); it has run for the period January 2003–June 2006. Its overall rationale is to use the techniques of data assimilation, DA (e.g. Kalnay, 2003) to develop a European capability for chemical and UV forecasting, and provide analyses for coupled climate/chemistry studies. The objectives of the ASSET project are: (1) assess the strategies for exploiting research satellite data by the Numerical Weather Prediction (NWP) community, and (2) using this data to investigate the distribution and variability of atmospheric chemical species. To address these objectives, the ASSET project brought together experts from all aspects of the DA problem: NWP systems; chemical DA; numerical modelling; meteorology; retrieval theory; DA theory; EO measurements; data analysis; and data management.

The ASSET partners brought to the project experience with different DA systems, DA techniques and assimilated data (see http://darc.nerc.ac.uk/asset). The DA systems used by the ASSET partners to assimilate constituent data from Envisat fell into three categories: NWP DA systems based on dynamical models, often General Circulation Models (GCMs); chemical DA systems based on Chemical Transport Models (CTMs) driven by off-line NWP analyses or short-term forecasts; coupled DA systems (usually a GCM coupled with a CTM). In this paper, we focus on the first two systems. The DA techniques used in the ASSET project included three- and four-dimensional variational data assimilation (3D-Var and 4D-Var, respectively), and Kalman Filter (KF) methods (Kalnay, 2003). Another variational approach used in the ASSET project was 3D-FGAT (First Guess at the Appropriate Time), a variant of 3D-Var. The Envisat data assimilated into the DA systems involved in ASSET consisted of level 1 data (radiances or line densities) and level 2 data (height-resolved or total-column retrievals) from the following instruments: AATSR (Advanced Along Track Scanning Radiometer), GOMOS (Global Ozone Monitoring by Occultation of Stars), MIPAS (Michelson Interferometer
for Passive Atmospheric Sounding), and SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY).

This paper brings together in one publication selected highlights of the ASSET project. Some of the work is described elsewhere in the literature in more detail, but this paper also includes much work that will not be published separately, and all of the results are explained in a wider context. Section 2 describes ASSET project highlights. Section 3 summarizes achievements and results of the ASSET project (many of them firsts in their field), and presents recommendations.

2 ASSET project highlights

In Sect. 2.1 we discuss the impacts of assimilation of height-resolved humidity and temperature data from MIPAS, and of direct assimilation of limb radiances from MIPAS. In Sect. 2.2 we discuss the stratospheric distribution of NO_x and NO_y (NO_x = NO + NO_2; NO_y = NO_x + HNO_3 plus other related chemical species), and the extraction of temperature information from constituent information, using analyses derived from assimilation of GOMOS data. In Sect. 2.3 we objectively evaluate the assimilation of MIPAS and SCIAMACHY ozone data. Sections 2.1–2.3 focus mainly on the stratosphere, but in Sect. 2.4, the assimilation of tropospheric constituents from Envisat is discussed. Section 2.5 discusses enhanced retrievals of MIPAS data. Section 2.6 discusses how the data produced by the ASSET project is archived and disseminated.

Table 1 provides a list of the DA systems participating in the ASSET ozone intercomparison (ASSIC) project (Sect. 2.3; Geer et al., 2006a). Geer et al. (2006b) provides further details of the ozone photochemistry and heterogeneous parametrizations used in the ASSIC project. Some of these systems were also used to study the assimilation of humidity and temperature, and the assimilation of limb radiances (Sect. 2.1).
2.1 Assimilation of MIPAS data

2.1.1 Assimilation of humidity retrievals

The daily variability of the water vapour field in the stratosphere is poorly known. Gaining an improved knowledge of this variability is very desirable, as upper troposphere/lower stratosphere (UTLS) water vapour is known to play an important role in many aspects of meteorology, including radiation, dynamics, chemistry and climate change (SPARC, 2000). The operational assimilation of humidity data in the stratosphere by NWP centres is limited by the availability of suitable data (Simmons et al., 1999). The reduction in specific humidity by six orders of magnitude from the surface to 0.1 hPa, also makes the assimilation problem considerably more difficult. Due to assimilation problems, both the Met Office (N. B. Ingleby, personal communication, 2003) and ECMWF (Hólm et al., 2002) have made to date ad hoc fixes to constrain the stratospheric humidity field. However, with the availability of high quality humidity data from Envisat, with high spatial and temporal density, it is now appropriate to revisit the issue of stratospheric humidity assimilation. Here, we present preliminary results from the assimilation of MIPAS humidity profiles into three different DA systems. This is the first study to perform such a comparison.

MIPAS humidity profiles are available from 12–60 km. The estimated error standard deviation for v4.61 is 10–20% near the tropopause, 3–10% in the 15–40 km layer, and 10–20% in the 40–60 km layer; the total error (random plus systematic, i.e., bias) is 20–25% near the tropopause, 15–20% in the 15–40 km layer, and 20–50% in the 40–60 km layer (Raspollini et al., 2006). The main systematic errors are associated with spectroscopy and horizontal temperature errors. Comparison of the MIPAS humidity profiles with balloon and aircraft data (Oelhaf et al., 2004), ground-based radiometer and lidar data (Pappalardo et al., 2004) and HALOE (Halogen Occultation Experiment), SAGE II (Stratospheric Aerosol and Gas Experiment) and POAM III (Polar Ozone and Aerosol Measurement) satellite data (Weber et al., 2004) shows good agreement between 15 km and 30 km. However, above 30 km the MIPAS retrievals have a positive
bias of up to 20% compared to the satellite data and of 7–15% compared to ground-based radiometer and lidar data. In the UTLS region the MIPAS retrievals have a small negative bias compared to the balloon and aircraft data.

Three different groups participating in the ASSET project have assimilated MIPAS water vapour profile data (v4.61): the Met Office, ECMWF and BASCOE (BIRA-IASB). Two are NWP systems (Met Office and ECMWF); BASCOE assimilates data into a CTM. Their DA systems are summarized in Table 1. The BASCOE DA system is only concerned with stratospheric humidity analyses and hence can use water vapour mixing ratio as the control variable, whereas the Met Office, as well as ECMWF, perform a combined stratospheric and tropospheric humidity analysis. In these experiments, the Met Office uses normalized specific humidity as the humidity control variable throughout the atmosphere. ECMWF uses normalized specific humidity in the stratosphere, and normalized relative humidity in the troposphere. The concept of normalization to make the background errors flow dependent and Gaussian was formulated by Hólm et al. (2002). Dee and da Silva (2003) highlighted the negative impact of temperature observations for humidity assimilation and, following their recommendations, any correlation between temperature and specific humidity is removed from the control variable in the stratosphere of both the Met Office and ECMWF systems. The Met Office has tested other formulations of the control variable, and these results appear later in this sub-section.

The Met Office and ECMWF generated their background error covariance matrices using the NMC method (Parrish and Derber, 1992) and ensemble method (Fisher, 2003), respectively. The BASCOE background error covariances are much simpler, with no vertical or horizontal error correlations and standard deviations equal to 20% of the background humidity field. Although there are no error correlations in the BASCOE system, information from MIPAS observations is still spread via the observation operator. The operator averages the information from the eight grid points surrounding the observation point; the relatively coarse grid resolution in the BASCOE CTM also helps to spread this information. For both ECMWF and the Met Office, the assimilation
schemes used are experimental and are different from the operational schemes, where humidity assimilation in the stratosphere is respectively switched off or constrained. The Met Office and ECMWF systems parametrize both water vapour production by methane oxidation in the stratosphere and water vapour loss through photolysis in the mesosphere. In the BASCOE system, these processes are modelled explicitly.

The humidity analyses from all three groups have varying accuracies that depend on the DA system, the level and the latitude. Figure 1 shows the monthly mean zonal water vapour analyses for September 2003 for the ECMWF and BASCOE systems. The analyses were interpolated onto a common grid to enable a simpler comparison. The common grid is the same as that used in the ASSIC project (Sect. 2.3; Geer et al., 2006a): 3.75° longitude × 2.5° latitude, with 37 fixed pressure levels.

The monthly mean analyses show good agreement with the UARS reference atmosphere for September (http://code916.gsfc.nasa.gov/Public/Analysis/UARS/urap/home.html). A number of well-known features can be seen in the stratospheric analyses from BASCOE and ECMWF. These include the very dry tropical tropopause and dehydration of the Antarctic winter polar vortex (SPARC, 2000). The role of the Brewer-Dobson circulation on the distribution of water vapour is reflected in the upward and poleward propagation of the dry air entering the stratosphere through the tropical tropopause. Methane oxidation is responsible for the relatively moist upper stratosphere and lower mesosphere. The Brewer-Dobson circulation transports this moist air downwards within the winter hemisphere polar vortex.

Between the tropopause and 1 hPa, the zonal mean monthly analyses for the BASCOE and ECMWF systems are reasonably similar. The BASCOE analyses show a drier UTLS region at most latitudes, whereas the ECMWF analyses show a more distinct dry tropical tropopause region. Consequently, the vertical gradient in specific humidity in the lower stratosphere is stronger in the BASCOE analyses. The southern hemisphere polar vortex is drier in the BASCOE analyses. Above 1 hPa the zonal mean specific humidity fields vary quite considerably between the two systems. In this region, the ECMWF analyses are ~2 ppmv moister than the BASCOE analyses, which
appear more realistic when compared to the UARS reference atmosphere. BASCOE analyses are \(\sim 5\%\) lower than MIPAS data in the lower mesosphere, but the corresponding ECMWF analyses are \(\sim 10\%\) greater. However, the ECMWF analyses are 25–30\% too low compared to the uppermost MIPAS layer at 0.2–0.1 hPa. It appears that the ECMWF analyses aim to find a compromise between these conflicting biases, in agreement with the expected background error correlations in the vertical. Most of the differences between the two analyses can be explained by the fact that BASCOE does not assimilate any MIPAS data below 95 hPa and above 0.2 hPa (data outside these regions is model generated). Influences from the troposphere and mesosphere are therefore excluded. The lack of any horizontal error correlations in the BASCOE assimilation scheme appears not to be a problem due to the high spatial and temporal frequency of the MIPAS humidity data.

The Met Office has investigated the impact of varying the control variable in the assimilation of MIPAS humidity data. The objective is to develop a humidity control variable that has the desirable properties that it is usable in both the troposphere and the stratosphere, it has approximately Gaussian background errors, that temperature and humidity increments are decoupled, and that allows realistic vertical error correlations. To achieve this, the Met Office have combined the ideas of Dee and da Silva (2003) and Hólm et al. (2002), and defined a normalized specific humidity variable. The impact of the normalization is to produce a considerably better conditioned background error covariance matrix and consequently the minimization in the 3D-Var algorithm is much faster. The removal of the influence of temperature increments reduces spurious upper stratospheric increments.

Figure 2 compares mean MIPAS and Met Office analysed profiles of specific humidity for the 60° S–90° S region on 25 September 2002. This is five days after the start of the assimilation experiment. Three different experiments are shown where the humidity control variable is either relative humidity (RH), normalized RH or normalized specific humidity. All three experiments show fairly reasonable specific humidity profiles below 5 hPa. However, at higher levels the fit to the MIPAS observations is less good, with
the analyses being consistently too dry. The experiment with the normalized specific humidity control variable has a more reasonable lower mesospheric specific humidity, but is still too dry when compared to the MIPAS observations.

The reasons for the Met Office’s poor upper stratosphere and lower mesosphere assimilation are not fully understood. It is possible that unrealistic vertical correlations in the background error covariance matrix in the upper stratosphere and lower mesosphere may explain the poor specific humidity profiles in this region and their excessive variability. The normalized specific humidity control variable experiment, where the influence of temperature increments on the control variable is removed, has profiles with a reduced and more realistic standard deviation than the other two trials in the upper stratosphere, especially at around 4 hPa. The improvement in tropopause specific humidity seen in the two trials with the normalized control variable (RH or specific humidity) may result from the normalized variable handling better the steep gradient in specific humidity across the tropopause. However, this improvement is not seen at all latitudes.

This section has given a brief overview of the humidity assimilation activities performed and results gained as part of the ASSET project. Evidence of a positive impact of humidity assimilation into DA systems (both operational and research) is provided. A more detailed intercomparison of the different humidity analyses is currently underway.

2.1.2 Assimilation of temperature retrievals

In this sub-section, results from experiments to assimilate Envisat temperature retrievals are shown. Evaluation studies have shown that GOMOS temperature retrievals have large biases and are not ready to be validated (Goutail and Bazureau, 2004), while icing on the instrument has made SCIAMACHY temperature retrievals impossible (Piters et al., 2006). Therefore the temperature assimilation experiments were carried out using height-resolved retrievals (v4.61) from MIPAS only. The observations have been assimilated using both the ECMWF and Met Office DA systems used in the ASSIC project (see Table 1).
MIPAS temperature profiles (v4.61) were monitored passively in the ECMWF system by Dethof et al. (2004). The study indicates a bias of between –1 K and –6 K in the mesosphere and near the stratopause, when compared with Met Office or ECMWF analyses, and between –2 K and 0 K when compared to National Centers for Environmental Prediction (NCEP) analyses. In the upper stratosphere, the MIPAS temperatures show a bias of between –2 K (compared to Met Office analyses) and 4 K (compared to NCEP and ECMWF analyses). In the lower stratosphere the MIPAS temperatures have a positive bias of 0.5–2.5 K compared to the other datasets.

ECMWF ran a set of experiments for a 43-day period during August–September 2003. The control experiment assimilates all available operational data plus GPS (Global Positioning Satellite) radio occultation bending angles from CHAMP (CHAlIening Minisatellite Payload) (Healy and Thépaut, 2006). The test experiment also assimilates height-resolved MIPAS temperature, humidity and ozone retrievals. Comparison of the results against MIPAS observations shows that assimilating MIPAS data has little impact on the analysis below the mid stratosphere, but that above ~5 hPa the bias between analysis and MIPAS retrievals, and the standard deviation of the analysis departures, is reduced. An example of this appears in Fig. 3, which shows the bias and standard deviation for the region 60° S–90° S. The bias of both the background and analysis in the upper stratosphere reduces from 4–8 K in the control run to 2–4 K in the test run. The standard deviation of the background departures also reduces compared to the control experiment by ~2 K in this region. Such reductions appear at all latitudes outside the tropics, which suggest that the information introduced through MIPAS is retained in the system.

The ECMWF analyses were also compared against independent data from HALOE. These results generally support the findings from the comparisons with MIPAS observations described above. The bias of the analysis against HALOE, and its standard deviation, is reduced above ~5 hPa when MIPAS data are assimilated, although the reduction in standard deviation is only a few tenths of degree K. These results have the caveats that the number of available HALOE profiles for the comparison is much
smaller than for MIPAS and that at levels below ~6 hPa the HALOE temperature data are not retrievals but NCEP analyses.

MIPAS assimilation experiments were also run with the Met Office 3D-Var system for August 2003 (here, no MIPAS humidity data were assimilated). In contrast to the ECMWF results, the assimilation of MIPAS temperature results in a slight degradation in the Met Office temperature analyses. This can be seen in Fig. 4, where the analyses from this experiment and a similar one in which no MIPAS temperatures are assimilated, are compared with HALOE data. The bias with respect to HALOE generally increases in the upper stratosphere and mesosphere with MIPAS temperature assimilation, especially in the northern hemisphere mid latitudes. Standard deviations of these biases (not shown) are also larger.

A clue as to why this degradation takes place in the Met Office analyses comes from examination of operational Met Office temperature analyses between August and November 2003. In October 2003, the Met Office forecast model changed from an Eulerian dynamical core to a semi-Lagrangian one (Davies et al., 2005). It was necessary to change the background error covariances when the dynamical core was changed. After the change, a clear degradation appears in the analysis profile near 10 hPa, in the form of a jagged cold bias, together with a slight increase in the jaggedness of the profile at other levels. This suggests that the background error covariance used may be the source of the problem. The covariances are generated using the NMC method (Parrish and Derber, 1992), but for the new model the variances had to be constrained above the 10 hPa level in order to make the assimilation system more robust. In retrospect, it appears that the way in which this was done may have introduced noise into the background error covariances near the 10 hPa level.

This noisiness may explain the degradation in the analyses when MIPAS temperatures are used, since jaggedness in the individual profiles mentioned above is maintained or sometimes worsened when MIPAS temperatures are assimilated. To test this, another pair of experiments was performed using background error covariances calculated using the method described by Polavarapu et al. (2005) (the so-called Canadian
Quick, CQ, covariances). The CQ covariances are much smoother than the NMC covariances. Results from experiments with and without MIPAS temperatures and using the CQ covariances are also shown in Fig. 4. Above 10 hPa the biases for the CQ runs are smaller than for the corresponding NMC runs. In addition, with CQ covariances, the impact of the MIPAS temperature assimilation is neutral or slightly positive at almost all locations and the clear negative impacts seen in northern hemisphere mid latitudes with the NMC covariances have been removed.

In summary, the results from ECMWF indicate that assimilation of MIPAS temperature retrievals helps to reduce biases in the temperature analyses above \( \sim 5 \) hPa, whereas little impact is found below \( \sim 10 \) hPa. This seems reasonable, since at upper levels the analysis is less well constrained from radiosondes or satellite radiance observations than below \( \sim 10 \) hPa, and the bias correction of the higher-peaking satellite channels is also less well characterized than at lower levels. The results from the Met Office underline the importance of using appropriate background error covariances in a DA system. When potential problems were identified and addressed, by replacing the NMC covariances with CQ covariances, the degradation caused by the assimilation of MIPAS temperatures vanished. The absence of the positive impact seen at upper levels in the ECMWF analyses may perhaps be related to further differences in background error covariances at these levels in the two systems, or perhaps different biases in the background fields.

### 2.1.3 Limb radiance assimilation

Direct radiance assimilation has been very successful for the assimilation of temperature and humidity information from nadir sounding instruments, and it is therefore used operationally at most major NWP centres (e.g. Saunders et al., 1999; McNally et al., 2006). Prompted by this success, the ASSET project applied for the first time the radiance assimilation framework to the assimilation of limb radiances. Variational data assimilation schemes allow the direct assimilation of radiances, without the additional step of performing retrievals off-line (e.g. Andersson et al., 1994). The main
advantages are that radiance assimilation provides a fairly generic framework for using radiances together with all other observations and the latest background data, such that the analysis/retrieval problem is better constrained. Radiance assimilation also avoids the need to account for complex error characteristics in retrievals, arising, for instance, through the use of a short-term forecast as a priori, or from other assumptions in the retrieval algorithm. At the same time, a challenge commonly encountered in radiance assimilation is that model-simulated and observed radiances almost always show systematic deviations or biases (e.g. Harris and Kelly, 2001). While some of these may be due to biases in the model fields, a large proportion is usually attributed to so-called “radiance biases”, i.e., biases arising from errors in the instrument characterization (e.g. spectral response functions), the calibration, the spectroscopy, or other aspects of the forward model. As DA schemes assume unbiased data, such biases need to be removed. The development of suitable ways to account for these biases is an area of active research in the case of nadir radiance assimilation (e.g. Dee, 2004).

Assimilation of limb radiances was applied to clear-sky emitted infrared radiances from MIPAS in the ECMWF system (e.g. Rabier et al., 2000). This required the development of a fast radiative transfer model, its tangent linear and adjoint (Bormann et al., 2005), the selection of a suitable subset of MIPAS data to be used in the assimilation (Bormann and Healy, 2006), and the modification of the ECMWF system to be able to deal with the limb-viewing geometry. The study also touched on a number of other novel and experimental aspects, such as the extraction of ozone information through direct assimilation of ozone-affected radiances, and first experiences with combined tropospheric and stratospheric humidity analyses, made possible through work by Hólm et al. (2002). The developments for the radiance assimilation also allow continuous monitoring of MIPAS radiances against model equivalents, opening, for example, new possibilities of characterizing the temporal stability of the instrument. Results of the assimilation of MIPAS radiances are described in detail in Bormann and Thépaut (2006) and Bormann et al. (2006); here we only give a brief overview of the main findings.

First trials over a period of 43 days in August/September 2003 demonstrate the
feasibility of assimilating information from the MIPAS instrument by means of direct radiance assimilation. In these experiments we use an observation operator which assumes local horizontal homogeneity for the radiative transfer calculations, and assimilate clear-sky radiances from 260 selected pseudo-channels over channel-specific tangent altitude ranges. Overall, the assimilation is drawing well to the limb radiances, without significantly degrading the fit to other observations. This is an important result as it suggests that the MIPAS data and our assimilation approach are consistent with the rest of the observing network and its use in the DA system.

The assimilation has a considerable impact on the mean temperature, humidity, and ozone analyses in the stratosphere, upper troposphere, and lower mesosphere (Fig. 5). For instance, mean differences between analyses with and without MIPAS radiances show oscillating structures in the vertical, especially over the higher latitudes. Independent HALOE retrievals agree well with the features present in the assimilation with MIPAS radiances (but the HALOE data provide only limited coverage of 60° N–71° N; Equator–45° S for the period in question). The radiance assimilation corrects erroneous temperature oscillations in the analyses; such oscillations are often referred to as “stratospheric ringing”, and they are a common problem in stratospheric data assimilation. Also, comparison with HALOE, SAGE II or POAM III data suggests that the radiance assimilation corrects a dry bias otherwise present throughout the stratosphere (Fig. 6a). For ozone, results are more mixed. Comparisons to independent data indicate improved ozone fields over the 60° N–90° N region (Fig. 6b), but with mean increments which are too broad in the vertical over the tropics (not shown).

Our results show considerable sensitivity to how the correction of biases in MIPAS radiances is handled. The presence of radiance biases in MIPAS data becomes apparent when MIPAS radiances are assimilated without bias correction. In this case, inconsistencies occur in the analysis biases for some MIPAS radiances whose weighting functions peak at similar altitudes, suggesting that these biases cannot be accounted for by biases in the model fields. The development of a bias correction for MIPAS is hampered by the presence of biases in the model fields. This makes it inappropriate
to assume that the model fields without MIPAS data are unbiased when deriving the bias correction, as is commonly done for nadir radiances (Harris and Kelly, 2001). This is especially true for channels sensitive to humidity or ozone in the stratosphere. The situation is very different to the case of nadir radiances sensitive to temperature over the much better observed troposphere, for which in situ data such as radiosondes and the presence of many different observations make cross-calibration of biases more feasible. Current experimentation with MIPAS radiances uses the so-called $\gamma/\delta$-method which scales optical depths in the radiative transfer model with a channel-specific $\gamma$, and models the remaining bias with a constant $\delta$ (Watts and McNally, 2004). More details on the derivation of the bias correction and its influence can be found in Bormann et al. (2006).

The assimilation of MIPAS radiances has also been extended to the use of a 2-D radiative transfer model (Bormann and Healy, 2006) which takes into account the horizontal gradients within the limb viewing plane (Bormann et al., 2006). Our studies show that the assumption of horizontal homogeneity can introduce a considerable forward model error. Use of a 2-D radiative transfer model leads to smaller “First Guess” departures for radiances from lower tangent altitudes or strongly absorbing channels. These translate into smaller analysis increments, especially for ozone and humidity around the polar vortices. However, statistics against independent observations show only relatively small improvements compared to using a 1-D radiative transfer model, mainly in the UTLS region.

The developments for the MIPAS limb radiance assimilation allow a range of different aspects to be studied in greater detail, and some of these will be further described in upcoming papers. This includes a comparison of assimilating retrievals versus assimilating radiances. Also, the developments could be adapted for other passive limb-sounding instruments, such as the Microwave Limb Sounder (MLS) on the EOS Aura satellite.
2.2 Stratospheric distribution of chemical species

We focus on the nitrogen oxides NO$_x$ and NO$_y$, which are of primary importance in controlling stratospheric ozone amounts. In the middle stratosphere, reactions involving NO$_x$ and NO$_y$ form the primary catalytic O$_3$ destruction cycle. In the lower stratosphere, NO$_x$ radicals moderate O$_3$ destruction by combining with hydrogen (HO$_x$) and halogen (ClO$_x$, BrO$_x$) radicals involved in catalytic ozone destruction. It is therefore important to understand quantitatively NO$_x$ chemistry in order to estimate the chemical ozone budget and assess the impact of NO$_x$ perturbations from, e.g., aircraft emissions or increasing N$_2$O emissions on stratospheric ozone levels. An additional motivation for assimilating short-lived species such as NO$_x$ is the strong temperature-dependence of their chemistry. In the same way as the variability of tracer fields can provide information on winds, the variability of short-lived species could provide information on temperature. One possible application is the use of temperature as a control variable in a chemical DA system. Note that, up to now, the ability to extract temperature information from chemical measurements has not been tested.

The assimilation of short-lived chemical species (chemistry timescales $\ll$ transport timescales) such as NO$_x$ is more challenging than the assimilation of chemical tracers (chemistry timescales $>$ transport timescales). This is because concentrations of short-lived species vary on timescales from less than a minute to one day, and hence detailed treatment of fast chemistry is required for simulating this variability. This added level of complexity partly explains why DA systems for short-lived chemical species are less common than for tracers.

We present summaries of two case studies that illustrate work done on the assimilation of NO$_x$ data within ASSET; a subsequent publication will describe an intercomparison of NO$_2$ analyses from CTMs participating in the ASSET project. The first study is a test of NO$_x$ chemistry using a 4D-var photochemical assimilation system (Marchand et al., 2003). The second study concerns preliminary results in the extraction of temperature information from NO$_x$ observations. Both studies use height-resolved constituent
data from the GOMOS instrument on board Envisat. GOMOS is a stellar occultation instrument making mostly night-time measurements of NO$_2$, O$_3$ and, for the first time, NO$_3$ (Marchand et al., 2004). The nitrate radical (NO$_3$) is an important intermediate in the establishment of the partitioning between the NO$_x$ and NO$_y$ reservoir species. The NO$_x$ night-time chemistry is understood to be relatively simple in the stratosphere, with NO$_3$ concentrations controlled by temperature, and by the NO$_2$ and O$_3$ concentrations.

In the first study, GOMOS chemical data are assimilated into a CTM forced with ECMWF analyses. The objective is to test our understanding of NO$_x$ night-time chemistry. The stratospheric photochemical scheme is standard and takes into account heterogeneous chemistry on sulphuric acid particles assuming a background aerosol loading. Rate constants for the photochemical reactions are as in Sander et al. (2003). The CTM (a photochemical box model) is coupled to a 4D-Var scheme that assumes that the model is perfect.

O$_3$ and NO$_2$ measurements from GOMOS are assimilated simultaneously. In all cases studied, the analysed O$_3$ and NO$_2$ fields match the corresponding O$_3$ and NO$_2$ GOMOS measurements within 8.4 $\times$ 10$^{-5}$% and 0.0093%, respectively. The analysed NO$_3$ field (i.e., NO$_3$ calculated by the model after assimilation of GOMOS NO$_2$ and O$_3$ data) is then compared to the corresponding GOMOS NO$_3$ measurement. A correlation plot of GOMOS NO$_3$ versus analysed NO$_3$ is shown for 296 tropical cases in Fig. 7. The linear regression slope is 0.98 ($\pm$0.04) which is not statistically different from the 1-1 line. Since NO$_3$, NO$_2$ and O$_3$ are known to be strongly coupled chemically, this good agreement suggests that the NO$_x$ chemistry scheme in the photochemical model is essentially correct and that this set of GOMOS measurements is chemically self-consistent. More details can be found in Marchand et al. (2004).

CTMs calculate the chemical composition of the stratosphere from chemical rate constants, and from off-line winds and temperatures from meteorological analyses. It is interesting to see to what extent the reverse procedure is possible, i.e., deriving winds and temperatures from the chemical composition of the stratosphere. Some fast chemistry processes are very sensitive to temperature changes, e.g., NO$_3$ night-time
chemistry. For a 1 K change, the NO$_3$ concentration typically changes by more than 10%. This suggests that temperature information could be inferred from changes in the NO$_3$ concentration.

In the second study, we derive temperature information directly from night-time NO$_3$ and O$_3$ concentrations assuming steady-state conditions for NO$_3$. This steady-state hypothesis is generally valid in the lower and mid stratosphere. Using recommended values for the chemical rate constants, we derive temperature fields and evaluate them against ECMWF temperatures. The biases between these temperatures and GOMOS-derived temperatures are found to be less than 5 K throughout most of the lower and mid stratosphere. It is also possible to eliminate most of the bias by varying slightly the reaction rate constants within the errors reported for the recommended kinetic data (Sander et al., 2003). Figure 8 shows an example for several months of GOMOS data in 2003, after adjustment of the chemical rate constants. Generally, the biases are very small and the root-mean-squared (RMS) statistic of the biases ranges between 2 K and 6 K depending on the star and period considered. The brighter the star, the smaller is the error in the GOMOS measurements and, thus, the smaller is the RMS statistic of the temperature bias. This result suggests that NO$_X$ data contain useful temperature information.

2.3 Objective evaluation of ozone analyses

Given the attention ozone has received over the past decade, both for NWP and for studies of chemical distributions in the stratosphere (Rood 2005), ASSET undertook an ozone intercomparison project (ASSIC) to provide an objective evaluation of ozone analyses produced using different DA systems and techniques. Most systems assimilated a common ozone observational dataset, i.e., MIPAS, though some assimilated SCIAMACHY. The resulting analyses were evaluated by comparison against each other, ozone analyses from outside the project, and independent observations (i.e., not assimilated). In this section, we describe the highlights of the ASSIC project and put them into context. Further details can be found in Geer et al. (2006a). One strong
motivation for the ASSIC project is that, by confronting these various DA systems and DA techniques with the newly available Envisat observations, it is possible both to gain an understanding of their strengths and weaknesses, and to make new developments. Such an intercomparison also provides insight into ozone assimilation strategies, which we discuss later. There have been a number of previous intercomparisons between the ozone fields in different CTMs (e.g. Bregman et al., 2001; Roelofs et al., 2003), but to our knowledge the ASSIC project is the first time ozone analyses have been compared.

The ASSIC project involved eleven sets of analyses from seven different DA systems (two systems based on NWP GCMs, and five systems based on CTMs; see Table 1). The NWP models did not include feedback between the analysed ozone field and the radiation field. The DA systems are summarized in Table 1. Figure 9 provides a colour key of the different systems whose analyses are evaluated in Figs. 10–11. If ozone chemistry is included in the assimilation system, it is done either by highly detailed photochemical schemes, or via a parametrization, often known as a Cariolle scheme (e.g. Cariolle and Déqué, 1986). A Cariolle scheme is a linearization of ozone photochemistry around an equilibrium state, using parameters derived from a more detailed chemical model.

Most of the analyses focus on the stratosphere, but the scope of the ASSIC project spans from the troposphere to the mesosphere. Analyses are interpolated from their native resolution onto a common grid and then compared to independent ozone data from HALOE (Russell et al., 1993; http://haloedata.larc.nasa.gov), ozonesondes (Komhyr et al., 1995; Thompson et al., 2003a, b) and TOMS (Total Ozone Mapping Spectrometer; McPeters et al., 1998), and to MIPAS (Fischer and Oelhaf, 1996; ESA, 2004). The comparison was done on a common grid rather than in observation space for organizational reasons; Geer et al. (2006a) show that the error incurred in this approach is not significant. Most data, figures, and code are publicly available via the project website (http://darc.nerc.ac.uk/asset/assic).

Analyses were compared for the period July–November 2003. This period was chosen because of the availability of high quality MIPAS ozone data (Geer et al., 2006a;
Lahoz et al., 2006; Raspollini et al., 2006). Statistics were built up from the difference between analyses and observations \((a \sim o)\) differences. The statistics were binned into the following regions: 90° S–60° S; 60° S–30° S; 30° S–30° N; 30° N–60° N; 60° N–90° N. Statistics were binned monthly; also for the entire period 18 August–30 November 2003 (before 18 August 2003, the DARC analyses were not adequately spun up).

Because ozone amounts vary by several orders of magnitude throughout the atmosphere, the statistics were normalized with respect to the climatology of Fortuin and Kelder (1998) in the stratosphere and Logan (1999) in the troposphere, and displayed as a percentage. In this way, all regions in the atmosphere are given approximately equal weight.

Figure 10 shows that, through most of the stratosphere (50–1 hPa) \(a \sim o\) biases are usually within ±10% compared to the HALOE instrument. Similar results are obtained against ozonesonde data for levels in the lower stratosphere, 100–10 hPa (not shown). Biases and standard deviations in the \(a \sim o\) differences are larger in the UTLS, in the troposphere, the mesosphere, and the Antarctic ozone hole region. In these regions, some analyses do substantially better than others, and this is mostly due to differences in the models. At the tropical tropopause, many analyses show positive biases and excessive structure in the ozone fields, likely due to shortcomings in assimilated tropical wind fields and a degradation in MIPAS data at these levels. In the troposphere (below 100 hPa) some analyses show quite substantial biases compared to ozonesonde observations (not shown). No ozone profiles below ∼400 hPa are assimilated in any of the analyses, and only one analysis is designed to assimilate tropospheric chemical species (MOCAGE-PALM Reprobus – Table 1).

In the Southern Hemisphere ozone hole, only the analyses which correctly model heterogeneous ozone depletion are able to reproduce the ozone destruction over the Pole (Fig. 11). These analyses are those using Cariolle scheme versions 1.2/2.1, which include a term to take account of heterogeneous chemistry (e.g. operational ECMWF system with MIPAS; MOCAGE-PALM Cariolle scheme – see Table 1), or those using comprehensive chemical schemes (e.g. MOCAGE-PALM Reprobus; BASCOE v3q33...
There are two other points worth mentioning: First, ECMWF operational analyses only capture the full ozone depletion during October when they began assimilating MIPAS ozone data for the first time, the benefit coming from the relatively high vertical resolution of MIPAS, and the fact that before this only limited ozone data were assimilated (Table 1); Second, the original BASCOE analyses (v3d24) were found to perform poorly in the ozone hole in the intercomparison – the scheme was improved and the new analyses (v3q33) performed better, demonstrating the value of intercomparisons for identifying shortcomings. Most of the analyses (except KNMI TEMIS, who assimilate SCIAMACHY total column ozone data) at this level show too high ozone in November compared to ozonesondes (Fig. 11). This is thought to be due to the relatively broad resolution of MIPAS and SCIAMACHY ozone profile data, thus the analyses show an influence from the much higher ozone amounts at levels above, where the polar vortex has already broken down.

In the upper stratosphere and mesosphere (levels above 5 hPa) some implementations of linear ozone photochemistry schemes cause large biases. However, these are easily remedied (Geer et al., 2006b). The diurnal cycle of mesospheric ozone is not captured, except by the one system that includes a detailed treatment of mesospheric chemistry (BASCOE – Table 1).

The conclusions drawn from the ASSIC project are that, in general, with current DA systems, in regions of good data quality and coverage, similarly good ozone analyses are obtained regardless of the DA method (KF; 3D- and 4D-Var; 3D-FGAT; direct inversion), or the model (NWP or CTM). This reflects the generally good quality of the MIPAS ozone observations. There were areas where some models performed better than others, and in general the improved performance could be explained by better modelling of transport and chemistry.

One way of understanding the behaviour of the ozone analyses is to compare ozone photochemical relaxation timescales with the timescale on which Envisat observations are available. Ozone has photochemical relaxation timescales of $O(1$ month) in the UTLS, and $O(1$ day) in the upper stratosphere; in the mesosphere it has a diurnal
cycle. Envisat provides observations with daily global coverage, but the revisit cycle at a particular location is \( O(\text{days}) \).

Thus, chemical biases can in principle be adjusted by the DA system in the UTLS, and the chemistry representation is not generally crucial (although the ozone hole is an exception due to the role of heterogeneous chemistry – see Geer et al., 2006a). In the upper stratosphere (10–1 hPa), as chemical timescales become faster, the influence of the ozone observations in a DA system can become increasingly limited, and the chemistry representation becomes increasingly important. Thus, it may be argued that in this region it is better to exclude chemistry than to model it inappropriately. However, this will only work where observational data coverage is very good. As an example, the Juckes and MIMOSA systems, which do not incorporate chemistry, compare as well to independent data in the upper stratosphere as the BASCOE analyses, which incorporate a detailed chemistry model. Many of the analyses that use the linearized Cariolle scheme approach to chemistry, such as those from DARC/Met Office and ECMWF, show much poorer agreement with independent data. These problems are mainly due to avoidable biases in the schemes, but there are also fundamental limitations with the approach. Hence, in the presence of the generally good quality and coverage of the MIPAS ozone observations in the upper stratosphere, the assimilation systems with no chemistry at all can do better than those with linearized chemistry.

Analyses based on SCIAMACHY total column ozone show a similar performance to the MIPAS analyses. The KNMI TEMIS set up (Segers et al., 2005a) produces reasonably realistic ozone profile shapes and stratospheric variability. However, significant model biases were detected in the upper stratosphere, mainly resulting from the way the Cariolle scheme is implemented (Geer et al., 2006a). Separately, SCIAMACHY limb profiles (Segers et al., 2005b) were assimilated, although they were only available in quantity in October and November during the intercomparison period. Hence, results from these analyses do not appear in Fig. 10, and the low bias of these analyses in Fig. 11 in August and September can only be attributable to biases in the KNMI model background. During October and November the SCIAMACHY profile analysis was
clearly worse than analyses that used MIPAS profiles only in the northern hemisphere lower stratosphere, where comparison with ozonesonde data (not shown) showed a 20% negative bias. This bias is likely due to problems in the shape of the SCIAMACHY profiles.

Using the analyses to compare MIPAS and independent observations indirectly, and treating MIPAS observations as point retrievals, we can evaluate MIPAS ozone data: it is \( \sim 5\% \) higher than HALOE ozone above 30 hPa, and \( \sim 10\% \) higher than ozonesondes and HALOE ozone in the lower stratosphere (100–30 hPa). This indicates MIPAS ozone data has a positive bias of 5–10\% in the stratosphere and mesosphere against ozonesondes and HALOE data.

The ASSIC project also provides some clues on ozone assimilation strategies. The choice of system for ozone DA depends on a range of considerations, including history, familiarity with a particular approach or cost. If a central concern is the production of weather forecasts, it makes sense to add ozone to enhance a pre-existing NWP system (e.g. ECMWF). While, if the main focus is on chemistry, there are strong arguments to build ozone assimilation into a CTM with sophisticated chemistry, taking the meteorological input as given (e.g. BIRA-IASB). If the focus is primarily on ozone itself, an alternative approach is to use a transport model, driven by pre-existing dynamical fields, in combination with a simplified chemistry scheme. An example is KNMI, who have developed semi-operational SCIAMACHY ozone analyses and forecasts (http://www.temis.nl).

The ASSIC results suggest that for current DA systems, provided there is good data quality and coverage, similar quality of ozone analyses is obtained whether one uses a DA system based on an NWP model or a CTM. However, the cost of the different DA systems can be an issue. For example, GCM-based analyses require substantially more computer power than the CTM approach, though ozone assimilation is a relatively small additional cost when included in an existing NWP system (see, e.g., Geer et al. 2006a).
2.4 Distribution of tropospheric pollutants

As part of the ASSET project, the spatial and temporal distribution of tropospheric pollutants was studied. A strong motivation is the need to monitor and forecast air quality, for example, as part of the EC/ESA Global Monitoring for the Environment and Security (GMES) initiative. Because the assimilation of tropospheric constituents provides an objective framework for monitoring and forecasting tropospheric pollutants, one ASSET partner (University of Köln) studied the impact of assimilating tropospheric constituents from Envisat into a DA system based on a CTM. The results from this study are summarized in this section.

The assimilation of tropospheric constituents has to take account of two issues: the general paucity of global tropospheric measurements from satellites, and the role played by emissions.

Most satellite instruments are currently unable to provide height-resolved profiles of constituents below tropopause levels. Furthermore, the presence of clouds often prevents retrievals of tropospheric columns of constituents, even when spectral characteristics of constituent absorption theoretically make this possible. For example, although MIPAS can theoretically retrieve information down to heights of 7 km in the absence of clouds (Spang et al., 2005), its ability to sound the troposphere is severely impaired by the presence of clouds. In the absence of clouds, MIPAS profiles (e.g. of ozone and HNO$_3$) are available for the upper half of the troposphere, and SCIAMACHY provides integrated tropospheric column information: e.g., global fields of NO$_2$ and SO$_2$ (see http://www.temis.nl).

Besides limited (in space and time) satellite tropospheric measurements, there are also surface measurements of the troposphere, mostly from in situ instruments operated routinely by the environmental protection agencies, and generally deployed only in populated regions. There is also a very sparse network of ozone radiosondes. Airborne in situ data are provided by the MOZAIC (Measurement of ozone and water vapour by Airbus in-service aircraft) project, with routine observations of ozone, water
vapour and CO concentrations from a limited number of commercial aircraft.

In contrast to stratospheric constituent DA and general meteorological DA, the evolution of the tropospheric model state is not primarily controlled by the initial state. Instead, emissions are a strong controlling factor, and exert a direct influence over short timescales (ranging from seconds to days). Furthermore, currently, emission rates are not sufficiently well known. Thus, emission rates must be considered as another parameter to be optimized in the data assimilation process. Tropospheric DA must also take account of the differences in spatial scale between satellite data retrievals and point-like emissions.

In 4D-Var, where a cost function penalizes the discrepancies between observations and an a priori state, a generalization with respect to emissions thus needs to be implemented. This can be done in the incremental formulation of 4D-Var by augmenting the state vector by including deviations from an emission inventory, as well as deviations from a background chemical state (see, e.g., Elbern and Schmidt, 2001, 2002).

A 4D-Var experiment for European-scale tropospheric gas phase chemistry has been performed at the University of Köln for the period June–November 2003, with the June simulation providing the CTM spin-up period. The DA system, including the formulation of the background error covariances, is described in Nieradzik and Elbern (2006). The DA configuration has a time window of four hours (08:00–12:00 UTC) to include Envisat satellite data with a late morning overpass over Europe. After assimilation, a 24-h forecast is made, starting at 08:00 UTC; the analysis produced by the assimilation is the initial field, and an emission rate correction factor is applied. Numerical experiments suggest that ∼12 iterations are sufficient to ensure convergence to the observations. With this configuration, 4D-Var is ∼5 times faster than real time. After assimilation, an a posteriori analysis is performed (Talagrand, 2003).

We now discuss two examples that illustrate the benefits of the assimilation of Envisat tropospheric data. In the first example, MIPAS retrievals for the upper troposphere (Mengistu Tsidu et al., 2004; Glatthor et al., 2006) are assimilated. Figure 12 shows first guesses and analyses for HNO₃ and O₃. Because 4D-Var includes the time dimen-
sion, the temporal discrepancy between the initial state (08:00 UTC) and the MIPAS observations (10:30 local solar time) is accounted for. The top scatter plot indicates poorly modelled HNO$_3$ variability prior to assimilation, likely due to insufficient interaction between the tropospheric model and the stratosphere, a source region for NO$_y$. After assimilation, this is corrected. Similarly, the bottom scatter plot indicates that for ozone, the analyses are an improvement on the model forecast.

In the second example, SCIAMACHY NO$_2$ column data from KNMI/BIRA-IASB are assimilated together with surface in situ observations (Fig. 13). The observation operator is based on the averaging kernels provided with the retrieval (Eskes and Boersma, 2002). The impact of the SCIAMACHY data can be observed over the Belgium-Netherlands region in Fig. 13. However, to interpret Fig. 13, it should be borne in mind that the satellite overpass time (approximately 09:30 local solar time) and the analysis time (08:00 UTC) are different, and that NO$_2$ is highly variable over a timescale of 2–3 h. Furthermore, local in situ observations that are assimilated also have an impact on the analyses. Comparison of the model forecast and the analyses against the SCIAMACHY data show that the analyses are an improvement over the forecasts.

This section describes initial efforts to assimilate tropospheric constituents at the University of Köln. To our knowledge, these efforts are the first of their kind. Subsequent publications will describe the extension of the 4D-Var method to include other research satellite data, as well as other important tropospheric species, e.g., aerosols.

2.5 Enhanced retrievals of MIPAS data

MIPAS measures the atmospheric limb emission spectrum in the frequency interval 680–2410 cm$^{-1}$ over the altitude range 6–68 km. The spacing between measurements is $\sim$3 km through the stratosphere, but larger above. MIPAS typically samples a volume $\sim$3 km in the vertical, $\sim$30 km in azimuth and $\sim$300 km along the line of sight. Precise values will differ slightly with species and between retrievals: the true vertical resolution is described by an averaging kernel (Rodgers, 2000). The MIPAS measured spectra are analysed by the ESA ground processor to provide height-resolved profiles.
of pressure, temperature, and six key atmospheric species (known as “target species”): H$_2$O, O$_3$, HNO$_3$, CH$_4$, N$_2$O and NO$_2$. Of these MIPAS measurements, humidity, temperature (Sect. 2.1) and ozone (Sect. 2.3) have been assimilated into NWP and CTM systems, and several tropospheric constituents, including ozone, HNO$_3$ and NO$_2$ have been assimilated into a CTM system (Sect. 2.4).

The MIPAS spectra also contain information on several additional species (known as “non-target species”), which are of interest for studies of the distribution of chemical species in the stratosphere and troposphere, and for assimilation into CTM systems. We now discuss work in ASSET to provide enhanced retrievals of MIPAS data, with particular attention to non-target species.

The ESA ground processor that provides the target MIPAS species is based on an optimized retrieval model, ORM (Ridolfi et al., 2000; Carli et al., 2004) explicitly developed to provide accurate results in near real time. The ORM implements a global fit algorithm (Carlotti et al., 1988) by using approximations that speed up the retrieval process and meet very stringent run-time requirements. Because of these approximations, some components of the systematic (i.e., bias) error may become significant; for example, the assumption of a horizontally homogeneous atmosphere may affect the error budget. In the presence of strong horizontal gradients along the instrument line of sight, and whenever the atmospheric state is far from climatology, the ORM will tend to produce inaccurate retrievals.

To overcome limitations in the ESA ground processor, three novel retrieval algorithms have been developed to analyse MIPAS species with improved accuracy: (1) a modified ORM with the capability for retrieving non-target species; (2) a 2-D retrieval algorithm (Geo Multi-Target Retrieval, GMTR) with the capability to simultaneously retrieve several species; and (3) a retrieval processor including several enhancements to the ESA ground processor. Algorithms (1)–(3) were developed at CNR-IFAC, the University of Bologna and IMK, Karlsruhe, respectively. We now discuss these algorithms in turn.

The ORM was modified to allow the retrieval of the following MIPAS non-target species: CFC-11, CFC-12, ClONO$_2$ and N$_2$O$_5$. Each species is retrieved sequentially
after the retrieval of pressure, temperature and the six MIPAS target species. The retrieved profiles are characterized by their variance-covariance matrix (produced by the retrieval algorithm) and by their averaging kernels, available from the CNR-IFAC website: http://www.ifac.cnr.it/retrieval/auxiliary.html. There are plans to assimilate these data into the BASCOE DA system (Table 1).

Figure 14a shows global maps of MIPAS N$_2$O$_5$ retrieved by the modified ORM averaged over the period 20 September–4 October 2003. The maps in Fig. 14a refer to averages over day and night conditions at altitudes of 24 km, 33 km and 39 km. The N$_2$O$_5$ fields have accuracy (random plus systematic error) of $\sim$25%.

The GMTR algorithm (Carlotti et al., 2001, 2006; Dinelli et al., 2004) is an enhancement on the ORM. The enhancements are: (i) the capability to model the horizontal variability of the atmosphere and perform a 2-D retrieval along a full MIPAS orbit using tomographic inversion; (ii) the capability to simultaneously retrieve several species and atmospheric state parameters.

These enhancements help reduce the systematic error in the retrievals. The former adds the capability to model atmospheric horizontal structures and improves the retrieval accuracy in the presence of strong horizontal gradients (e.g. at the polar vortex edge), while the latter helps avoid systematic errors arising from spectral interference between species.

The GMTR algorithm also allows a suitable compromise between the spatial resolution and the precision of the retrievals (Ridolfi et al., 2004). In particular, it has been shown that it is possible to double the horizontal resolution of the 2-D retrieval with respect to the standard 1-D retrieval if one accepts a factor of two degradation in the retrieval precision.

Figure 14b shows the comparison between MIPAS HNO$_3$ fields retrieved by ORM and GMTR, at 24 km altitude and averaged over the period 20 September–4 October 2003. The largest differences between the two algorithms are in regions where atmospheric meridional variability is greatest (e.g. near the Antarctic polar vortex). The HNO$_3$ fields (ORM and GMTR) generally have an estimated accuracy of 6%, but in the
region at the edge of the polar vortex, where meridional variability is large, the GMTR-derived fields are slightly more accurate than the ORM-derived fields (of order tenths of a percent).

In contrast to the ORM and GMTR algorithms, the MIPAS IMK processor is designed to retrieve as many chemical species as possible. The IMK processor has also been used to extend the altitude range of MIPAS retrievals; it now includes the upper troposphere and the mesosphere/thermosphere. These upper troposphere retrievals have been used for studies of the distribution of tropospheric pollutants (Sect. 2.4).

The main differences between the IMK data processor and the ORM algorithm are discussed in von Clarmann et al. (2003a); the former has a different radiative transfer forward model (KOPRA; Stiller et al., 2000), constrained least squares inversion on a fixed fine altitude grid (von Clarmann et al., 2003b), a non local thermodynamic equilibrium treatment (Funke et al., 2005), and inclusion of cloud parameters (Höpfner et al., 2006a, b). The MIPAS IMK data products currently include tangent altitude, pressure, temperature, 19 constituents (including H$_2$O, O$_3$, CH$_4$, N$_2$O, HNO$_3$, NO$_2$, NO and N$_2$O$_5$), cloud coverage and cloud top altitudes, and the chemical composition of clouds. The MIPAS IMK products have been used to study interesting meteorological conditions such as the ozone hole split of September 2002 (Glatthor et al., 2005).

Figure 15 shows a field of MIPAS ozone retrieved at 8 km altitude with the IMK data processor, averaged over one week. The white colour indicates data gaps due to cloud coverage. Figure 15 shows that MIPAS ozone data can be used to track the transcontinental transport of pollution (note the relatively high values to the East of North America).

2.6 Distribution of ASSET datasets

The ASSET project has generated several data products from a variety of DA systems and satellite retrieval processors. Coming from different communities, the partners previously used different data formats and did not generally follow a common standard for metadata definition. To effectively disseminate the ASSET data products to a large
EO community and avoid users having to deal with various data formats and definitions, it was decided to implement a common project database using a unified data format. This would enable data sharing between partners during the project and provide a permanent open archive for the products after the end of ASSET.

The Envisat Cal/Val database at NILU was chosen as a platform for the ASSET database; it provided an established standard for data formatting, archiving and sharing. A metadata guidelines document (Bojkov et al., 2002) was available at the start of the project and its guidelines were adapted and expanded to address the needs of ASSET. The Envisat Cal/Val database code was also copied and developed into a stand-alone database.

A full technical description of the database has been given elsewhere (Vik et al., 2005), and only a brief overview is given here. Data and information about the data are available through a web interface, http://nadir.nilu.no/asset/. Tools for formatting data into HDF (Hierarchical Data Format) files are available at this website and files are checked for errors and consistency and automatically inserted into the database after data upload. The web portal enables users to search for data through a set of filter criteria and bulk downloads of selected data are possible. It is also possible to browse files for data content and visualize data on-line.

The datasets included in the ASSET database include a selection from the DA experiments that were performed as part of the project. The total volume of all data produced is very large and only analyses considered of interest for external users are stored in the database. Eventually, this will include total columns and values on pressure levels from all DA systems participating in the ASSIC project (Sect. 2.3). MIPAS non-standard products are also available (Sect. 2.5). The data files are self-descriptive, but additional information on the various products is available at the project website. Data use rules and requests for acknowledgments are given for each data file.

The ASSET database is open to the public and there is no need to register. Users have access to search for and download data, but a personal account is needed for those who wish to upload additional data to the system. The database will be main-
tained as part of NILU’s Atmospheric Database for Interactive Retrieval (NADIR).

3 Conclusions

The ASSET project has used data assimilation to bring together a broad range of EO expertise, including: NWP systems; chemical DA; numerical modelling; meteorology; retrieval theory; DA theory; EO measurements; data analysis; and data management. This expertise has been applied to several overarching themes involving the assimilation of Envisat data: enhancement of analyses derived from NWP models; studies of the distribution of stratospheric chemical species by assimilation into CTM systems; objective assessment of the quality of ozone analyses; studies of the spatial and temporal evolution of tropospheric pollutants; enhanced retrievals for study of the distribution of tropospheric and stratospheric chemical species, and assimilation into CTM systems; and data archival and dissemination. A particular strength of ASSET is that the inter-comparison of analyses has enabled us to learn much more about the strengths and weaknesses of each participants’s DA system than if each participant had assessed their own system independently.

The ASSET project has had a number of achievements, some of them firsts in the field:

– One of the first evaluations of the impact of assimilation of height-resolved troposphere-stratosphere humidity and temperature data into NWP and CTM systems, demonstrating a positive impact of assimilation of humidity. In contrast, temperature only showed positive impact above 5 hPa in the ECMWF system; below this level, there is much existing operational data, and impact is neutral. The study also showed the difficulty of getting appropriate background error covariances in the stratosphere, and the detrimental effect of temperature biases in the stratopause region.

– The first implementation of an algorithm to assimilate limb radiances into a NWP
system, demonstrating the feasibility of direct assimilation of limb radiances.

- One of the first evaluations of the impact of assimilation of NO\textsubscript{x} data into a CTM system, including a first demonstration of the feasibility of extracting temperature information from the assimilation of short-lived chemical species.

- The first intercomparison of ozone analyses, demonstrating that for current DA systems, in general, in regions of good data quality and coverage, similarly good results are obtained regardless of the assimilation method, or the DA system (NWP or CTM). In general, the first priority for improving ozone DA systems is improved modelling of chemistry and transport.

- A first study of the assimilation of height-resolved constituent data into a CTM system designed to study tropospheric pollutants, demonstrating the feasibility of extracting information on tropospheric pollution from Envisat data.

- Enhanced retrievals of MIPAS data, providing information beyond MIPAS target species.

- The construction of a database to disseminate the results from the ASSET project to the EO community.

ASSET has shown that DA is invaluable for the use of constituent measurements. Data assimilation not only enables one to fill in gaps between observations in an intelligent, and statistically justifiable, manner, but also allows the use of heterogeneous measurements. The incorporation of a numerical model allows information to be propagated forward in time, enabling the combination of measurements available at different times and locations. When properly applied, and as shown by the ASSET project, DA can thus add value to observations and models, compared to the information that each can supply on their own.

Based on its results and achievements, the ASSET project strongly recommends that ESA (and other space agencies) position data assimilation at the centre of their satellite
activities, from pre-launch mission planning to post-launch value-adding activities: the objective evaluation of the incremental value of future satellites (e.g. using Observing System Simulation Experiments, OSSEs); the objective evaluation of current satellites (e.g. using Observing System Experiments, OSEs); and adding value to satellite data (by combining information from observations and models).

Acknowledgements. WAL is funded by the NERC Data Assimilation Research Centre (DARC). The work described in this paper is funded by ASSET. ASSET is an EU-funded Framework V project (contract EVK2-CT-2002-00137). The COST 723 action (http://www.cost723.org) and the International Space Science Institute (ISSI; http://www.issibern.ch) provided support for this work, including funds for ASSET project meetings.

References

Bregman, A., Krol, M.C., Teyssedre, H., et al.: Chemistry-transport model comparison with

Carli, B., Alpaslan, D., Carlotti, et al.: First results of near real time retrievals from MI-


derived from measurements of SCIAMACHY on Envisat, Atmos. Chem. Phys. Discuss., 5, 4429–4475, 2005,
http://www.atmos-chem-phys-discuss.net/5/4429/2005/.


   http://www.atmos-chem-phys.net/6/1549/2006/.


Table 1. List of DA systems participating in the ASSIC project (Sect. 2.3). The Met Office, ECMWF MIPAS and BASCOE v3q33 systems were also used to study the assimilation of humidity and temperature, and the assimilation of limb radiances (Sect. 2.1). See Geer et al. (2006a, b) for more details of the DA systems. SBUV is the Solar Backscatter UV radiometer; GOME is the Global Ozone Monitoring Experiment.

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Winds</th>
<th>Scheme</th>
<th>Ozone observations</th>
<th>Ozone photochemistry</th>
<th>Heterogeneous ozone chemistry</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECMWF operational</td>
<td>NWP</td>
<td>GCM</td>
<td>4D-Var</td>
<td>SBUV, GOME total columns, MIPAS v4.59 from 07/10/03</td>
<td>Cariolle v1.2</td>
<td>T&lt;195 K term</td>
</tr>
<tr>
<td>ECMWF MIPAS</td>
<td>NWP</td>
<td>GCM</td>
<td>4D-Var</td>
<td>As ECMWF operational, but throughout ASSIC period</td>
<td>Cariolle v1.2</td>
<td>T&lt;195 K term</td>
</tr>
<tr>
<td>DARC/Met Office KNMI TEMIS</td>
<td>NWP</td>
<td>GCM</td>
<td>3D-Var</td>
<td>MIPAS v4.61, SCIAMACHY TOSOMI total columns (Eskes et al., 2005)</td>
<td>Cariolle v1.0</td>
<td>Cold tracer</td>
</tr>
<tr>
<td>KNMI SCIAMACHY profiles</td>
<td>CTM</td>
<td>ECMWF</td>
<td>Sub-optimal KF</td>
<td>SCIAMACHY profiles, IFE Bremen v1.6</td>
<td>Cariolle v1.0</td>
<td>Cold tracer</td>
</tr>
<tr>
<td>BASCOE v3d24 (BIRA-IASB)</td>
<td>CTM</td>
<td>ECMWF</td>
<td>4D-Var</td>
<td>MIPAS v4.61</td>
<td>57 species, Mesospheric chemistry</td>
<td>PSCBox</td>
</tr>
<tr>
<td>BASCOE v3q33 (BIRA-IASB)</td>
<td>CTM</td>
<td>ECMWF</td>
<td>4D-Var</td>
<td>MIPAS v4.61</td>
<td>As BASCOE v3d24</td>
<td>PSC parametrization</td>
</tr>
<tr>
<td>MOCAGE-PALM/Cariolle (Météo-France/CERFACS)</td>
<td>CTM</td>
<td>Arpege</td>
<td>3D-FGAT</td>
<td>MIPAS v4.61</td>
<td>Cariolle v2.1</td>
<td>T&lt;195 K term</td>
</tr>
<tr>
<td>MOCAGE-PALM/Reprobus (Météo-France/CERFACS)</td>
<td>CTM</td>
<td>Arpege</td>
<td>3D-FGAT</td>
<td>MIPAS v4.61</td>
<td>Reprobus (Lefèvre et al., 1994), Tropospheric chemistry</td>
<td>Carslaw et al. (1995)</td>
</tr>
<tr>
<td>MIMOSA (Service d’Aeronomie) Juckes (2006) (RAL)</td>
<td>CTM</td>
<td>ECMWF</td>
<td>Sub-optimal KF</td>
<td>MIPAS v4.61</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>CTM</td>
<td>ECMWF</td>
<td>Direct inversion</td>
<td>MIPAS v4.61</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>
Fig. 1. Monthly zonal mean specific humidity analyses for September 2003 for BASCOE (upper plot) and ECMWF (lower plot). MIPAS water vapour profiles have been assimilated in both cases. Blue denotes relatively low specific humidity values; red denotes relatively high specific humidity values. Units: ppmv.
Fig. 2. Met Office mean (left-hand plot) and standard deviation (right-hand plot) of specific humidity profiles for 25 September 2002 over the 60° S–90° S region. Black: MIPAS observations; red: analyses using a relative humidity (RH) control variable; blue: analyses using a normalized RH control variable; green: analyses using a normalized specific humidity control variable. Units: ppmv.
Fig. 3. Left-hand plot: Bias between MIPAS temperature retrievals and the ECMWF background (solid) and analysis (dotted) for the test experiment with assimilation of MIPAS retrievals (black) and the control experiment without assimilation of MIPAS data (grey). Statistics cover the region 60° S–90° S for the period 1–29 September 2003. Middle plot: As for left-hand plot, but for the standard deviation of the background and analysis departures. Right-hand plot: Number of MIPAS observations assimilated. Units for left-hand and middle plots: K.
Fig. 4. Mean Met Office minus HALOE analysed temperature for MIPAS temperature and ozone assimilation (green), MIPAS ozone assimilation (red), MIPAS temperature and ozone assimilation with “Canadian Quick” (CQ, see text) background error covariances (blue) and MIPAS ozone assimilation with CQ background error covariances (orange). Left-hand plot: 60° S–30° S; centre plot: 30° S–30° N, right-hand plot: 30° N–60° N. The numbers in brackets indicate the HALOE/analysis coincidences within each latitude bin. Units: K.
Fig. 5. (a) Zonal mean temperature differences between the ECMWF experiments with and without assimilation of MIPAS radiances. Contour interval is 0.5 K, with positive values shown by solid black contour lines and negative values shown by dashed grey lines. (b) Same as (a), but for humidity (relative to the experiment without MIPAS radiance assimilation), with a contour interval of 8%. (c) Same as (a), but for ozone volume mixing ratio with a contour interval of 0.1 ppmv. Positive values indicate that the experiment with assimilation of MIPAS radiances has higher values than that without assimilation of MIPAS radiances. Units: (a) K; (b) percent; (c) ppmv.
Fig. 6. Comparison of ECMWF analyses for the period 1–29 September 2003 against 195 SAGE II retrievals over the 60° N–90° N region. Solid lines show the retrieval minus analysis difference (positive differences indicate that the analyses have a negative bias against SAGE II data); dotted lines show the standard deviation of the differences between the retrieval and the analysis, both relative to the mean retrieval. Red lines indicate the experiment without assimilation of MIPAS data, black lines show statistics with MIPAS radiance assimilation. (a) Statistics for humidity analyses. (b) Statistics for ozone analyses. Units: percent.
Fig. 7. GOMOS NO$_3$ measurements as a function of analysed NO$_3$. Two periods are considered: (diamonds) 5–6 December 2002, and (filled circles) 30 January–2 February 2003. NO$_3$ measurement random errors are indicated by the vertical lines. The linear regressions for the different periods are given by the dashed lines. The 1-1 line is represented by the thin black line. Units: ppbv. Based on Marchand et al. (2004).
Fig. 8. Mean difference (crosses) between GOMOS-derived temperature and ECMWF temperature as a function of (top) altitude in km, and (bottom) latitude in degrees. The standard deviations (dotted vertical lines) are also provided. To aid visualization of these diagrams, 3 K standard deviations (horizontal bars) are also plotted. The GOMOS data is for the first four months of 2004. Units: K.
Fig. 9. Colour key used in Figs. 10–11.
Fig. 10. Mean of analysis minus HALOE differences, normalized by climatology, for the period 18 August–30 November 2003. See Fig. 9 for colour key. The numbers in brackets indicate the HALOE/analysis coincidences within each latitude bin. Units: percent. Based on Geer et al. (2006a).
**Fig. 11.** Comparison of analyses and ozonesondes at the South Pole, 68 hPa over the period August–November 2003. See Fig. 9 for colour key. Units: ppmm. Based on Geer et al. (2006a).
Fig. 12. Left hand plots: Results from the assimilation of IMK MIPAS retrievals for 8 November 2003, 47° N, 13° W; HNO₃ (top) and ozone (bottom). Observations are given by the red error bars, first guess by the black broken lines; analyses by the blue lines. Right hand plots: Scatter diagram showing performance of the assimilation; HNO₃ (top) and ozone (bottom). Red dots: first guess 24-h forecasts, based on an earlier assimilation cycle; blue dots: 4d-var analyses. Units: ppbv.
Fig. 13. SCIAMACHY tropospheric NO$_2$ columns for 12 September 2003 retrieved by KNMI and assimilated as averaging kernels. Top left panel gives the model equivalent for the first guess at 08:00 UTC; top right panel gives the 4D-Var analysis for the same time; bottom left panel gives the retrieved total columns and their footprint at approximately 09:30 local solar time. Red indicates relatively high values; blue relatively low values. Units: number of molecules.cm$^{-2}$. 

12822
Fig. 14. (a) Maps of MIPAS N$_2$O$_5$ retrieved by ORM, averaged over the period 20 September 4 October 2003 (and for day and night conditions) at altitudes of 24 km, 33 km and 39 km. Red indicates relatively high values; blue relatively low values. Units: ppmv; (b) Comparison between ORM and GMTR produced fields of MIPAS HNO$_3$ at 24 km altitude, and averaged over the period 20 September 4 October 2003. Red indicates relatively high values; blue relatively low values. Units: ppmv.
Fig. 15. MIPAS ozone at 8 km altitude as retrieved with the IMK data processor; the data has been averaged over one week in July 2003. The white colour indicates data gaps due to cloud coverage. Red indicates relatively high values; blue relatively low values. The statistical uncertainty of these weekly mean mixing ratios is $\sim 10$ ppbv, depending on the actual sample size. Units: ppbv.