A comparative analysis of UV nadir-backscatter and infrared limb-emission ozone data assimilation

Rossana Dragani

European Centre for Medium-Range Weather Forecasts, Shinfield Park, RG2 9AX, Reading, UK

Correspondence to: Rossana Dragani (rossana.dragani@ecmwf.int)

Abstract. This paper presents a comparative assessment of ultra violet nadir-backscatter and infrared limb-emission ozone profile assimilation. The Meteorological Operational Satellite A (MetOp-A) Global Ozone Monitoring Experiment 2 (GOME-2) nadir and the ENVISAT Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) limb profiles, generated as part of the European Space Agency Climate Change Initiative, were individually added to a reference set of ozone observations

5   and assimilated in the European Centre for Medium-Range Weather Forecasts (ECMWF) data assimilation system (DAS). The two sets of resulting analyses were compared with that from a control experiment, only constrained by the reference dataset, and independent, unassimilated observations.

Comparisons with independent observations show that both datasets improve the stratospheric ozone distribution. The changes inferred by the limb-based observations are more localized and, in places, more important than those implied by

10   the nadir profiles, albeit they have a much lower number of observations. A small degradation (up to 0.25 mg/kg for GOME-2 and 0.5 mg/kg for MIPAS) is found in the tropics between 20 and 30 hPa. In the lowermost troposphere below its vertical coverage, the limb data is found able to modify the ozone distribution with changes as large as 60%. Comparisons of the ozone analyses with sonde data show that at those levels the assimilation of GOME-2 leads to about 1 Dobson Unit (DU) smaller root mean square error (RMSE) than that of MIPAS. However, the assimilation of MIPAS can still improve the quality of the ozone

15   analyses, and - with a reduction in the RMSE up to about 2 DU - outperform the control experiment thanks to its synergistic assimilation with total column ozone data within the DAS.

High vertical resolution ozone profile observations are essential to accurately monitor and forecast ozone concentrations in a DAS. This study demonstrates the potential and limitations of each dataset and instrument type, as well as the need for a balanced future availability of nadir and limb sensors, and long-term plans for limb-viewing instruments.

20   # 1   Introduction

Since the discovery of its global decline in early 1980s (Farman et al., 1985), ozone has attracted the interest of both the scientific community and policy-makers (e.g. WMO, 2014a, b, and earlier assessments), as well as of the general public. Such an interest is driven and justified by the crucial role ozone plays in the chemistry and in the thermal structure of the atmosphere: a change in the amount of ozone could lead to a warming/cooling of the Earth (depending on the altitude where the change

25   occurs) and affect its climate (e.g. McLinden and Fioletov, 2011), as well as human life.

The concern for the ozone decline - at an observed rate for the global total column of 2.5% between the 1980s and early 1990s (WMO, 2014a) - led to the signature of an international treaty (the 1997 Montreal Protocol and subsequent amendments) to regulate the release in the atmosphere of Ozone-Depleting Substances (ODSs, e.g. SPARC, 1998; Randel and Wu, 2007). The 2014 WMO (acronyms not defined in the text can be found in table 1) assessment of ozone depletion stresses that although recent studies agree that the start of the 21st century represented a turning point in the global total column ozone trend, which now sees a slow increase in its abundance, it is not yet clear whether the current global increase can be attributed to a reduction in the amount of ODSs (WMO, 2014a). Projections of the ozone future evolution seem to agree that the ozone amount will recover towards the stratospheric levels registered before the 1980s by the end of the current century (e.g. Barnes et al., 2014). However, they do not agree on when such a recovery will be achieved (Velders and Daniel, 2014). Thus, the attention in closely monitoring the ozone evolution remains high.

Satellites have been critically important in delivering valuable information to continuously monitor this important atmospheric gas over the last four decades. The first satellite ozone measurements date back to the early 1970s when the nadir-pointing BUV instrument was launched on board of the Nimbus-4 satellite (Heath et al., 1973) followed by a successful number of similar instruments (SBUV and SBUV/2) on Nimbus-7 in 1978 and several NOAA platforms (Bhartia et al., 1996) that stretch up to present. In addition, a large variety of instruments to measure ozone have been launched since 1978 and many more are planned for the next ten to fifteen years. A rather complete account of the past and planned missions able to deliver ozone measurements can be found at www.wmo-sat.info/oscar/gapanalyses?view=108. This wealth of ozone observations offer the opportunity to derive a record of over forty years to study the ozone variability and changes, as well as derive trends. However, they are highly inhomogeneous in viewing geometry (nadir, limb, occultation, or a combination of them), in observed spectral range (UV, visible, near-infrared, thermal infrared, microwave) and spectral resolution, in spatial coverage (GEO vs. LEO platforms) and spatial resolution (ranging from a few hundred kilometres to a few kilometres).

A powerful way to integrate and exploit such a wealth of observations in a way that is consistent with their uncertainties is within a Data Assimilation System (DAS) with many successful examples covering the medium-range weather forecasting and reanalysis available in the literature (e.g. Daley, 1991; Courtier et al., 1994; Veersé and Thépaut, 1998; Rabier et al., 2000; Kalnay, 2003; Dee et al., 2011). To exploit data synergy, guarantee redundancy, thus resilience to sudden changes in the observing system, and to be less sensitive to limitations of a particular instrument design, the general tendency in both NWP and reanalysis productions is to assimilate as many observation types as possible (e.g. Uppala et al., 2005; Dee et al., 2011; Dragani, 2011). However, there are also cases where precise choices on data selection are made from the outset. For instance, in the latest NASA/GMAO reanalysis (MERRA-2), McCarty et al. (2015) explain that their ozone analyses were constrained using the SBUV ozone products (version 8.6 McPeters et al., 2013) until 2004 when the ozone assimilation was switched to a combination of ozone products retrieved from MLS (Froidevaux et al., 2008) and OMI (Bhartia, 2002) both flying on board the Aura platform (Schoeberl et al., 2006). The former is a limb-emission sensor providing ozone profiles retrieved from microwave measurements; the latter is a nadir backscatter instrument providing total column ozone from measurements in the visible and ultraviolet spectral range.

An analysis of the Earth Observation capability planned for the next decade up to about 2025 shows a generally good temporal coverage from nadir-looking instruments with sensitivity to ozone, on either LEO or GEO platforms. Many of these instruments are or will be operated as part of operational missions thus ensuring long term data provision (for instance the forthcoming Sentinels 4 and 5). The same cannot be said for instruments with a limb viewing geometry. These satellite instru-

5   ments are rarely operational and their combined data record show significant gaps. Yet, limb sounders are greatly important when coming to ozone monitoring as their high vertical resolution can lead to significant differences when deriving trends. Furthermore, when designed to have a sensitivity to spectral ranges such as the infrared (like in the case of the ENVISAT MIPAS sensor, Dubock et al. (2001); Fischer et al. (2008)) or sub-millimitre spectral range (like in the case of the Aura MLS sounder), these instruments can provide measurements in night-time conditions thus overcoming one of the most severe lim-

10   itation of UV-Visible measurements. Many studies have been published in the literature to assess the impact of limb ozone data, particularly from the MIPAS and MLS instruments. The results from the first assimilation trial using the MIPAS ozone profiles within the ECMWF DAS were discussed by Dethof (2003). In their results, MIPAS assimilation was found able to improve the quality of the ozone analyses at high latitudes in the winter hemisphere, in general, and provide a better characterization of the Antarctic ozone hole, in particular. These results were confirmed by later studies. For instance, Wargan et al.

15   (2005) discussed the improvements in the ozone analyses, particularly in the polar night region and below the ozone maximum. Geer et al. (2006) were able to reproduce accurately the unusual event of the ozone hole split that occurred in September 2002 in the Met Office system. Dragani (2013) repeated the Dethof study using MIPAS retrievals obtained after a severe instrumental problem triggered changes in the MIPAS spectral specifications. Their results confirmed that the assimilation of MIPAS ozone data could substantially improve the quality of the ECMWF ozone analyses compared to a baseline only constrained by

20   UV-Visible ozone data. A number of studies focused on the assimilation of the MLS ozone profiles recording improvements in the stratospheric ozone distribution (e.g. Jackson, 2007; Feng et al., 2008; Stajner et al., 2008). In particular, Stajner et al. (2008) also discussed how the assimilation of the MLS ozone profiles can be exploited synergistically with total column ozone from OMI (also on board of the NASA's Aura satellite) to improve the tropospheric ozone column analyses, in addition to the vertical ozone distribution.

25   All these studies agree that limb measurements are valuable, yet the need for these observations does not match the long term availability and plans for limb instruments. Thus, do the results from the available studies and the above considerations, as well as the pragmatic decision taken in constructing the MERRA-2 reanalysis call for more sustained and longer term plans for limb measurements than currently available?

By presenting a comparative assessment of the impact on the ECMWF ozone analyses of assimilating either nadir or limb

30   ozone profiles to the same ozone baseline, this study aims at addressing the above question. The present work focuses specifically on the point of view of data assimilation with potential implications for a large part of the ozone research spectrum spanning from NWP to climate reanalysis, from air quality to stratospheric trends and variability.

The paper is structured as follows: the observations used in the present study are presented in section 2; the DAS set-up is discussed in section 3. A preliminary assessment of the data quality prior the assimilation is discussed in section 4 while the

35   assimilation results are analysed in section 5. Concluding remarks and recommendations are summarized in section 6.

2 The ozone datasets

This section focuses on the two ozone products that were used to address the question raised in section 1, namely the MetOp-A GOME-2 instrument and the ENVISAT MIPAS sounder. Additional ozone information was used to constrain all assimilation experiments. This is described in section 3.

5   Launched on board of the EUMETSAT MetOp satellites, GOME-2 (Callies et al., 2000; Munro et al., 2006) is one of the new generation European instruments. The first two GOME-2 instruments were part of the MetOp-A (October 2006) and MetOp-B (September 2012) payloads, respectively. A third one is scheduled to be launched on MetOp-C in 2017. This series of instruments continue the long-term monitoring of atmospheric ozone started by the ERS-2 GOME and ENVISAT SCIAMACHY instruments, whilst generally characterized by much smaller pixels than their predecessors to make their observations useful

10   for air quality forecasting (e.g. Hao et al., 2014). Like its predecessors, GOME-2 is an optical spectrometer that measures the Earth's backscattered radiance and extraterrestrial solar irradiance, in the ultraviolet and visible part of the spectrum (240-790 nm). Its high spectral resolution (between 0.2-0.4 nm) permits to obtain an excess of 4000 spectral points from four detector channels per individual measurement. The first two channels (covering the 240-315 nm and 310-403 nm spectral regions) are important for ozone retrieval.

15   The second dataset was retrieved from the MIPAS measurements (Fischer et al., 2008). This was a Fourier transform spectrometer launched in 2002 on a Sun synchronous polar orbit on board of ENVISAT. The instrument measured thermal emission at the atmospheric limb in the mid-infrared spectral range between 4.15 and 14.6 μm (or 680-2275 cm^{-1}), thus permitting vertical profile retrieval of several minor atmospheric constituents. It was initially operated with a high spectral resolution of 0.025 cm^{-1} reduced to 0.0625 cm^{-1} in January 2005 to resume operations after instrumental problems occurred in March

20   2004. The reduced spectral resolution led to a proportional reduction in the measurement time from 4.5 sec to 1.8 sec that was exploited to increase the number of measured spectra in each scan in order to have a finer vertical limb grid in the UTLS region and an altitude range coverage from 6 to 70 km. The reduction in the measurement time was also exploited to improve the horizontal resolution between two contiguous limb scan measurements. The instrument was operated until April 2012 when communication with the satellite was lost.

25   This study focuses on the assimilation of the GOME-2 and the MIPAS ozone profiles retrieved by the ozone consortium (O3-CCI hereafter) created as part of the ESA Climate Change Initiative (CCI, Plummer, 2009, www.esa-cci.org). The CCI programme was established, on one hand, in response to the GCOS call for climate-quality satellite data, and, on the other hand, to realise the full potential of the ESA global Earth Observation (EO) archive. CCI aims at providing stable, long-term, satellite-based ECV data products to support climate modellers and researchers. Particular attention is paid in characterizing

30   the observation uncertainty and providing comprehensive, fully traceable information on calibration and validation, long term algorithm maintenance, data curation and reprocessing.

The O3-CCI retrieval scheme for nadir ozone profiles was initially developed at RAL for the ERS-2 GOME instrument (Munro et al., 1998). It is based on an optimal estimation algorithm, which combines measurements and an *a priori* in a way consistent with their error covariance matrices (Rodgers, 2000). It uses three sequential steps to retrieve and improve the ozone

information from the Hartley (266-307 nm) and Huggins (323-335 nm) bands. These provide altogether between 5 and 6 degrees of freedom for signal (Rodgers, 2000)). The *a priori* is derived from the McPeters et al. (2007) climatology while the temperature and pressure profiles are taken from the ERA-Interim reanalysis (Dee et al., 2011). An empirical correction is included in the CCI product to address the instrument throughput degradation at the shorter UV wavelengths (Lang et al., 2009). Particular attention was paid in characterizing the various sources of uncertainty in the resulting retrieval. A detailed overview of this product is presented by Miles et al. (2015).

The O3-CCI retrieval scheme for limb ozone retrievals was jointly developed by the Institut fur Meteorologie und Klimaforschung (IMK) and the Instituto de Astrofísica de Andalucía (IAA) (von Clarmann et al., 2003, 2009). The scheme makes use of all the measurements within the 740-800 cm^{-1} and 1060-1110 cm^{-1} spectral ranges after filtering out cloud-contaminated spectra. It assumes local thermodynamic equilibrium conditions, normally verified in the troposphere and most of the stratosphere in the selected spectral regions (Echle et al., 2000). In this study we made use of the version fv0003 dataset. A full description can be found in Sofieva et al. (2013).

Due to differences in spectral ranges and viewing geometries, the two instruments used in the present study offered different horizontal (Figure 1) and vertical (Figure 2) coverages. The geographical distribution of the GOME-2 data (top panel of Figure 1) varies with latitude and time with the daily data count that ranges from a few tens to a few hundreds over a 2° latitudinal band. In contrast, MIPAS offers a more homogenous data coverage (bottom panel of Figure 1) albeit a much lower data count than GOME-2 with a MIPAS:GOME-2 data count ratio that ranges from 1:2 up to 1:40.

The vertical coverage and vertical resolution of the two products are schematically shown in Figure 2. Overall, the O3-CCI nadir profiles (NPO_3 hereafter) are provided on 19 vertical levels spanning the atmosphere from surface up to 0.01 hPa. The O3-CCI limb ozone profiles (LPO_3 hereafter) are also derived on a fixed vertical grid consisting of 32 vertical levels spanning the region of the atmosphere from 0.05 down to 300 hPa.

3 The data assimilation system

The data assimilation system used here is a low resolution version of the ECMWF Integrated Forecasting System (IFS). IFS is a comprehensive atmospheric forecasting system that simulates the dynamics, thermodynamics and composition of the Earth's atmosphere and interacting parts of the Earth-system. It includes three components: a global spectral atmospheric model, an ocean wave model, and an ocean model that simulates the ocean circulation and sea ice.

At the time of writing, the global high resolution spectral model uses a resolution truncation of T_l1279, which corresponds to a linear reduced Gaussian grid of 16 km grid spacing, and 137 vertical levels spanning from surface to 0.01 hPa (corresponding to an altitude of about 80km). The data assimilation is performed using a four-dimensional variational (4D-Var) data assimilation scheme (Rabier et al., 2000) formulated in terms of increments (Courtier et al., 1994; Veersé and Thépaut, 1998) and with a 12-hour assimilation window. At the time of writing, over seventy different data sources were received and monitored daily from satellite alone. Variational quality control and first-guess checks are carried out for all assimilated data. A variational bias correction scheme (VarBC, Dee, 2005) is used to automatically detect and correct for observation systematic biases. VarBC is

formulated for all observations as linear regressions of a number of bias predictors that are observation type, and sensor dependent, as well as geographically varying. The coefficients of those regressions are part of the state vector and computed during the 4D-Var minimization, thus updated every assimilation cycle. Detailed information on the ECMWF system is available at www.ecmwf.int/en/research/modelling-and-prediction.

5   In this forecast model and analysis system, ozone is fully integrated (Dethof and Hólm, 2004). The ozone model is parameterized according to an updated version of the Cariolle and Déqué (1986) scheme (CD86 hereafter). In the CD86 parameterization, the ozone continuity equation is expressed as a linear relaxation towards a photochemical equilibrium for the local value of the ozone mixing ratio, the temperature, and the overhead ozone column. In addition, an ozone destruction term that depends on the equivalent chlorine content for the actual year is included to parameterize the heterogeneous chemistry. The

10 variational bias correction scheme originally introduced to automatically detect and correct for observation systematic biases in the radiances (Auligné et al., 2007) was later extended to retrieved ozone products (Dragani, 2009). For relevance in the discussion, it is noted that accounting for the vertical sensitivity of any retrieved product as provided by the data averaging kernels (AKs) is currently not possible in the IFS. Thus, retrieved observations, including ozone data, are assimilated with a box-car approximation (i.e., with perfect AKs). With such an approximation, the vertical spread of the ozone information provided by

15 the assimilated ozone observations depends on the background error variances and covariance for ozone. This is particularly the case of the assimilation of TCO_3. An example of background error profile and vertical correlation matrix obtained from Dragani and McNally (2013) is given in Figure 3. This figure shows that the ozone background error is largest in the stratosphere between 20 and 80 hPa (panel a), implying that the ozone increments generated by the assimilation of TCO_3 products is most likely spread over this region of the atmosphere. Similar considerations are less straightforward when assimilating

20 profiles, as the region of the atmosphere where an impact can be made also depends on the observation error characteristics.

3.1 Experiment set-up

Three assimilation experiments (a control, referred to as Exp/Ctrl, and two perturbation experiments) were run for the period July-October 2008 using a low resolution version of the IFS with an horizontal truncation of T_l511, which corresponds to about 40 km grid resolution, and 91 vertical levels spanning the atmosphere from surface up to 0.01hPa.

25 With the exception of observations sensitive to ozone, the three experiments assimilated the same set of observations consisting in those available in the ECMWF archive for the period under study. Regarding the ozone observations, the Exp/Ctrl was constrained with ozone sensitive radiances in the infrared (from HIRS, IASI, and AIRS sounders as described in Dragani and McNally (2013)) and ozone retrievals. The latter consisted of a total column ozone (TCO_3) product retrieved at KNMI from the ENVISAT SCIAMACHY measurements (Brinksma, 2004; Eskes et al., 2005; Antón et al., 2011), and ozone

30 retrievals from NOAA-16, -17 and -18 SBUV/2 instruments (Bhartia et al., 2012). These observations were all taken from the ECMWF operational data archive. It is noted that the SBUV/2 data produced by NOAA as 21-level ozone profiles were converted into a six-layer product (Top-1 hPa, 1-2 hPa, 2-4 hPa, 4-8 hPa, 8-16 hPa and 16 hPa-surface) at ECMWF.

The two perturbed experiments were then run with exactly the same configuration and observing system of Exp/Ctrl plus the assimilation of either the MetOp-A GOME-2 NPO_3 product or the ENVISAT MIPAS LPO_3 dataset. In the remainder of this paper, the former perturbed experiment will be referred to as Exp/GOME2 while the latter will be referred to as Exp/MIPAS.

4 Data quality analysis

A preliminary data analysis was performed for both datasets prior to the assimilation. The purpose was to determine the level of agreement between observed and modelled ozone fields, and, due to their importance in data assimilation, provide a preliminary assessment of the observation uncertainties. This is motivated by the fact that NWP-based systems tend to be conservative when using new observations. This can be achieved either by being particularly selective on the data that are actually assimilated (for instance through first-guess checks that depend on the level of discrepancy between model and observations) or by limiting their impact by inflating their provided uncertainties, which determine the weight the observations themselves have on the analyses. Although significant progress has been made in understanding and characterizing most sources of error in the observations, as well as in following best practise, limitations in those estimates still exist. This is either because some of the sources of uncertainty are particularly difficult to characterize or because they are unknown. Of these two aspects, the former is investigated by examining the observation-minus-analysis (O-A) residuals (or simply analysis departures) that measure the discrepancy between the observations and colocated analyses. Here, the analyses were taken from Exp/Ctrl and thus not constrained by the observations under assessment. The latter aspect is instead investigated by comparing the assumed (i.e. provided) observation uncertainty with an estimate derived with the Desroziers et al. (2005) method. Under the hypotheses of optimality in the analysis and separation of scales between background and observation, the observation uncertainty (meant as the standard deviation) can be estimated using the first-guess and analysis residuals from the observations.

Time series of the global mean analysis departures for both instruments are presented in Figure 4. To ease the comparison, these are provided in terms of integrated column over the region between 0.05 and 100 hPa, which is covered by both instruments. Despite the fact that the analyses used to derive the statistics are constrained by UV-retrieved ozone measurements, GOME-2 is the instrument that shows the largest mean discrepancy, with O-A differences being positive during the four month period assessed and a mean value of about 5DU (about 1.5% of the global mean ozone column over that region, estimated as 90% of a global TCO_3 value of 350DU). The reason for this is still under investigation at the time of writing. The MIPAS measurements indicate that during July-August the global mean ozone analyses are about 5DU too high (they were 10DU too low based on the GOME-2 data). Here, the discrepancy between the two instruments is very likely related to different coverage of the two instruments, particularly over the high latitudes in the SH, as shown in Figure 1. During September-October, the O-A residuals for the two instruments are more similar and they both indicate an underestimation of the ozone analyses of about 5DU above 100hPa. The first-guess check implemented in the IFS discards all observations that, after they successfully pass the data quality control, show discrepancies from the background of 30DU or more over the column. Figure 4 shows that on average the observations from both instruments are well within such a threshold, although it is noted that individual observations might have shown residuals from the background larger than 30DU.

The second aspect, which has a bearing in the level of impact of the two instruments on the ozone analyses, is the characterization of the observation uncertainty. An estimate for both instruments was derived using first-guess and analysis residuals from the observations as discussed in Desroziers et al. (2005). Figure 5 shows the relative difference between the estimated and provided uncertainty for the GOME-2 NPO_3 product. The provided uncertainty appears to be larger than that estimated at

5   all latitudinal bands from surface up to about 5hPa while in the upper stratosphere it is smaller than the estimated uncertainty, especially at high latitudes in the summer hemisphere (the Northern Hemisphere, NH, in this case) and at midlatitudes in the winter hemisphere (i.e. the Southern Hemisphere, SH, in this case). Despite the differences between provided and estimated uncertainties appearing to be rather large, they only represent up to about 4% of the observation values (not shown). From a data assimilation perspective, an uncertainty overestimation, here seen at most levels from surface up to about 5hPa, is gener-

10   ally desirable to account for the effect of vertical correlations in the observation error that are normally neglected in most data assimilation systems, including the IFS. The only consequence of such an overestimation would be to limit the impact of the corresponding data. The underestimation of the observation uncertainty in the upper stratosphere is more of a concern. This is because underestimated uncertainties increase the impact of the corresponding observations and if that is associated with poor quality data the assimilation can lead to negative impact on the analyses and forecasts. As the ozone abundance above 5hPa

15   substantially reduces with height, it is expected that if there was a detrimental effect on the ozone analyses as a consequence of assimilating these observations without any correction on their uncertainty, this should generally be negligible or at worst small. Thus, it was decided not to apply any correction to the GOME-2 observation uncertainties above 5hPa.

The vertical cross-section of the difference between the estimated and the provided uncertainty for the MIPAS LPO_3 is shown in Figure 6. In the lower and middle stratosphere, the MIPAS uncertainty appears to be generally overestimated compared with

20   its estimated equivalent, with the exception of the high latitudes in the winter hemisphere (SH). In contrast, the provided field is on average smaller than the estimated uncertainty in the upper troposphere, especially in the winter hemisphere, and in the upper stratosphere. The underestimation of the MIPAS uncertainty in the upper troposphere and in the upper stratosphere can potentially produce a negative impact on the analyses and forecasts. Using the same argument discussed above, any detrimental effect in the upper stratosphere is expected to be generally from negligible to small. Thus, no correction is made at those levels.

25   The possible underestimation of the MIPAS uncertainty in the upper troposphere mostly affects the mid and high latitudes in the SH. It is noted that the period under consideration is July-October, thus winter and spring times in the SH. A data assimilation system strongly constrained by UV observations can show limitations in polar night conditions, while facing an important transition in springtime, when the UV data availability at these latitudes slowly restarts (as shown for instance in the top panel of Figure 1). Thus, it could be argued that the model-based estimate of the MIPAS uncertainty at high latitudes in the SH is

30   less reliable than that provided at other latitudinal bands. A hint of a small underestimation of the uncertainty compared to its model equivalent in the middle troposphere in the southern midlatitudes can also be found for GOME-2 (Figure 5). Arguably, this could also point to problems in the assumed background errors in this region of the atmosphere and time of the year. Based on these considerations, it was decided to test the assimilation of both datasets without applying any correction on the provided uncertainties.

5 Assimilation results and discussion

5.1 Impact on the analyses

The comparison of the Exp/GOME2 and Exp/MIPAS analyses with those from Exp/Ctrl is shown in Figure 7. Both instruments contribute to an increase in the extra-tropics total column ozone, although with different amounts. In contrast, their impact differs in the tropical region. Here, the GOME-2 assimilation contributes to an overall reduction of up to 10 DU in the mean TCO$_3$ (top panel of Figure 7) while that of MIPAS shows a small increase of up to about 2 DU (bottom panel of Figure 7).

The mean vertical cross-section of the ozone analysis differences shown in Figure 8 also depicts a consistent picture between the two datasets at most levels and latitudes, particularly in the extra-tropics. To account for the difference in the order of magnitude of the analysis residuals in the troposphere and stratosphere, the vertical cross-sections were also scaled to a mean ozone profile computed from the Exp/Ctrl ozone analyses over the three month period under consideration and shown in the bottom panels of Figure 8. On average, both datasets tend to increase the ozone analyses in a thick layer between 20 and 70 hPa with differences of up to 2 mg/kg from the Exp/Ctrl. Over this layer, the largest difference in terms of impact between the two instruments can be seen at high latitudes in both hemispheres where the assimilation of the GOME-2 dataset seems to add more ozone with respect to the Exp/Ctrl than done by the assimilation of the MIPAS profiles. With a mean difference of about 1 mg/kg over the 20-70 hPa layer, the estimated mean column change produced by either dataset in the middle stratosphere adds up to an increase of around 24 DU.

In the upper stratosphere and in the UTLS region, the assimilation of either the GOME-2 or the MIPAS ozone profiles leads to a reduction of the ozone analyses of an amount that depends on the instrument and latitudinal band. The largest difference is noticeable in the tropics, in the region of the ozone mixing ratio maximum around 15 hPa. This could also explain the discrepancy found in Figure 7 in the tropics.

In the lower and middle troposphere, both the assimilation of MIPAS and that of GOME-2 profiles tend to remove ozone in the extra-tropics. This is also the case for the tropical troposphere in the case of MIPAS while the assimilation of GOME-2 leads to a tropospheric ozone increase in this latitudinal band.

An aspect that emerges from Figure 8 (bottom panels) is that the assimilation of the MIPAS ozone profiles can substantially modify the ozone distribution in the lower and middle troposphere with changes as large as 60%. Such a change occurs despite the fact that a limb instrument vertical coverage does not normally extend to pressure levels below about 300 hPa, and the observation uncertainty in the upper troposphere is normally larger than in the stratosphere, thus having a lower weight in the data assimilation system. Yet, the structure of the changes are consistent in sign and often in amplitude with those inferred by the GOME-2 nadir ozone profiles.

5.2 Comparisons with independent data

To determine whether the changes introduced by the assimilation of either the GOME-2 or the MIPAS ozone profiles represented an improvement or a degradation, the quality of the ozone analyses from the experiments described above was assessed in terms of the analysis agreement with independent, unassimilated ozone observations. Ozone profiles retrieved from MLS

(version 3.04) and ozone sondes available at the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) were used as independent ozone references for the stratosphere and troposphere/lower stratosphere, respectively.

The MLS observations offers a near global horizontal coverage from about 82S to 82N. The used vertical range spans from about 316 hPa up to 0.01 hPa, which also coincides with the model top, and a vertical resolution of about 3 km. Froidevaux et al. (2008) compared the previous v2.2 MLS ozone retrievals with matching ozone data from ground-based and satellite observations, and found that the differences are generally within 5% in most of the stratosphere, although residuals of 10-20% were found in the lower stratosphere. Livesey et al. (2013) reports differences between the v3.3 and v2.2 MLS ozone profiles typically within 1-2% in the stratosphere and lower mesosphere.

The ground-based stations that are included in the WOUDC archive are spatially inhomogeneous, with the highest availability over Europe and North America. The archive includes measurements performed with both electrochemical concentration cell (ECC), and Brewer Mast (BM) sonde types. The ECC precision was estimated by Komhyr et al. (1995) to be within ±5% in the vertical range between 200 and 10 hPa, between -14% and +6% above 10hPa and between -7% and +17% below 200 hPa. The same order of precision was found by Steinbrecht et al. (1998) for the BM sondes. A data check is used to discard all soundings with a pressure burst either at or below 40 hPa.

5.2.1 Methodology

The comparisons were performed in two stages. First, the 3D ozone analysis closest in time to the independent observation was interpolated at the observation location. This gives a temporal mismatch of up to 3 hours between the observation sensing time and the analysis valid time. The second stage takes care of the vertical interpolation. This is done by degrading the profile with the highest vertical resolution to the coarsest grid. Only the levels spanning the region of the atmosphere encompassed by both datasets are used. In the comparisons with MLS, the coarsest grid is represented by that of the MLS data with its vertical resolution of about 3 km while in the comparisons with the ozone sondes the model has the lowest vertical resolution. The comparisons with MLS are displayed as the vertical cross-section of the change in the analysis fit to the observations due to the addition of either GOME-2 or MIPAS to the reference observing system assimilated in the control experiment. Such a change, Δ, is defined as follows:

$$\Delta = \left| \mathrm{STAT}\left(\mathrm{MLS} - \mathrm{Analyses}^{(\mathrm{Exp/PERT})}\right) \right| - \left| \mathrm{STAT}\left(\mathrm{MLS} - \mathrm{Analyses}^{(\mathrm{Exp/CTRL})}\right) \right| \tag{1}$$

where $\mathrm{STAT}()$ can be either the mean or the standard deviation. For either statistics, a negative value of Δ means that the analyses from a given perturbed experiment, generically labelled above as Exp/PERT, fits MLS observations better than those from Exp/Ctrl, thus leading to an improvement. In contrast, a positive value of Δ is associated to a degradation in the ozone analyses.

The comparisons with the ozone sondes are instead shown in terms of mean RMS residuals, RMSE, between the sonde profiles and the colocated analyses from the three experiments (Exp/Ctrl, Exp/GOME2, Exp/MIPAS). Thus the smaller is the RMSE, the better is the analysis fit to the sonde measurements. The level of agreement between the three sets of ozone

analyses and the ozone sondes is also discussed in terms of a Mean Common Area score Factor (MCAF). The Common Area score Factor (CAF, e.g. as proposed by Dragani et al., 2015) is an indicator of how well the analysis profile fits the observation profile. It is defined as

$$CAF = \frac{\int_p min([\mathbf{O_3}^{(observation)}(p), \mathbf{O_3}^{(model)}(p)])dp}{\int_p max([\mathbf{O_3}^{(observation)}(p), \mathbf{O_3}^{(model)}(p)])dp} \quad (2)$$

5   where $\mathbf{O_3}$ is the ozone vertical profile at the observation location. The integrals are computed over the observation and model common vertical pressure grid. The score varies between zero and one with a CAF value of one implying a perfect fit to the observations. The MCAF score is based on a similar idea as the CAF, but averaged over a number of profiles and stations, for instance here it is computed for all profiles available over given latitudinal bands (the high latitudes in the SH and NH; the midlatitudes in the SH and NH; the tropics). It is important to bear in mind that, due to the sonde vertical coverage, the CAF

10 and MCAF scores are most representative of the part of the atmosphere spanning from the surface up to about 10 hPa at most.

In all comparisons, the ozone analyses were spatially colocated with the independent observations allowing a maximum of three hour time lag. To avoid any spin-up effect, only the results computed over the August-October 2008 period are presented.

5.2.2 Analysis and discussion

Figure 9 gives an indication of the improvement in the fit of the Exp/GOME2 and Exp/MIPAS analyses to MLS compared

15 to that of the Exp/Ctrl analyses, based on equation 1 and presented in terms of both the mean (top panels) and the standard deviation (bottom panels). Despite the large difference in data counts, the two datasets produce similar large scale structures in the analysis fit to MLS at all latitudes for most levels in the middle to lower stratosphere in both the mean (top panels) and the standard deviation (bottom panels). By contrasting Figures 9 and 8, it results that both instruments generally lead to stratospheric improvements in the analysis agreement to MLS at most latitudes, particularly in the atmospheric layer between

20 10 and 70 hPa, with a reduction in the mean residuals from MLS of up to about 2 mg/kg in the middle stratosphere. In the tropical region, both instruments lead to a slightly degraded fit (up to about 0.5 mg/kg) of the ozone analyses to MLS between 20 and 30 hPa. Such a degradation, which is more pronounced in the case of MIPAS than in that of GOME-2, is balanced by an improved fit in the tropical regions above and below the 20-30 hPa layer. With changes up to 1 mg/kg, the improvement in the tropics outside the 20-30 hPa layer is more important for the nadir instrument than for the limb. At high latitudes, the

25 nadir instrument triggers a reduction in the level of agreement with MLS. Here, the changes are up to about 0.5 mg/kg in the NH (summer hemisphere) and about double in the SH (winter hemisphere). Negligible to small changes (typically less than 0.25 mg/kg) are triggered in the upper stratosphere with a sign that depends on the latitudinal band. Overall, panels **a)** and **b)** of Figure 9 highlights a complementarity of the two observing systems in impacting the corresponding stratospheric ozone analyses. The bottom panels of Figure 9 show that the assimilation of either instrument leads to a reduction in the standard

30 deviation in the extra-tropical middle stratosphere, although when assimilating MIPAS the impact is slightly larger (thus better) in places than in the case of GOME-2 assimilation. A small degradation can be found in the standard deviations when GOME-2 is assimilated between 20 and 40 hPa at latitudes southern than 60°S - where the fit to MLS was degraded also in the mean -,

as opposed to a substantial improvement determined by the assimilation of MIPAS. This increased noise given by the Nadir, UV-based observations is an indication of the difficulty of this observation type to provide a strong constraint to the ozone analyses and forecasts at high latitudes in the SH during winter-spring time. Furthermore, in contrast to what is found in the mean changes, here the vertical coverage of the impacted region is lower for MIPAS than that for GOME-2. Those differences
5   can, arguably, be related to the instrument vertical sensitivity associated with the different viewing geometry. Averaging kernels (AKs) that sharply peak at the level of maximum sensitivity for the limb sensor (e.g. Ceccherini and Ridolfi, 2002) as opposed to broad and highly overlapping AKs for GOME-2 (Miles et al., 2015) allow the former to generate much more localized changes than the latter can do.

     The conclusions drawn by assessing the change in the level of stratospheric agreement between the analyses and the MLS
10 profiles implied by the assimilation of either the GOME-2 or MIPAS ozone profiles are confirmed by the comparisons with ozone sondes. The MCAF scores are computed over five latitudinal bands and divided into two contributions for the strato-sphere and troposphere. For simplicity, the 100 hPa pressure level has been used as the separation level for the two contributions. The stratospheric and tropospheric scores for the midlatitudes in the NH, characterized by the largest availability of sonde pro-files, are presented in Figures 10 and 11, respectively. The results at the other four latitudinal bands generally confirm those at
15 midlatitudes in the NH. Overall, both the MIPAS and GOME-2 ozone profiles produce stratospheric analyses that are in better vertical agreement with the ozone sonde profiles than those from Exp/Ctrl, with the Exp/MIPAS score being slightly higher than that of Exp/GOME2 (Figure 10). The MCAF scores are generally 20% lower in the lowermost part of the atmosphere up to 100 hPa (Figure 11) than they are in the region above (Figure 10). Although, both instruments lead to an improved agreement of the corresponding analyses to sonde profiles compared to the Exp/Ctrl, the differences with the Exp/Ctrl scores
20 reduce significantly. In particular, both GOME-2 and MIPAS appear to produce similar levels of agreement, especially in the tropics (not shown). In the extra-tropics, Exp/GOME2 is marginally better in the NH (summer hemisphere) while Exp/MIPAS is slightly better in the SH (winter hemisphere, not shown).

     Why are the scores of the two Exp/PERT analyses so close in the troposphere? Based on Figure 2, is good vertical coverage with low vertical resolution (like for GOME-2) equivalent to having a poor vertical coverage but with high vertical resolution
25 where available (like for MIPAS) in the troposphere? If so, what does it happen in the lowermost troposphere, in particularly below the level of MIPAS availability of roughly 400 hPa?

     An answer to these additional questions can be found by analysing the vertically resolved fit of the three sets of analyses to the ozone sondes (Figure 12). As explained in section 5.2.1, this is expressed in terms of the RMS difference between the sonde profiles and colocated analyses, thus the smaller such a difference is, the higher the level of agreement between the analysed and
30 observed ozone profiles becomes. Figure 12 in general confirms the results over the stratosphere discussed above, including the difficulty of the limb data in the tropics at pressure levels within the 20-30 hPa layer shown in panel **b)** of Figure 9. In the upper troposphere (down to about 400 hPa), Exp/MIPAS still shows a higher level of agreement to the sondes than Exp/GOME2. In the lower troposphere (below 400 hPa), the Exp/GOME2 assimilation generally outperforms that of Exp/MIPAS. However, in this region the assimilation of MIPAS LPO$_3$ still improves the ozone analyses compared to the control experiment, despite
35 a vertical coverage for the limb sounder only down to about 400hPa. This positive result is a consequence of exploiting

the synergy between the LPO$_3$ and the total column product that was assimilated in all experiments. Figure 3 (courtesy of Dragani and McNally (2013)) shows that the ozone background error variances are largest in the stratosphere between 20 and 80 hPa, implying that the ozone increments generated by the assimilation of total column products alone would most likely be spread over this region of the atmosphere. However, in this case, the combination of improved stratospheric ozone concentration obtained thanks to the assimilation of the LPO$_3$ observations, and the synergistic assimilation of the latter with TCO$_3$ products provides an indirect constraint on the ozone analyses at levels below the limb vertical coverage.

These results would suggest that the vertical resolution does matter particularly in a region like the UTLS. The observation vertical coverage, although important, matters to a less extent as an impact can be made in unobserved regions if the synergy with other observations can be exploited within the data assimilation system.

6 Concluding remarks and recommendations

The present study aims at providing a comparative analysis of UV nadir-backscatter and infrared limb-emission ozone profile data assimilation, and thus draw conclusions on how the viewing geometry impacts the ability of these two classes of sensors in constraining and improving the quality of the resulting analyses. The MERRA-2 reanalysis already represents an example where a precise, pragmatic choice was made from the outset in whether to assimilate limb ozone profiles in place of UV nadir-backscatter ozone profiles, despite the tendency typical in NWP and reanalysis of using as many observations as possible. This consideration called for a more quantitative assessment of the relative impact on and ability to improve the quality of the resulting ozone analyses of these two classes of observations. An analysis of the Earth Observation capability planned for the next ten to fifteen years shows a generally good temporal coverage from nadir-looking instruments measuring the backscattered solar radiation in the UV(-Vis) spectral range. Many of these instruments are or will be operated as part of operational missions thus ensuring long term data provision. In contrast, over the same time-frame the availability of measurements from limb-viewing sensors is, at best, poor, and normally confined to research satellites with limited long-term continuation, if any at all. Thus, the results are of primary interest to space agencies as input in shaping future plans for Earth Observation, and have implications for a large part of the ozone research spectrum spanning from Numerical Weather Prediction (NWP) to climate reanalysis, from air quality to stratospheric trends and variability.

The main findings from this study are:

- On average, the two datasets produce changes in the vertical distribution of the ozone analyses with similar large-scale patterns, in particular an increase of the ozone analyses up to 2 mg/kg compared to the control experiment at levels between 20 and 70 hPa and an ozone amount reduction in the extra-tropical lower and middle troposphere. Comparisons with independent observations indicate that those changes improve the agreement between the resulting ozone analyses and independent observations.

- In the middle stratosphere, the changes induced by the assimilation of the MIPAS limb ozone profiles are more localized than those implied by the assimilation of the GOME-2 nadir profiles.

- Differences in the impact were noticeable, especially in the tropics and at high latitudes. In the tropics, both instruments lead to degraded fit to MLS (up to about 0.5 mg/kg increase in the residuals) in the layer between 20 and 30 hPa, and to improved fit (up to about 1 mg/kg reduction in the residuals) above and below that layer. A reduction in the level of agreement between the ozone analyses and the MLS observations is also triggered by the GOME-2 assimilation at high latitudes where the mean residuals are increased of up to about 0.5 mg/kg in the NH (summer hemisphere) and about 1 mg/kg in the SH (winter hemisphere).

- The MCAF scores show that both datasets lead to improved ozone analyses compared to the control experiment (about 20% higher MCAF score in the stratosphere and about 10% higher MCAF in the troposphere than the MCAF from Exp/Ctrl). But while the assimilation of MIPAS is better than that of GOME-2 in the stratosphere, the scores in the troposphere (here simply referred to as the layer below 100 hPa) are much similar to each other despite the fact that the limb vertical coverage does not normally extend below about 400 hPa

- In the upper troposphere (down to about 400 hPa), the assimilation of MIPAS profiles outperforms that of GOME-2 based on the fit to the ozone sondes. The former shows RMSE values between 1 and 2 DU lower than those derived for the latter.

- In the lowermost troposphere, the assimilation of the MIPAS ozone profiles was found able to substantially modify the ozone distribution with changes as large as 60%. Comparisons with ozone sondes profiles show that, although not always better than that of GOME-2, the assimilation of the MIPAS ozone profiles improves the fit to the independent data by reducing the RMSE of up to 2 DU compared to that of the Exp/Ctrl analyses.

The results presented in this paper highlight the complementarity of the two observing systems and confirmed how the instrument characteristics (observed spectral range and viewing geometry) shape the observation ability in constraining the ozone analyses. Overall, both types of observations can improve the vertical distribution of the stratospheric ozone analyses that are too high above 20 hPa and too low below 25-30 hPa in the control experiment. The assimilation of the nadir observations proves to be more successful than that of the limb observations in the tropical stratosphere. In contrast, the assimilation of the limb ozone profile is essential at high latitudes and in the upper troposphere. Only small differences between the two perturbation experiments were seen in the lowermost troposphere that is not sampled by limb sensors but where nadir instruments, like GOME-2, can provide useful information (Miles et al., 2015). On one hand the limb limited sensitivity can only be realized here if their synergy with other ozone observations (in particular total column ozone products) can be exploited within the data assimilation system; on the other hand such a synergy cannot completely replace the improvement produced by the assimilation of nadir-based ozone profiles.

To accurately monitor and forecast ozone concentrations in a data assimilation system, high vertical resolution ozone profile observations are essential. With only total column ozone measurements, analyses would depend too much on the model background and the background error covariances with analysis increments that would be most likely located in the region of the atmosphere where the background error is largest.

Limb observations are essential to bring the improvements in the ozone analyses necessary for example to assess long-term stratospheric changes from reanalysis productions, especially at high latitudes, but also in the upper troposphere and across the tropopause. However, in the lowermost part of the atmosphere important for air quality, nadir-based profiles are still the most important source of information.

5 The present study has demonstrated the potential that each dataset and instrument type has in its own right. The findings discussed in this study would indicate the need for a more balanced availability between nadir and limb sensors than currently exists thus supporting revised EO plans to improve the availability and long-term continuation of limb viewing instruments in the forthcoming years.

Acknowledgements. The author was supported by the European Space Agency through the Climate Change Initiative / Climate Modelling
10 User Group (CCI/CMUG) project. A special thank-you goes to Dr. Stephen English (ECMWF) and Dr. Roger Saunders (The Met Office) for comments and useful discussion. Anabel Bowen (ECMWF) skilfully improved the quality of the figures. The ESA-funded Ozone-CCI consortium is thanked for the provision of the CCI GOME-2 and MIPAS ozone profiles. The MLS ozone data used in this study for the independent validation were the reprocessed (version 3.04) ozone profiles retrieved from the NASA Distributed Active Archive Center at the Goddard Space Flight Center. The ozone sondes were retrieved from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC).

References

Antón, M., Kroon, M., López, M., Vilaplana, J. M., Bañón, M., van der A, R., Veefkind, J. P., Stammes, P., and Alados-Arboledas, L.: Total ozone column derived from GOME and SCIAMACHY using KNMI retrieval algorithms: Validation against Brewer measurements at the Iberian Peninsula, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2011JD016436, 2011.

Auligné, T., McNally, A., and Dee, D.: Adaptive bias correction for satellite data in a numerical weather prediction system, Q. J. R. Meteorol. Soc., 133, 631–642, 2007.

Barnes, E. A., Barnes, N. W., and Polvani, L. M.: Delayed Southern Hemisphere Climate Change Induced by Stratospheric Ozone Recovery, as Projected by the CMIP5 Models, J. Climate, 27, 852–867, doi:http://dx.doi.org/10.1175/JCLI-D-13-00246.1, 2014.

Bhartia, P.: OMI Algorithm Theoretical Basis Document, Tech. Rep. version 2, NASA, available at http://eospso.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-02.pdf, 2002.

Bhartia, P., McPeters, R., Mateer, C., Flynn, L., and Wellemeyer, C.: Algorithm for the estimation of vertical profiles from the backscattered ultraviolet technique, J. Geophys. Res., 101, 18 793–18 806, 1996.

Bhartia, P. K., McPeters, R. D., Flynn, L. E., Taylor, S., Kramrova, N. A., Frith, S., Fisher, B., and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm, Atmos. Meas. Tech. Discuss., 5, 5913–5951, 2012.

Brinksma, E.: Validation of SCIAMACHY TOSOMI ozone columns with ground-based data. Available from www.temis.nl/protocols/O3total.html, Tech. rep., KNMI, 2004.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 - Metop's Second-Generation, Sensor for Operational Ozone Monitoring, ESA Bulletin 102, ESA, 2000.

Cariolle, D. and Déqué, M.: Southern Hemisphere Medium–Scale Waves and Total Ozone Disturbances in a Spectral General Circulation Model, J. Geophys. Res., 91, 10 825–10 846, 1986.

Ceccherini, S. and Ridolfi, M.: Averaging Kernels for MIPAS near real time retrievals, Tech. Rep. TN-IFAC-OST0201, IFAC-CNR, available at http://www.ifac.cnr.it/retrieval/documents/AK_report.pdf, 2002.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Q. J. R. Meteorol. Soc., 120, 1367–1388, 1994.

Daley, R.: Atmospheric Data Analysis, Cambridge University Press, 1991.

Dee, D.: Bias and data assimilation, Q. J. R. Meteorol. Soc., 131, 3323–3343, 2005.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, 2011.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. R. Meteorol. Soc., 131, 3385–3396, 2005.

Dethof, A.: Assimilation of ozone retrievals from the MIPAS instrument on board ENVISAT, Tech. Memo. 428, Eur. Cent. for Medium-range Weather Forecasts, Shinfield Park, Reading, UK, 2003.

Dethof, A. and Hólm, E.: Ozone assimilation in the ERA-40 reanalysis project, Q. J. R. Meteorol. Soc., 130, 2851–2872, 2004.

Dragani, R.: Variational bias correction of satellite ozone data. Technical report R43.8/RD/0934, Tech. rep., ECMWF, available from R. Dragani (rossana.dragani@ecmwf.int), 2009.

Dragani, R.: On the quality of the ERA-Interim ozone reanalyses: Comparisons with satellite ozone data, Q. J. R. Meteorol. Soc., 137, 1312–1326, 2011.

Dragani, R.: Validation of the reprocessed MIPAS and SCIAMACHY retrievals using ERA-Interim, and one-year assimilation of MIPAS ozone profiles at ECMWF, Tech. rep., ECMWF, 2013.

Dragani, R. and McNally, A. P.: Operational assimilation of ozone-sensitive infrared radiances at ECMWF, Q.J.R. Meteorol. Soc., pp. 2068–2080, doi:10.1002/qj.2106, 2013.

Dragani, R., Abdalla, S., Engelen, R. J., Inness, A., and Thépaut, J.-N.: Ten years of ENVISAT observations at ECMWF: a review of activities and lessons learnt, Q. J. R. Meteorol. Soc., 141, 598–610, doi:10.1002/qj.2380, 2015.

Dubock, P. A., Spoto, F., Simpson, J., Spencer, D., Schutte, E., and Sontag, H.: The Envisat Satellite and Its Integration, ESA bulletin, 106, 26–45, available at http://www.esa.int/esapub/bulletin/bullet106/bul106_2.pdf, 2001.

Echle, G., von Clarmann, T., Dudhia, A., Flaud, J.-M., Funke, B., Glatthor, N., Kerridge, B., López-Puertas, M., Martín-Torres, F. J., and Stiller, G. P.: Optimized spectral microwindows for data analysis of the Michelson Interferometer for Passive Atmospheric Sounding on the Environmental Satellite, Appl. Opt., 39, 5531–5540, 2000.

Eskes, H. J., van der A, R. J., Brinksma, E. J., Veefkind, J. P., de Haan, J. F., and Valks, P. J. M.: Retrieval and validation of ozone columns derived from measurements of SCIAMACHY on Envisat, Atmos. Chem. Phys. Discuss, 5, 4429–4475, 2005.

Farman, J. C., Grewe, V., Hein, R., Schnadt, C., Bruhl, C., and Steil, B.: Large loss of total ozone in Antarctica reveal seasonal ClOx/NOx interaction, Nature, 315, 207–210, 1985.

Feng, L., Brugge, R., Hólm, E. V., Harwood, R. S., O'Neill, A., Filipiak, M. J., Froidevaux, L., and Livesey, N.: Four-dimensional variational assimilation of ozone profiles from the Microwave Limb Sounder on teh Aura satellite, J. Geophys. Res., 113, D15S07, doi:10.1029/2007JD009 121, 2008.

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151–2188, 2008.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., and1 R. E. Cofield, B. J. D., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., , and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurement, J. Geophys. Res., 113, 2008.

Geer, A. J., Peubey, C., Bannister, R., Jackson, D. R., Lahoz, W. A., Migliorini, S., O'Neill, A., and Swinbank, R.: Assimilation of stratospheric ozone from MIPAS into a global general circulation model: the September 2002 vortex split, Q. J. R. Meteorol. Soc., 132, 231–257, 2006.

Hao, N., Koukouli, M. E., Inness, A., Valks, P., Loyola, D. G., Zimmer, W., Balis, D. S., Zyrichidou, I., Van Roozendael, M., Lerot, C., and Spurr, R. J. D.: GOME-2 total ozone columns from MetOp-A/MetOp-B and assimilation in the MACC system, Atmospheric Measurement Techniques, 7, 2937–2951, doi:10.5194/amt-7-2937-2014, http://www.atmos-meas-tech.net/7/2937/2014/, 2014.

Heath, D. F., Mateer, C. L., and Krueger, A. J.: The Nimbus-4 Backscatter Ultraviolet (BUV) atmospheric ozone experiment: two years' operation, Pure and Applied Geophysics, 106, 1238–1253, 1973.

Jackson, D. R.: Assimilation of EOS MLS ozone observations in the Met Office data-assimilation system, Q. J. R. Meteorol. Soc., 133, 1771–1788, 2007.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Komhyr, W. D., Barnes, R. A., Borthers, G. B., Lathrop, J. A., Kerr, J. B., and Opperman, D. P.: Electrochemical concentration cell ozonesonde performance evaluation during STOIC 1989, J. Geophys. Res., 100, 9231–9244, 1995.

Lang, R., Munro, R., Livschitz, Y., Dyer, R., and Lacan, A.: GOME-2 FM3 Long-Term In-Orbit Degradation - Basic Signatures After 2nd Throughput Test, Technical Report EUM.OPS-EPS.DOC.09.0464, EUMETSAT, 2009.

Livesey, N. J., Logan, J. A., Santee, M. L., Waters, J. W., Doherty, R. M., Read, W. G., Froidevaux, L., and Jiang, J. H.: Interrelated variations of O_3, CO and deep convection in the tropical/subtropical upper troposphere observed by the Aura Microwave Limb Sounder (MLS) during 2004-2011, Atmospheric Chemistry and Physics, 13, 579–598, doi:10.5194/acp-13-579-2013, http://www.atmos-chem-phys.net/13/579/2013/, 2013.

McCarty, W., Considine, D., Lee, T., Randles, C., Coy, L., Wargan, K., Bosilovich, M., and Gelaro, R.: Use of Satellite Observations in NASA Reanalyses: MERRA-2 and Future Plans, Technical report, NASA CGMS-43 NASA-WP-06, v1, 2015.

McLinden, C. A. and Fioletov, V.: Quantifying stratospheric ozone trends: Complications due to stratospheric cooling, Geophys. Res. Lett., 38, L03 808, doi:10.1029/2010GL046012, 2011.

McPeters, R. D., Lebow, G. J., and Logan, J. A.: Ozone climatological profiles for satellite retrieval algorithms, J. Geophys. Res., 112, D05 308, doi:10.1029/2005JD006823, 2007.

McPeters, R. D., Bhartia, P. K., Haffner, D., Labow, G. J., and Flynn, L.: The version 8.6 SBUV ozone data record: An overview, J. Geophys. Res. Atmos., 118, 8032–8039, doi:10.1002/jgrd.50597, 2013.

Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. A. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, Atmos. Meas. Tech., 8, 385–398, doi:doi:10.5194/amt-8-385-2015, 2015.

Munro, R., Siddans, R., Reburn, W. J., and Kerridge, B. J.: Direct measurement of tropospheric ozone distributions from space, Nature, 392, 168–171, 1998.

Munro, R. a. M. E., Anderson, C., Cllies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Albinana, A.: GOME-2 on MetOp, in: The 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, 12-16 June 2006, ISBN 92-9110-076-5, EUMETSAT, 2006.

Plummer, S.: The ESA Climate Change Initiative - Description, Technical Report EOP-SEP/TN/0030-09/SP, ESA, available from http://46.19.32.34/files/ESACCIDescription.pdf, 2009.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics, Q. J. R. Meteorol. Soc., 126, 1143–1170, 2000.

Randel, W. J. and Wu, F.: A stratospheric ozone profile data set for 1979-2005: Variability, trends, and comparisons with column ozone data, J. Geophys. Res., 112, D06 313, doi:10.1029/2006JD007339, 2007.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding. Theory and Practice, Atmospheric, Oceanic and Planetary Physics, World Scientific Publishing and Co. Pte. Ltd., PO Box 128, Farrer Road, Singapore 912805, 2000.

Schoeberl, M. R., Douglass, A., Hilsenrath, E., Bhartia, P., Beer, R., Waters, J., Gunson, M., Froidevaux, L., Gille, J., Barnett, J., Levelt, P., and DeCola, P.: Overview of the EOS Aura mission, IEEE Trans. Geosci. Remote Sens., 44, 1066 – 1074, 2006.

Sofieva, V. F., Rahpoe, N., Tamminen, J., Kyro, E., Kalakoski, N., Weber, M., Laeng, A., von Clarmann, T., Stiller, G., Lossow, S., Degenstein, D., Bourassa, A., Adams, C., Roth, C., Lloyd, N., Bernath, P., Hargreaves, R. J., Urban, J., Murtagh, D., Hauchecorne, A., Roozendael,

M. V., Kalb, N., and Zehner, C.: Harmonized dataset of ozone profiles from satellite limb and occultation measurements, Earth Syst. Sci. Data, 5, 349–363, 2013.

SPARC: SPARC/IOC/GAW Assessment of trends in the vertical distribution of ozone, no. 1 in SPARC Report, pp. 1–289, 1998.

Stajner, I., Wargan, K., Pawson, S., Hayashi, H., Chang, L.-P., Hudman, R. C., Froidevaux, L., Livesey, N., Levelt, P. F., Thompson, A. M., Tarasick, D. W., St"ubi, R., Andersen, S. B., Yela, M., K"onig-Langlo, G., Schmidlin, F. J., and Witte, J. C.: Assimilated ozone from EOS-Aura: Evaluation of the tropopause region and tropospheric columns, J. Geophys. Res., 113, D16S32, doi:10.1029/2007JD008 863, 2008.

Steinbrecht, W., Shwartz, R., and Claude, H.: New Pump Correction for Brewer-Mast Ozone Sonde: Determination from Experiment and Instrument Intercomparisons, J. Atm. O. Tech., 15, 144–156, 1998.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quart. J. Roy. Meteor. Soc., 131, 2961–3012, 2005.

Veersé, F. and Thépaut, J.-N.: Multiple truncation incremental approach for four-dimensional variational data assimilation, Q. J. R. Meteorol. Soc., 124, 1889–1908, 1998.

Velders, G. J. M. and Daniel, J. S.: Uncertainty analysis of projections of ozone-depleting substances: mixing ratios, EESC, ODPs, and GWPs, Atmos. Chem. Phys., 14, 2757–2776, doi:10.5194/acp-14-2757-2014, 2014.

von Clarmann, T., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Tsidu, G. M., Milz, M., Steck, T., Stiller, G. P., Wang, D. Y., Fischer, H., Funke, B., Gil-López, S., and López-Puertas, M.: Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), J. Geophys. Res., 108, 4736, doi:10.1029/2003JD003602, 2003.

von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U. andGlatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, H2O, O3, HNO3, CH4, N2O, ClONO2 and ClO from MIPAS reduced resolution nominal mode limb emission measurements, Atmos. Meas. Tech., 2, 159–175, doi:10.5194/amt-2-159-2009, 2009.

Wargan, K., Stajner, I., Pawson, S., Rood, R. B., and Tan, W.: Assimilation of ozone data from the Michelson Interferometer for Passive Atmospheric Sounding, Q. J. R. Meteorol. Soc., 131, 2713–2734, 2005.

WMO: Scientific assessment of ozone depletion: 2014, Tech. rep., Geneva, Switzerland, global Ozone Res. Monit. Proj. Rep. 55, 2014a.

WMO: Assessment for Decision-Makers: Scientific assessment of ozone depletion: 2014, Tech. rep., Geneva, Switzerland, global Ozone Res. Monit. Proj. Rep. 56, 2014b.

Table 1. List of acronyms and abbreviations used in this paper.

Acronym	Definition
AIRS	Advanced InfraRed Sounder
BUV	Backscatter Ultra Violet
ECV	Essential Climate Variable
ENVISAT	ENVIronmental SATellite
EOS	Earth Observing System
ERS-2	European Remote Sensing 2
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GCOS	Global Climate Observing System
GEO	Geostationary
GMAO	Global Modeling and Assimilation Office
GOME-2	Global Ozone Monitoring Experiment 2
HIRS	High-resolution Infrared Radiation Sounder
IASI	Infrared Atmospheric Sounding Interferometer
IR	InfraRed
LEO	Low Earth Orbit
MERRA	Modern Era Retrospective-analysis for Research and Applications
MetOp	Meteorological Operational Satellite
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLS	Microwave Limb Sounder
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
OMI	Ozone Monitoring Instrument
RAL	Rutherford Appleton Laboratory
SBUV	Solar Backscatter Ultra Violet
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
TOMS	Total Ozone Mapping Spectrometer
UARS	Upper Atmosphere Research Satellite
UTLS	Upper Troposphere and Lower Stratosphere
UV	Ultra Violet
UV-Vis	Ultra Violet and Visible
WOUDC	World Ozone and Ultraviolet Radiation Data Centre
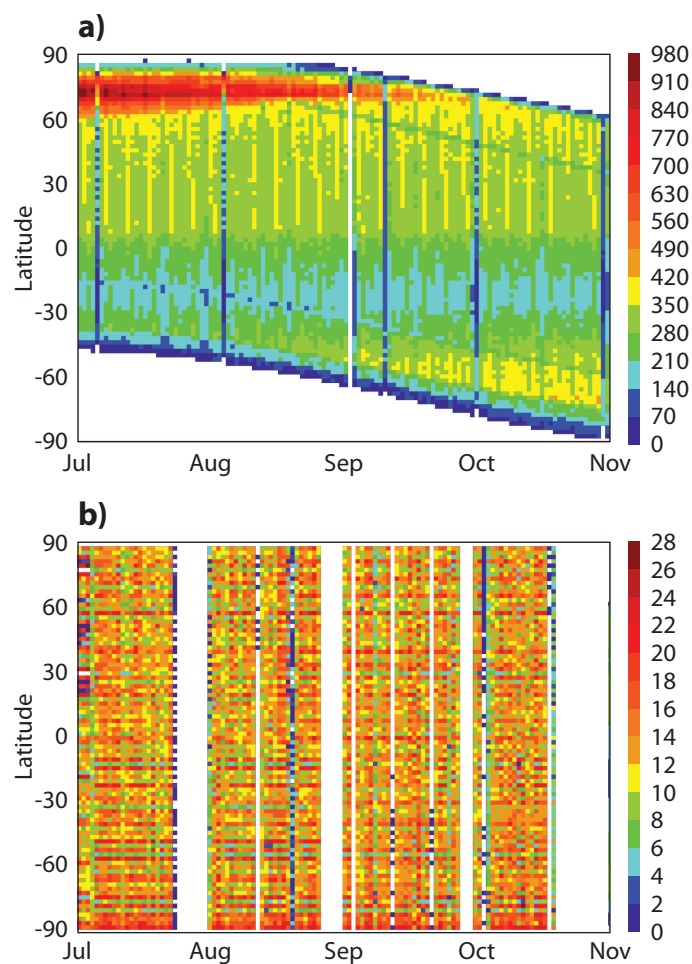
Figure 1. Data coverage and daily count binned in 2° latitudinal bands for the CCI GOME-2 (top panel) and the CCI MIPAS (bottom panel) data. The two colour scales are different. The colour scale used for GOME-2 (panel **a**) varies from 0 (dark blue) to 980 (dark red) observations with a step of 70; the one used for MIPAS ranges from 0 (dark blue) to 28 (dark red) observations with a step of 2.
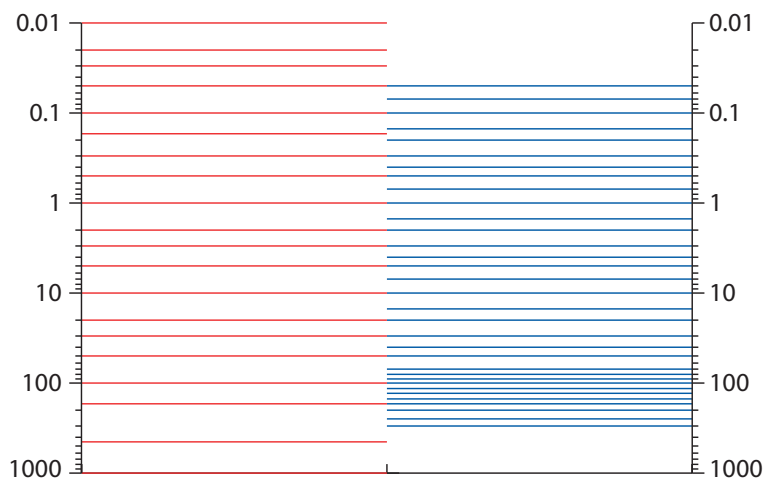
Figure 2. Schematic of the vertical coverage and vertical resolution provided by the CCI GOME-2 (red lines) and the CCI MIPAS (blue lines) retrieval algorithms. The vertical axis represents pressure in hPa. The surface pressure level for GOME-2 is for illustration purposes drawn at 1000 hPa.
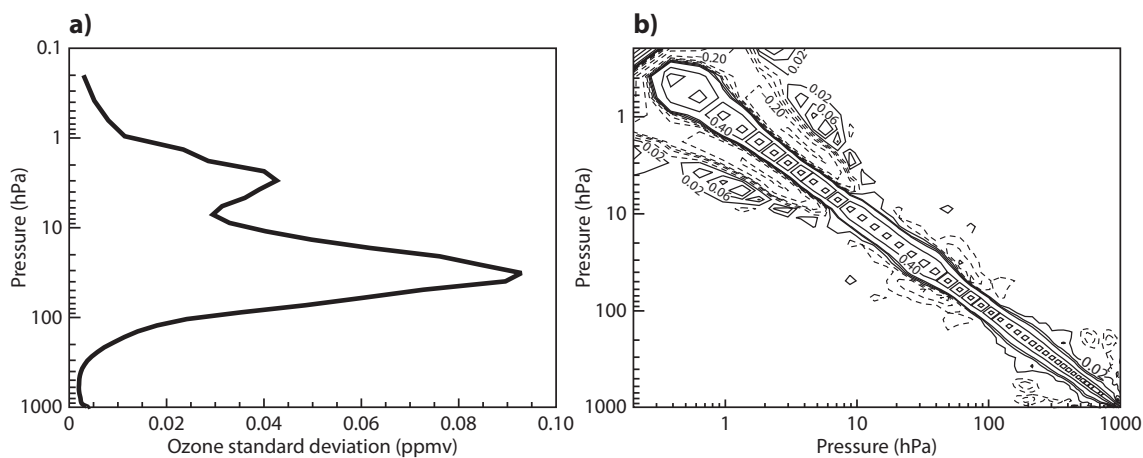
Figure 3. Example of ozone background error standard deviation profile (left panel) and vertical correlation matrix (right panel) for the ozone background errors. Data are in mixing ratio (parts per million by volume, ppmv). In the right panel, negative correlations are plotted as dashed lines, and the interval is 0.04 ppmv between -0.1 and 0.1 ppmv, and 0.2 for larger absolute values. Source Dragani and McNally (2013).
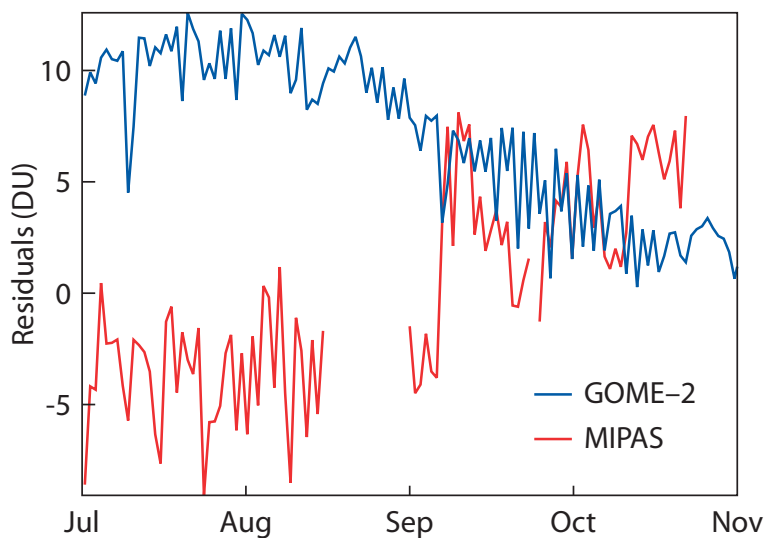
Figure 4. Time series of the global mean analysis departures for GOME-2 NPO3 (blue line) and MIPAS LPO3 (red line) over the period Jul-Oct 2008. Statistics are shown as integrated columns computed over the GOME-2 and MIPAS common vertical grid spanning from 0.05 to 100 hPa. Data are in DU.
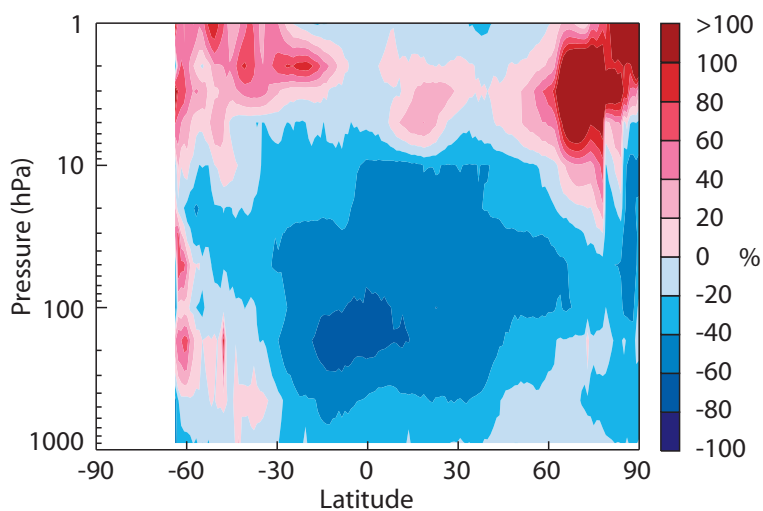
Figure 5. Vertical cross-section of the difference between the estimated and the provided uncertainty relative to the provided uncertainty for GOME-2 NPO3 over the period Jul-Oct 2008. Negative (positive) values in blue (red) colours mean that the provided uncertainty is larger (smaller) than that estimated with the Desroziers et al. (2005) method. Data are in %.
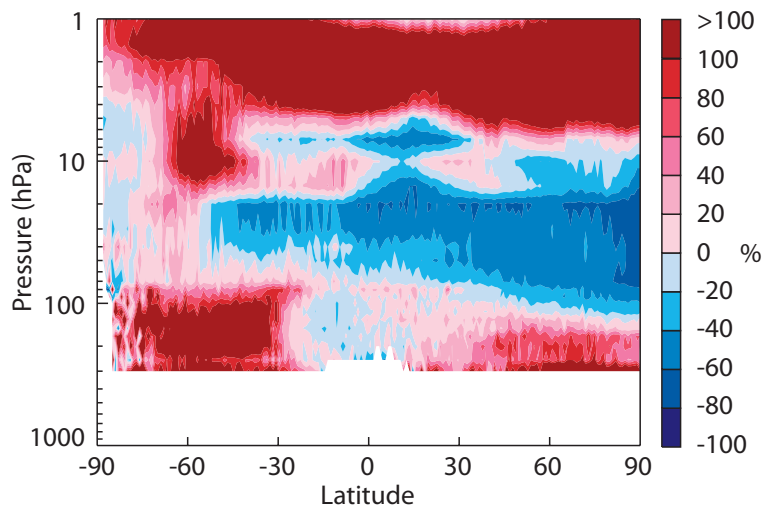
Figure 6. As in Figure 5 but for the CCI MIPAS LPO$_3$ data.
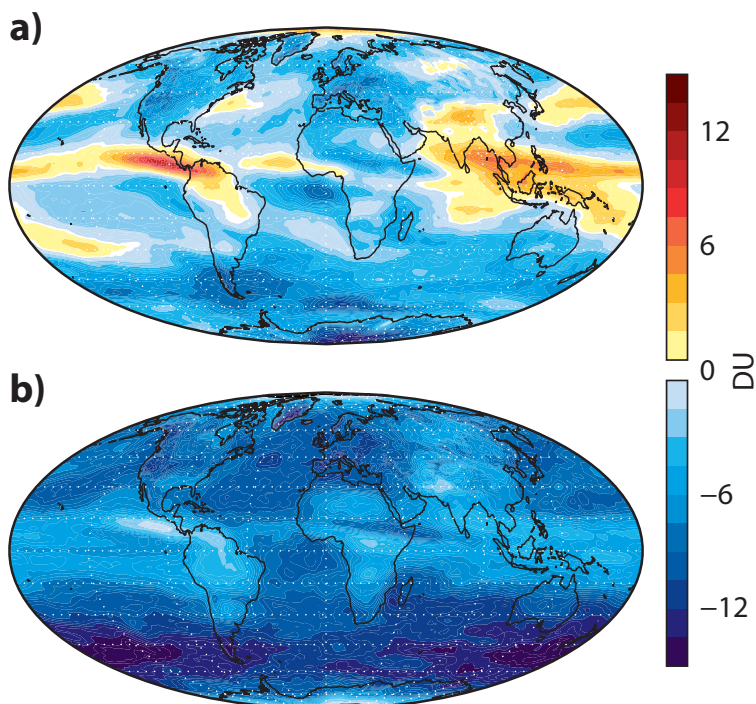
Figure 7. TCO$_3$ mean differences between the analyses from Exp/Ctrl and those from Exp/GOME2 (top) and Exp/MIPAS (bottom) computed for the period Aug-Oct 2008. Negative (positive) values in blue (red) colours mean that the additional instrument is increasing (decreasing) the TCO$_3$ amount compared to that obtained from the ozone datasets assimilated in Exp/Ctrl. Data are in DU. The colour scale ranges from -15 DU (dark blue) to +15 DU (dark red) with a step of 1.5 DU.
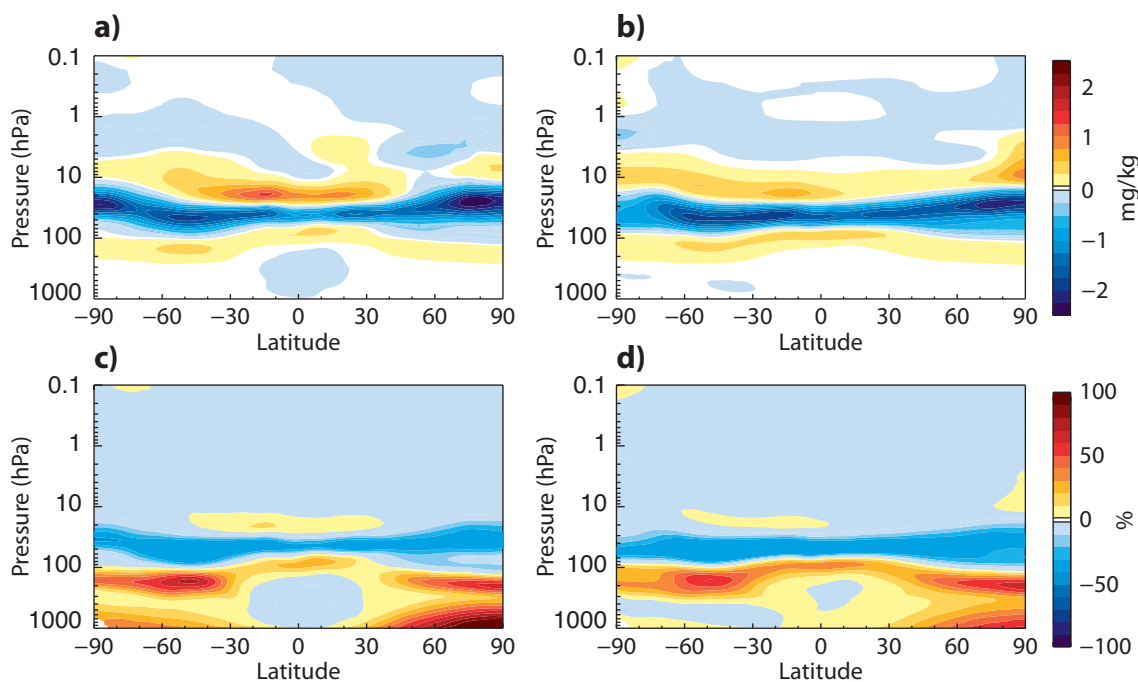
Figure 8. Top panels: Zonal mean temporal mean differences between the ozone analyses from Exp/Ctrl and the ozone analyses from Exp/GOME2 (top left panel) and Exp/MIPAS (top right panel) computed for the period Aug-Oct 2008. Negative (positive) values in blue (red) colours mean that the additional instrument is increasing (decreasing) the ozone mixing ratio amount compared to that obtained from the ozone datasets assimilated in Exp/Ctrl. The white regions are where the differences are within ± 0.1 mg/kg. Data are in mg/kg. The colour scale ranges from -2.5 mg/kg (dark blue colour) to +2.5 mg/kg (dark red colour) with a step of 0.25 mg/kg, with the exception of the yellow and pale blue that start at 0.1 and -0.1 mg/kg, respectively, and represent a step of 0.15 mg/kg. **Bottom panels:** Like in the top panels but for the relative differences computed with respect to a temporal mean global mean ozone profile obtained from the Exp/Ctrl analyses. Data are in %. The colour scale ranges from -100 % (dark blue colour) to +100 % (dark red colour) with a step of 10%.
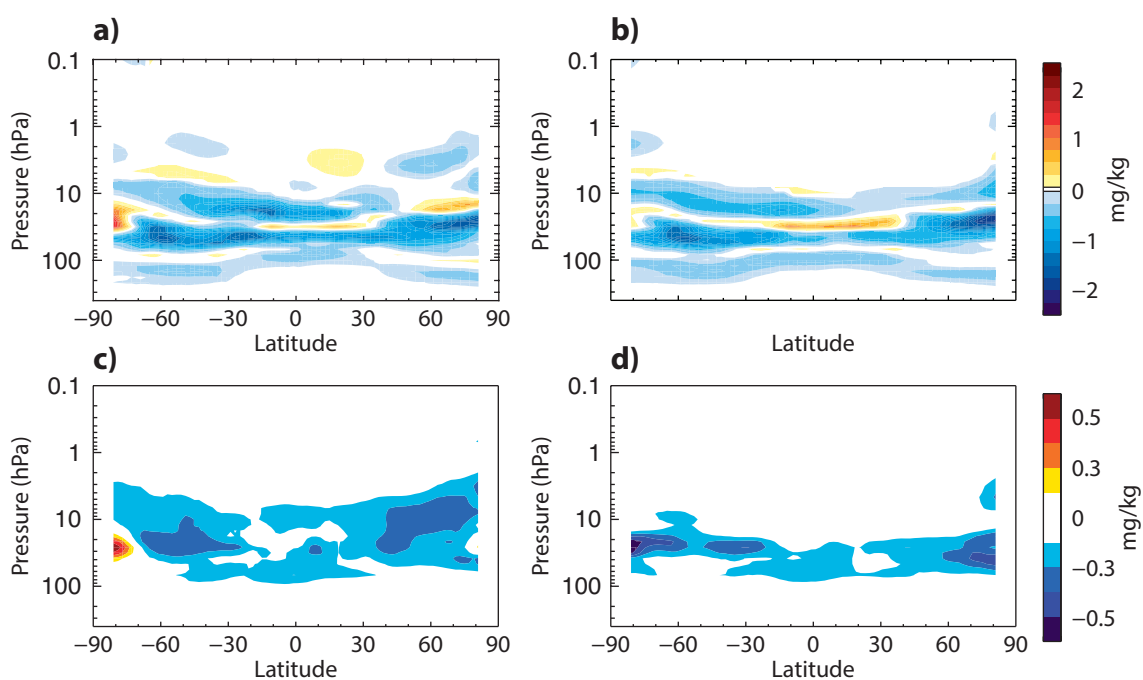
Figure 9. Top panels: Change in the zonal mean differences between the MLS retrievals and collocated ozone analyses from the perturbation experiment compared to the Exp/Ctrl for Aug-Oct 2008. The perturbation experiment is Exp/GOME2 in the left panel, and Exp/MIPAS in the right panel. Negative (positive) values in blue (red) colours indicate a decrease (an increase) in the mean in the Control and thus a degraded fit to MLS ozone profiles. Data are in mg/kg. The colour scale ranges from -2.5 mg/kg (dark blue colour) to +2.5 mg/kg (dark red colour) with a step of 0.25 mg/kg. **Bottom panels:** As a), but for the standard deviation of differences. The colour scale ranges from -0.6 mg/kg (dark blue colour) to +0.6 mg/kg (dark red colour) with a step of 0.1 mg/kg.
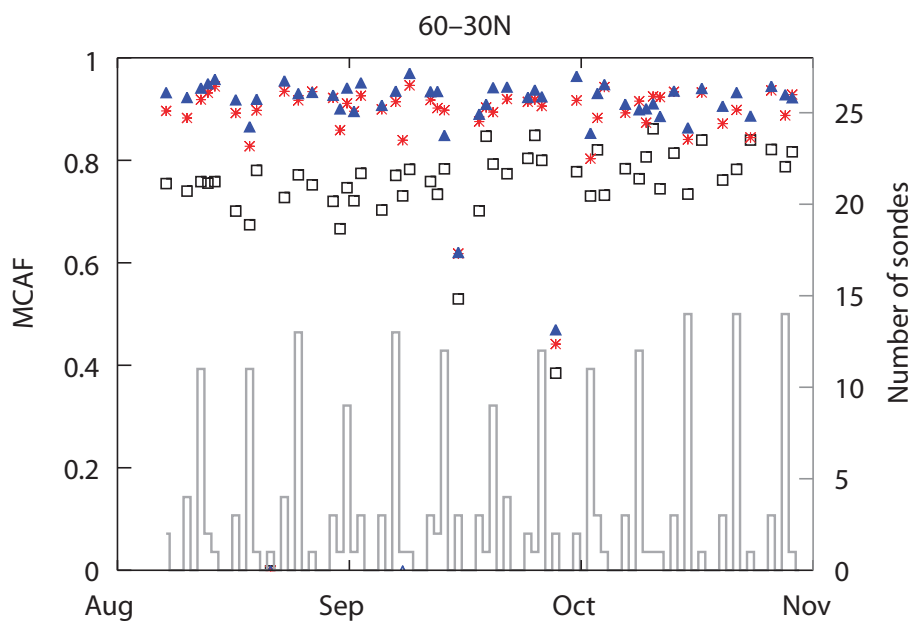
Figure 10. Time series of the stratospheric contribution (considered as that at pressure levels lower than 100 hPa) to the Mean Common Area Fraction (MCAF) score computed for different sonde profiles and the three sets of ozone analyses. The analyses were taken from Exp/Ctrl (black squares), and the two Exp/PERT experiments assimilating also the reprocessed GOME-2 NPO_3 (red asterisks) and MIPAS LPO_3 (blue triangles). The plot refers to the midlatitudes in the NH. The grey histogram in each panel shows the number of daily sondes included in the MCAF score and refers to the right hand side axis.
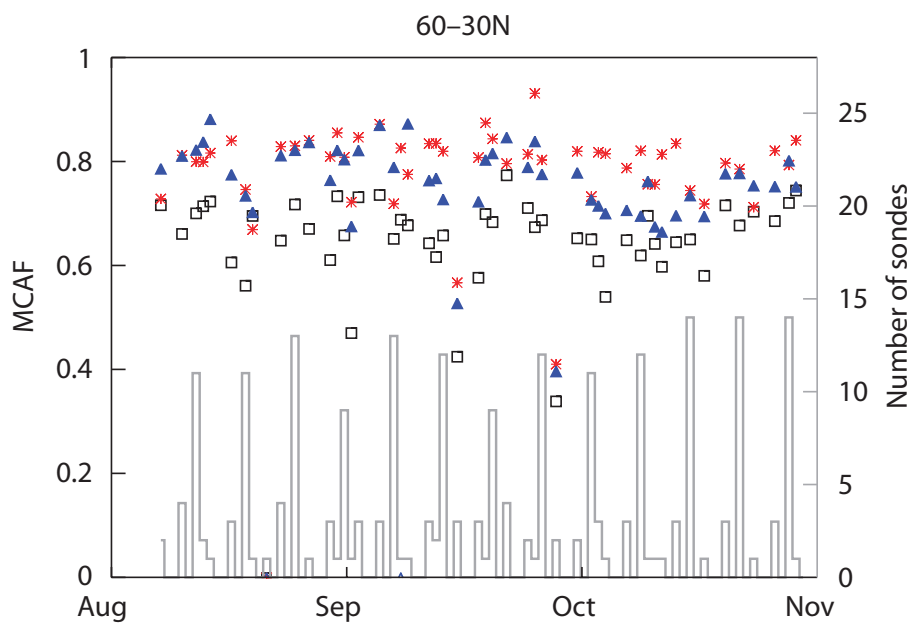
Figure 11. As Figure 10, but for the tropospheric contribution.
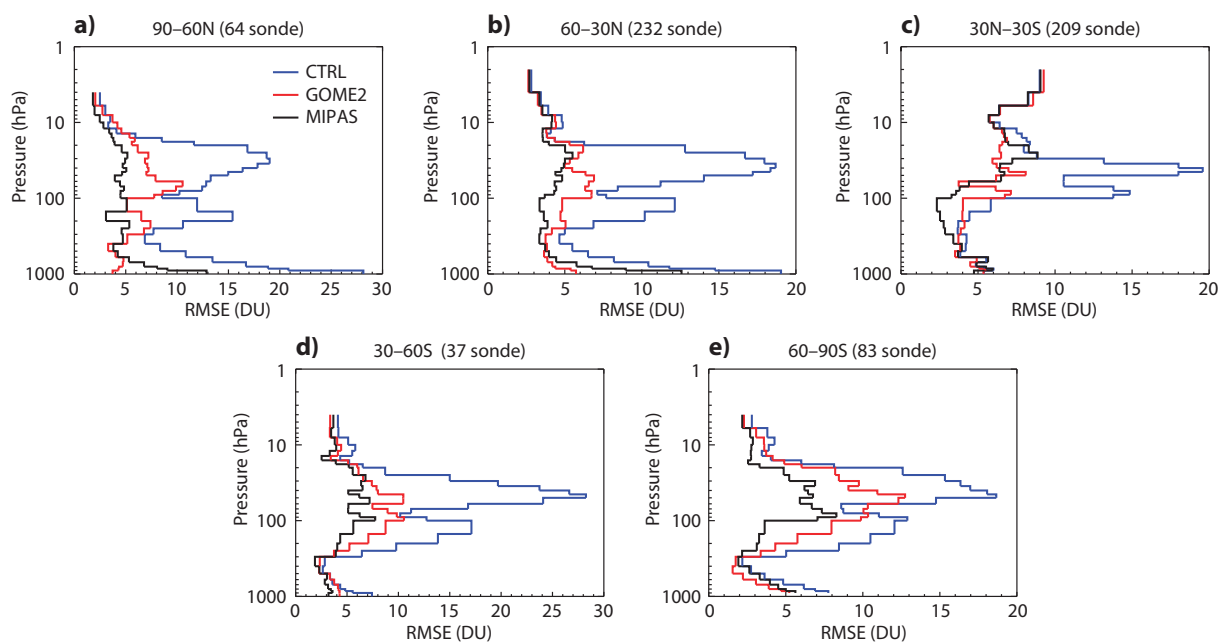
Figure 12. Fit of the ozone analyses from three experiments to ozone sondes given in terms of the RMSE over five latitudinal bands. The comparisons were computed by averaging over Aug - Oct 2008. The three analyses were taken from Exp/Ctrl (blue lines), and the two Exp/PERT experiments assimilating also the reprocessed GOME-2 NPO$_3$ (red lines) and MIPAS LPO$_3$ (black lines) from the Ozone-CCI. The latitudinal band each panel refers to and the number of ascents included in the average can be found in the corresponding panel title. Data are in Dobson Unit (DU).