

Corresponding author's response accompanying the revised manuscript for the paper by Lefever et al. (ACPD, 2014)

This response consists in the reply to the first referee (pages 2-12), the reply to the second referee (pages 13-20) and the marked-up difference between the original and revised manuscripts (pages 21-end).

We have added a new figure (fig.12) to provide a major clarification (in reply to 1st general comment by reviewer 1). We have removed figure 4 from the original manuscript, because reviewer 1 correctly identified large interpolation errors in this figure. We have re-written large parts of the introduction, conclusions and abstract to address the reviewers' comments. We hope that this will give more impact to our study.

Despite all these changes we would like to stress that there is no new finding or new outcome in the revised manuscript – only major clarifications, a better explanation on the interest of the paper, and an increased focus on input datasets in the title, abstract and conclusions.

S. Chabrillat, 14 December 2014

Reply to review 1 of paper by Lefever et al. (ACPD, 2014)

Dear reviewer,

The authors would like to thank you for your in-depth review of our manuscript. The interpretation of our validation results in the original manuscript was misleading, which led you on a false track. We have removed this obsolete interpretation of our validation results, we re-wrote the interpretation of our sensitivity tests and our conclusions, and clarified the small but numerous imprecisions of the original text. Below is a point-by-point response to your comments (copied in italics); paragraphs highlighted in bold font have been inserted into the revised manuscript.

General comments

- In my view, the main weakness of the study is in the interpretation of the results presented. The authors don't explore the reasons behind the differences between the assimilation systems. If explanations are given they are often speculative (phrases like 'this may be due to. . .' are used). For instance, BASCOE appears to have the best performance and IFS-MOZART is worse despite the fact that it assimilates more data. Why? Section 4.1 discusses alternating bias patterns in IFS-MOZART; Section 5.2 ascribes unstable performance of IFS-MOZART to assimilation of UV/VIS but offers no explanation as to why these observations would degrade the analysis. Section 6.2 demonstrates that turning off UV/VIS is beneficial but, again, it doesn't explain why.*

We thank the reviewer for this comment, as it helped us realize that we had to rephrase several sentences of the paper, clarify our conclusions and most importantly show the results of our last sensitivity test. Indeed, the last sentence in section 6.2 of our original manuscript read:

“An additional IFS-MOZART experiment assimilating, besides MLS V3, also the other column products (as defined in Table1) shows that the analysis is well constrained by MLS alone in the stratosphere, while it is beneficial to have the combination of profile and total column data in the troposphere.”

This refers to a last sensitivity test where IFS-MOZART assimilates MLS v3 down to 215 hPa *and* the UV/VIS data from SBUV/2, OMI and SCIAMACHY. We have added in the revised manuscript one last plot to show how these three IFS-MOZART runs compare with ozonesondes in March 2011:

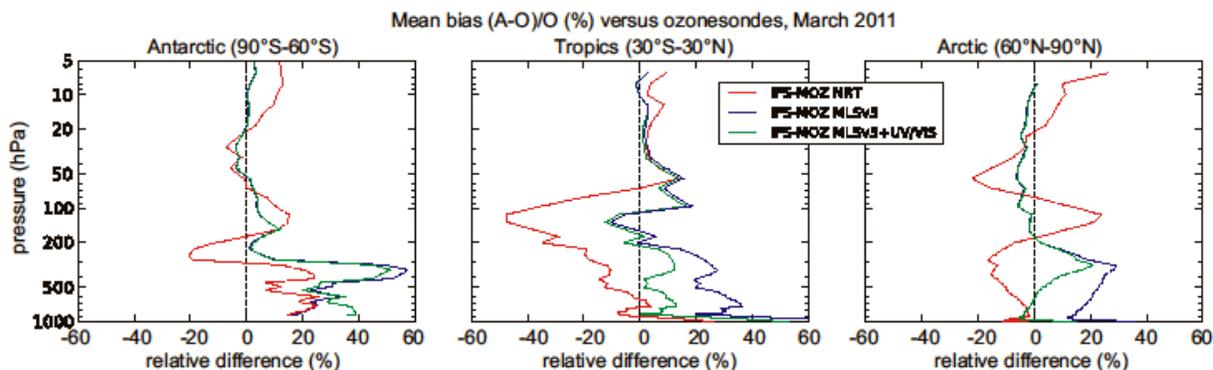


Fig. 12. Mean biases of three ozone analyses by IFS-MOZART using O3 sonde profiles as reference, for March 2011. Results are shown for the Antarctic (left), Tropics (center) and Arctic (right) latitude bands using the IFS-MOZART NRT analyses (red lines), the offline experiment assimilating only MLS v3

(blue lines) and another offline experiment assimilating MLS v3 and the UV/VIS observations (green lines). See text for details.

These results are explained in section 6.2 of the revised manuscript, which clarifies the interpretation of our results:

To identify the exact cause of the large improvement in IFS-MOZART analyses, we ran a last sensitivity test with IFS-MOZART assimilating the usual set of UV/VIS data (OMI and SCIAMACHY total columns; SBUV/2 partial columns) in addition to the offline MLS v3 dataset. As can be seen in figure 12, the bad performance of IFS-MOZART NRT was not due to the assimilation of UV/VIS data but rather to the assimilation of the MLS v2 NRT data. If the MLS v3 and UV/VIS observations are assimilated together (green curves), the quality of the ozone analyses delivered by IFS-MOZART improves: tropospheric ozone is improved over the previous sensitivity test assimilating only MLS v3 (blue curves), and the simultaneous assimilation of UV/VIS observations does not degrade the analysis of stratospheric ozone.

The worse performance of IFS-MOZART NRT is probably not due to the earlier version of the MLS dataset either, because our sensitivity test with BASCOE (Fig. 11, blue lines) shows that the analyses of MLS v2.2 SCI (left) performed nearly as well as the analyses of MLS v3 (right) despite the usage of an earlier version of BASCOE. Hence the better performance of BASCOE NRT is primarily due to its assimilation of MLS v2.2 SCI down to 150 hPa, while IFS-MOZART had to assimilate MLS v2.2 NRT which was not valid (and filtered out) below 68 hPa. This subtle difference in configuration is due to an operational constraint: IFS-MOZART had to be run closer to real-time and could not wait 3 extra days for the distribution of MLS SCI.

The words “despite the usage an earlier version of BASCOE” are explained in section 4.4.2 of the revised manuscript (new paragraph, lines 638-645, added in reply to your specific comment about P12495 L19).

We have also re-written the end of our conclusions (p.26 l. 856-863 of the revised manuscript) to reflect the important messages to be drawn from these sensitivity tests:

From a system design point of view, the sensitivity tests performed in section 6.2 deliver important conclusions:

- **All systems used in MACC require profile data to provide a good vertical distribution of stratospheric ozone.**
- **This profile data must include the lower stratosphere.**
- **IFS-MOZART is able to assimilate limb profiles and nadir products successfully. The profiles constrain well the stratosphere, allowing the partial and total columns (by UV/VIS instruments) to constrain well the troposphere.**

These conclusions have also been added to the abstract of the revised manuscript, and the title was slightly modified to reflect the importance of the input datasets:

Copernicus Atmospheric Service for stratospheric ozone, 2009-2012: Validation, systems intercomparison and roles of input datasets

At the face of it this is surprising because SBUV agrees quite well with MLS – if its averaging kernels are taken into account. Are they used in the assimilation systems? If not then one has to contend with large smoothing errors inherent in nadir data. In the case of SBUV it may help to interpret this in the context of [Kramarova et al., 2013]. I think that the interplay between different input data types is an important aspect of chemical data assimilation and this study is the right place to dive into these issues - but none of them gets much (if any) attention. Addressing this point can be considered a major revision of the manuscript.

As explained above, the assimilation of SBUV/2 was not responsible for the worse performance of IFS-MOZART NRT. However we acknowledge that the original manuscript did not describe clearly the SBUV/2 dataset assimilated in MACC. We have re-written section 2.1.4 as follows:

SBUV/2 is a series of seven remote sensors on NOAA weather satellites (Mc Peters et al., 2013), of which three were assimilated by IFS-MOZART during the period investigated here (September 2009 to September 2012): NOAA-17 and NOAA-18 during the whole period; NOAA-19 after 2011-06-22. Bhartia et al. (2013) describe the two latest versions of the SBUV/2 retrievals: v8 which was available during the period investigated here, and v8.6 which was released more recently. While SBUV v8.6 includes the averaging kernels (AK) for each retrieved profile, these were not available in the v8 BUFR data used operationally at ECMWF. Hence we used the same procedure as first described for ERA-40 (Dethof and Holm, 2004): in order to decrease unwanted vertical correlations between errors at different levels, the thirteen layers of the original SBUV v8 retrievals were combined at ECMWF over six thick layers (0.1--1 hPa, 1--1.6 hPa, 1.6--4.1 hPa, 4.1--6.4 hPa, 6.4--16 hPa, 16 hPa--surface). Among the resulting partial ozone columns, the last one contributes most to the total columns.

Table 1 has been corrected accordingly. The following sentence was inserted in the conclusions:

The newer SBUV/2 v8.6 profiles are distributed over 21 layers and each profile is distributed with its matrix of Averaging Kernels. Kramarova et al. (2013) illustrated the importance of using properly this information. While it is planned to implement SBUV/2 Averaging Kernels in the MACC NRT system at ECMWF, our last sensitivity test shows that this improvement was not necessary to assimilate successfully SBUV/2 v8 after a vertical re-gridding over 6 thick layers.

-
2. *This paper will be of interest to scientists less familiar with data assimilation. It is important to make it easy to read for someone who is not fluent in DA terminology. The word ‘model’ is used in several places where a better term would be ‘data assimilation system’. There are phrases like ‘models which assimilate data’ or (P12482) ‘. . .the model underestimates total ozone. . .’ where you are really talking about analysis. Be more precise. Models don’t assimilate anything!*

In the revised manuscript, all occurrences of ‘model’ have been replaced with ‘Data Assimilation System’, its acronym ‘DAS’, ‘Chemical Data Assimilation System’, or ‘ozone analysis’ when the output dataset is discussed. The word ‘model’ is now used only to describe the forward model component of each DAS.

3. *Explain the advantages of assimilating data as opposed to simply using retrieved observations. In particular, the conclusions section says that the four analyses perform well where data are present and not so well where there are no data. The reader may wonder why you need assimilation at all? Why not just use observational data?*

We have re-written the relevant paragraph in our introduction (p.2-3, l.52-66) to better explain the interest of constituent analyses compared to observational datasets:

Satellite observations of stratospheric composition are retrieved with varying spatial and temporal resolutions which depend on the instrument design, the retrieval strategy and the circumstances of its operational use. Data Assimilation Systems can process these datasets (Lahoz and Errera, 2010) to deliver, at regular time intervals, analyses which are meshed on a two-dimensional grid (total column) or on a three-dimensional grid (vertically resolved field). The spatial and temporal gradients in these analyses are expected to reflect dynamical and chemical processes rather than the details of the observing system. This feature is exploited in several studies of the photochemistry of the middle atmosphere, especially in the polar regions (see e.g. Robichaud et al, 2010; Lahoz et al, 2011; Sagi et al, 2014).

Thanks to their gridded and instantaneous description of the atmospheric composition, chemical analyses enable short-range to middle-range forecasts (Flemming et al., 2011) and are much easier to use and to interpret than satellite observations. The resulting "snapshot" maps show stratospheric composition at a specific time and are routinely used to monitor the evolution of the ozone layer, e.g. above the Antarctic (Antarctic Ozone Bulletins distributed by WMO/GAW: <http://www.wmo.int/pages/prog/arep/gaw/ozone/index.html>).

We could not find in our conclusions any paragraph stating specifically that "*the four analyses perform well where data are present and not so well where there are no data*". In any case the end of our conclusions has been re-written (see above our reply to the first comment).

While our study does not intend to explore specifically the added value of Chemical Data Assimilation in the stratosphere and upper troposphere, this area has indeed gained interest: Richard Ménard (Enviroment Canada) and Quentin Errera (BIRA-IASB) have recently formed at the International Space Science Institute a Study Group to explore precisely this topic (<http://www.issibern.ch/program/teams.html#Teams2014>).

-
4. *Figures: The labels are too small to read without strong magnification. Almost all plots have that problem.*

We apologize for this technical issue. It looks like figures 2, 4, 5, 8 and 10 were erroneously resized to allow them to fit the special page layout of ACPD. We have corrected all these figures (except for figure 4 which has been removed from the revised manuscript – see below) and we will make sure with the production office that all figures are easily readable in the published version of the revised paper.

Specific comments

Abstract. Too long. In particular, when you state a result you do not need to provide explanation. For example, on P12463 L20, everything after the comma can be deleted.

We have re-written most of the abstract, using the new conclusion text given here. This revised abstract is 33 lines long (length in original manuscript was 45 lines) and takes into account the guideline reminded by the reviewer.

P12464 L1 'may be related'. That's speculative. I would delete this part of the sentence.

This has been done.

P12465 L12 'information of satellite observations' → information obtained from satellite observations.

This has been done.

P12468 L4. '(...) with a vertical resolution of about 3 km.' Add 'in the stratosphere'. It's actually closer to 6 km in the mesosphere.

This has been done.

Section 2.1.3. Please clarify. Do SACADA and TM3DAM use both, total ozone and profiles from GOME-2? How many independent pieces of information do the profiles have?

As stated in the title of this section and summarized in Table 1 ("TC" versus "PC" in column "Assim. Data"), the GOME-2 data assimilated by SACADA and TM3DAM consists only in total columns. One sentence has been added at the end of section 2.1.3 to clarify this:

This study used only GOME-2 total columns: the GOME-2 ozone profiles were not assimilated.

P12469 L 23. '(: :) of which four are currently still operational: NOAA-16/17/18/19' Please, double check that. As far as I know NOAA-18 SBUV is not operational (<http://www.oso.noaa.gov/poesstatus/spacecraftStatusSummary.asp?spacecraft=18>) and NOAA-16 SBUV data are not usable at the moment.

Indeed this had to be corrected. As explained above, section 2.1.4 has been re-written and now starts with the following sentence where the correct dates and satellite numbers are given:

SBUV/2 is a series of seven remote sensors on NOAA weather satellites (McPeters et al., 2013), of which three were assimilated by IFS-MOZART during the period investigated here (September 2009 to September 2012): NOAA-17 and NOAA-18 during the whole period; NOAA-19 after 2011-06-22.

Section 2.1.4. As for GOME-2, I suggest discussing the number of independent pieces of information and/or smoothing errors inherent in ozone partial columns retrieved from nadir observations. For SBUV there is a good analysis of smoothing errors in Kramarova et al., 2013. Typically, the actual information content in the vertical is much less than suggested by the number of layers. This applies to GOME-2 as well.

This is addressed in our reply to the first general comment. Since this issue is not the cause for the poorer performance of the NRT analyses by IFS-MOZART, we think that it is not necessary to

discuss the number of independent pieces of informations and/or smoothing errors in SBUV/2. This is not necessary for GOME-2 either since we assimilated only its total columns. Repeating our reply above, the following sentence has been inserted in the conclusions:

The newer SBUV/2 v8.6 profiles are distributed over 21 layers and each profile is distributed with its matrix of Averaging Kernels. Kramarova et al. (2013) illustrated the importance of using properly this information. While it is planned to implement SBUV/2 Avergaing Kernels in the MACC NRT system at ECMWF, our last sensitivity test shows that this improvement was not necessary to assimilate successfully SBUV/2 v8 after a vertical re-gridding over 6 thick layers.

P12470 L6. The MLS Data Quality Document lists two components of error: precision and accuracy. Are both taken into account in these estimates – and in assimilation?

P12470 L6 refers to the validation of MLS v2 by Froidevaux et al. (2008). The abstract reads:

“The uncertainty estimates are often of the order of 5%, with values closer to 10% at the lowest stratospheric altitudes”.

These estimates are obtained from comparison with observations by other instruments, hence should be considered as typical biases. We think that this can not be compared quantitatively with the estimations of accuracy and precision in the MLS Data Quality Document because those are obtained “upstream” by the retrieval team. In our revised manuscript each sentence about the validation of an input dataset has been moved to the dataset description. With respect to MLS v2.2 we now state more clearly:

Froidevaux et al. (2008) estimated from comparisons with other instruments that the the MLS v2 ozone profiles have an uncertainty of the order of 5 % in the stratosphere, with values closer to 10% at the lowest stratospheric altitudes.

P12469 L24. Version 8 SBUV retrievals are given on 21 layers between 1000hPa – TOA. Does IFS-MOZART combine the layers somehow to get six? How is it done?

See above: SBUV version 8.6 gives retrievals on 21 layers but we used version 8 retrievals given 13 layers. In any case the reviewer guessed correctly: this dataset is combined into 6 layers prior to assimilation. Repeating the reply above to 1st general comment:

(...) we used the same procedure as first described for ERA-40 (Dethof and Holm, 2004): in order to decrease unwanted vertical correlations between errors at different levels, the thirteen layers of the original SBUV v8 retrievals were combined at ECMWF over six thick layers (0.1–1 hPa, 1–1.6 hPa, 1.6–4.1 hPa, 4.1–6.4 hPa, 6.4–16 hPa, 16 hPa–surface).

P12470 L8-10. Again, MLS errors are reported as precision and accuracy. Is the combined error 10%? When you say that most biases disappear in v3.4 you are talking about the accuracy component. Please, be more precise here.

The original manuscript is indeed not clear as it probably confused the input data of our sensitivity tests (MLS v3.4) with the output of these tests. There is unfortunately no validation paper similar to Froidevaux et al. (2008) but for MLS v3. So after reporting about the MLS v2 validation (see reply above), the revised manuscript now states:

Sensitivity tests were performed with IFS-MOZART, BASCOE and SACADA using the offline MLS v3 dataset (see section 6.2). The accuracy and precision of these retrievals (Livesey et al., 2013b) are very

similar to those reported for MLS v2 (Livesey et al. , 2013a) so the uncertainties of MLS v3 are expected to be at least as small as those reported for MLS v2.

P12472 L2. *'The system also includes a parameterization of the effects of Polar Stratospheric Clouds (PSCs) on the gas-phase species'. This is vague. What kind of parameterization? Does it include catalytic ozone destruction? While Table 2 provides references to model chemistry schemes it would help if it were stated explicitly whether or not each of these models includes heterogeneous ozone loss on PCS (as it is done in the case of IFS-MOZART). This is important because the paper talks about the representation of ozone holes in the analyses.*

For IFS-MOZART and SACADA, clear descriptions of the heterogeneous chemistry and PSC parameterization can be found in Stein et al. (2013) and in Elbern et al. (2010), respectively. With respect to BASCOE, the vague sentence on P12472 L2 has been replaced by the following paragraph:

Heterogeneous reactions on the surface of Polar Stratospheric Clouds (PSC) particles are explicitly taken into account. The BASCOE version used here adopts a simple cold-point temperature parameterization to represent the surface area available for these reactions: type Ia (Nitric Acid Trihydrate) PSCs are set to appear at temperatures between 186 K and 194 K with a surface area density of 10^{-7} cm²/cm³. At gridpoints colder than 186K they are replaced by type II PSCs (i.e. water ice particles) with a surface area density of 10^{-6} cm²/cm³.

P12472 L14. *'In contrast, IFS-MOZART assimilates other satellite instruments apart from Aura MLS, but those are measuring only ozone as species relevant for stratospheric chemistry.' I don't understand this sentence.*

This sentence was not necessary and has been removed in the revised manuscript.

P12475, L13. *'The background variance is set to 50 %.' Please, state this more clearly: The background error variance is 50% of what? The units of variance are the square of the units of the field itself. You can't get, say, ppmv² by taking 50% of ppmv. The proportionality coefficient has to be a dimensioned factor.*

This statement has been corrected:

The background error standard deviation is set to 50% of the background field, which is quite low...

P12476 L26. Define μ

Brewer data at $\mu > 3$ were filtered out, where μ is the increase in the ozone optical path length due to the obliquity of the sun's rays (Brewer, 1973).

P12482 L4. *Looking at Figure 2, the two large positive excursions in SACADA are punctuated by sharp dips before the onset of the ozone hole. This likely indicates some underlying errors in the model chemistry. I think this warrants more discussion than what is provided in the following subsection, which basically simply explains that there are no data in the polar night.*

It was found that the sharp dips in SACADA Total Ozone Columns at Syowa are due to corrupted datafiles on 2011-07-09, 2011-07-10, 2012-06-29, 2012-06-30 and 2012-07-01. Figure 2 has been re-plotted with re-processed SACADA output files.

P12482 L20 positive biased ! positively biased

This has been corrected.

P12483 L9. 'Latitudes not covered by observations can therefore only be influenced via tracer transport.' – and chemistry.

This has been corrected.

P12483 L24-27. Does this mean that only the two sonde measurements directly below and above a model level were used in the interpolation? Doing it this way will introduce a lot of noise. Assuming that assimilated ozone at a given layer represents the mean mixing ratio within that layer, the best strategy would be to integrate the portion of a sonde profile that falls within that layer and compare that (pressure weighted) integral with the mixing ratio value from assimilation. Looking at Figure 4, there is a lot of scatter in the sonde data vs. assimilation. I wonder if some of it is due to the way the data are interpolated. I suggest that the sonde – assimilation comparisons be repeated using mean (mixing ratio integrated w.r.t. pressure and divided by delta p) sonde observations within each layer.

The reviewer is correct: the interpolation used only two sonde measurements directly below and above a model level. We agree that this method introduced some noise in the sonde observations. Figure 4 shows sonde data from individual profiles, so the scatter may very well be due to the imprecise vertical interpolation. **Figure 4 does not add much information compared with figure 3, so we have removed it from the manuscript as well as the two paragraphs which discussed it** (P12484 L3 and L26).

It should be noted that for the time-averaged vertical profiles, the sonde data was correctly re-gridded to the vertical grids of the analyses in a mass-conserving manner (algorithm described by Langerock et al., 2014). Hence the figures 3, 4, 10 and 12 of the revised manuscript are not problematic.

P12487 L5 'Mixing ratios are appropriately scaled by an altitude independent factor using the model's one ozone profile.' I'm not sure if I understand. Do you mean that the background (guess) profile is scaled proportional to the ozone column increment by applying the transpose of the observation operator? People who are not well versed in data assimilation will wonder if this is some ad hoc trick or if it is part of the mathematics of the best linear unbiased estimation.

This sentence was not clear indeed. You understood it correctly, and this is a result of the observation operator using the modelled mixing ratios as input (not an ad hoc trick). But in practice the assimilated profiles will not be simply scaled with respect to the first guess, because SACADA has a flow-dependent background error correlation (Elbern et al., 2010) and also because it is a 4D-VAR system which transports the information from each observation away from its location. So we preferred to **delete this confusing sentence** and add one word on line 8:

“We find that, in the case of SACADA, the lack of information constraining **the shape of** the ozone profile leads primarily to...”

P12487 L26. I wouldn't say that the tropopause location is hard to define. We just go ahead and define it. The hard part is to correctly represent the sharp ozone gradients near the tropopause.

Indeed. We have re-written this sentence and the next one:

These sharp ozone gradients in the Upper Troposphere-Lower Stratosphere (UTLS) are very difficult to represent in three-dimensional models and probably require a very fine vertical resolution (Conside et al., 2008). Furthermore relative differences are amplified in this region due to its low ozone abundance.

P12490 L22+. How do these standard deviations compare to the ACE-FTS errors?

Standard deviations are on average around 6–7%. **This is only slightly larger than the relative mean difference between ACE-FTS and coincident MLS profiles, reported by Dupuy et al. (2009, table 7) as +4.7%.**

P12491 L6. ‘Anticorrelation between levels’. Three levels are shown. Which two are anticorrelated? From the plot it looks like maybe 50 hPa and 100hPa but it’s not very clear. What is the correlation coefficient?

This has been corrected in the revised manuscript:

..., interpolation at specific pressure levels (10, 50, and 100 hPa) reveals alternatingly positive and negative biases in the vertical for IFS-MOZART, both in the Arctic and in the Antarctic, especially during ozone hole events (Fig. 7),...

P12491 L14-28. I don’t see how the missing MLS data can explain the differences between BASCOE and IFS-MOZART given that, as you say, the effect in March is minimal, there are no ACE data in all of April, and both systems were assimilating MLS for most of May. It’s more likely that the difference arises from either different chemistry schemes or the fact that IFS-MOZART assimilated UV/VIS data the whole time and BASCOE did not. Some of the UV data may have influenced the composition inside the vortex either directly or through transport. This harks back to my general comment (1).

Indeed this interpretation was not correct. Only the sensitivity tests discussed in section 6.2 can explain the significantly worse performance of IFS-MOZART NRT in March 2011, and these tests show that the assimilation of UV/VIS data is not the culprit (see our reply to your first general comment). This whole paragraph has been replaced by a much simpler and shorter statement:

In the Arctic, the biases of SACADA and IS-MOZART become largest in March 2011. The obvious explanation is the occurrence of exceptionally low ozone abundances in the Arctic during this period (Manney et al., 2011). This event will be discussed in section 6.

The first paragraph of section 6.2 has been replaced as follows:

On 26 March 2011 Aura MLS stopped sending data and resumed normal operations on 19 April 2011. BASCOE ran freely (unconstrained CTM mode) during this time, and started again to assimilate MLS as soon as observations came back. IFS-MOZART assimilated only UV/VIS observations from 26 March 2011 until 10 May 2011, when the assimilation of Aura MLS was switched back on. Unfortunately, ACE-FTS did not collect any measurements in the Arctic during April 2011 (see Sect. 2.4.4).

These uncontrolled modifications of the observing system led us to explore in a more systematic manner the impact of the assimilated observations on the quality of the analyses. We chose a one-month period with the Arctic ozone depletion already well underway while MLS and ACE-FTS were still scanning the area, i.e. the month of March 2011. We first defined three new experiments with IFS-MOZART, BASCOE and SACADA assimilating the same dataset: Aura MLS version 3.3 offline ozone, keeping all observations down to 215 hPa. BASCOE was not allowed to assimilate any other species than ozone. To allow a short spin-up period of about one week, the three systems were started on 25 February from the BASCOE analysis delivered in NRT for that date.

P12492 L5-8. '...may be due to the fact...' – very speculative. Does this mean that UV/VIS data degrade the analysis? Why is that? See my general comment (1)

This confusing text has been removed from the revised manuscript.

P12492 L21-23. Again, can you explain why UV/VIS degrade the performance?

No, the assimilation of UV/VIS observations did not degrade the performance of IFS-MOZART. This confusing text has been removed from the revised manuscript.

P12493 L26-28. Please re-write. You don't have to go into details here but I don't think it's fair to say that ozone depleting gases are 'trapped' in PSCs. If anything is trapped it is nitric acid. The PSCs particles serve as a surface for heterogeneous reactions, which convert the reservoir species into ClO and Cl₂. Then, in the presence of sunlight, active chlorine catalytically destroys ozone.

Inside the vortex, the air masses were cold enough to allow PSC particles to condense. Heterogeneous reactions took place at the surface of these particles, converting chlorine reservoir molecules HCl and ClONO₂ into chemically active ClO and Cl₂. Hence catalytic destruction of ozone could start as soon as sunlight came back to illuminate these air masses.

P12495 L10+. You mean underestimation of mixing ratio not ozone depletion, right? It's clear from the plot but not from the context.

At the level where ozone depletion is maximum ($\theta \approx 485$ K), we see that the depletion is much too severe in IFS-MOZART NRT analyses and completely absent in SACADA NRT analyses.

P12495 L19. '(..) may be related': are there any other differences between the two runs of BASCOE? If there aren't than the slightly better performance is due to different versions of MLS (not 'may be').

This may be due to two different causes: the assimilation of MLS offline v3.3 instead of MLS offline v2.2, and/or an improvement in the pre-processing of the ECMWF wind fields which drive the transport in BASCOE. Indeed the BASCOE version used in NRT suffered from aliasing errors in the input wind fields, leading to some erroneous noise in the horizontal distribution of chemical tracers (Fig. 8).

P12495 L21-22. Either the extended range of MLS or the absence of other ozone data

In the revised manuscript this paragraph has been replaced by some new text which is explained above.

References added in the revised manuscript

- Bhartia, P. K., McPeters, R. D., Flynn, L. E., Taylor, S., Kramarova, N. A., Frith, S., Fisher, B., and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm, *Atmos. Meas. Tech.*, 6, 2533–2548, doi:10.5194/amt-6-2533-2013, 2013.
- Brewer, A. W.: A replacement for the Dobson spectrophotometer?, *Pure Appl. Geophys.*, 106, 919–927, doi: 10.1007/BF00881042, 1973.
- Considine, D. B., Logan, J. A., and Olsen, M. A.: Evaluation of near-tropopause ozone distributions in the Global Modeling Initiative combined stratosphere/troposphere model with ozonesonde data, *Atmos. Chem. Phys.*, 8, 2365–2385, doi:10.5194/acp-8-2365-2008, 2008.
- Dethof, A. and Hólm, E. V.: Ozone assimilation in the ERA-40 reanalysis project, *Q. J. Roy. Meteor. Soc.*, 130, 2851–2872, 2004.
- Kramarova, N. A., Bhartia, P. K., Frith, S. M., McPeters, R. D., and Stolarski, R. S.: Interpreting SBUV smoothing errors: an example using the quasi-biennial oscillation, *Atmos. Meas. Tech.*, 6, 2089–2099, doi:10.5194/amt-6-2089-2013, 2013.
- Lahoz, W. A., Errera, Q., Viscardy, S., and Manney, G. L.: The 2009 stratospheric major warming described from synergistic use of BASCOE water vapour analyses and MLS observations, *Atmos. Chem. Phys.*, 11, 4689–4703, doi:10.5194/acp-11-4689-2011, 2011.
- Langerock, B., De Mazie`re, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, *Geoscientific Model Development Discussions*, 7, 8151–8178, doi:10.5194/gmdd-7-8151-2014, 2014.
- Livesey, N., Read, W., Lambert, A., Cofield, R., Cuddy, D., Froidevaux, L., Fuller, R., Jarnot, R., Jiang, J., Jiang, Y., Knosp, B., Kovalenko, L., Pickett, H., Pumphrey, H., Santee, M., Schwartz, M., Stek, P., Wagner, P., Waters, J., and Wu, D.: EOS Aura/MLS Version 2.2 and 2.3 Level 2 data quality and description document, Jet Propulsion Laboratory, Pasadena, California, http://mls.jpl.nasa.gov/data/v2_data_quality_document.pdf, 2013a.
- McPeters, R. D., Bhartia, P. K., Haffner, D., Labow, G., and Flynn, L.: The version 8.6 SBUV ozone data record: an overview, *J. Geophys. Res.*, 118, 8032–8039, doi:10.1002/jgrd.50597, 2013.
- Robichaud, A., Ménard, R., Chabrilat, S., de Grandpré, J., Rochon, Y. J., Yang, Y., and Charette, C.: Impact of energetic particle precipitation on stratospheric polar constituents: an assessment using monitoring and assimilation of operational MIPAS data, *Atmos. Chem. Phys.*, 10, 1739–1757, doi:10.5194/acp-10-1739-2010, 2010.
- Sagi, K., Murtagh, D., Urban, J., Sagawa, H., and Kasai, Y.: The use of SMILES data to study ozone loss in the Arctic winter 2009/2010 and comparison with Odin/SMR data using assimilation techniques, *Atmos. Chem. Phys. Discuss.*, 14, 7889–7916, doi:10.5194/acpd-14-7889-2014, 2014.

References to correct in revised manuscript

Livesey et al. (2011) becomes Livesey et al. (2013b):

- Livesey, N., Read, W., Froidevaux, L., Lambert, A., Manney, G., Pumphrey, H., Santee, M., Schwartz, M., Wang, S., Cofield, R., Cuddy, D., Fuller, R., Jarnot, R., Jiang, J., Knosp, B., Stek, P., Wagner, P., and Wu, D.: EOS Aura/MLS Version 3.3 and 3.4 Level 2 data quality and description document, Jet

Propulsion Laboratory, Pasadena,
http://mls.jpl.nasa.gov/data/v3_data_quality_document.pdf, 2013b.

California,

Stein (2013) becomes Stein et al. (2013):

Stein, O., Huijnen V. and Flemming, J.: Model description of the IFS-MOZART and IFS-TM5 coupled systems, MACC-II project deliverable D_55.4, ECMWF, [http://www.gmes-atmosphere.eu/documents/maccii/deliverables/grg/MACCII GRG DEL D 55.4 IFS-MOZART and TM5.pdf](http://www.gmes-atmosphere.eu/documents/maccii/deliverables/grg/MACCII_GRG_DEL_D_55.4_IFS-MOZART_and_TM5.pdf) , 2013.

Reply to review 2 of paper by Lefever et al. (ACPD, 2014)

Dear reviewer,

The authors would like to thank you for your review of our manuscript – even and especially when it becomes somewhat provocative. We did our best to address all your comments, although some are very general and difficult to address indeed. Below is a point-by-point response to your comments (copied in italics); paragraphs highlighted in bold font have been inserted into the revised manuscript.

General comments

- 1. The whole set-up of the MACC stratospheric system strikes me as strange and more explained by politics than by science. Why are there 4 models at all, and why these? I can see the difference between using MLS NRT and MLS science data, but why is there not just a delayed mode IFS run using MLS science data?*

As the reviewer probably knows, the set-up of the MACC stratospheric system was primarily the result of the different approaches adopted by the European Union and the European Space Agency with respect to the monitoring of atmospheric composition. While the details of this “politics” are not interesting in ACP, the different priorities and operational constraints applied to the four Data Assimilation Systems (DAS) described here resulted in different configurations which explain our results. For example, separate IFS-MOZART runs in delayed mode are generally avoided due to their large computing costs.

Most of the introduction has been re-written to better explain this context and the motivations for selecting and configuring the four systems discussed here. After an explanation about the interest of constituent analyses (see reply to third general comment by Reviewer 1), we have included the following paragraphs:

For ten years, the development of these monitoring and forecasting abilities has been the primary goal of a series of European projects. The European Union project MACC-II (Monitoring Atmospheric Composition and Climate – Interim Implementation) was the third in a series of projects funded since 2005 to build up the atmospheric service component of the Global Monitoring for Environment and Security (GMES)/Copernicus European programme (Peuch et al., 2013). In this paper, the terminology “MACC” refers to both the MACC and MACC-II projects. The final goal of MACC is to cover all aspects of atmospheric dynamics and chemistry with one global DAS based on an operational Numerical Weather Prediction (NWP) system.

Two coupled systems were created in MACC: IFS-TM5 and IFS-MOZART (Flemming et al., 2009; Stein et al., 2013). These coupled dynamics-chemistry DAS are run at ECMWF in NRT for monitoring present and near-future atmospheric conditions up to 5 days ahead, through analyses and forecasts of carbon monoxide (CO), formaldehyde (HCHO), nitrogen oxides (NO_x i.e. NO+NO₂), sulphur dioxide (SO₂) and ozone (O₃). They were both designed to deliver in one run a complete and self-consistent picture of atmospheric chemistry and dynamics and both solve explicitly a complete set of photochemical reactions relevant to tropospheric chemistry. The description of photochemistry in IFS-MOZART also includes the halogen species, the reactions of interest in the stratosphere, and a parameterization of the heterogeneous reactions responsible for ozone depletion in the polar lower stratosphere.

For European-scale analyses relevant to Air Quality applications, MACC organized successfully an ensemble of limited-area Chemistry Data Assimilation (CDA) systems (Gauss et al., 2012). A similar approach was adopted to deliver global analyses of stratospheric and total column ozone through the MACC stratospheric ozone service (<http://www.copernicus-stratosphere.eu>). Besides IFS-MOZART, this service uses three independent CDA systems in order to identify model weaknesses and aid in the improvement of the main system. These three systems are BASCOE (Errera et al., 2008; Viscardy et al., 2010), SACADA (Elbern et al., 2010), and TM3DAM (Eskes et al., 2003; van der A et al., 2010). These three systems first delivered monitoring services for the programme PROMOTE (PROtocol MOniToring for the GMES Service Element Atmosphere - <http://www.gse-promote.org>) which was funded by the European Space Agency from 2004 until 2009. They are run at the centers where they were designed, use offline analyses of atmospheric dynamics, and have more relaxed operational constraints than the NRT runs of IFS-MOZART and IFS-TM5 at ECMWF.

The TM3DAM system is specifically designed to generate a long-term ozone column dataset: the ozone multi-sensor reanalysis (MSR), which document the day-to-day variability and allow trend studies trends in total ozone over more than 30 years. Contrarily to IFS-MOZART, BASCOE and SACADA were developed specifically to study and monitor stratospheric chemistry. Their adjoint models include photochemistry, allowing these 4D-VAR systems to deliver multi-variate analyses which should provide a more self-consistent chemical analysis of the stratosphere than possible with IFS-MOZART. Until now BASCOE and SACADA have assimilated only one instrument at a time and BASCOE processed only vertical profiles from limb-scanning instruments. In view of its advanced modelling of transport and background error covariances, it was decided to assimilate with SACADA only total ozone columns. This sub-optimal configuration was meant to increase the variety of assimilation set-ups in the “mini-ensemble” test the quality of 3-D ozone analyses by an advanced 4D-VAR system in the absence of limb profilers, and to and it was decided to configure BASCOE differentiate their products through the assimilation of differen datasets. This is used as a reference to test the MACC system and to guide developments.

In this paper we compare the ozone analyses delivered in NRT by these four systems over the three-year period September 2009–September 2012, using as reference several datasets of independent observations: groundbased instruments, balloon soundings and a solar occultation satellite instrument. We also explore the roles of the input datasets in the outcome of this exhaustive validation. Our study is similar to the ASSET intercomparison of ozone analyses (Geer et al., 2006; Lahoz et al., 2007), with some major differences: here the DAS were configured primarily to satisfy operational constraints and deliver NRT products (and in the case of IFS-MOZART to deliver several tropospheric products in addition to stratospheric ozone); they assimilated a large variety of datasets while ASSET used only observations from Envisat (Environmental Satellite); and the investigated period is much longer (3 years instead of 5 months).

If the BASCOE model outperforms the other models so clearly in most respects, what is then the added value of SACADA and TM3DAM? Which of the data sets is the user supposed to use, and what are possible applications for such assimilation systems?

The BASCOE DAS actually does not outperform SACADA when both systems assimilate the same input dataset (see fig. 11 of revised manuscript) but the analyses delivered in NRT by BASCOE indeed outperformed those delivered by SACADA due to different configuration choices. This has been clarified in the revised version of the manuscript (see our reply to first comment by reviewer 1).

The outcome of this study was not known at the beginning of MACC and in our opinion it deserves proper documentation, i.e. all four systems must be included. The beginning of the conclusion was re-written to provide a good rationale for the inclusion of these four systems and to provide guidance to the users:

Four ozone data assimilation systems (DAS) have been run continuously and simultaneously during several years. These DAS have very different designs (offline or online dynamics; grid set-up; specification of background error covariances) and were set-up very differently with respect to the assimilated datasets. In this paper we seized this opportunity, first to provide an intercomparison and validation of the resulting analyses, and second to investigate the causes of their very different biases.

This study shows what can be achieved in Near Real Time (NRT) with state-of-the-art DAS for stratospheric ozone and provides guidance to the users of the resulting analyses. Among the three sets of vertically resolved NRT analyses of stratospheric ozone, those delivered by BASCOE had the best overall quality. This is due primarily to the focus of BASCOE on stratospheric observations retrieved from limb sounders, and to more relaxed operational constraints allowing it to wait for the delivery of the best input dataset available.

TM3DAM is based on a sequential Kalman Filter algorithm and does not model stratospheric chemistry explicitly. It aims only to provide total columns of ozone by making optimal use of the ozone column measurements from UV-Vis satellite sounders, with very small biases between the analyses/forecasts and satellite datasets. It was shown that TM3DAM is a good reference to test the ability of the three other systems to produce accurate ozone column amounts.

The low quality of the analyses delivered in NRT by SACADA is a good indication of the drawbacks to expect from current CDA systems when they are configured to assimilate total ozone columns only. This should be considered as a worst-case scenario in a future situation where no limb sounder would be available and no proper effort would be invested to assimilate correctly vertical profiles retrieved from nadir-looking instruments.

Finally, while IFS-MOZART did not deliver the best NRT analyses in this intercomparison, it still has the potential to deliver the best analyses (figures 12 and 13). Official reviews of international monitoring capacities (e.g. WMO, 2011b) expect an imminent lack of ozone-profiling capabilities at high vertical resolution. Contrarily to the BASCOE version used here, IFS-MOZART should be able to adapt to this situation thanks to its demonstrated ability to assimilate several instruments simultaneously.

Why not use measurements directly or one dedicated model having good stratospheric chemistry such as BASCOE?

Reviewer 1 made a similar comment about the direct usage of observations: see our reply (general comment 3) which results in two new introductory paragraphs. The forward model in BASCOE can not be used to replace its analyses with chemical forecasts, even though it is driven by dynamical analyses. This is now by a new paragraph in section 2.2.2:

If we use the BASCOE forward Chemistry-Transport Model (CTM) with no constraining observations, the stratospheric ozone fields diverge after a few weeks from our analyses. These results are similar to those found with IFS-MOZART by Flemming et al. (2011). In the case of the BASCOE CTM this is due to the absence of tropospheric processes and surface emissions which prevents proper exchanges with the troposphere; and to the parameterization of PSC surface area density which lacks any memory of the coldness experienced by polar air masses. This last issue was discussed by Lindenmaier

et al. (2011) using the coupled model GEM-BACH which inherited its photochemistry and PSC parameterization from BASCOE.

2. *The one message that I will remember from this paper is that all systems perform well where they have assimilated the right data and perform disappointingly weak where there is no data assimilated or not the right one used (in this case O3 profiles from an IR limb profiler). This is an important message with large implications for the planning of future satellite missions...*

We thank the reviewer for highlighting this essential point. We will make it the last item in our conclusions and in our abstract:

- **When they assimilate the same dataset with good quality and large observational density, BASCOE, IFS-MOZART and SACADA deliver very similar performance despite their very different designs. The quality of modern ozone analyses depends primarily on the assimilated data. This conclusion has large implications for the planning of future satellite missions.**
-

... but could have been delivered in a much shorter manuscript which would have been read by many more people.

In retrospect we would have preferred indeed to split this study into a validation paper (to publish e.g. in GMD) and a shorter paper about system design and intercomparison. Unfortunately the MACC series of projects is coming to an end, and we do not have the time or resources to re-submit two papers. We believe that the large effort documented here is in any case worthy of publication in ACP.

3. *For the same reason, large parts of the manuscript feel repetitive even if they are using different data sets for comparison – as BASCOE is strongly constrained by MLS, and MLS is in good agreement with other measurements (ground-based, sondes, ACE-FTS), comparison of BASCOE results with different validation datasets comes down to a repetition of MLS validation. The same is true for the column assimilating systems – if SCIAMACHY / GOME-2 columns are as good as stated in the text, one would hope that the assimilation system will agree well with other observations in those locations where this data is assimilated.*

While one would certainly hope for such a positive outcome from multi-instrument validation, it was still necessary to prove that MLS constrains BASCOE (and can constrain SACADA and IFS-MOZART) strongly enough to allow it. This had not been done yet with NRT analyses (to the best of our knowledge). We think that the teams responsible for observations could benefit from reading our paper, as it would guide them in using such analyses to interpret their measurements or plan their campaigns.

In summary and somewhat provocatively, I think this is a detailed, thorough and well written validation study but I do not see a lot of readers for it.

We have re-written large parts of the introduction, conclusions and abstract to address your comments as well as those by the first reviewer. We think that this should give more impact to our study. We have even slightly modified the title to reflect the increased focus of the revised manuscript on input datasets:

Copernicus Atmospheric Service for stratospheric ozone, 2009-2012: Validation, systems intercomparison and roles of input datasets

So we do hope that more than a few readers will read this paper, in whole or in part.

Minor comments

- I find the first part of the introduction a bit arbitrary and even confusing in some places. For example, I don't think that interest in stratospheric ozone and the measurement systems really started only after the ozone hole was discovered (the first TOMS and SBUV instruments were launched way before that, as were the studies about ozone depletion by gas-phase chemistry involving ClOx and NOx and the use of ozone as tracer for stratospheric transport). Also I think that the first PROMOTE project actually preceded GEMS. It might also be worthwhile to mention that data assimilation has a longer tradition for meteorological models than for models of the chemical composition.*

The introduction has been nearly completely re-written, with more focus on satellite data and CDA. Keeping only the beginning of the first paragraph (until P12464 L26), we now continue with

Data assimilation determines a best possible state for a system using observations and short range forecasts. This process was first developed to enable Numerical Weather Prediction (NWP; e.g. Lorenz et al., 1986). Systematic satellite measurements of ozone started in the late 1970s with the series of Total Ozone Mapping Spectrometer (TOMS) and Solar Backscatter Ultraviolet Instrument (SBUV) instruments. The discovery of the Antarctic ozone hole in 1985 (Farman et al., 1985) led to the development of improved satellite instruments to observe the composition and dynamics of the stratosphere. While these instruments played a key role in the discovery of the physical processes responsible for the ozone hole (e.g. Solomon, 1999), the variety of sounders monitoring the ozone layer soon opened a new application in the field of data assimilation: Chemical Data Assimilation (CDA), or more properly constituent data assimilation (Rood et al., 1989; Lahoz and Errera, 2010).

Satellite observations of stratospheric composition are retrieved with varying spatial and temporal resolutions which depend...

The introduction now goes on to explain the advantages of CDA (see reply to 3rd general comment of reviewer 1), then with a short history of the DAS used in this study and finishes with the motivation for the study itself (see reply to your 1st general comment above).

- Aura satellite: The statement that Aura provides coverage between 82S and 82N is misleading as coverage depends also on the swath width of a satellite instrument and therefore measurements by instruments such as OMI cover all latitudes, at least when there is enough light.*

“from 82°S to 82°N” (P12467 L19) has been removed from the revised manuscript.

- *SCIAMACHY spatial resolution should be separated between nadir (32 x 60 km²) and limb*
In the revised manuscript, the horizontal resolution of SCIA nadir has been corrected. Since SCIA limb is not used, the mention of its vertical resolution has been deleted.
- *GOME-2 nadir profiles are mentioned which is confusing as they are not used later. It might also be good to mention Metop-B in this context.*
The text has been clarified (see similar specific comment by reviewer 1). In this section we prefer to discuss only the satellite instruments which were actually assimilated in this study.
- *SBUV-2 – I would replace “larger precision” by “lower precision”*
This has been done.
- *SAOZ / DOAS description is mixed up as it is not clear which parts of this are common to all UV/vis instruments and which are specific to SAOZ. To my knowledge, the only important difference is that there is a large network of similar SAOZ instruments while the other UV/vis instruments tend to be designed and operated by individual research groups.*
DOAS is a retrieval technique which can be used for many instruments in the UV/Vis spectral range, including those of the SAOZ family. Among the three UV/Vis instruments described in this section, all use DOAS. Two instruments (at Zhigansk and Scoresby Sund) have the SAOZ design but one (at Harestua) does not. This has been clarified in the text.
- *Alert comparison (figure 2 and section 3.1) – I find differences of 50 DU for summer in the Arctic quite a lot and wonder what the reason for such large discrepancies could be in systems assimilating measurements.*
We do not see any differences as large as 50 DU with observations. As discussed in the text the systems disagree a lot during spring 2010 and spring 2011. TM3DAM and IFS-MOZART actually agree, while SACADA and BASCOE deliver lower TOC. For BASCOE this is not surprising since it does not assimilate TOC at all (see text). The issues with SACADA assimilation of TOC are discussed in section 3.4.
- *page 12487, line 6 and page 12498, line 1. I find this use of “the models one ozone profile” confusing and would suggest to replace by “the model’s own profile”*
See replies to specific comments of reviewer 1: these sentences (P12487 L6 and P12498 L1) are confusing and not necessary; they have been deleted.
- *page 12490 line 22: similarly => similar*
..., the results by IFS-MOZART and BASCOE are very similar.
- *page 12493, last sentences: The description of ozone depletion through heterogeneous processes is not quite correct, please rewrite.*
This has been done (see new text in replies to specific comments of reviewer 1).

- *many figures are very small and have even smaller labels and axis numbers. Please enlarge.*

We apologize for this technical issue. It looks like figures 2, 4, 5, 8 and 10 were erroneously resized to allow them to fit the special page layout of ACPD. We will make sure with the production office that all figures are easily readable in the published version of the revised paper.

References to add in revised manuscript

- Gauss, M., Valdebenito, A., Benedictow, A., et al.: Report on end-to-end analysis for the seven individual production lines and for ensemble production, MACC-II project deliverable D_108.1, ECMWF, http://www.copernicus-atmosphere.eu/documents/maccii/deliverables/ens/MACCII_ENS_DEL_D_108.1_201301_METE_OFR.pdf, 2012.
- Lahoz, W. A., Geer, A. J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Khattatov, B., Marchand, M., Massart, S., Peuch, V.-H., Rharmili, S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H. E., Vik, A. F., and von Clarmann, T.: The Assimilation of Envisat data (ASSET) project, *Atmos. Chem. Phys.*, 7, 1773-1796, doi:10.5194/acp-7-1773-2007, 2007.
- Lindenmaier, R., K. Strong, R. L. Batchelor, P. Bernath, S. H. Chabrilat, M. Chipperfield, W. H. Daffer, J. R. Drummond, W. Feng, A. I. Jonsson, F. Kolonjari, G. L. Manney, C. A. McLinden, R. Ménard, and Walker, K. A.: A study of the Arctic NO_y budget above Eureka, Canada, *J. Geophys. Res.*, 116, D23302, doi:10.1029/2011JD016207, 2011.
- Lorenc, A. C.: Analysis methods for numerical weather prediction, *Q. J. Roy. Met. Soc.*, 112, 1177-1194, doi:10.1002/qj.49711247414, 1986.
- Rood, R. B., Allen, Dale J., Baker, W. E., Lamich, D. J. and Kaye, J. A.: The Use of Assimilated Stratospheric Data in Constituent Transport Calculations. *J. Atmos. Sci.*, 46, 687–702, doi:10.1175/1520-0469(1989)046<0687:TUOASD>2.0.CO;2, 1989.
- Van der A, R., Allaart, M. and Eskes, H.: KNMI multi-sensor reanalysis version 2: 33 year ozone record, MACC-II project deliverable D_54.3, ECMWF, http://www.copernicus-atmosphere.eu/documents/maccii/deliverables/grg/MACCII_GRG_DEL_D54.3_May2014_VanderA_final.pdf, 2014.

References to correct in revised manuscript

Stein (2013) becomes Stein et al. (2013):

- Stein, O., Huijnen V. and Flemming, J.: Model description of the IFS-MOZART and IFS-TM5 coupled systems, MACC-II project deliverable D_55.4, ECMWF, http://www.gmes-atmosphere.eu/documents/maccii/deliverables/grg/MACCII_GRG_DEL_D_55.4_IFS-MOZART_and_TM5.pdf, 2013.

Copernicus ~~atmospheric service~~ Atmospheric Service for stratospheric ozone, 2009-2012: ~~validation and~~ Validation, systems intercomparison and roles of four ~~near real-time analyses, 2009–2012~~ input datasets

K. Lefever¹, R. van der A², F. Baier³, Y. Christophe¹, Q. Errera¹, H. Eskes², J. Flemming⁴, A. Inness⁴, L. Jones⁴, J.-C. Lambert¹, B. Langerock¹, M. G. Schultz⁵, O. Stein⁵, A. Wagner⁶, and S. Chabrillat¹

¹Belgian Institute for Space Aeronomy, BIRA-IASB, Ringlaan 3, 1080 Brussels, Belgium

²Royal Dutch Meteorological Institute, KNMI, P.O. Box 201, 3730 AE De Bilt, the Netherlands

³German Aerospace Center, DLR, Muenchner Str. 20, 82234 Wessling, Germany

⁴European Centre for Medium-range Weather Forecasts, ECMWF, Shinfield Park, Reading RG2 9AX, UK

⁵Research Center Jülich, FZJ, Institute for Energy and Climate Research, IEK-8: Troposphere, Wilhelm-Johnen-Straße, 52428 Jülich, Germany

⁶Deutscher Wetterdienst, DWD, Meteorologisches Observatorium Hohenpeissenberg, Albin-Schwaiger-Weg 10, 82383 Hohenpeissenberg, Germany

Correspondence to: S. Chabrillat (Simon.Chabrillat@aeronomie.be)

Abstract. This paper evaluates ~~the performance and discusses the quality~~ of the stratospheric ozone analyses delivered in near real time by the MACC (Monitoring Atmospheric Composition and Climate) project during the 3 year period between September 2009 and September 2012. Ozone analyses produced by four different ~~chemistry transport models and data assimilation techniques are~~
5 ~~examined~~ Chemical Data Assimilation (CDA) systems are examined and compared: the ECMWF Integrated Forecast System (IFS) coupled to MOZART-3 (IFS-MOZART), the BIRA-IASB Belgian Assimilation System for Chemical Observations (BASCOE), the DLR/RIU Synoptic Analysis of Chemical Constituents by Advanced Data Assimilation (SACADA), and the KNMI Data Assimilation Model based on Transport Model version 3 (TM3DAM). The assimilated satellite ozone
10 retrievals differed for each system: SACADA and TM3DAM assimilated only total ozone observations, BASCOE assimilated profiles for ozone and some related species, while IFS-MOZART assimilated both types of ozone observations.

~~The stratospheric ozone analyses are compared to independent ozone observations from ground-based instruments, ozone sondes and the ACE-FTS (Atmospheric Chemistry Experiment—Fourier Transform~~
15 ~~Spectrometer) satellite instrument. All analyses show~~ All analyses deliver total column values which ~~are generally in good agreement with groundbased~~ agree well with ground-based observations (biases <5 %) and have a realistic seasonal cycle. ~~The only exceptions are found for BASCOE which systematically underestimates~~, except for BASCOE analyses which underestimate total ozone in

the Tropics with about 7–10% all year long by 7 to 10% at Chengkung (Taiwan, 23.1°N/121.365°E),
20 resulting from the fact that BASCOE does not include any tropospheric processes, and for SACADA
which overestimates, and SACADA analyses which overestimate total ozone in the absence of UV
observations for the assimilation.

Due to the large weight given to column observations in the assimilation procedure, IFS-MOZART
is able to reproduce total column observations very well, but polar night regions by up to 30%.
25 The validation of the vertical distribution is based on independent observations from ozone sondes
and the ACE-FTS (Atmospheric Chemistry Experiment – Fourier Transform Spectrometer) satellite
instrument. It can not be performed with TM3DAM which is designed only to deliver analyses of
total ozone columns. Vertically alternating positive and negative biases compared to ozonesonde and
ACE-FTS satellite data are found in the vertical IFS-MOZART analyses as well as an overestimation
30 of 30 to 60 % in the polar lower stratosphere during polar ozone depletion events. The assimilation of
near real-time (NRT) Microwave Limb Sounder (MLS) profiles which only go down to 68 hPa SACADA
underestimates lower stratospheric ozone by up to 50 hPa is not able to correct for the deficiency of
the underlying MOZART model, which may be related to the applied meteorological fields.

Biases of BASCOE compared to ozonesonde or ACE-FTS ozone profiles do not exceed % during
35 these events above the South Pole and overestimates it by approximately the same amount in the
Tropics. The three-dimensional analyses delivered by BASCOE are found to have the best quality
among the three systems resolving the vertical dimension, with biases not exceeding 10% over
the entire vertical stratospheric range, thanks to the good performance of the model in ozone hole
conditions and the assimilation of offline MLS profiles going down to 215 hPa % all year long, at all
40 stratospheric levels and in all latitude bands, except in the tropical lowermost stratosphere.

TM3DAM provides very realistic total ozone columns, but is not designed to provide information
on the vertical distribution of ozone.

Compared to ozonesondes and ACE-FTS satellite data, SACADA performs best in the Arctic, but
shows large biases (>50%) for ozone in the lower stratosphere in the Tropics and in the Antarctic,
45 especially during ozone hole conditions.

This study shows that ozone analyses with realistic total ozone column densities do not necessarily
yield good agreement with the observed ozone profiles. It also shows the large benefit obtained from
The Northern Spring 2011 period is studied in more detail to evaluate the ability of the analyses to
represent the exceptional ozone depletion event which happened above the Arctic in March 2011.
50 Offline sensitivity tests are performed during this month and indicate that the differences between
the forward models or the assimilation algorithms are much less important than the characteristics
of the assimilated datasets. They also show that IFS-MOZART is able to deliver realistic analyses
of ozone both in the troposphere and in the stratosphere, but this requires the assimilation of a
single limb-scanning instrument (Aura-MLS) with a high density of observations. Hence even
55 state-of-the-art models of stratospheric chemistry still require observations from nadir-looking instruments

as well as the assimilation of ~~limb observations for a correct representation of the vertical distribution of ozone in the~~ profiles which are well resolved vertically and extend into the lowermost stratosphere.

1 Introduction

60 The presence of a high-altitude ozone layer in the atmosphere, which protects ~~Earth the~~ Earth system against the harmful ultraviolet (UV) light from the Sun, was first determined in the 1920s from observations of the solar UV spectrum. Systematic measurements of stratospheric ozone using ozonesondes started in the late 1950s (Solomon et al., 2005). At that time, the development of satellites just started, the first one (Sputnik) being launched in 1957.

65 Systematic satellite measurements of ozone started in the late 1970s with the series of Total Ozone Mapping Spectrometer (TOMS) and Solar Backscatter Ultraviolet Instrument (SBUV) instruments. ~~While the interest in stratospheric ozone started off as purely scientific curiosity (??), the attitude changed after the~~ The discovery of the ~~Antarctic~~ ozone hole in 1985 (Farman et al., 1985) ~~, which led to the build-up of a comprehensive monitoring program consisting of balloon-borne ozonesondes, Brewer-Dobson spectrometers, and various satellite instruments.~~ ~~development of improved satellite instruments to observe the composition and dynamics of the stratosphere. These instruments played a key role in the discovery of the physical processes responsible for the ozone hole (e.g. Solomon, 1999).~~

~~The emergence and increasing capabilities of computer facilities opened at the same time the path to numeric modeling, which has permitted simulating the behavior of the atmosphere and predicting its evolution under specific conditions. The chemical processes involving ozone-depleting substances, in particular the heterogeneous reactions in~~ Data assimilation determines a best possible state for a system using observations and short range forecasts. This process was first developed to enable Numerical Weather Prediction (NWP; e.g. Lorenc, 1986) . In view of the ~~polar regions, which lead to the ozone hole, are by now well known and documented in previous studies (e.g. Solomon, 1999) and ozone assessments (e.g. ?)~~ planned increase in the number and variety of sounders monitoring the ozone layer, the last years of the 1980's saw the appearance of a new application for data assimilation: Chemical Data Assimilation (CDA), or more properly constituent data assimilation (Rood et al., 1989; Lahoz and Errera, 2010) .

85 ~~Satellite observations of stratospheric composition are retrieved with varying spatial and temporal resolutions which depend on the instrument design, the retrieval strategy and the circumstances of its operational use. Data Assimilation Systems can process these datasets (Lahoz and Errera, 2010) to deliver, at regular time intervals, analyses which are meshed on a two-dimensional grid (total column) or on a three-dimensional grid (vertically resolved field). The spatial and temporal gradients in these analyses are expected to reflect dynamical and chemical processes rather than the details of the~~

observing system. This feature is exploited in several studies of the photochemistry of the middle atmosphere, especially in the polar regions (see e.g. Robichaud et al., 2010; Lahoz et al., 2011; Sagi et al., 2014).

Through data assimilation (?), chemistry transport models (CTM) can process information of satellite observations and use them to update the modeled 3-D chemistry field, in order to match real-world observations as close as possible and deliver more realistic datasets than pure model forecasts. Thanks to their gridded and instantaneous description of the atmospheric composition, chemical analyses enable short-range to middle-range forecasts (Flemming et al., 2011) and are much easier to use and to interpret than satellite observations. The resulting analyses of atmospheric composition enable the monitoring and forecasting of the distribution of “snapshot” maps show stratospheric composition at a specific time and are routinely used to monitor the evolution of the ozone layer, e.g. above the Antarctic (Antarctic Ozone Bulletins distributed by WMO/GAW; <http://www.wmo.int/pages/prog/arep/gaw/ozone> and other key constituents in the earth atmosphere (Lahoz and Errera, 2010) [/index.html](#)).

Succeeding to GEMS (?; Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in-situ data), followed by PROMOTE (PROtocol MOniToring for the GMES Service Element: Atmosphere) and MACC, the EU (European Union) FP7 (Framework Project) For ten years, the development of these monitoring and forecasting abilities has been the primary goal of a series of European projects. The European Union project MACC-II (Monitoring Atmospheric Composition and Climate – Interim Implementation) is/was the third in a series of projects funded since 2005 to build up the atmospheric service component of the Global Monitoring for Environment and Security (GMES)/Copernicus European programme (?)(Peuch et al., 2014). In this paper, the terminology “MACC” refers to both the MACC and MACC-II projects. MACC provides NRT tropospheric and stratospheric data. The final goal of MACC is to cover all aspects of atmospheric dynamics and chemistry with one global DAS based on an operational Numerical Weather Prediction (NWP) system.

Two coupled systems were created in MACC: IFS-TM5 and IFS-MOZART (Flemming et al., 2009; Stein et al., 2013). These coupled dynamics-chemistry DAS are run at ECMWF in NRT for monitoring present and near-future atmospheric conditions up to 5 days ahead (Stein et al., 2012), as well as a reanalysis for the years 2003–2012 (Inness et al., 2013). The MACC stratospheric ozone service () deals with the evolution of stratospheric ozone and related species, and provides regular stratospheric ozone analyses and forecasts. The service currently uses a set of four independent chemical assimilation systems :-, through analyses and forecasts of carbon monoxide (CO), formaldehyde (HCHO), nitrogen oxides (NO_x i.e. NO+NO₂), sulphur dioxide (SO₂) and ozone (O₃). They were both designed to deliver in one run a complete and self-consistent picture of atmospheric chemistry and dynamics and both solve explicitly a complete set of photochemical reactions relevant to tropospheric chemistry. The description of photochemistry in IFS-MOZART (Flemming et al., 2009; Stein et al., 2012), an operational implementation of also includes the halogen species, the reactions of interest in the

stratosphere, and a parameterization of the heterogeneous reactions responsible for ozone depletion in the polar lower stratosphere.

130

For European-scale analyses relevant to Air Quality applications, MACC organized successfully an ensemble of limited-area CDA systems (Gauss et al., 2013). A similar approach was adopted to deliver global analyses of stratospheric and total column ozone through the MACC stratospheric ozone service (<http://www.copernicus-stratosphere.eu>). Besides IFS-MOZART, this service uses three independent CDA systems in order to identify model weaknesses and aid in the improvement

135

of the main system. These three systems are BASCOE (Errera et al., 2008; Viscardy et al., 2010), SACADA (Elbern et al., 2010), and TM3DAM (Eskes et al., 2003; van der A et al., 2010). (Eskes et al., 2003; van der A et al., 2010)

140

These three systems first delivered monitoring services for the programme PROMOTE (PROtocol MOniToring for the GMES Service Element Atmosphere - <http://www.gse-promote.org>) which was funded by the European Space Agency from 2004 until 2009. They are run at the centers where they were designed, use offline analyses of atmospheric dynamics, and have more relaxed operational constraints than the NRT runs of IFS-MOZART and IFS-TM5 at ECMWF.

145

~~To ensure the reliability and suitability of the data for decision makers and other users, assessing and documenting the quality of the delivered data products is of key importance.~~ The TM3DAM system is specifically designed to generate a long-term ozone column dataset: the ozone Multi-Sensor Reanalysis (MSR), which document the day-to-day variability and allow trend studies trends in total ozone over more than 30 years. Contrarily to IFS-MOZART, BASCOE and SACADA were developed specifically to study and monitor stratospheric chemistry. Their adjoint models include photochemistry, allowing these 4D-VAR systems to deliver multi-variate analyses which should provide a more self-consistent chemical analysis of the stratosphere than possible with IFS-MOZART.

150

Until now BASCOE and SACADA have assimilated only one instrument at a time and BASCOE processed only vertical profiles from limb-scanning instruments. In view of its advanced modelling of transport and background error covariances, it was decided to assimilate with SACADA only total ozone columns. This sub-optimal configuration was meant to increase the variety of assimilation set-ups in the “mini-ensemble” test the quality of 3-D ozone analyses by an advanced 4D-VAR system in the absence of limb profilers, and to and it was decided to configure BASCOE differentiate their products through the assimilation of differen datasets. This is used as a reference to test the MACC system and to guide developments.

155

160

~~In this paper, we assess the ability of the MACC stratospheric ozone service to simulate stratospheric ozone for the 3-period~~ we compare the ozone analyses delivered in NRT by these four systems over the three-year period September 2009–September 2012. ~~This intercomparison paper does not aim directly at model issues, but rather documents the overall quality of the MACC stratospheric service, i.e. a full-blown intercomparison of the models is out of the scope of this paper.~~ Intercomparison 2012, using as reference several datasets of independent observations: groundbased instruments, balloon soundings and a solar occultation satellite instrument. We also explore the roles of the

165 input datasets in the outcome of this exhaustive validation. Our study is similar to the ASSET
intercomparison of ozone analyses ~~has been done before by Geer et al. (2006) who evaluated the~~
~~performance of different ozone data assimilation systems based on (Geer et al., 2006; Lahoz et al., 2007),~~
with some major differences: here the DAS were configured primarily to satisfy operational constraints
and deliver NRT products (and in the case of IFS-MOZART to deliver several tropospheric products
170 in addition to stratospheric ozone); they assimilated a large variety of datasets while ASSET used
only observations from Envisat (Environmental Satellite) only. ~~Here we intercompare for the first~~
~~time analyses based on data from different satellites, either total ozone column or profile observations~~
~~or a combination of both;~~ and the investigated period is much longer (3 years instead of 5 months).

The next section describes the different analyses in the MACC stratospheric ozone service and the
175 reference observations used for their validation. Section 3 contains the evaluation of the total ozone
columns based on Brewer/Dobson observations, while the vertical distribution of ozone is assessed
in Sects. 4 and 5, through comparison with ozonesondes and ACE-FTS satellite data, respectively.
In Sect. 6, we assess the performance of the MACC analyses during an event of exceptional nature:
the Arctic ozone hole 2011. We additionally investigate the influence of the assimilated dataset on
180 the performance of the analyses for one month covered by this event: March of 2011. The final
section provides a summary and conclusions.

2 Data

The MACC stratospheric ozone service currently consists of four independent systems, running rou-
tinely on a daily basis, with a maximum delay of 4 days between data acquisition and delivery of the
185 analyses: IFS-MOZART (1 day), BASCOE (4 days), SACADA (2 days), and TM3DAM (2 days).
This section gives a detailed description of the analyses: the observations that were assimilated, the
underlying ~~chemistry transport model~~ atmospheric composition models, the applied data assimila-
tion algorithms, and the way the different ~~models~~ DAS deal with background error statistics. Table 1
summarizes the satellite retrievals of ozone that were actively assimilated by the four ~~models~~ DAS of
190 the MACC stratospheric ozone service, while an overview of the system specifications can be found
in Table 2. This section additionally includes a description of the datasets used in the validation of
the four analyses.

2.1 Assimilated observations

2.1.1 Aura satellite: OMI total columns and MLS profiles

195 Aura is NASA's (National Aeronautics and Space Administration) third large Earth Observing Sys-
tem (EOS) mission, flying in a sunsynchronous nearly-polar orbit since 9 August 2004, aiming at the
provision of trace gas observations for climate and air pollution studies (Schoeberl et al., 2006). Due
to its nearly-polar orbit, Aura is able to provide a nearly global latitude coverage ~~from 82S to 82N~~. It

has four instruments onboard, amongst which the Ozone Monitoring Instrument (OMI, Levelt et al.,
200 2006) and MLS (Waters et al., 2006), which provide complementary information.

The OMI instrument is a nadir-viewing imaging spectrometer, measuring the solar radiation backscattered by the Earth's atmosphere and surface in the ultraviolet to visible (UV-VIS) wavelength range, providing total ozone columns with a horizontal resolution of $13 \text{ km} \times 24 \text{ km}$ at nadir. ~~OMI ozone data are~~ This dataset is delivered in near real-time and was validated using Brewer and
205 Dobson spectrophotometer ground-based observations (Balis et al., 2007). OMI also provides nadir ozone profiles, but these have not been ~~used~~ assimilated.

MLS is a limb-viewing microwave radiometer, providing some 3500 daily vertical profile measurements of several atmospheric parameters, such as ozone (O_3), nitric acid (HNO_3), water vapour (H_2O), hydrochloric acid (HCl), hypochlorous acid (HOCl), and nitrous oxide (N_2O) from about
210 8 to 80 km (0.02 hPa to 215 hPa) with a vertical resolution of about 3 km in the stratosphere and a horizontal resolution of 200–300 km (Waters et al., 2006). As a microwave remote sensing sounder, MLS also provides observations during the polar night, which has a positive impact on ozone analyses during the onset of the ozone hole.

MLS ozone data are delivered in near real-time by NASA/JPL (Jet Propulsion Laboratory), with
215 a latency of only 2 to 4 h, whereas a scientific dataset, containing additionally non-ozone species, is delivered with a delay of 4 days. The former dataset is used for the assimilation of ozone by IFS-MOZART (see Sect. 2.2.1), whereas the latter is used by BASCOE (see Sect. 2.2.2) for the assimilation of O_3 , HNO_3 , H_2O , HCl, and N_2O (v2.2). The useful range of both datasets differs: NRT ozone profiles ~~are were~~ only recommended for scientific use at pressure levels ~~68–0.20.2–68~~ hPa (~~except~~
220 ~~for the calculation of ozone columns where the 100 pressure level may be used as a lower limit~~), while the offline MLS dataset ~~can could~~ be used for the entire pressure range ~~215–0.020.02–215~~ hPa.

Froidevaux et al. (2008) estimated from comparisons with other instruments that the the MLS v2 ozone profiles have an uncertainty of the order of 5% in the stratosphere, with values closer to 10% at the lowest stratospheric altitudes. These lower stratospheric biases mostly disappear with
225 the improved MLS v3.4 data (Livesey et al., 2013b), which have a useful range of 261 to 0.1 hPa. Sensitivity tests were performed with IFS-MOZART, BASCOE and SACADA using the offline MLS v3 dataset (see section 6.2). The accuracy and precision of these retrievals (Livesey et al., 2013b) are very similar to those reported for MLS v2 (Livesey et al., 2013a) so the uncertainties of MLS v3 are expected to be at least as small as those reported for MLS v2.

230 2.1.2 Envisat satellite: SCIAMACHY total columns

SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) is a UV-VIS-NIR (Near InfraRed) imaging spectrometer onboard ESA's Environmental Satellite (Envisat) launched on 1 March 2002. SCIAMACHY observed earthshine radiance in limb and nadir viewing geometry and solar and lunar light transmitted through the atmosphere in occultation view-

235 ing geometry. While spectrometers such as MLS are able to provide ozone profiles over the poles throughout the year, UV-VIS instruments such as SCIAMACHY are limited to periods with sufficient solar radiation. On the other hand, they can attain much higher spatial resolution. SCIAMACHY ~~data have a nadir total columns have a~~ horizontal resolution of typically 32 km × 21560 km and ~~a limb vertical resolution of 3.~~ were extensively validated against groundbased measurements (Eskes et al., 2005).

240 After having operated five years beyond the planned mission lifetime of five years, all communication with the Envisat satellite was lost on 8 April 2012. IFS-MOZART ~~assimilates assimilated~~ SCIAMACHY total ozone columns until the last date (7 April 2012). To have a clean monthly mean, it was decided to reprocess TM3DAM for the first days of April using GOME-2 from 1 April 2012 ~~,~~ onwards. Due to a better global coverage within one day for GOME-2 (SCIAMACHY attains global coverage in 6 days), leading to an improved performance, the official MACC NRT product for SACADA had already switched from SACADA-SCIAMACHY to SACADA-GOME2 on 28 October 2011.

2.1.3 MetOp-A satellite: GOME-2 total columns

250 The GOME-2 (Global Ozone Monitoring Experiment-2) instrument carried onboard EUMETSAT's (European Organisation for the Exploitation of Meteorological Satellites) Meteorological Operational Satellite MetOp-A (launched in October 2006) continues the long-term monitoring of atmospheric trace gases by ESA's (European Space Agency) ERS-2 (European Remote sensing Satellite-2) GOME. It is a nadir-viewing UV-VIS scanning spectrometer, which is able to achieve global coverage within one day (Munro et al., 2006). Total columns are provided with a horizontal resolution of 80 km × 40 km, ~~while ozone profiles are given as partial ozone columns on varying pressure levels (40 levels between surface and 0.01).~~ GOME-2 ~~NRT products total ozone columns~~ are available about two hours after sensing ~~and were validated against groundbased measurements by Loyola et al. (2011).~~ Ozone profiles are also retrieved from this instrument but this study used only the total columns.

2.1.4 NOAA satellite: SBUV-2 partial columns

~~The SBUV-2 radiometer onboard the NOAA (National Oceanic and Atmospheric Administration) satellites continues the long-term record started by NASA's SBUV onboard Nimbus-7. SBUV-2 is a SBUV/2 is a series of seven remote sensors on NOAA weather satellites (McPeters et al., 2013), of which four are currently still operational: NOAA-16/17/18/19. It provides partial columns mostly at stratospheric levels. Six partial columns are three were assimilated by IFS-MOZART, located between 0.1 and 1013.~~

~~The above data have been extensively validated against groundbased measurements: OMI by Balis et al. (2007), MLS by Froidevaux et al. (2008), SCIAMACHY by Eskes et al. (2005), GOME-2~~

270 by Loyola et al. (2011) and SBUV-2 by ? . Total column estimates are generally in agreement within
during the period investigated here (September 2009 to September 2012): NOAA-17 and NOAA-18
during the whole period; NOAA-19 after 2011-06-22. Bhartia et al. (2013) describe the two latest
versions of the SBUV/2 (OMI, SCIAMACHY, GOME-2, MLS) , except for SBUV-2, which has
a somewhat larger precision of 5–15. The uncertainty estimates of the MLS v2.2 ozone profile in the
275 stratosphere are often of the order of 5, with values closer to 10 at the lowest stratospheric altitudes.
Tests with the improved MLS v3.4 data (?), which have a useful range of 261 to retrievals: v8
which was available during the period investigated here, and v8.6 which was released more recently.
While SBUV v8.6 includes the averaging kernels (AK) for each retrieved profile, these were not
available in the v8 BUFR data used operationally at ECMWF. Hence we used the same procedure
280 as first described for ERA-40 (Dethf and H6lm, 2004) : in order to decrease unwanted vertical
correlations between errors at different levels, the thirteen layers of the original SBUV v8 retrievals
were combined at ECMWF over six thick layers (0.1 show that a lot of those biases in the lower
stratosphere disappear—1 hPa, 1–1.6 hPa, 1.6–4.1 hPa, 4.1–6.4 hPa, 6.4–16 hPa, 16 hPa–surface).
Among the resulting partial ozone columns, the last one contributes most to the total columns.

285 2.2 ~~Models description~~ Description and setup of the data assimilation systems

2.2.1 IFS-MOZART

Within the GEMS project, the Integrated Forecast System (IFS), operated by ECMWF, was extended
to be able to simulate and assimilate the abundance of greenhouse gases (Engelen et al., 2009),
aerosols (Morcrette et al., 2009; Benedetti et al., 2009), as well as tropospheric and stratospheric
290 reactive gases (Flemming et al., 2009; Inness et al., 2009; Stein et al., 2012) from satellite retrieval
products. Satellite observations for the following reactive gases can be assimilated: O₃, nitrogen
dioxide (NO₂), carbon monoxide (CO), formaldehyde (HCHO), and sulphur dioxide (SO₂), but
only the former three were assimilated in the operational analysis discussed in this paper. The
assimilation window of IFS-MOZART is 12 h.

295 The version of IFS-MOZART used here was described in detail by Stein et al. (2013). To provide
concentrations and chemical tendencies of the reactive gases, the IFS was coupled to a CTM using
the coupling software OASIS4 (~~Ocean Atmosphere Sea Ice Soil, ?~~) (Ocean Atmosphere Sea Ice Soil: Redler et al., 2010) .
The IFS computes only the transport of the forementioned reactive gases while the coupled CTM
provides the chemical tendencies due to chemical conversion, deposition and emission.

300 The CTM selected to deliver analyses of stratospheric ozone for the MACC global monitoring and
forecast sytem is MOZART-3 (Kinnison et al., 2007; Stein et al., 2012) because it simulates both
tropospheric and stratospheric chemistry, including the catalytic destruction of ozone in the lower
polar stratosphere. Inness et al. (2009) give a detailed description of the applied procedure for the
assimilation of atmospheric constituents in IFS-MOZART.

305 During the period studied here, the IFS was run at T159L60, where T159 denotes an expansion
to wavenumber 159 in the spherical-harmonic representation used by the model (corresponding to
approximately 125 km horizontal resolution at the equator), and L60 denotes a vertical grid com-
prising 60 hybrid-pressure levels extending from 0.1 hPa down to the surface. This run uses IFS
version (“cycle”) 36R1. The CTM component, MOZART-3, used the same 60 vertical levels and
310 a regular longitude-latitude grid with $1.875^\circ \times 1.875^\circ$ horizontal resolution. Its chemical scheme
includes 115 species interacting through 325 reactions (Stein et al., 2013).

The following satellite O_3 data were simultaneously assimilated (see Table 1): partial columns by
NOAA SBUV-2, total columns by Aura OMI and Envisat SCIAMACHY, and profiles by Aura MLS
down to 68 hPa. Note that all ozone data assimilated in IFS-MOZART are NRT products. Hence
315 the MLS dataset used here (v2.2) is the product delivered 2 to 4 h after measurement, in contrast to
the data assimilated by BASCOE (see Sect. 2.2.2).

The IFS-MOZART version described here was run daily (experiment f93i) from 1 September
2009 till 30 September 2012, which determined the period considered in this paper.

2.2.2 BASCOE

320 BASCOE (Errera et al., 2008) is a 4D-Var system developed at the Belgian Institute for Space Aeron-
omy, BIRA-IASB. Based on a stratospheric CTM, BASCOE assimilates satellite retrievals of O_3 ,
 H_2O , HNO_3 , HCl , $HOCl$, and N_2O , gathered by MLS. The assimilation window is 24 h, while
BASCOE produces output every three hours. The CTM includes 57 species that interact using 143
gas-phase reactions, 48 photolysis reactions and 9 heterogeneous reactions. ~~The system also includes~~
325 ~~a parameterization of the effects-~~

~~Heterogeneous reactions on the surface of Polar Stratospheric Clouds (PSCs) on the gas-phase~~
~~species-~~

~~In this setup, the CTM is PSC particles are explicitly taken into account. The BASCOE version~~
~~used here adopts a simple cold-point temperature parameterization to represent the surface area~~
330 ~~available for these reactions: type Ia (Nitric Acid Trihydrate) PSCs are set to appear at temperatures~~
~~between 186 K and 194 K with a surface area density of $10^{-7} \text{ cm}^2/\text{cm}^3$. At gridpoints colder than~~
~~186K they are replaced by type II PSCs (i.e. water ice particles) with a surface area density of 10^{-6}~~
 ~~cm^2/cm^3 .~~

~~When the BASCOE forward Chemistry-Transport Model (CTM) is run with no constraining~~
335 ~~observations, the stratospheric ozone fields become less realistic after a few weeks or months,~~
~~depending on the region. These results are similar to those found with IFS-MOZART by Flemming et al. (2011).~~

~~In the case of the BASCOE CTM this is due to the absence of tropospheric processes and surface~~
~~emissions which prevents proper exchanges with the troposphere; and to the parameterization of~~
~~PSC surface area density which lacks any memory of the coldness experienced by polar air masses.~~

340 ~~This last issue was discussed by Lindenmaier et al. (2011) using the coupled model GEM-BACH~~

which inherited its photochemistry and PSC parameterization from BASCOE.

For the MACC stratospheric ozone service, the BASCOE DAS is driven by the ECMWF operational 6 hourly analyses (winds, temperature and surface pressure). BASCOE is run at a horizontal resolution of 3.75° longitude by 2.5° latitude, ~~and applies and uses~~ a vertical hybrid-pressure grid ~~with comprising~~ 37 levels, ~~extending from 0.1 down to the surface,~~ most of them lying in the stratosphere. ~~It As the driving meteorological analyses, this vertical grid extends from 0.01 hPa down to the surface.~~ BASCOE does not include any tropospheric processes and is therefore not expected to produce a realistic chemical composition below the tropopause, resulting in larger systematic error biases for the total columns and in the lower stratosphere.

~~The system is driven by the ECMWF operational 6hourly analyses (winds, temperature and surface pressure).~~ Both BASCOE and IFS-MOZART analyses assimilate Aura MLS data, but while IFS-MOZART uses the NRT retrievals v2.2 of ozone only, BASCOE uses the standard scientific, offline retrievals (level-2) v2.2 including five other species which are available with a delay of typically four days. ~~In contrast, IFS-MOZART assimilates other satellite instruments apart from Aura MLS, but those are measuring only ozone as species relevant for stratospheric chemistry.~~ BASCOE was configured to filter out ozone observations below 150 hPa.

2.2.3 SACADA

Within the project SACADA, a 4D-Var scheme has been developed by the Rhenish Institute for Environmental Research at the University of Cologne and partners (Elbern et al., 2010) aiming at the assimilation of atmospheric Envisat data using state-of-the-art numerical methods. This system has been implemented for operational use at the Deutsches Zentrum für Luft- und Raumfahrt, DLR, who deliver routinely daily (12 h UT) trace gas analyses based on Envisat SCIAMACHY ozone columns since March 2010. In parallel, another SACADA service assimilates MetOp-A GOME-2 total column data since January 2008. In research mode, SACADA has been successfully applied to other satellite- and groundbased observations (Elbern et al., 2010; Schwinger and Elbern, 2010; Baier et al., 2013).

The SACADA system uses an icosahedral grid (i.e. 20 equilateral triangles) on sigma-pressure levels with an approximate resolution of 250 km. The vertical grid consists of 32 model levels extending from 7 to 66 km altitude (440 to 0.1 hPa). The tropospheric ozone column is prescribed from the TOMS V8 climatology. Like IFS-MOZART and BASCOE, SACADA applies a comprehensive stratospheric chemistry scheme (see Table 2). The NRT service additionally provides information on the following unconstrained species: HNO₃, H₂O, and HCl. ~~Contrary to the other models applied for MACC, the SACADA CTM is driven by the~~ Unlike the other CDA systems used in MACC, SACADA is not driven directly by winds and temperature from the IFS NWP system; it takes these input fields from the meteorological forecast system GME (Majewski et al., 2001), run at DLR. GME is started from ECMWF analyses data daily at 0 h UTC and provides its own

24 h forecasts. The SACADA 4D-Var assimilation uses an assimilation window of 24 h. Note that SACADA products are delivered on a standard latitude-longitude grid with 3.75° by 2.5° resolution from 147 to 0.3 hPa altitude.

380 Here, we investigate two independent SACADA NRT products for two consecutive time intervals (see Table 1). NRT delivery started on 4 March 2010 with ~~SCIAMACHY-SACADA V2.0~~[SACADA 2.0 assimilating SCIAMACHY observations of total ozone columns \(version 5\)](#). After 28 October 2011, ~~SACADA was~~[SACADA was upgraded to version 2.4 and](#) switched to the GOME-2 instrument (~~SACADA V2.4 retrieval version GDP 4.1~~), which has a better ~~data coverage daily. Total columns were only provided with the data product from 13 January 2011 on.~~[daily data coverage than SCIAMACHY.](#)

385

2.2.4 TM3DAM

The TM3DAM data assimilation system is based on the TM3/TM5 tracer transport model and is driven by operational 6 hourly meteorological fields from ECMWF. The main purpose of TM3DAM is the generation of 30–45 year reanalyses of total ozone based on all available satellite datasets (van der A et al., 2010), but in MACC it has also been operated to provide real time analyses and forecasts. TM3 contains parameterized schemes for the description of stratospheric gas-phase and heterogeneous ozone chemistry.

390

The assimilation scheme in TM3DAM is based on a simplified Kalman-filter assimilation approach, with a time and space dependent error covariance, but with fixed correlations (Eskes et al., 2003), which considerably reduces the computational cost. The TM3DAM assimilation code has been updated as described in van der A et al. (2010). The ~~model runs system is run~~[at a global horizontal resolution of 3° longitude by 2° latitude](#). It applies a vertical hybrid-pressure grid, consisting of 44 levels extending from 0 hPa to the surface (1013 hPa). From the upper troposphere upwards, the layers coincide with the ~~ECMWF-60-layer model version~~[vertical grid used at ECMWF](#).

400

TM3DAM assimilates near real-time level-2 total ozone column data from Envisat/SCIAMACHY until the end of March 2012 and switched to MetOp-A/GOME-2 after all communication with the Envisat satellite was lost on 8 April 2012. NRT production of daily analyses (valid at 21 h UT) in the framework of MACC started on 16 March 2010. Only total columns are available. Besides the daily analyses, TM3DAM also generates daily forecasts for up to 9 days ahead. The Observation-minus-Forecast (OmF) statistics show that the bias of the system compared to the individual satellite measurements is typically less than 1 % for a forecast period of 1 day.

405

2.3 Comparison of ozone background errors

The specification of the background error covariance matrix (e.g., Kalnay, 2002) is one of the most difficult parts of an assimilation system as assimilation errors are never observed directly, they can only be estimated in a statistical sense. Each of the considered analyses has a different way of

410

dealing with background error statistics. In IFS-MOZART, the background error covariance matrix is given in a wavelet formulation (Fisher, 2006), allowing both spatial and spectral variations of the horizontal and vertical background error covariances (Inness et al., 2013). For ozone, the background error correlations were derived from an ensemble of forecast differences, using a method proposed by Fisher and Andersson (2001). The background error standard deviation profiles and the horizontal and vertical correlations can be found in Fig. 1 of Inness et al. (2009).

BASCOE analyses use a diagonal background error correlation matrix \mathbf{B} with a fixed error usually between 20 % to 50 % of the background field, 30 % in this version. The diagonal setup of \mathbf{B} implies that spatial correlations are neglected. Spatial correlations help to spread the information from the data into the model. As mentioned by Errera et al. (2008), they can be neglected in first approximation if the spatial coverage of the assimilated observations and their vertical resolution are comparable to the model-DAS resolution. This is the case here, where a maximum of three days of MLS observations are necessary to constrain all BASCOE grid points. Note that spatial correlation on the \mathbf{B} -matrix has been implemented recently in BASCOE (Errera and Ménard, 2012), following the method by Hollingsworth and Lönnberg (1986).

The SACADA 4D-Var assimilation uses a flow dependent parameterisation of the background error covariance matrix with a diffusion approach (Weaver and Courtier, 2001). The basic idea is to formulate covariances by Gaussians and approximate these Gaussians by integration of the diffusion operator over some specified time. Horizontal and vertical background error correlation lengths are fixed to 600 km and 3 km, respectively. The background variance-standard deviation is set to 50% of the background field, which is quite low and allows the observations to have a strong impact on results.

In the parameterized Kalman filter approach of TM3DAM, the forecast error covariance matrix is written as a product of a time independent (i.e. fixed) correlation matrix and a time dependent diagonal variance (Eskes et al., 2003). All aspects of the covariance matrix, including the time dependent error growth and correlation length, are carefully tuned on the basis of OmF (Observation minus Forecast) statistics. In the total ozone product a realistic time dependent error bar is provided for each location and time.

2.4 Reference ozone data

2.4.1 Brewer/Dobson observations

To assess the condition of the ozone layer, one frequently uses the total column of ozone. Roughly 150 ground stations perform total ozone measurements on a regular basis. Data are submitted into the World Ozone and UV Data Center (WOUDC), operated by Environment Canada (<http://www.woudc.org>), as part of the Global Atmosphere Watch (GAW) programme of the World Meteorological Organization (WMO). The observations are predominantly taken with Dobson and Brewer

UV spectrophotometers at about 60 and 70 stations respectively, but WOUDC also includes observations from UV-VIS DOAS spectrometers.

Even though Dobson and Brewer instruments are based on the same general measurement principle, previous studies have identified a seasonal bias of a few percent between their midlatitude total ozone column measurements, Brewer measurements being in slightly better agreement with satellite data than Dobson measurements. In the Northern Hemisphere, Dobson instruments exhibit a +1 % bias compared to Brewer instruments and the bias exhibits a seasonal cycle which is not the case for Brewer instruments (Scarnato et al., 2009; Lerot et al., 2013). Similar conclusions hold for the Southern Hemisphere. Since the Brewer network has not such a good coverage in the Southern Hemisphere, however, we use the Dobson instruments as a reference in the Antarctic, keeping in mind this +1 % bias compared to Brewer instruments (i.e. we did not correct the Dobsons for this bias, but instead used the original data).

In order to assess the quality of the total ozone columns (TOC) delivered by the 4 systems, we selected three stations from the WOUDC database for which the time coverage for this three year period was sufficiently large (red squares in Fig. 1: a high northern latitude station, Alert (82.49° N, 62.42° W, data gathered by the Meteorological Service of Canada), a tropical station, Chengkung (23.1° N, 121.365° E, data gathered by the Central Weather Bureau of Taiwan), and a southern latitude station, Syowa (69° S, 39.58° E, data gathered by the Japan Meteorological Agency). As indicated above, we used the observations gathered by the Brewer instruments at Alert and Chengkung, and those gathered by the Dobson spectrophotometer for Syowa. For Alert, we used the data for both Brewer instruments 019 (MKII) and 029 (MKV). The Brewer instrument (#061) at Chengkung is of type MKIV. Brewer data at $\mu > 3$ were filtered out, where μ is the increase in the ozone optical path length due to the obliquity of the sun's rays (Brewer, 1973). The Dobson instrument (#119) at Syowa was replaced on the 1st of February 2011 by a new Beck model (#122).

2.4.2 SAOZ/DOAS observations

Conventional techniques for measuring ozone in the UV, such as Dobson spectrometers, are inapplicable for SZA Solar Zenith Angles (SZA) larger than about 80°. The SAOZ (Système d'Analyse par Observation Zenithale, Pommereau and Goutail (1988)), measuring the absorption by the atmosphere Zenith-sky UltraViolet-VISible (UV-VIS) spectroscopy allows measurements of the atmospheric absorption of scattered sunlight at the zenith sky. This is the only instrument which is type of ground-based instruments able to measure continuously and at all latitudes up to the polar circle in winter. SAOZ are designed to allow observations of and total columns twice a day during twilight (sunrise and sunset). The retrieval is based on Differential Optical Absorption Spectroscopy (DOAS), a technique which has recently been. This technique was validated by Van Roozendaal et al. (1998) through comparisons with Dobson measurements and was recently improved by Hendrick et al. (2011). The accuracy of the zenith-sky UV-VIS spectroscopy for the measurement of total ozone has been

investigated by Van Roozendael et al. (1998) by comparing against Dobson measurements

The SAOZ (Système d'Analyse par Observation Zenithale, Pommereau and Goutail (1988)) instruments
485 belong to this family and have a standardized design which allows observations of NO₂ and O₃ total
columns twice a day during twilight (sunrise and sunset). As a general result, the SAOZ O₃ measure-
ments are between 2–8 % higher than the Dobson ones, with a scatter of about 5 % in midlatitudes
and increasing at higher latitudes.

For the “Arctic ozone hole 2011” case study in this paper (see Sect. 6), the total ozone columns
490 by the four analyses were compared with data received by three UV-VIS zenith-sky instruments
at Arctic locations, which are part of the Network for the Detection of Atmospheric Composi-
tion Change (NDACC, <http://www.ndacc.org>): ~~Zhigansk (Russia, 66.8° N, 123.42198° E), Harestua (Norway, 60° N, 21.9811° W), Zhigansk (Russia, 66.8° N, 123.4° E), and Scoresby Sund (Greenland, 70.49° N, 21.9811° W).~~ Data at Harestua
495 were gathered by BIRA-IASB (van Roozendael et al., 1995), the other two (SAOZ) E). The instru-
ments at Zhigansk and Scoresby Sund have the SAOZ design and are owned by LATMOS/CNRS
(Laboratoire Atmosphères, Milieux, Observations Spatiales/Centre National de Recherche Scien-
tifique) respectively by CAO (Central Aerological Observatory) and by DMI (Danish Meteorological
Institute) while the instrument at Harestua has an improved design and is operated by BIRA-IASB
500 (van Roozendael et al., 1995).

2.4.3 Ozonesonde profiles

Balloon-borne ozonesondes measure the vertical distribution of ozone concentrations up to an al-
titude of about 35 km. The observed ozonesonde profiles are archived by NDACC, WOUDC, and
the Southern Hemisphere ADditional OZonesondes network (SHADOZ, <http://croc.gsfc.nasa.gov/shadoz/>).
505 The majority of soundings (85 %) are performed with Electrochemical Concentration Cell
(ECC) sondes, while the remaining part consists of Brewer-Mast, Indian and Japanese Carbon-Iodine
sondes. Optimally treated, ECC sondes yield profiles with random errors of 3–5 % and overall un-
certainties of about 5 % in the stratosphere (Smit et al., 2007; Deshler et al., 2008; Stübi et al., 2008;
Hassler et al., 2013). Other sonde types have somewhat larger random errors of 5–10 % (Kerr et al.,
510 1994; Smit et al., 1996).

We use ozone observations gathered by balloon sondes at 38 locations, taken from the above-
mentioned databases for the period September 2009 to September 2012: 12 in the Arctic, 19 in the
Tropics, and 7 in the Antarctic (see Fig. 1). For each latitude band, we picked out one station which
is representative for the general behaviour in this latitude band and for which the time coverage for
515 this three year period was sufficiently large, for a more detailed discussion: the Arctic station at Ny-
Ålesund (79° N, 12° E), the equatorial station at Nairobi (1.27° S, 36.8° E), and the Antarctic station
at Neumayer (70.65° S, 8.25° W) (red dots in Fig. 1). Data are provided by the Alfred-Wegner
Institute in Potsdam, Germany (for Ny Alesund and Neumayer) and by MeteoSwiss in Payerne,

Switzerland (for Nairobi).

520 **2.4.4 ACE-FTS satellite data**

ACE-FTS is one of the two instruments on the Canadian satellite mission SCISAT-1 (first Science Satellite), ACE (Bernath et al., 2005). It is a high spectral resolution Fourier Transform Spectrometer operating with a Michelson interferometer. Vertical profiles of atmospheric parameters such as temperature, pressure and volume mixing ratios of trace constituents are retrieved from the occul-
525 tation spectra, as described in (Boone et al., 2005), with a vertical resolution of maximum 3–4 km. Level 2 ozone retrievals (version 3.0) are used as an independent reference dataset to validate the ozone profiles of the MACC stratospheric ozone system.

It must be noted that the low spatio-temporal sampling of ACE-FTS (due to the solar occultation technique) does not deliver profiles in all latitude bands for each month. There are also two periods
530 during the year where there are no measurements for a duration of almost 3 weeks due to the fact that the spacecraft is in constant sunlight: June and December (Hughes and Bernath, 2012). There are four periods per year, lasting about 1 month (Northern Hemisphere: April, June, August, December; Southern Hemisphere: February, June, October, December) with no occultation poleward of 60° (see Fig. 4 of Hughes and Bernath (2012)). At very high β angles (i.e. the angle between the orbital plane
535 of the satellite and the Earth-sun direction $> 57^\circ$), it is common practise to skip more than half of the available measurement opportunities to avoid exceeding onboard storage capacities and overlapping command sequences. Therefore, the amount of observations in the Tropics is significantly lower than in the polar regions.

The previous version of these retrievals (version 2.2) was extensively validated against 11 other
540 satellite instruments, ozonesondes and several types of ground-based instruments (Dupuy et al., 2009). This version reports more ozone than most correlative measurements from the upper troposphere to the lower mesosphere. Dupuy et al. (2009) found a “slight positive bias with mean relative differences of about 5 % between 15 and 45 km. Tests with a preliminary version of the next generation ACE-FTS retrievals (version 3.0) have shown that the slight positive stratospheric bias has
545 been removed.” Adams et al. (2012) additionally present an intercomparison of ACE ozone profiles (both versions 2.2 and 3.0) against groundbased observations at Eureka, confirming that the new ACE-FTS v3.0 and the validated v2.2 partial ozone columns are nearly identical, with mean relative difference of $0.0 \pm 0.2\%$ for v2.2. minus v3.0.

Standard deviations for levels where there are fewer than 20 observations are omitted for reasons
550 of non-representativeness.

3 Validation of total ozone columns

We intercompare for the first time analyses based on data from different satellites and of different types: partial/total ozone columns, profile observations or a combination of both. For an optimal interpretation of the validation results, it is important to keep in mind that SACADA and TM3DAM
555 exclusively assimilated total ozone columns, but while TM3DAM delivers only total ozone columns as output product, SACADA also provides ozone profiles. BASCOE exclusively assimilated vertical profiles of ozone (besides other species) and IFS-MOZART used a combination of total columns, partial columns and vertical profiles from various instruments.

In this section, we discuss the results obtained for the validation of the total ozone columns against
560 Brewer observations at Alert (Arctic) and Chengkung (Tropics), and against Dobson observations in Syowa (Antarctic). The TOC datasets from the four systems were interpolated to the latitude and longitude of these stations. The resulting time series are shown in Fig. 7, side by side with the corresponding observed groundbased data. ~~Remember that there are no SACADA total ozone products before 13 January 2011. Even though the time coverage at other locations was often poorer, the drawn conclusions are very similar (not shown).~~
565

3.1 Alert (Arctic)

The seasonal O_3 cycle at Alert is very similar each year. The only deviations from usual behaviour of the total ozone columns occur, e.g., in November 2009, when an airmass with exceptionally high ozone passed over Alert, and in February–March 2011, when 30 % of the total ozone column above
570 Alert was destroyed by the end of March. The latter event will be studied in detail in Sect. 6 as a separate case study.

All four analyses match each other and the observed total ozone columns very closely. Peak-to-peak difference in TOC are of the order of 250 Dobson Units (D.U.), with maximum values reached during boreal winter and spring as a result of poleward and downward transport of ozone-rich air by
575 the large-scale Brewer–Dobson circulation (Brewer, 1949; Dobson, 1956; Weber et al., 2011).

The only significant differences among the analyses occur during the O_3 maximum in northern spring (where mutual differences of maximum 50 D.U., about 10 %, are observed) and during the Arctic ozone hole season, where SACADA delivers TOC values which are about 75 D.U. (20 %) above the other analyses. Unfortunately, this coincides exactly with the periods where reliable
580 groundbased observations are missing due to the lack of sun light.

3.2 Chengkung (Tropics)

The ozone columns in the Tropics are lower (between 240 and 330 D.U.) due to the large-scale ascent of tropospheric low-ozone air and the higher incidence of solar radiation. Ozone maxima are reached in April each year, after which ozone is decreasing slowly until the beginning of November

585 and more rapidly afterwards. The lowest values are seen in December, January, and February (DJF), when the upwelling part of the Brewer–Dobson circulation is strongest. Ozone is recovering very rapidly from January till April. This seasonality is in general well reproduced by all analyses. The ozone recovery is slower in the analyses than in the observations and the observed ozone maxima are never reached.

590 IFS-MOZART, SACADA, and TM3DAM mutually differ by 2 % at most, and have biases of the order of maximum 5 % below the Brewer observations. BASCOE systematically underestimates total ozone by 20 D.U. throughout the year (about 7–10 %). As ~~we will see later~~discussed below (Sect. 4.4.2), this is due to the underestimation of ozone in the lower stratosphere(~~see Fig. 7 and the discussion in Sect. 4.5).~~

595 3.3 Syowa (Antarctic)

At the Antarctic station Syowa, the local spring-time ozone hole is evident, with values below 200 D.U. during the months September, October, and November (SON). The total ozone columns are reduced by up to 50 %, from approximately 300 D.U. during austral summer and autumn down to 150 D.U. during the austral spring season.

600 ~~Again, the seasonality~~The seasonal cycle of total ozone is very well reproduced by ~~all analyses.~~the IFS-MOZART, BASCOE and TM3DAM ~~deliver total columns which are~~analyses with results very close to each other (biases < 2%). Due to the loss of Envisat in April 2012, the differences between IFS-MOZART and TM3DAM become slightly larger. Before this incident, both IFS-MOZART and TM3DAM assimilated SCIAMACHY data, but afterwards, TM3DAM switched to
605 GOME-2, while IFS-MOZART continued to assimilate observations from SBUV/2, OMI, and MLS.

SACADA exhibits strong positive biases from observations during austral winters, right before the onset of the ozone hole (up to 30 % in 2012). Closer inspection of SACADA analyses shows that these larger differences coincide with missing SCIAMACHY and GOME-2 observations during polar night when solar zenith angles are close to or in excess of 90°. ~~This~~While this coverage effect
610 should especially influence ~~models systems~~which assimilate data from UV-instruments only.~~The,~~the TM3DAM system is found less vulnerable to data gaps than SACADA ,as it performs very well under the same circumstances.

3.4 Discussion of SACADA total column results

All analyses show a realistic seasonal cycle in all three latitude bands and total ozone column values
615 which are generally in very good agreement with independent observations, with the exception of SACADA during polar night. Differences between IFS-MOZART, BASCOE, and TM3DAM are usually within 5 %. Only a few exceptions were identified, i.e. larger mutual differences (up to 10 %) are found at high altitudes during polar night, and for BASCOE in the Tropics, where the ~~model system~~underestimates total ozone by 7–10 %.

620 In contrast to these three analyses, SACADA total ozone results deviate strongly from observa-
tions during certain episodes. There is a general tendency in SACADA results for ~~positive~~positively
biased ozone columns during the winter months at high latitudes compared to Alert and Syowa
station data in the Northern and Southern Hemisphere, respectively. Backscatter UV-instruments
provide no information for zenith angles above 90° . As recommended for SCIAMACHY data ver-
625 sion 3 (Lerot et al., 2007), only observations with zenith angles up to 75° were used. Thus, no
SCIAMACHY data were assimilated until May 2011 at the latitudes of Alert station (82.49° N).
Accordingly, at Syowa station (69° S), SCIAMACHY data was not processed from end of March
until end of September ~~2011. Data quality improved with SCIAMACHY version 5 (which is used
for this analysis, even though recommendations for version 3 were maintained) and the coverage~~
630 ~~would significantly improve with zenith angles up to 80° .~~

From 28 October 2011 onwards, GOME-2 observations were assimilated by SACADA up to
zenith angles of 90° . In this case, the instrument is blind from mid September 2011 to April 2012 at
Alert, and from mid April to mid September at Syowa. These time periods correlate in general well
with the positive bias anomalies in ozone columns found in SACADA results. The area of impact of
635 a total column observation on assimilation results is limited by the background correlation matrix,
which uses a horizontal correlation radius of 600 km. Latitudes not covered by observations can
therefore only be influenced via tracer transport and chemistry. In summary, we conclude that ~~the~~
~~biases evident, these large biases~~ reflect a general tendency of the SACADA model to overestimate
total ozone in polar night regions. Since its assimilation set-up was limited to UV-Vis observations,
640 these could not constrain the erroneous model results at high latitudes.

4 Validation of the vertical distribution of stratospheric ozone against ozonesondes

In this section, we discuss the results obtained for the validation of the ozone profiles against
ozonesonde observations at Ny Alesund (Arctic), Nairobi (Tropics), and Neumayer (Antarctic).

In order to compare the ozone fields from the 3 systems with the observed ozonesonde data, the
645 ~~modelled ozone fields analyses~~ were first linearly interpolated to the geographical location of the
launch sites. Even though sondes may drift long distances during their ascent, especially within the
polar vortex, this often significant horizontal movement was disregarded, as tracking information
is not always available. As a next step, the two analysis profiles preceding and following the mea-
surement closest in time were linearly interpolated to the time of observation. Since the ozonesonde
650 profiles have a much higher vertical resolution than ~~their modelled counterparts~~the analyses, the
ozonesonde data have been ~~log-linearly interpolated on vertically re-gridded to~~ the coarser pressure
grid of the DAS, degrading the observations to the lower resolution of the ~~model~~DAS (Langerock et al., 2014).
Figure 3 shows time series of the monthly mean ozone bias profiles with respect to the ozonesondes
at the selected sites for each of the three MACC systems.

655 ~~As the ozone number density reaches its stratospheric maximum between 20 and 30 altitude, depending on latitude (?), we used the 50 level as indicator for the system performance in the middle stratosphere. Figure ?? shows the time series of the ozone volume mixing ratio at this pressure level, allowing to identify the seasonal cycle of ozone in the middle stratosphere and to intercompare the performance of the different analyses. This time, no time interpolation has been done. Instead, 5 day~~
660 ~~moving averages of the ozone analyses have been calculated, to exclude any severe influence of differences in the original output frequency.~~

4.1 Arctic – Ny-Ålesund

The seasonal cycle at Ny-Ålesund is very well reproduced by the three analyses. Biases at Ny-Ålesund are generally smaller than 20 % for all MACC analyses throughout the stratosphere (Fig. 37).
665 The time series of the ozone profiles shows alternating behavior in the vertical for IFS-MOZART, persistent over the entire 3 year period, with positive biases in the lower (below 70 hPa) and upper (above 20 hPa) stratosphere and no or only slightly negative biases (mostly 5–10 %) in the middle stratosphere. The performance of BASCOE is stable throughout the stratosphere and for the entire 3 year period, with biases mostly less than 5 %. Largest biases over the whole period for
670 IFS-MOZART (–20 to –30 % between 50 and 70 hPa) and for SACADA (> 50 % between 35 and 65 hPa) are found for March 2011. While the ozone hole simulated by IFS-MOZART is too deep, SACADA simulates an Arctic ozone hole which is not deep enough. This special event will be discussed in detail in Sect. 6. Until March 2011, SACADA mainly overestimates ozone over the entire altitude range, while middle stratospheric ozone is mostly underestimated afterwards.

675 ~~At 50 (Fig. ??), IFS-MOZART and BASCOE deliver very similar results (differences < 5%) throughout the 3 year period, in good agreement with the sonde data at this altitude, while SACADA analyses show a tendency to slightly higher mixing ratios until March 2011 and slightly lower mixing ratios during the period September 2011 to January 2012. The error is about 10. From northern spring 2012 onwards, all assimilation results are very close.~~

680 ~~The depletion of ozone due to the Arctic ozone hole in northern spring 2011 is very prominently present (Fig. ??). IFS-MOZART and BASCOE reproduce this extraordinary behaviour very well.~~

4.2 Tropics – Nairobi

The O₃ bias profile time series (Fig. 37) now displays a changing performance in the vertical for all three analyses. Lower stratospheric ozone is underestimated by more than 40 % by both IFS-
685 MOZART and BASCOE (below 80 hPa and below 100 hPa, respectively) throughout the year. For BASCOE, this is followed by a small pressure range just above (between 75 and 90 hPa) where ozone values are overestimated by more than 50 %. The results for the remaining middle to upper part of the stratosphere are almost identical to the observed ozonsonde values at Nairobi, although with a tendency to overestimate O₃ by IFS-MOZART (< 20%). SACADA, on the other hand,

690 overestimates ozone below 40 hPa with more than 50 %, while its performance is usually very good
above. Between July 2011 and May 2012, however, SACADA underestimates O₃ with up to 20 %,
and even 30 % from September to December 2011, in the pressure range between 10 and 30 hPa.

This is reflected in the time series at 50 hPa: all analyses overestimate O₃. Whereas IFS-
MOZART and BASCOE produce very similar results at 50 hPa (IFS-MOZART slightly above
695 BASCOE) with biases of about 15 % compared to the ozonesonde data at Nairobi, and even up
to 30 % in the period August–November 2010, SACADA shows ozone values which are at least
35 % higher than the other two analyses, while the seasonality is well reproduced. The discontinu-
ity in the SACADA products from 6 to 7 September 2010, is due to resumption of the assimilation
after a period where SACADA ran freely (July–September 2010) due to a data gap in the assimi-
700 lated SCIAMACHY. As mentioned before for the total ozone columns, the SACADA analysis tends
to drift in the absence of UV observations for the assimilations. Once resumed, the assimilation
reduces the mismatch with the other two analyses from 60 % down to only 10 %.

4.3 Antarctic – Neumayer

The O₃ bias profile time series show that the biases are smallest and most stable for BASCOE
705 (usually less than 10 %). IFS-MOZART on the other hand has an annually recurrent pattern, over-
estimating O₃ with more than 50 % between roughly 70 and 150 hPa each Antarctic ozone hole
season, from September till December, while underestimating ozone between 30 and 60 hPa in
September. This indicates that IFS-MOZART has problems with a correct simulation of the ozone
depletion. This is a known problem of the underlying MOZART CTM in the MACC configuration
710 which cannot be completely fixed by the data assimilation (Flemming et al., 2011; Inness et al.,
2013), especially because the assimilated profile only gives information down to 68 hPa. MOZART
performs better with WACCM meteorology (Kinnison et al., 2007), which indicates that the chemi-
cal parameterizations are sensitive to the meteorological fields that are used to drive transport in the
models. SACADA has problems to correctly simulate the ozone concentration in the lower strato-
715 sphere (below 80 hPa) . While the ozone hole depth of 2010 is underestimated (positive bias), the
corresponding ozone depletion in 2011 and 2012 is overestimated by more than 50 %. This is re-
lated to the premature onset and end of the ozone depletion as predicted by the model, which is
reflected also in the ozone values at 50 Pa. Apart from this, the observed ozone values at Neumayer
at 50 hPa are in general well reproduced by the three analyses of the MACC system. IFS-MOZART
720 and BASCOE differ only little from each other.

4.4 Discussion

4.4.1 SACADA results

In our evaluation, SACADA is the only chemical data assimilation system with full chemistry which assimilates total column ozone only. Ozone columns are assimilated by constraining the ~~model system's~~ ozone column first guess at the satellite footprint. ~~Mixing ratios are appropriately scaled by an altitude independent factor using the model's one ozone profile.~~ We find that, in the case of SACADA, the lack of information constraining the shape of the ozone profile leads primarily to an overestimation of ozone in the lower stratosphere as can be seen, e.g., in Fig. 3-7 in comparison to station Nairobi (1.27° S). The excess ozone in the lower stratosphere leads to an underestimation at higher altitudes above 30 hPa (see also Fig. 4). The standard deviations between the MACC systems and the ozonesondes are largest for SACADA. We conclude that total column assimilation does not sufficiently constrain the ~~model system's~~ ozone profile.

4.4.2 IFS-MOZART and BASCOE results

Biases are mostly smaller than 10 % for IFS-MOZART and BASCOE in the middle to upper stratosphere. IFS-MOZART has problems with a correct representation of the vertical distribution of ozone. Often, over- and underestimations are alternating in the vertical. Biases are highest in austral spring during the Antarctic ozone hole season. Also during March 2011, when the first documented significant ozone hole in the Arctic occurred (Manney et al., 2011), somewhat larger differences are found. While IFS-MOZART and BASCOE deliver quite similar results, BASCOE profiles have a more stable behavior at all altitudes and during events of extraordinary behavior, as proven by its behavior during the Arctic and Antarctic ozone hole seasons. Largest biases occur, for both systems, in the lower stratosphere in the Tropics.

This can be partially explained by the strong gradients in ozone near the tropopause, which is located at higher altitudes in the Tropics than at the poles. ~~The exact location of the tropopause is very hard to define in the models. Ozone concentrations are rapidly decreasing and tend to increase the relative differences.~~ These sharp ozone gradients in the Upper Troposphere-Lower Stratosphere (UTLS) are very difficult to represent in three-dimensional models and probably require a very fine vertical resolution (Considine et al., 2008). Furthermore relative differences are amplified in this region due to its low ozone abundance.

For BASCOE, two more elements play a role in the worse performance in the lower tropical stratosphere: the low vertical resolution and aliasing errors in the horizontal wind fields, which are larger close to the UTLS and which lead to noise in the horizontal distribution of chemical tracers. This bug has been corrected in an upgraded version τ , which is running operationally since the beginning of 2013. The vertical grid of the ~~model is improved and extended~~ system is improved, from 37 levels (~~model top at 0.1~~) to 91 levels (~~model top at 0.01~~) and λ , with a much finer resolution

in the UTLS region. Comparison between both versions shows that O₃ values become ~~lower~~smaller around 80 hPa and ~~higher~~larger lower down (which would thus correct the currently large biases in these regions).

The larger biases for IFS-MOZART in the lower stratosphere globally (i.e. not only at the Tropics, but also at the poles, especially in the Antarctic) ~~may~~ also result from the fact that the useful range of the NRT MLS v2.2 data ~~is~~was restricted to levels above 68 hPa, which means that it ~~includes~~included no profile information below that pressure level, in contrast to BASCOE, which ~~assimilates~~assimilated the offline MLS v2.2 dataset down to 150 hPa. Tests with the improved NRT MLS v3.4 data ~~(?)~~, which have a useful range of 261 to 0.1 (Livesey et al., 2013b), which can be used down to 261 hPa and which are used in a later version of IFS-MOZART, show that a lot of those biases in the lower stratosphere disappear (see ~~also~~ Sect. 6.2).

To illustrate that the selected stations at each latitude band are representative for the results at all stations and that the same conclusions hold in general, we additionally show the mean ozone profiles and ozone bias profiles for the MACC analyses compared to all considered ozonesonde measurements in each latitude band (see Fig. 1), averaged over the entire 3 year period from September 2009 to September 2012 (Fig. 4). On average, all analyses agree with the sondes mostly to within $\pm 10\%$ above 70 hPa. Larger biases are observed for IFS-MOZART in the upper stratosphere (above 10 hPa) at the poles and in the lower stratosphere with overall biases reaching 30 % in the Antarctic and -40% at the equator, and for BASCOE below 150 hPa.

Standard deviations between the MACC systems and the ozonesondes are smallest for BASCOE, and only slightly higher for IFS-MOZART, usually between 10 and 20 %, except for the region below 70 hPa in the Tropics. The standard deviations for IFS-MOZART are higher in the area between 60 and 100 hPa in the Tropics, and between 100 hPa and 200 hPa in the Antarctic.

4.4.3 Influence of the temporal and horizontal resolution

SACADA data are sampled only once a day (at 12 h UT), IFS-MOZART 6 hourly, and BASCOE data 3 hourly. This may affect its performance when compared to ozone sondes. To exclude the effect of temporal resolution, we have degraded the temporal resolution of both IFS-MOZART and BASCOE to the temporal resolution of SACADA.

Relative differences between the fine and the coarse temporal resolution datasets are usually less than 2 %, but can be as high as 10 % for some months, and at some altitudes without any clear pattern (~~figures not shown~~). The effect on the standard deviation of the differences when using the 24 h resolution dataset for all three analyses ~~appears to be minimal (Fig. 4). Only is not significant except~~ in the lower tropical stratosphere ~~, somewhat larger differences (compared to their original temporal resolution) in the standard deviations for IFS-MOZART and BASCOE appear, but they never reach the high values of the SACADA data (figures not shown).~~

On the other hand, also a lower horizontal resolution may lead to larger standard deviations. BAS-

COE and SACADA have, however, the same horizontal resolution (3.75° by 2.5°), which is coarser than for IFS-MOZART (1.875° by 1.875°). This illustrates that the differences in standard deviations between the MACC systems is not exclusively dependent of the temporal nor the horizontal resolution.

5 Validation of the vertical distribution of stratospheric ozone against ACE-FTS

Additionally to the groundbased and ozonesonde data, the MACC ozone analyses have been compared to independent ACE-FTS satellite observations. The comparison between the measurements by ACE-FTS and the analysis output is performed in the following manner. The analyses, first re-gridded to a common $1^\circ \times 1^\circ$ grid, are collocated with the ACE-FTS data in space (horizontally and vertically) and time through linear interpolation. Since SACADA results are only provided every 24 h, we assume a constant composition throughout the day. Monthly mean biases of the spatial-temporal collocated data are calculated for 5 latitude bins, using 25 pressure bins based on the standard UARS fixed pressure grid (i.e., six pressure levels per decade which corresponds approximately to 2.5 km). These monthly mean biases and their associated standard deviations can be displayed as time series of the monthly mean bias (e.g. Figs. 5 and 6) or as vertical profiles (e.g. Fig. 7).

In view of the problems to constrain the ~~model's ozone profile from SACADA three-dimensional ozone field using only~~ total column assimilation, ~~shown earlier,~~ we will still ~~include the results for SACADA show the SACADA results~~ in the figures ~~;~~ but we will not include ~~the model these analyses~~ in the discussion.

5.1 Partial ozone columns

The time series of the standard deviations in Fig. 5 gives a global view of how well the analyses are performing against the satellite data. The standard deviations are averaged over the entire globe (90° S– 90° N) and over the entire stratospheric area of interest (200–5 hPa). As shown earlier, when comparing with groundbased and ozonesonde observations, ~~the results by~~ IFS-MOZART and BASCOE ~~perform very similarly~~ are very similar. Standard deviations are on average around 6–7% ~~;~~ ~~while the highest values~~ ~~This is only slightly larger than the relative mean difference between ACE-FTS and coincident MLS profiles, reported by Dupuy et al. (2009, table 7) as +4.7%. The largest standard deviations~~ are found around March and August each year.

Binning into a stratospheric pressure layer (100–5 hPa for the Tropics, and 200–5 hPa for all other latitude bands) shows small overall mean biases for both systems. Individual monthly mean biases for IFS-MOZART and BASCOE always remain below 5%, which shows that these analyses have an overall stable behaviour (figure not shown).

825 5.2 Ozone at predefined pressure levels

Even though partial columns indicate an excellent and stable behavior for both IFS-MOZART and BASCOE, interpolation at specific pressure levels (10, 50, and 100 hPa) reveals ~~an anticorrelation of the alternatingly positive and negative~~ biases in the vertical for IFS-MOZART, both in the Arctic and ~~in the~~ Antarctic, especially ~~in association with ozone hole conditions during ozone hole events~~ (Fig. 6), which was also seen earlier in the comparison with O₃ sondes (Fig. 37). These vertical oscillations in bias compensate to deliver correct (assimilated) partial or total columns (see Sect. 3).

In the Arctic, biases are, for all analyses, largest in March 2011. Biases remain low for BASCOE (<10 %), but attain values up to 20 % for IFS-MOZART. One obvious explanation is the occurrence of extreme conditions in the Arctic ~~, with the appearance of the first Arctic ozone hole~~ Manney et al. (2011). ~~However, another issue may have an influence on the system performance in this period: on the 26 March 2011, Aura MLS stopped sending data until the 18 April. Since these data are assimilated both by IFS-MOZART and BASCOE, it means that BASCOE is running freely during this period, while IFS-MOZART is only assimilating partial or total ozone columns. The effect is, however, expected to be minimal for March, as it concerns only 5 days (out of 30~~ for which collocations with the ACE-FTS measurements are found) in the monthly mean. ~~Since the assimilation of Aura MLS data was only reactivated in IFS-MOZART on the 10 May 2011, while BASCOE restarted to assimilate Aura MLS as soon as the data were back online (i.e. on 19 April), the effect of missing profile information is expected to become more apparent in April. Unfortunately, ACE-FTS did not collect any measurements in the Arctic in April 2011 (see Sect. 2.4.4),~~ but the earlier comparison with sondes (Fig. 3) does give an indication that the larger biases disappear in April 2011. ~~This illustrates that it are indeed the extreme Arctic conditions which lie at the basis of the larger biases in March. (Manney et al., 2011). This event will be discussed in section 6.~~

The same conclusions can be drawn for the Antarctic during the yearly ozone hole conditions. Biases for BASCOE still remain within 10 %, but are more pronounced for IFS-MOZART than in the Arctic, especially in the lower stratosphere (100 hPa), where relative differences up to almost 50 % in 2011 and even 60 % in 2010 are found in September, even now that Aura MLS data are available for assimilation. ~~The stable behaviour of BASCOE in time may be due to the fact that BASCOE assimilates the same data all-year-round (day and night time), while IFS-MOZART also~~ assimilate UV-VIS data which are not available at all times of the year at the poles. ~~Unless the chemistry scheme of MOZART can be improved to better simulate ozone hole conditions without data assimilation, the assimilation of vertical profiles is essential for the stratospheric forecasting system.~~

5.3 Seasonal mean ozone profiles

860 Figure 7 shows seasonally averaged relative ozone biases for austral spring and boreal winter, for the three consecutive years in the studied period. BASCOE has a stable performance compared to ACE-FTS data throughout the stratosphere, very similar each year, but slightly underestimating ozone with an average of 5 % in the Arctic. While the biases vary between -10 and 0 % in austral spring 2010, the variability is larger (biases between -15 and $+5$ %) in austral spring 2011. The
865 seasonal mean biases of IFS-MOZART again illustrate the oscillating behaviour of the profiles, both in the Arctic and Antarctic. Antarctic biases appear to be three times as large as those in the Arctic, and largest for the first year. ~~The stable performance of BASCOE in the vertical can be explained by the fact that BASCOE is only assimilating profile observations, whereas IFS-MOZART highly relies on (partial and total) column observations.~~

870 6 Arctic ozone hole event 2011

6.1 Case study

Besides the overall performance of the different analyses, we want to evaluate the ability of the MACC system to capture special events, such as the yearly recurrent Antarctic ozone holes, or the exceptional Arctic ozone hole in northern winter/spring 2011 (Manney et al., 2011). Long-lasting
875 exceptionally cold conditions prevailing over the Arctic, together with man-made ozone-depleting compounds lingering in the atmosphere, caused the destruction of almost 40 % of stratospheric ozone by the end of March (Manney et al., 2011). In this section, we address the performance of the MACC system during this particular event. Throughout the previous discussions, we have already shown that biases with respect to observations are largest at the peak of the ozone hole (i.e. March 2011),
880 illustrating that most systems have difficulties to correctly simulate such an unexpected event. In view of the fact that SACADA did not assimilate SCIAMACHY data at high northern latitudes before May 2011 (see earlier), we have omitted the discussion of SACADA results in this particular case study (results are only shown in Fig. 10).

Figure 9 shows the evolution of the ozone depletion, as simulated by IFS-MOZART and BASCOE
885 in the north pole vortex at 485 K potential temperature (~ 20 km, ~ 50 hPa) during the month March 2011. The vortex is determined by the potential vorticity (PV). Two contours of scaled PV (sPV) delimit the outer and inner vortex edges, respectively using an sPV of 1.410^{-4} (as in Manney et al., 2011) and $1.710^{-4} \text{ s}^{-1}$.

Manney et al. (2011) showed that, in February–March 2011, the barrier to transport at the Arctic
890 vortex edge was the strongest in either hemisphere for the last ~ 30 years. This barrier isolates the cold air in the vortex, preventing it from mixing with air in the mid-latitudes, causing a build-up of ozone, brought by long-range transport outside the vortex. ~~At the same time, the return of sunlight~~

~~after the polar night, releases the ozone-destroying gases trapped in the polar stratospheric cloud (PSC) particles within the vortex, breaking down ozone into individual oxygen molecules. Inside~~
895 ~~the vortex, the air masses were cold enough to allow PSC particles to condense. Heterogeneous~~
~~reactions took place at the surface of these particles, converting chlorine reservoir molecules HCl and~~
~~ClONO₂ into chemically active ClO and Cl₂. Hence catalytic destruction of ozone could start as~~
~~soon as sunlight came back to illuminate these air masses.~~

From late February/early March 2011 on, reduced levels of ozone are observed inside the vortex
900 and the ozone hole starts to develop. The largest chemical loss was recorded on the 26 March. At
that time, a stretched vortex is covering Scandinavia and North-West-Russia.

As seen in Fig. 97, IFS-MOZART and BASCOE provide very similar results, BASCOE values
being slightly higher and slightly noisier than the IFS-MOZART ones. The slightly higher noise is
reflected in the standard deviation (Fig. 5), but has been corrected in a later version of BASCOE (see
905 earlier).

We compared the IFS-MOZART, BASCOE, and TM3DAM analyses with data received by 3
UV-VIS SAOZ/DOAS instruments at Arctic locations, which are part of the NDACC: Zhigansk
(Russia, 66.8° N, 123.4° E), Harestua (Norway, 60° N, 11° E, DOAS), and Scoresby Sund (Green-
land, 70.49° N, 21.98° W) (see Sect. 2.4.2). From Fig. 97, we see that on 1, 13, and 26 March,
910 respectively only Zhigansk, only Scoresby Sund, and only Harestua are located inside the winter-
time polar vortex. The observations are well reproduced by the three analyses, with a tendency to
a slight overestimation.

Comparison with ozone soundings at Ny-Ålesund (Spitsbergen), which is always located within
the polar vortex, shows that both IFS-MOZART and BASCOE could correctly reproduce the ozone
915 hole conditions with relative biases mostly less than 10 % in the stratosphere (Fig. 10). IFS-MOZART
has, however, problems with a correct simulation of the vertical profile, when the ozone depletion is
strongest (in March 2011) and alternating positive and negative biases up to 30 % can be seen.

6.2 Influence of the assimilated data on the performance of the analyses

~~We suspect that the significantly larger biases for the NRT stratospheric ozone products delivered by~~
920 ~~SACADA and to a lesser extent by On 26 March 2011 Aura MLS stopped sending data and resumed~~
~~normal operations on 19 April 2011. BASCOE ran freely (unconstrained CTM mode) during this~~
~~time, and started again to assimilate MLS as soon as observations came back. IFS-MOZART in~~
~~the vertical, compared to BASCOE are mainly due to the assimilated dataset of MLS, i.e. the~~
~~fact that BASCOE assimilated Aura-MLS v2.2 Offline ozone which is valid down to 261 (besides~~
925 ~~MLS -, and -) while IFS-MOZART assimilated Aura-MLS v2.2 NRT ozone down to 68 (and various~~
~~column obs) and SACADA assimilated only total column data from UV observations. To test this~~
~~hypothesis, we assimilated only UV/VIS observations from 26 March 2011 until 10 May 2011, when~~
~~the assimilation of Aura MLS was switched back on. Unfortunately, ACE-FTS did not collect any~~

measurements in the Arctic during April 2011 (see Sect. 2.4.4).

930 These uncontrolled modifications of the observing system led us to explore in a more systematic manner the impact of the assimilated observations on the quality of the analyses. We chose a one-month period with the Arctic ozone depletion already well underway while MLS and ACE-FTS were still scanning the area, i.e. the month of March 2011. We first defined three new experiments in which all three models assimilate the same data, namely with IFS-MOZART, BASCOE and 935 SACADA assimilating the same dataset: Aura MLS version 3.3 offline ozone down to 261 (i. e. no additional species such as H_2O , O_3 , and NO_2). We focused on the period considered in the Arctic ozone hole 2011 case study and hence, the models were run for one month (March 2011) with a CO_2 keeping all observations down to 215 hPa. BASCOE was not allowed to assimilate any other species than ozone. To allow a short spin-up period of about one week, the three systems were started on 25 February 940 from the BASCOE analysis delivered in NRT for that date.

Figure 11 (left) shows the mean bias and standard deviations of the differences between the NRT (i.e. the original) analyses and ACE-FTS observations, keeping only the (~ 200) ACE profiles within the North Pole vortex, with the vortex edge calculated with an sPV of $> 1.7e^{-4} \text{s}^{-1}$. At ozone hole level ($\theta \approx 485$ the level where ozone depletion is maximum ($\theta \approx 485$ K), we see the large 945 underestimation that the depletion is much too severe in IFS-MOZART and the absence of ozone depletion in SACADA. This figure is to be compared with NRT analyses and completely absent in SACADA NRT analyses. Figure 11 (right) shows the results of the three new offline experiments assimilating the same dataset: now all analyses perform very similarly.

To identify the exact cause of the large improvement in IFS-MOZART analyses, we ran a last 950 sensitivity test with IFS-MOZART assimilating the usual set of UV/VIS data (OMI and SCIAMACHY total columns; SBUV/2 partial columns) in addition to the offline MLS v3 dataset. As can be seen in figure 12, the bad performance of IFS-MOZART NRT was not due to the assimilation of UV/VIS data but rather to the assimilation of the MLS v2 NRT data. If the MLS v3 and UV/VIS observations are assimilated together (green curves), the quality of the ozone analyses delivered by IFS-MOZART 955 improves: tropospheric ozone is improved over the previous sensitivity test assimilating only MLS v3 (blue curves), and the simultaneous assimilation of UV/VIS observations does not degrade the analysis of stratospheric ozone.

The worse performance of IFS-MOZART NRT is probably not due to the earlier version of the MLS dataset either, because our sensitivity test with BASCOE (Fig. 11(right), which shows the 960 results of the three new offline experiments, blue lines) shows that the analyses of MLS v2.2 SCI (left) performed nearly as well as the analyses of MLS v3 (right) despite the usage of an earlier version of BASCOE. Hence the better performance of BASCOE NRT is primarily due to its assimilation of MLS v2.2 SCI down to 150 hPa, while IFS-MOZART had to assimilate MLS v2.2 NRT which was not valid (and filtered out) below 68 hPa. This subtle difference in configuration is 965 due to an operational constraint: IFS-MOZART had to be run closer to real-time and could not wait

3 extra days for the distribution of MLS SCI.

It is now possible to interpret the slight differences between the performances of the three systems assimilating the same ~~dataset~~ MLS dataset (Fig. 11, right). ~~Now, all analyses perform very similarly.~~ The biases and standard deviations are smallest for IFS-MOZART, ~~which might be related to the~~ probably thanks to its higher horizontal resolution. The standard deviations for SACADA are slightly larger than the ones for IFS-MOZART and BASCOE, which ~~may be due to the~~ is due to its lower time sampling (24 h output frequency instead of 6 h for the two other). ~~The~~ Finally, the BASCOE experiment delivers smaller ~~bias~~ biases and standard deviations than ~~for the original~~ the original NRT analysis. This may be ~~related to~~ due to two different causes: the assimilation of MLS ~~offline v3.3 instead of MLS offline v2.2,~~ and/or an improvement in the pre-processing of the ECMWF wind fields which drive the transport in BASCOE. Indeed the BASCOE version used in NRT suffered from aliasing errors in the input wind fields, leading to some erroneous noise in the horizontal distribution of chemical tracers (Fig. 7).

~~The results for IFS-MOZART show that the increased vertical range of MLS V3 is beneficial compared to the more limited range of the NRT MLS V2 data. An additional IFS-MOZART experiment assimilating, besides MLS V3, also the other column products (as defined in Table 1) shows that the analysis is well constrained by MLS alone in the stratosphere, while it is beneficial to have the combination of profile and total column data in the troposphere.~~

7 Conclusions

~~This paper presents the NRT stratospheric ozone service delivered in the framework of the MACC project. The service is based on four independent Four ozone data assimilation systems : IFS-MOZART, BASCOE, SACADA, and TM3DAM. Two of them (SACADA and TM3DAM) assimilate only total columns, consecutively SCIAMACHY and GOME-2. The IFS-MOZART analysis is also based on SCIAMACHY data (at least until the end of Envisat), but additionally assimilates total columns of OMI, partial columns of SBUV-2, and profiles of MLS. The delayed, but more extended MLS profile dataset is the unique input for the BASCOE analyses. This paper presents the validation results of these ozone analyses against groundbased observations, ozonesondes and ACE-FTS satellite retrievals for the period September 2009 to September 2012. All intercomparisons show consistent results. Data assimilation seems to work well in all systems. When ozone columns are assimilated, columns are well reproduced. Assimilation of profiles does lead to a good representation of profiles. Model differences only show up where the systems are not well constrained by the observations, and this is where BASCOE performs best.~~

~~All analyses show a realistic seasonal cycle in the three considered latitude bands and total column values are generally in very good agreement with independent Brewer/Dobson observations. The UTLS is the most difficult range to model for all analyses. Large biases are found especially in the~~

Tropics, where tropospheric air masses reach higher altitudes.

1005 On the other hand, each system has its own strengths and weaknesses. IFS-MOZART assimilates the largest variety of satellite instruments simultaneously, going from partial and total ozone columns to ozone profiles. While large weight is given to total column observations in the assimilation
1010 procedure (hence the very good reproduction of total columns), the only profile information comes from the NRT MLS observations (DAS) have been run continuously and simultaneously during several years. These DAS have very different designs (offline or online dynamics; grid set-up; specification of background error covariances) and were set-up very differently with respect to the assimilated datasets. In this paper we seized this opportunity, first to provide an intercomparison
1015 and validation of the resulting analyses, which can only be used down to 68. Hence the input into the analysis is limited, leading to alternating positive and negative biases in the vertical compared to ozonesonde and ACE-FTS satellite data. Biases are largest (30 to 60) during ozone depletion events. This is a known problem of the underlying MOZART-CTM in the MACC configuration, which improves with WACCM meteorology, but which cannot be corrected by the assimilation due to missing profile information at these altitudes, and second to investigate the causes of their very different biases.

The BASCOE analysis is based exclusively on ozone profiles, in this case from MLS, without any changes for the entire period, leading to stable results throughout the year. The assimilation of MLS data gives information during the polar night when the UV instruments GOME-2, SBUV/2,
1020 SCIAMACHY and OMI can not observe the ozone field because there is no backscattered solar radiation. With its purely stratospheric design, larger systematic biases for the total ozone columns are expected, but still BASCOE delivers reliable TOCs. Only in the Tropics, slightly larger (negative) biases are observed. The exclusive assimilation of MLS profiles and the fact that BASCOE is specifically designed to deal with stratospheric chemistry processes leads to ozone profiles which are in very good agreement (biases less than 5–10) with the observed profiles delivered by ozonesondes
1025 and This study shows what can be achieved in Near Real Time (NRT) with state-of-the-art DAS for stratospheric ozone and provides guidance to the users of the ACE-FTS satellite observations throughout the entire stratosphere and for the entire period between September 2009 and September 2012.

1030 Compared to resulting analyses. Among the three sets of vertically resolved NRT analyses of stratospheric ozone, those delivered by BASCOE had the best overall quality. This is due primarily to the other three CTMs, TM3DAM is designed to assimilate only total column observations. A bias correction is applied to the satellite observations to reduce on average the bias with the surface Brewer focus of BASCOE on stratospheric observations retrieved from limb sounders, and to more relaxed operational constraints allowing it to wait for the delivery of the best input dataset available.
1035

TM3DAM is based on a sequential Kalman Filter algorithm and does not model stratospheric

chemistry explicitly. It aims only to provide total columns of ozone by making optimal use of the ozone column measurements from UV-Vis satellite sounders, with very small biases between the analyses/Dobson observations. We showed that forecasts and satellite datasets. It was shown that TM3DAM provides very realistic is a good reference to test the ability of the three other systems to produce accurate ozone column amounts.

The low quality of the analyses delivered in NRT by SACADA is a good indication of the drawbacks to expect from current CDA systems when they are configured to assimilate total ozone columns. Unfortunately, we could not compare the profiles with independent data because the 3-D model output is not provided only. This should be considered as a worst-case scenario in a future situation where no limb sounder would be available and no proper effort would be invested to assimilate correctly vertical profiles retrieved from nadir-looking instruments.

SACADA is the only chemical data assimilation system with full chemistry which (just like TM3DAM) assimilates total column ozone only, from one UV satellite instrument at a time. As long as reliable UV data for the assimilation are available, SACADA is able to reproduce observed total ozone columns. If the model is running unconstrained, large drifts in ozone are observed. In contrast to TM3DAM, SACADA does attempt to provide vertical profile information as well, by appropriately scaling the mixing ratios by an altitude independent factor using the models one ozone profile. The lack of information constraining the ozone profile leads primarily to an overestimation of ozone in. Finally, while IFS-MOZART did not deliver the best NRT analyses in this intercomparison, it still has the potential to deliver the best analyses (figures 11 and 12). Official reviews of international monitoring capacities (e.g. WMO, 2011), expect an imminent lack of ozone profiling capabilities at high vertical resolution. Contrarily to the BASCOE version used here, IFS-MOZART should be able to adapt to this situation thanks to its demonstrated ability to assimilate several instruments simultaneously.

From a system design point of view, the sensitivity tests performed in section 6.2 deliver important conclusions:

- All systems used in MACC require profile data to provide a good vertical distribution of stratospheric ozone.
- This profile data must include the lower stratosphere and in the Antarctic, especially during ozone hole conditions.

We have shown that assimilating total columns only provides total column values which are usually in good agreement with independent observations (SACADA, TM3DAM), but that these analyses are not able to provide reliable information about the ozone vertical distribution (SACADA). Even more, ozone analyses with realistic total ozone column densities (

- IFS-MOZART) are not necessarily in good agreement with the observed ozone profiles. This study illustrates that even state-of-the-art models of stratospheric chemistry still require the

1075 ~~assimilation of limb observations for a correct representation of the vertical distribution of ozone in the stratosphere. This is corroborated by the results of three experimental runs, in which is able to assimilate limb profiles and nadir products successfully. The profiles constrain well the stratosphere, allowing the partial and total columns (by UV/VIS instruments) to constrain well the troposphere.~~

1080 ~~– When they assimilate the same dataset with good quality and large observational density, BASCOE, IFS-MOZART, BASCOE, and SACADA assimilate the same dense MLS dataset over the entire stratosphere for the month March 2011. The profile results of the three analyses are now very close, with much reduced biases. The IFS-MOZART experiment exhibits now the best performance (lowest biases and standard deviations). This illustrates and SACADA deliver very similar performance despite their very different designs. The quality of modern ozone analyses depends primarily on the assimilated data. This conclusion has large implications for the planning of future satellite missions.~~

1090 ~~The newer SBUV/2 v8.6 profiles are distributed over 21 layers and each profile is distributed with its matrix of Averaging Kernels. Kramarova et al. (2013) illustrated the importance of using properly this information. While it is planned to implement SBUV/2 Averaging Kernels in the MACC NRT system at ECMWF, our last sensitivity test shows that this improvement was not necessary to assimilate successfully SBUV/2 v8 after a vertical re-gridding over 6 thick layers.~~

1095 ~~This study demonstrates the large benefit obtained from the assimilation of a single limb-scanning instrument (Aura MLS) with a high density of observations. Therefore, (Aura MLS). Therefore we can only share the serious concern about the lack of ozone-profiling capabilities at high vertical resolution at short term which has been expressed earlier on many occasions in official reviews of international monitoring capacities (e.g. WMO, 2011). Only when ozone profile information is available, are the CTMs able to capture events of scientific interest or events of exceptional nature.~~

1100 ~~Nadir UV instruments such as GOME-2 do bring profile information into the stratosphere, but with a limited vertical resolution of about 5. Therefore, there is currently only little interest in the assimilation of nadir profiles, but they may provide a solution in case we lose Aura in the short term, as expressed already in (WMO, 2011).~~

1105 *Acknowledgements.* This work ~~had been was~~ carried out in the framework of the MACC and MACC II projects. Both projects ~~are were~~ funded under the European Community's Seventh Research Framework Programme for research and technological development including demonstration activities. MACC ran from 1 June 2009 to 31 December 2011 and was followed up by MACC II with a two month overlap, 2011. MACC II ran from 1 November 2011 on, for a period of 33 months. MACC II is to 31 July 2014 and was funded under grant agreement CP-CSA 283576. The authors thank the MACC GRG and VAL teams for constructive discussions.

1110 We thank ~~all data providers (NASA, ESA, EUMETSAT) for granting access to the necessary satellite ozone products used in each of the MACC systems~~ the O3M SAF project of the EUMETSAT for providing the

GOME-2 ozone product used in this paper. We acknowledge the ~~Canadian Space Agency and science teams for providing observations from ACE-FTS, the main instrument on the Canadian satellite SCISAT-1.~~ We acknowledge the EVIVA (ENVISAT Value Adding for Continuous Monitoring of Atmospheric Trace Gases and Aerosols) project funded by the German Aerospace Center for processing and providing SCIAMACHY data
1115 on behalf of ESA. We also thank NOAA and NASA for granting access to the other satellite ozone products assimilated in the MACC systems, as well as the Canadian Space Agency and science teams for providing validation observations by ACE-FTS, an instrument on the Canadian satellite SCISAT-1.

For the use of groundbased Brewer/Dobson, SAOZ/DOAS, and ozonesonde data, we thank both the individual contributors and the projects and databases from which they were obtained: the World Ozone and Ultraviolet Radiation Data Centre (WOUDC, <http://www.woudc.org>), the Network for the Detection of Atmospheric Composition Change (NDACC, <http://www.ndsc.ncep.noaa.gov/data/>) and the Southern Hemisphere Additional OZonesondes network (SHADOZ, <http://croc.gsfc.nasa.gov/shadoz/>). We thank in particular the agencies who provided the data which were shown explicitly as a reference in this paper, i.e. the Meteorological Service of Canada, the Central Weather Bureau of Taiwan, the Japan Meteorological Agency, the Belgian Institute for
1120 Space Aeronomy, the National Centre of Scientific Research of France, the Central Aeronomy Observatory of Russia, the Danish Meteorological Institute, the Alfred-Wegener Institute of Potsdam, and MeteoSwiss of Switzerland.

The comparisons against ozone sondes have been obtained by means of validation tools developed in the NORS project (Demonstration Network Of ground-based Remote Sensing observations in support of the Copernicus Atmospheric Service) which was funded under grant agreement, funded by the European Community's Seventh
1130 Research Framework Programme under grant agreement 284421.

References

- Adams, C., Strong, K., Batchelor, R. L., Bernath, P. F., Brohede, S., Boone, C., Degenstein, D., Daffer, W. H., Drummond, J. R., Fogal, P. F., Farahani, E., Fayt, C., Fraser, A., Goutail, F., Hendrick, F., Kolonjari, F.,
1135 Lindenmaier, R., Manney, G., McElroy, C. T., McLinden, C. A., Mendonca, J., Park, J.-H., Pavlovic, B., Pazmino, A., Roth, C., Savastiouk, V., Walker, K. A., Weaver, D., and Zhao, X.: Validation of ACE and OSIRIS ozone and NO₂ measurements using ground-based instruments at 80° N, *Atmospheric Measurement Techniques*, 5, 927–953, doi:10.5194/amt-5-927-2012, 2012.
- Baier, F., Erbetseder, T., Elbern, H., and Schwinger, J.: Impact of different ozone sounding networks on a
1140 4D-Var stratospheric data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, doi: 10.1002/qj.2086, 2013.
- Balis, D., Kroon, M., Koukouli, M. E., Brinksma, E. J., Labow, G., Veeffkind, J. P., and McPeters, R. D.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, *Journal of Geophysical Research (Atmospheres)*, 112, D24S46,
1145 doi:10.1029/2007JD008796, 2007.
- Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, *Journal of Geophysical Research (Atmospheres)*, 114, D13205, doi: 10.1029/2008JD011115, 2009.
1150
- Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., ~~DeMaziDe Mazière~~, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, *Geophysical Research Letters*, 32, L15S01, doi:10.1029/2005GL022386, 2005.
1155
- ~~Bhartia~~
1160 ~~Bhartia~~, P. K., ~~McPetersMcPeters~~, R. D., ~~Mateer, C. L., Flynn, L. E., and Wellemeyer, C.:~~ ~~Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique, Journal of Geophysical Research, 101, 18793, 1996.~~ ~~Taylor, S., Kramarova, N. A., Frith, S., Fisher, B., and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm, Atmospheric Measurement Techniques, 6, 2533–2548, doi:10.5194/amt-6-2533-2013, http://www.atmos-meas-tech.net/6/2533/2013/, 2013.~~
- Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer, *Applied Optics* IP, 44, 7218–7231, doi:10.1364/AO.44.007218, 2005.
1165
- ~~Brasseur, G. P. and Solomon, S.: Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere, 2005.~~
- Brewer, A. W.: Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, *Quarterly Journal of the Royal Meteorological Society*, 75, 351–363, doi:
1170

10.1002/qj.49707532603, 1949.

[Brewer, A. W.: A replacement for the Dobson spectrophotometer?, *Pure and Applied Geophysics*, 106-108, 919-927, doi:10.1007/BF00881042, <http://dx.doi.org/10.1007/BF00881042>, 1973.](#)

1175 Cariolle, D. and Teyssède, H.: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations, *Atmospheric Chemistry & Physics*, 7, 2183-2196, 2007.

~~Cornu, A.: Observations de la limite ultraviolette du spectre solaire, *Comptes Rendus de l'Académie des Sciences (Paris)*, 89, 1879.~~

1180 [Considine, D. B., Logan, J. A., and Olsen, M. A.: Evaluation of near-tropopause ozone distributions in the Global Modeling Initiative combined stratosphere/troposphere model with ozonesonde data, *Atmospheric Chemistry and Physics*, 8, 2365-2385, doi:10.5194/acp-8-2365-2008, <http://www.atmos-chem-phys.net/8/2365/2008/>, 2008.](#)

Damski, J., Thölix, L., Backmann, L., Taalas, P., and Kulmala, M.: FinROSE – middle atmospheric chemistry transport model, *Boreal Env. Res.*, 12, 535-550, 2007.

1185 Deshler, T., Mercer, J. L., Smit, H. G. J., Stubi, R., Levrat, G., Johnson, B. J., Oltmans, S. J., Kivi, R., Thompson, A. M., Witte, J., Davies, J., Schmidlin, F. J., Brothers, G., and Sasaki, T.: Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes, *Journal of Geophysical Research (Atmospheres)*, 113, D04307, doi:10.1029/2007JD008975, 2008.

1190 [Dethf. A. and Hólm, E. V.: Ozone assimilation in the ERA-40 reanalysis project, *Quarterly Journal of the Royal Meteorological Society*, 130, 2851-2872, doi:10.1256/qj.03.196, <http://dx.doi.org/10.1256/qj.03.196>, 2004.](#)

Dobson, G. M. B.: Origin and Distribution of the Polyatomic Molecules in the Atmosphere, *Royal Society of London Proceedings Series A*, 236, 187-193, doi:10.1098/rspa.1956.0127, 1956.

1195 Dupuy, E., Walker, K. A., Kar, J., Boone, C. D., McElroy, C. T., Bernath, P. F., Drummond, J. R., Skelton, R., McLeod, S. D., Hughes, R. C., and many others: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), *Atmospheric Chemistry & Physics*, 9, 287-343, 2009.

Elbern, H., Schwinger, J., and Botchorishvili, R.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration, *Journal of Geophysical Research (Atmospheres)*, 115, D06302, doi:10.1029/2009JD011953, 2010.

Engelen, R. J., Serrar, S., and Chevallier, F.: Four-dimensional data assimilation of atmospheric CO₂ using AIRS observations, *Journal of Geophysical Research (Atmospheres)*, 114, D03303, doi:10.1029/2008JD010739, 2009.

1205 Errera, Q. and Ménard, R.: Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the belgian assimilation system for chemical observations (BASCOE), *Atmospheric Chemistry & Physics Discussions*, 12, 16763-16809, doi:10.5194/acpd-12-16763-2012, 2012.

Errera, Q., Daerden, F., Chabrilat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D.: 4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, *Atmospheric Chemistry & Physics*, 8, 6169-6187, 2008.

1210 Eskes, H. J., Velthoven, P. F. J. V., Valks, P. J. M., and Kelder, H. M.: Assimilation of GOME total-ozone

- satellite observations in a three-dimensional tracer-transport model, *Quarterly Journal of the Royal Meteorological Society*, 129, 1663–1681, doi:10.1256/qj.02.14, 2003.
- 1215 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, *Atmospheric Chemistry & Physics Discussions*, 5, 4429–4475, 2005.
- Farman, J. C., Gardiner, B. G., and Shanklin, J. D.: Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction, *Nature*, 315, 207–210, doi:10.1038/315207a0, 1985.
- 1220 Fisher, M.: Wavelet Jb - A new way to model the statistics of background errors., *ECMWF Newsletter* 106: 23-28, 2006.
- Fisher, M. and Andersson, E.: Developments in 4D-Var and Kalman Filtering., *ECMWF Technical Memorandum* 347, available at: <http://www.ecmwf.int/publications/library/do/references/list/14>, 2001.
- 1225 Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, *Geoscientific Model Development*, 2, 253–265, 2009.
- Flemming, J., Inness, A., Jones, L., Eskes, H. J., Huijnen, V., Schultz, M. G., Stein, O., Cariolle, D., Kinnison, D., and Brasseur, G.: Forecasts and assimilation experiments of the Antarctic ozone hole 2008, *Atmospheric Chemistry & Physics*, 11, 1961–1977, doi:10.5194/acp-11-1961-2011, 2011.
- 1230 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., Sunnicht, 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., Drouin, B. J., Cofield, R. E., 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 measurements, *Journal of Geophysical Research (Atmospheres)*, 113, D15S20, doi:10.1029/2007JD008771, 2008.
- 1235 [Gauss, M., Valdebenito, A., Benedictow, A., and et al.: Report on end-to-end analysis for the seven individual production lines and for ensemble production, MACC-II project deliverable D108.1, ECMWF, http://www.gmes-atmosphere.eu/documents/maccii/deliverables/ens/MACCII_ENS_DEL_D_108.1_201301_METEOF.pdf, 2013.](http://www.gmes-atmosphere.eu/documents/maccii/deliverables/ens/MACCII_ENS_DEL_D_108.1_201301_METEOF.pdf)
- 1240 Geer, A. J., Lahoz, W. A., Bekki, S., Bormann, N., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Juckes, M. N., Massart, S., Peuch, V.-H., Rharmili, S., and Segers, A.: The ASSET intercomparison of ozone analyses: method and first results, *Atmospheric Chemistry & Physics*, 6, 5445–5474, 2006.
- ~~Hartley, W.: On the absorption spectrum of ozone, *Journal of Chemical Society*, 39, 57–60, 1881.~~
- 1245 Hassler, B., Petropavlovskikh, I., Staehelin, J., August, T., Bhartia, P. K., Clerbaux, C., Degenstein, D., De Mazière, M., Dinelli, B. M., Dudhia, A., Dufour, G., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Granville, J., Harris, N. R. P., Hoppel, K., Hubert, D., Kasai, Y., Kurylo, M. J., ~~Kyrölä, E.~~, Lambert, J.-C., Levelt, P. F., McElroy, C. T., McPeters, R. D., Munro, R., Nakajima, H., Parrish, A., Raspollini, P., Remsberg, E. E., Rosenlof, K. H., Rozanov, A., Sano, T., Sasano, Y., Shiotani, M., Smit, H. G. J., Stiller, G., Tamminen, J., Tarasick, D. W., Urban, J., van der A, R. J., Veefkind, J. P., Vigouroux, C., von Clarmann, T., von Savigny,
- 1250 C., Walker, K. A., Weber, M., Wild, J., and Zawodny, J.: S12N overview paper: ozone profile measurements: techniques, uncertainties and availability, *Atmospheric Measurement Techniques Discussion*, 6, 9857–9938,

- 2013.
- Hendrick, F., Pommereau, J.-P., Goutail, F., Evans, R. D., Ionov, D., Pazmino, A., Kyrö, E., Held, G., Eriksen, P., Dorokhov, V., Gil, M., and van Roozendaal, M.: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations, *Atmospheric Chemistry & Physics*, 11, 5975–5995, doi:10.5194/acp-11-5975-2011, 2011.
- 1255 Hollingsworth, A. and Lönnberg, P.: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field, *Tellus Series A*, 38, 111, doi:10.1111/j.1600-0870.1986.tb00460.x, 1986.
- 1260 ~~Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J., Rayner, P., Peuch, V., Rouil, L., Schultz, M. G., and Simmons, A. J.: Toward a Monitoring and Forecasting System For Atmospheric Composition: The GEMS Project, *Bulletin of the American Meteorological Society*, 89, 1147–+, , 2008.~~
- Hughes, R. and Bernath, P.: ACE Mission Information for Public Data Release, Document Number: ACE-SOC 0022, 2012.
- 1265 Inness, A., Flemming, J., Suttie, M., and Jones, L.: GEMS data assimilation system for chemically reactive gases, ECMWF technical memorandum 587, 2009.
- Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrilat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, *Atmospheric Chemistry and Physics*, 13, 4073–4109, doi:10.5194/acp-13-4073-2013, 2013.
- 1270 Kalnay, E.: *Atmospheric Modeling, Data Assimilation and Predictability*, 2002.
- 1275 Kerr, J. B., Fast, H., McElroy, C. T., Oltmans, S. J., Lathrop, J. A., Kyrö, E., Paukkunen, A., Claude, H., ~~Khler~~Köhler, U., Sreedharan, C. R., Takao, T., and Tsukagoshi, Y.: The 1991 WMO international ozonesonde intercomparison at Vanscoy, Canada, *Atmos.-Ocean*, 32, 1994.
- Kinnison, D. E., Brasseur, G. P., Walters, S., Garcia, R. R., Marsh, D. R., Sassi, F., Harvey, V. L., Randall, C. E., Emmons, L., Lamarque, J. F., Hess, P., Orlando, J. J., Tie, X. X., Randel, W., Pan, L. L., Gettelman, A., Granier, C., Diehl, T., Niemeier, U., and Simmons, A. J.: Sensitivity of chemical tracers to meteorological parameters in the MOZART-3 chemical transport model, *Journal of Geophysical Research (Atmospheres)*, 112, D20302, doi:10.1029/2006JD007879, 2007.
- 1280 ~~Kramarova, N. A., Bhartia, P. K., Frith, S. M., McPeters, R. D., and Stolarski, R. S.: Interpreting SBUV smoothing errors: an example using the quasi-biennial oscillation, *Atmospheric Measurement Techniques*, 6, 2089–2099, doi:10.5194/amt-6-2089-2013, <http://www.atmos-meas-tech.net/6/2089/2013/>, 2013.~~
- Lahoz, W. and Errera, Q.: *Data Assimilation: Making Sense of Observations*, chap. Constituent Assimilation, p. 449, Springer, 2010.
- 1290 ~~Lahoz, W. A., Geer, A. J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Khattatov, B., Marchand, M., Massart, S., Peuch, V.-H., Rharmili, S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H. E., Vik, A. F., and von Clarmann, T.: The Assimilation of Envisat data (ASSET) project, *Atmospheric Chemistry and Physics*, 7, 1773–1796, doi:10.5194/acp-7-1773-2007, <http://www.atmos-meas-tech.net/7/1773/2007/>.~~

[//www.atmos-chem-phys.net/7/1773/2007/](http://www.atmos-chem-phys.net/7/1773/2007/), 2007.

1295 [Lahoz, W. A., Errera, Q., Viscardy, S., and Manney, G. L.: The 2009 stratospheric major warming described from synergistic use of BASCOE water vapour analyses and MLS observations, Atmospheric Chemistry and Physics, 11, 4689–4703, doi:10.5194/acp-11-4689-2011, http://www.atmos-chem-phys.net/11/4689/2011/, 2011.](#)

1300 [Langerock, B., De Mazière, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, Geoscientific Model Development Discussions, 7, 8151–8178, doi:10.5194/gmdd-7-8151-2014, http://www.geosci-model-dev-discuss.net/7/8151/2014/, 2014.](#)

1305 [Lerot, C., Van Roozendaal, M., van Geffen, J., van Gent, J., Fayt, C., and Spurr, R.: On the accuracy of GOME and SCIAMACHY total ozone measurements in polar regions, Proceeding 'Envisat Symposium 2007', Montreux, Switzerland, 23-27 April 2007 \(ESA SP-636, July 2007\), 2007.](#)

1305 [Lerot, C., Van Roozendaal, M., Spurr, R., Loyola, D., Coldewey-Egbers, M., Kochenova, S., van Gent, J., Koukouli4, D. Balis4., M., Lambert, J.-C., Granville, J., and Zehner, C.: Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat and GOME-2/MetOp-A., Journal of Geophysical Research \(Atmospheres\), 2013.](#)

1310 [Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The Ozone Monitoring Instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1093–1101, doi:10.1109/TGRS.2006.872333, 2006.](#)

[Lin, S. J. and Rood, R. B.: A fast flux form semi-Lagrangian transport scheme on the sphere, Monthly Weather Review, 124, 2046–2070, 1996.](#)

~~Livesey, N., Read, W., Froidevaux,~~

1315 [Lindenmaier, R., Strong, K., Batchelor, R. L., Lambert, A., Bernath, P. F., Chabrilat, S., Chipperfield, M. P., Daffer, W. H., Drummond, J. R., Feng, W., Jonsson, A. I., Kolonjari, F., Manney, ~~Manney~~-G. L., McLinden, C., Ménard, R., and Walker, K. A.: A study of the Arctic NOy budget above Eureka, Canada, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2011JD016207, http://dx.doi.org/10.1029/2011JD016207, 2011.](#)

1320 [PumphreyLivesey, H.N. J., Santeevan Snyder, MW., SchwartzRead, M., W. G., and WangWagner, S.P. A.: Retrieval Algorithms for the EOS Microwave Limb Sounder \(MLS\), IEEE Transactions on Geoscience and Remote Sensing, 44, ~~Cofeld~~, 1144–1155, doi:10.1109/TGRS.2006.872327, 2006.](#)

1325 [Livesey, N. J., Read, W., Lambert, A., Cofield, R., CuddyCuddy, D., FullerFroidevaux, L., Fuller, R., JarnotJarnot, R., JiangJiang, J., KnospJiang, Y., Knosp, B., StekKovalenko, L., Pickett, H., Pumphrey, H., Santee, M., Schwartz, M., Stek, P., WagnerWagner, P., and WuWaters, J., and Wu, D.: EOS ~~Aura/MLSVersion 3.3~~ Earth Observing System \(EOS\) Aura Microwave Limb Sounder \(MLS\) Version 2.2 and 2.3 Level 2 data quality and description document, ~~2011~~, Tech. rep., Jet Propulsion Laboratory, \[http://mls.jpl.nasa.gov/data/v2_data_quality_document.pdf\]\(http://mls.jpl.nasa.gov/data/v2_data_quality_document.pdf\), 2013a.](#)

~~Livesey~~

1330 [Livesey, N. J., ~~van Snyder~~Read, W. G., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Cofield, R. E., Cuddy, D. T., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Knosp, B. W., Stek, P. C., Wagner, ~~Read, W. G., and Wagner, P. A., and Wu, D. L.:~~ Retrieval Algorithms for](#)

- the EOS Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) V₃, *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1144–1155, 2006. [Version 3.3 and 3.4 Level 2 data quality and description document, Tech. rep., Jet Propulsion Laboratory, http://mmls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf](http://mmls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf), 2013b.
- 1335 Lorenc, A. C.: [Analysis methods for numerical weather prediction, Quarterly Journal of the Royal Meteorological Society](http://dx.doi.org/10.1002/qj.49711247414), 112, 1177–1194, doi:10.1002/qj.49711247414, <http://dx.doi.org/10.1002/qj.49711247414>, 1986.
- Loyola, D. G., Koukouli, M. E., Valks, P., Balis, D. S., Hao, N., van Roozendaal, M., Spurr, R. J. D., Zimmer, W., Kiemle, S., Lerot, C., and Lambert, J.-C.: The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation, *Journal of Geophysical Research (Atmospheres)*, 116, D07302, doi:10.1029/2010JD014675, 2011.
- 1340 Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W., and Baumgardner, J.: The operational global icosahedral-hexagonal gridpoint model GME: Description and high resolution tests, *Mon. Weather Rev.*, 130, 319–338, doi:10.1029/1998GL900152, 2001.
- 1345 Manney, G. L., Santee, M. L., Rex, M., Livesey, N. J., Pitts, M. C., Veefkind, P., Nash, E. R., Wohltmann, I., Lehmann, R., Froidevaux, L., Poole, L. R., Schoeberl, M. R., Haffner, D. P., Davies, J., Dorokhov, V., Gernandt, H., Johnson, B., Kivi, R., Kyrö, E., Larsen, N., Levelt, P. F., Makshtas, A., McElroy, C. T., Nakajima, H., Parrondo, M. C., Tarasick, D. W., von der Gathen, P., Walker, K. A., and Zinoviev, N. S.: Unprecedented Arctic ozone loss in 2011, *Nature*, 478, 469–475, doi:10.1038/nature10556, 2011.
- 1350 [McPeters, R. D., Bhartia, P. K., Haffner, D., Labow, G. J., and Flynn, L.: The version 8.6 SBUV ozone data record: An overview, Journal of Geophysical Research: Atmospheres](http://dx.doi.org/10.1002/jgrd.50597), 118, 8032–8039, doi:10.1002/jgrd.50597, <http://dx.doi.org/10.1002/jgrd.50597>, 2013.
- 1355 Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, *Journal of Geophysical Research (Atmospheres)*, 114, D06206, doi:10.1029/2008JD011235, 2009.
- Munro, R., Anderson, C., [and Callies, J. ~~and et al.~~ a.](#): GOME-2 on MetOp, in: *Atmospheric Science Conference*, vol. 628 of *ESA Special Publication*, 2006.
- 1360 Peuch, V.-H., Engelen, R., Simmons, A., Lahoz, W., Laj, P., and Galmarini, S., eds.: Monitoring atmospheric composition and climate, research in support of the Copernicus/GMES atmospheric service, *Atmos. Chem. Phys. Special Issue*, 2013, http://www.atmos-chem-phys.net/special_issue310.html, 2014.
- 1365 Pommereau, J. P. and Goutail, F.: O₃ and NO₂ ground-based measurements by visible spectrometry during Arctic winter and spring 1988, *Geophysical Research Letters*, 15, 891–894, doi:10.1029/GL015i008p00891, 1988.
- Prather, M. J.: Numerical advection by conservation of second-order moments, *Journal of Geophysical Research*, 91, 6671–6681, doi:10.1029/JD091iD06p06671, 1986.
- 1370 [Redler, R., Valcke, S., and Ritzdorf, H.: OASIS4 – a coupling software for next generation earth system modelling, Geoscientific Model Development](http://www.geosci-model-dev.net/3/87/2010/), 3, 87–104, doi:10.5194/gmd-3-87-2010, <http://www.geosci-model-dev.net/3/87/2010/>, 2010.

- 1375 [Robichaud, A., Ménard, R., Chabrilat, S., de Grandpré, J., Rochon, Y. J., Yang, Y., and Charette, C.: Impact of energetic particle precipitation on stratospheric polar constituents: an assessment using monitoring and assimilation of operational MIPAS data, *Atmospheric Chemistry and Physics*, 10, 1739–1757, <http://www.atmos-chem-phys.net/10/1739/2010/>, 2010.](http://www.atmos-chem-phys.net/10/1739/2010/)
- [Rood, R. B., Allen, D. J., Baker, W. E., Lamich, D. J., and Kaye, J. A.: The Use of Assimilated Stratospheric Data in Constituent Transport Calculations, *Journal of the Atmospheric Sciences*, 46, 687–702, doi:10.1175/1520-0469\(1989\)046\(0687:TUOASD\)2.0.CO;2, 1989.](https://doi.org/10.1175/1520-0469(1989)046(0687:TUOASD)2.0.CO;2)
- 1380 [Sagi, K., Murtagh, D., Urban, J., Sagawa, H., and Kasai, Y.: The use of SMILES data to study ozone loss in the Arctic winter 2009/2010 and comparison with Odin/SMR data using assimilation techniques, *Atmospheric Chemistry and Physics Discussions*, 14, 7889–7916, doi:10.5194/acpd-14-7889-2014, <http://www.atmos-chem-phys-discuss.net/14/7889/2014/>, 2014.](http://www.atmos-chem-phys-discuss.net/14/7889/2014/)
- 1385 Sander, S. P., Golden, D. M., and Kurylo, M. J., e. a.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation Number 15, JPL Publication 02-25, Jet Propulsion Laboratory, Pasadena, CA, USA., 2006.
- ~~Sandu, A. and Chai, T.: Chemical Data Assimilation—An Overview, *Atmosphere*, 2, 426–463, 2011.~~
- Scarnato, B., Staehelin, J., Peter, T., Gröbner, J., and Stübi, R.: Temperature and slant path effects in Dobson and Brewer total ozone measurements, *Journal of Geophysical Research (Atmospheres)*, 114, D24303, doi:10.1029/2009JD012349, 2009.
- 1390 Schoeberl, M. R., Douglass, A. R., Hilsenrath, E., Bhartia, P. K., Beer, R., Waters, J. W., Gunson, M. R., Froidevaux, L., Gille, J. C., Barnett, J. J., Levelt, P. F., and de Cola, P.: Overview of the EOS Aura Mission, *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1066–1074, doi:10.1109/TGRS.2005.861950, 2006.
- 1395 Schwinger, J. and Elbern, H.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: a posteriori validation of error statistics in observation space, *Journal of Geophysical Research (Atmospheres)*, 115, D18307, doi:10.1029/2009JD0131153, 2010.
- 1400 Smit, H. G. J., ~~Strater~~Sträter, W., Helten, M., Kley, D., Ciupa, D., Claude, H., ~~Khler~~Köhler, U., Hoegger, B., Levrat, G., Johnson, B., Oltmans, S. J., Kerr, J. B., Tarasick, D. W., Davies, J., Shitamichi, M., Srivastav, S. K., and Vialle, C.: JOSIE: The 1996 WMO international intercomparison of ozonesondes under quasi-flight conditions in the environmental chamber at Jülich, in *Atmospheric Ozone, Proceedings of the Quadrennial O3 Symposium, l’Aquila, Italy*, pp. 971–974, 1996.
- 1405 Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stubi, R., Schmidlin, F. J., Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., and Posny, F.: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), *Journal of Geophysical Research (Atmospheres)*, 112, D19306, doi:10.1029/2006JD007308, 2007.
- Solomon, S.: Stratospheric ozone depletion: a review of concepts and history, *Reviews of Geophysics*, 37, 275–316, 1999.
- 1410 Solomon, S., Portmann, R. W., Sasaki, T., Hofmann, D. J., and Thompson, D. W. J.: Four decades of ozonesonde measurements over Antarctica, *Journal of Geophysical Research (Atmospheres)*, 110, D21311, doi:10.1029/2005JD005917, 2005.

~~Stein, O.: Model description of the IFS-MOZ coupled system, MACC-II project deliverable GRG-D55.4, 2013.~~

1415 Stein, O., Flemming, J., Inness, A., Kaiser, J., and Schultz, M.: Global reactive gases forecasts and reanalysis in the MACC project, *J. Integr. Environ. Sci.*, 9, 57–70, doi:10.1080/1943815X.2012.696545, 2012.

~~Stein, O., Huijnen, V., and Flemming, J.: Model description of the IFS-MOZART and IFS-TM5 coupled systems, MACC-II project deliverable D55.4, ECMWF http://www.gmes-atmosphere.eu/documents/maccii/deliverables/grg/MACCII-GRG_DEL_D_55.4_IFS-MOZART_and_TM5.pdf, 2013.~~

1420 Stübi, R., Levrat, G., Hoegger, B., Viatte, P., Staehelin, J., and Schmidlin, F. J.: In-flight comparison of Brewer-Mast and electrochemical concentration cell ozonesondes, *Journal of Geophysical Research (Atmospheres)*, 113, D13302, doi:10.1029/2007JD009091, 2008.

~~Valeke, S. and Redler, R.: OASIS4 User Guide (OASIS4.0.2), PRISM Support Initiative, Technical Report No 4, 2006.~~

1425 van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Multi sensor reanalysis of total ozone, *Atmospheric Chemistry & Physics*, 10, 11 277–11 294, doi:10.5194/acp-10-11277-2010, 2010.

van Roozendaal, M., Hermans, C., Kabbadj, Y., and et al.: Ground-based measurements of stratospheric OClO, NO₂ and O₃ at Harestua, Norway (60°N, 10°E) during SESAME, European rocket and balloon programmes and related research, 370, 305, 1995.

1430 Van Roozendaal, M., Peeters, P., Roscoe, H. K., De Backer, H., Jones, A. E., Bartlett, L., Vaughan, G., Goutail, F., Pommereau, J.-P., Kyro, E., Wahlstrom, C., Braathen, G., and Simon, P. C.: Validation of Ground-Based Visible Measurements of Total Ozone by Comparison with Dobson and Brewer Spectrophotometers, *Journal of Atmospheric Chemistry*, 29, 55–83, 1998.

1435 van Roozendaal, M., Loyola, D., Spurr, R., Balis, D., Lambert, J.-C., Livschitz, Y., Valks, P., Ruppert, T., Kenter, P., Fayt, C., and Zehner, C.: Ten years of GOME/ERS-2 total ozone data - The new GOME data processor (GDP) version 4: 1. Algorithm description, *Journal of Geophysical Research (Atmospheres)*, 111, D14311, doi:10.1029/2005JD006375, 2006.

Viscardy, S., Errera, Q., Christophe, Y., Chabrilat, S., and Lambert, J.-C.: Evaluation of ozone analyses from UARS MLS assimilation by BASCOE between 1992 and 1997, *JSTARS* 3, pp. 190–202, 2010.

1440 von Barga, A., ~~Schröder~~Schröder, T., Kretschel, K., Hess, M., Lerot, C., Van Roozendaal, M., Vountas, M., Kokhanovsky, A., Lotz, W., and H., B.: Operational SCIAMACHY Level 1B-2 Off-line processor: total vertical columns for O₃ and NO₂ and cloud products., *Proceeding Envisat Symposium 2007, Montreux, Switzerland, 23-27 April 2007 (ESA SP-636, July 2007)*, 2007.

1445 Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., Labelle, R. C., Lam, J. C., Lee, A. K., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Vansnyder, W., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura Satellite, *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1075–1092, doi:10.1109/TGRS.2006.873771, 2006.

Table 1: Satellite retrievals of ozone that were actively assimilated by the four models of the MACC stratospheric ozone service. The Aura MLS data used by IFS-MOZART and BASCOE are not the same: IFS-MOZART used the MLS NRT retrievals of ozone only, while BASCOE uses the standard scientific, offline retrievals including five other species. PC stands for partial columns, TC for total columns and PROF for profiles. When two references are provided, the first refers to the satellite sensor, the second one to the retrieval algorithm.

Analysis	Satellite	Sensor	Provider	Version	Assim. data	Period	Reference
IFS-MOZART	NOAA- 16/17/18/19	SBUV/2	NOAA	V08-V8.0	PC (see text)	1 Sep 2009– 30 Sep 2012	? Bhartia et al. (2013)
	Aura	OMI	NASA/JPL	V003	TC	1 Sep 2009– 30 Sep 2012	Levelt et al. (2006)
	Envisat	SCIAMACHY	KNMI	TOSOMI v2.0	TC	1 Sep 2009– 7 Apr 2012	Eskes et al. (2005)
	Aura	MLS	NASA/JPL	V2.2, NRT	PROF, <68 hPa	1 Sep 2009– 30 Sep 2012	Waters et al. (2006) Livesey et al. (2006)
BASCOE	Aura	MLS	NASA/JPL	V2.2, SCI	PROF	1 Jul 2009– 30 Sep 2012	Waters et al. (2006) Livesey et al. (2006) Livesey et al. (2013a)
SACADA V2.0	Envisat	SCIAMACHY	DLR, on behalf of ESA	SGP-5.01	TC	5 Mar 2010– 27 Oct 2011	von Bargaen et al. (2007)
SACADA V2.4	MetOp-A	GOME-2	EUMETSAT	GDP 4.1	TC	28 Oct 2011– 30 Sep 2012	Loyola et al. (2011) van Roozendaal et al. (2006)
TM3DAM	Envisat	SCIAMACHY	KNMI	TOSOMI v2.0	TC	16 Mar 2010– 31 Mar 2012	Eskes et al. (2005)
	MetOp-A	GOME-2	DLR	GDP 4.x	TC	1 Apr 2012– 30 Sep 2012	Loyola et al. (2011) van Roozendaal et al. (2006)

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, *Quarterly Journal of the Royal Meteorological Society*, 127, 1815–1846, 2001.

1455 Weber, M., Dikty, S., Burrows, J. P., Garny, H., Dameris, M., Kubin, A., Abalichin, J., and Langematz, U.: The Brewer-Dobson circulation and total ozone from seasonal to decadal time scales, *Atmospheric Chemistry & Physics*, 11, 11 221–11 235, doi:10.5194/acp-11-11221-2011, 2011.

WMO: ~~Scientific Assessment of Ozone Depletion: 2010, Global Ozone Research and Monitoring Project Report No. 52, Geneva, Switzerland, 516 pp., 2011.~~

1460 ~~WMO: WMO/UNEP Report of the 8th ORM (Ozone Research Managers) Meeting of the Parties to the Vienna Convention for the Protection of the Ozone Layer (Geneva, Switzerland, 2-4 May 2011), 2011.~~ [2011](#)

Table 2: Specification of the characteristics of the four assimilation systems: IFS-MOZART, BASCOE, SACADA, and TM3DAM. The horizontal and vertical resolution have been abbreviated to Hor. and Vert. resol. respectively. Freq. stands for frequency, and Assim. for assimilation.

	IFS-MOZART	BASCOE	SACADA	TM3DAM
Hor. resol.	1.875° × 1.875°	3.75° × 2.5°	250 km	3° × 2°
Vert. resol.	60 layers up to 0.1 hPa	37 layers up to 0.1 hPa	32 layers between 7 and 66 km	44 layers between 0 and 1013 hPa
Output freq.	6 hourly	3 hourly	daily, at 12 h UT	daily, at 21 h UT
Meteo input	operational hourly meteo fields from IFS	operational 6 hourly meteo analyses from IFS	24 h GME forecast initialised by IFS analyses	operational 6 hourly meteo analyses from IFS
Advection scheme	flux form semi-Lagrangian (Lin and Rood, 1996)	flux form semi-Lagrangian (Lin and Rood, 1996)	semi-Lagrangian and upstream method	flux-based second order moments scheme (Prather, 1986)
Chemical mechanism	JPL-06 (Sander et al., 2006) with some modifications as described in Stein et al. (2013) 115 species 325 reactions	JPL-06 (Sander et al., 2006) Aerosols and PSCs 57 species 200 reactions	JPL-06 (Sander et al., 2006) Aerosols and PSCs (Damski et al., 2007) 48 species 177 reactions	Cariolle parameterization + cold tracer (2 species) (Cariolle and Teyssède, 2007)
Assim. method	4D-Var	4D-Var	4D-Var	Kalman Filter approach
Assim. window	12 h	24 h	24 h	24 h

1465 ~~Comparison between the TOC time series (5 day moving average) of the four analyses of the
MACC stratospheric ozone service (IFS-MOZART in red, BASCOE in blue, SACADA in green,
and TM3DAM in cyan) interpolated at a high northern latitude station (Alert, 82.49N, 62.42W),
a tropical station (Chengkung, 23.1N, 121.365E) and a southern latitude station (Syowa, 69S, 39.58E),
for the period September 2009 to September 2012. Black asterisks are 5day moving averages
of Brewer (for Alert and Chengkung) and Dobson (for Syowa) observations from the WOUDC
network. A conservative observational error of 2% is indicated by gray errorbars.~~

1470 ~~Comparison between the time series (5 day moving average) of ozone at 50by IFS-MOZART
(red), BASCOE (blue), and SACADA (green) interpolated at Ny-Ålesund (top panel, 78.92N, 11.93E),
Nairobi (middle panel, 1.27S, 36.8E) and Neumayer (bottom panel, 70.68S, 8.26W), for the period
September 2009 to September 2012. Black asterisks are 5 day moving averages (except for Nairobi
where observations were taken at least one week apart) of sonde observations from the NDACC
1475 network.~~

Location of observational sites

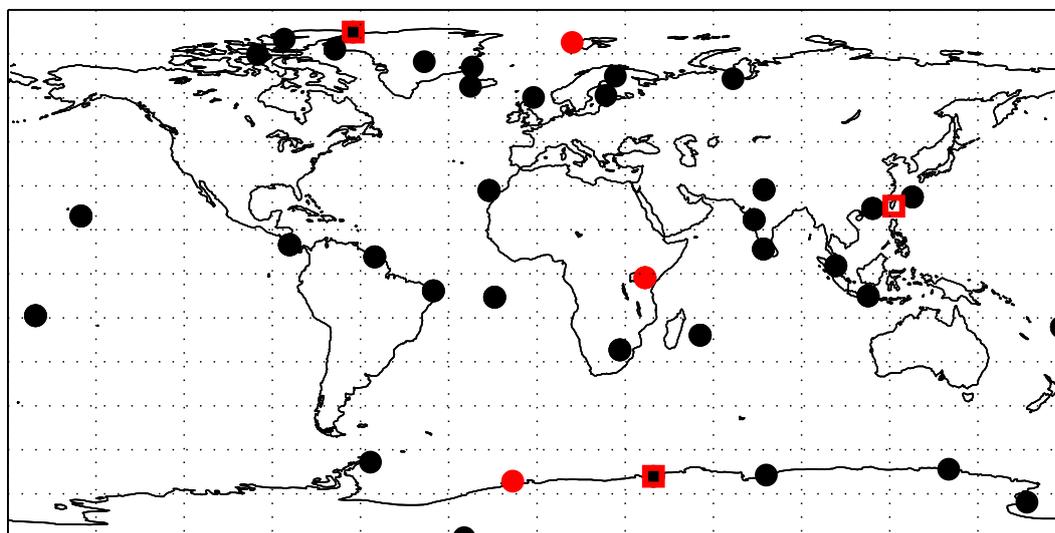


Fig. 1: Location of all stations used in this paper. O₃ sondes are indicated as filled black circles. The ones selected for a more detailed discussion have been marked in red: Ny-Ålesund (79° N, 12° E) in the Arctic, Nairobi (1.27° S, 36.8° E) in the Tropics, and Neumayer (70.65° S, 8.25° W) in the Antarctic. The three sites selected for the Total Ozone Column (TOC) discussion are indicated by the red squares.

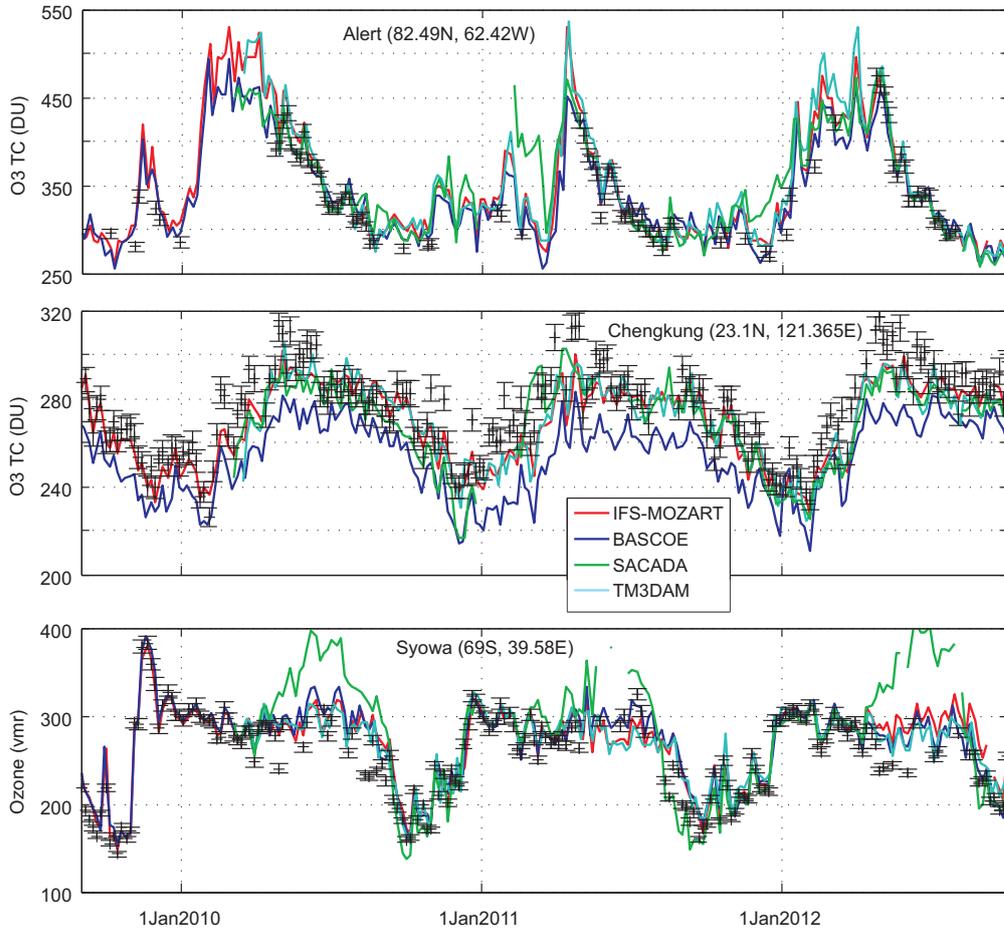


Fig. 2: Comparison between the TOC time series (5 day moving average) of the four analyses of the MACC stratospheric ozone service (IFS-MOZART in red, BASCOE in blue, SACADA in green, and TM3DAM in cyan) interpolated at a high northern latitude station (Alert, 82.49° N, 62.42° W), a tropical station (Chengkung, 23.1° N, 121.365° E) and a southern latitude station (Syowa, 69° S, 39.58° E), for the period September 2009 to September 2012. Black symbols are 5 day moving averages of Brewer (for Alert and Chengkung) and Dobson (for Syowa) observations from the WOUDC network. In order to indicate the observational uncertainty, the height of each symbol is set to 4% of the observed value.

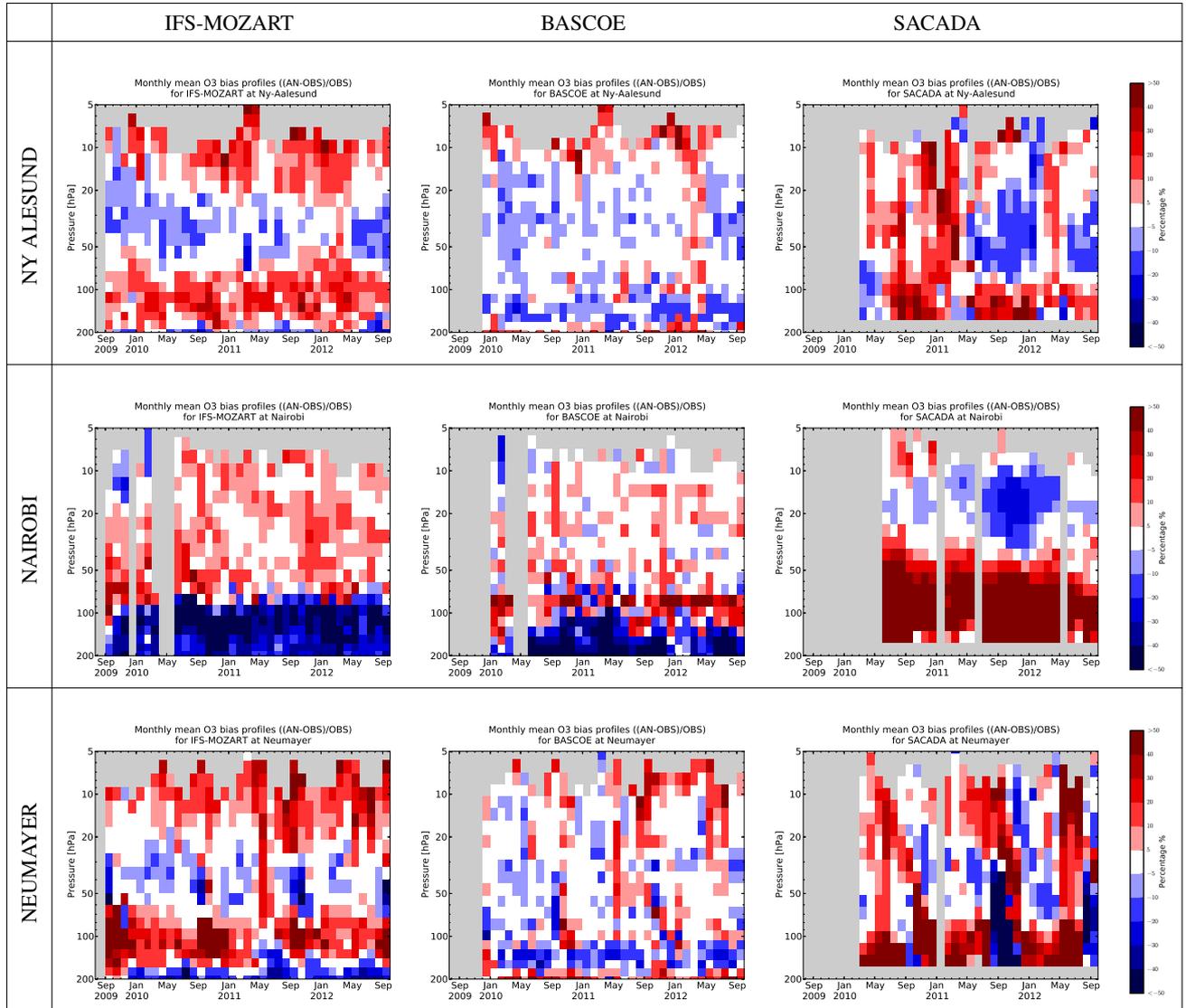


Fig. 3: Time series of monthly mean ozone biases (analysis minus observations) with respect to ozonesondes at Ny-Ålesund (top panel, 78.92° N, 11.93° E), Nairobi (middle panel, 1.27° S, 36.8° E) and Neumayer (bottom panel, 70.68° S, 8.26° W) for the period September 2009 to September 2012 in %. Left: IFS-MOZART, middle: BASCOE, right: SACADA.

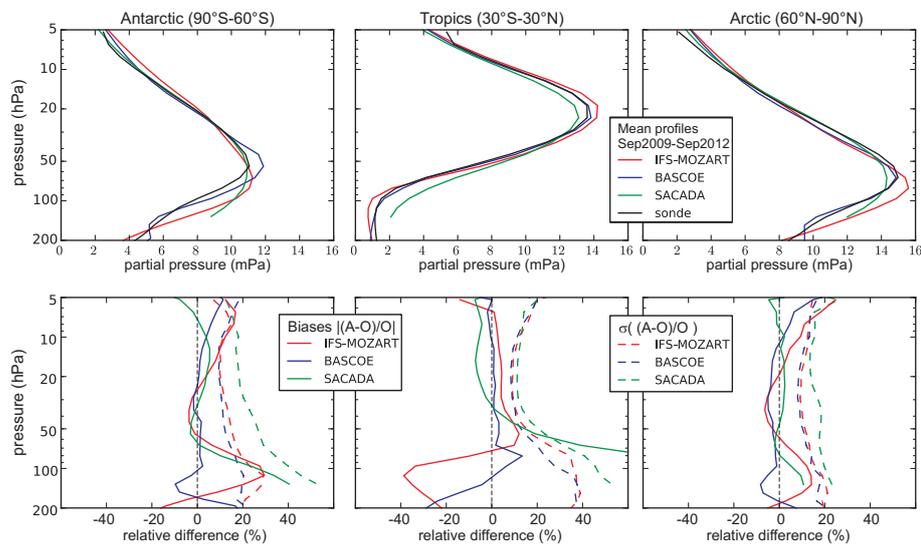


Fig. 4: **Average** Top row: mean ozone profiles (left top rows) as partial pressures in mPa from IFS-MOZART (red), BASCOE (blue), SACADA (green) and ozonesondes (black). Bottom row: mean and the average ozone bias (solid lines) and standard deviation (dashed lines) of th relative differences, in %, of these analyses against the ozonesondes (right), over the period September 2009 to September 2012. Dotted lines represent the standard deviation for BASCOE and IFS-MOZART with the temporal resolution degraded to the one of SACADA. See Sect. 4.4.3.

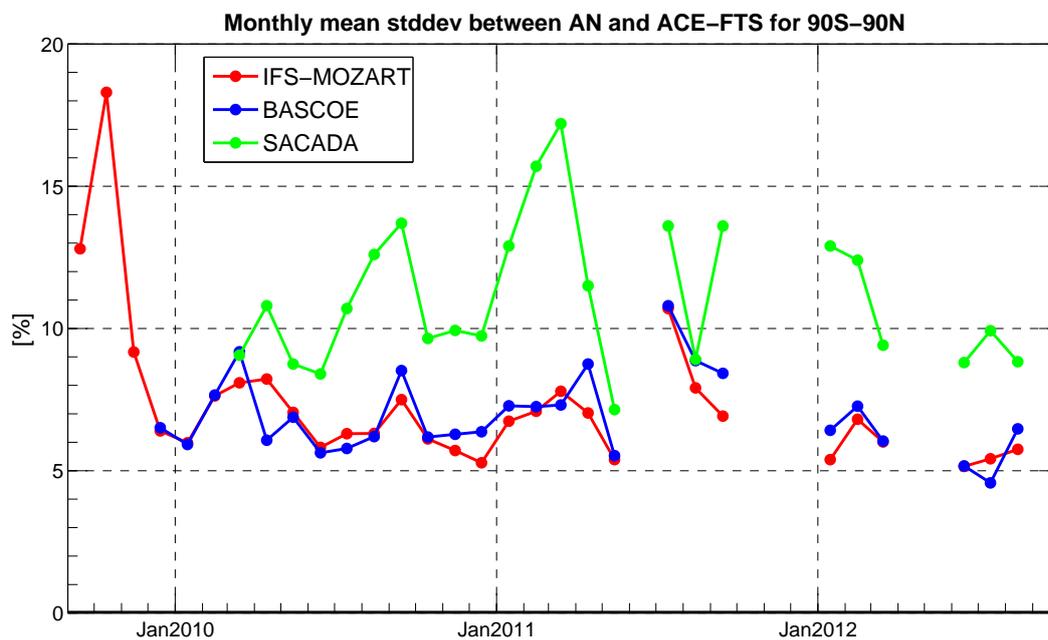


Fig. 5: Comparison of the global (i.e. from 90° S to 90° N) monthly mean standard deviation between IFS-MOZART (red), BASCOE (blue), and SACADA (green) with ACE-FTS (analysis minus observations) in %, for the [200,5]hPa pressure bin, for the period September 2009 to September 2012. Standard deviations for levels for which there are less than 20 observations are left out. Note that standard deviations are not weighted by the cosine of the latitude.

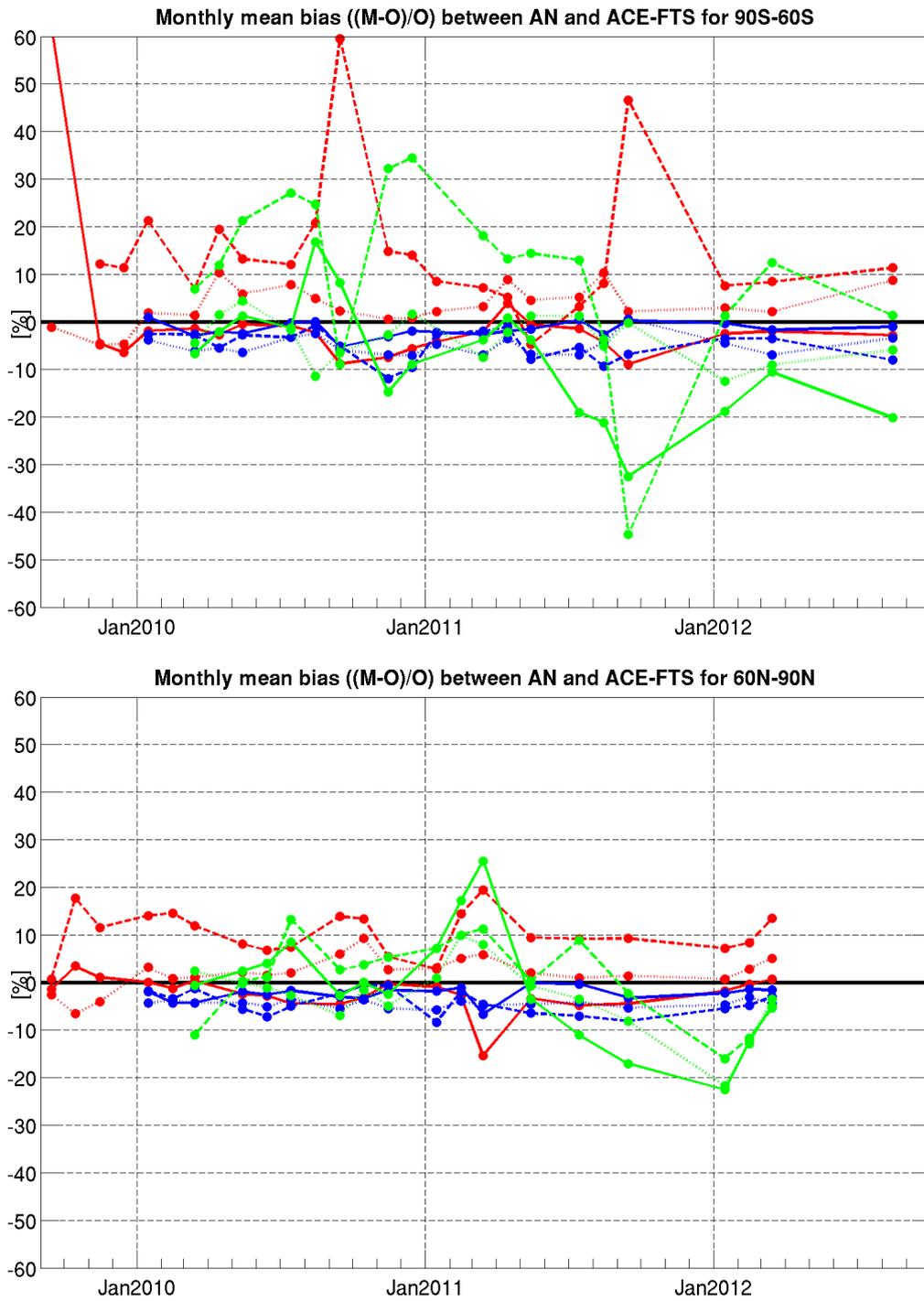


Fig. 6: Comparison of the monthly mean relative ozone biases between IFS-MOZART (red), BASCOE (blue), and SACADA (green) and ACE-FTS (analysis minus observations) in %, at 100 (dashed), 50 (full) and 10 hPa (dotted) for the period September 2009 to September 2012 for the Arctic-Antarctic (topleft) and the Antarctic-Arctic (bottomright).

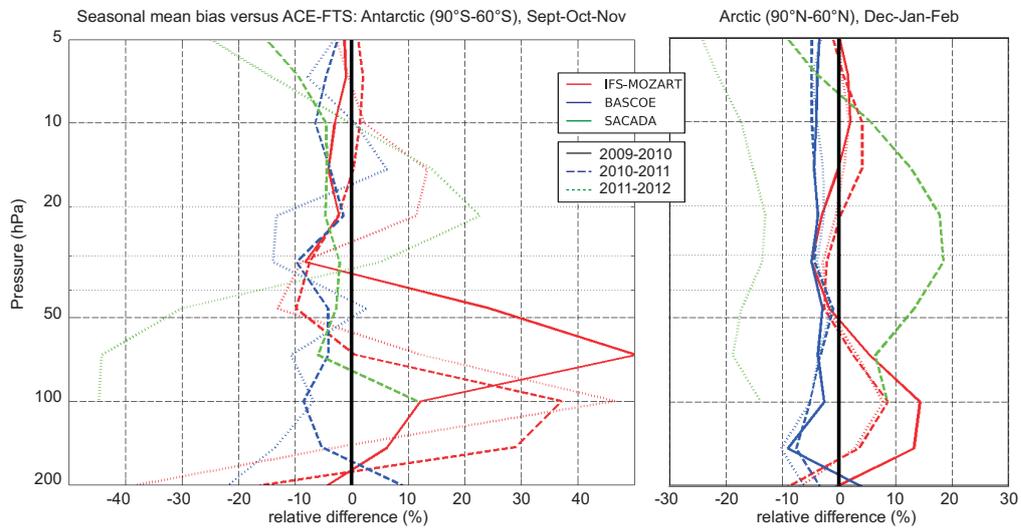


Fig. 7: Seasonally averaged relative ozone bias profiles of IFS-MOZART (red), BASCOE (blue), and SACADA (green) vs. ACE-FTS (analysis minus observations) in % for the ~~Arctic winter~~ Antarctic spring (DJF i.e. left: months ~~December~~September, ~~January~~October, and ~~February~~November) 2009–2010 (full), 2010–2011 (dashed), 2011–2012 (dotted) and ~~for the Antarctic spring~~ Arctic winter (SON i.e. right: months ~~September~~December, ~~October~~January, and ~~November~~February) 2009 in 2009–2010 (full lines), 2010–2010–2011 (dashed lines), and 2011–2011–2012 (dotted lines).

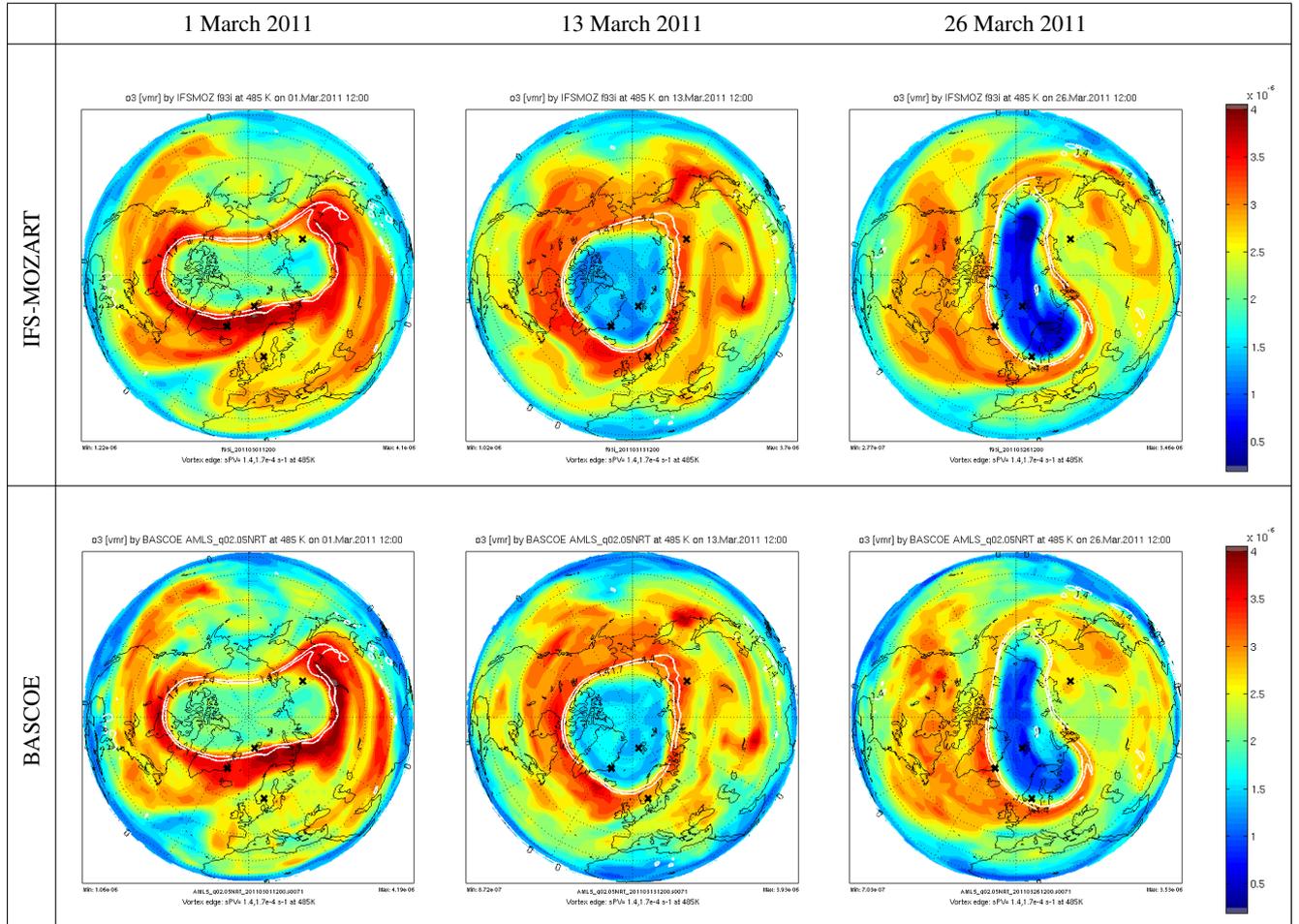


Fig. 8: Evolution of ozone volume mixing ratios at 485 K during March 2011 (left: 1 March, middle: 13 March, right: 26 March). Red/blue colours indicate respectively high/low ozone values. In white, the inner and outer polar vortex edges are indicated, calculated with an sPV of, respectively, $> 1.7e^{-4}s^{-1}$ and $> 1.4e^{-4}s^{-1}$. Top: IFS-MOZART, bottom: BASCOE. The location of three SAOZ stations used for the detailed total ozone column evaluation are indicated by black crosses: Zhigansk (Russia, 66.8° N, 123.4° E), Harestua (Norway, 60° N, 11° E), Scoresbysund (Greenland, 70.49° N, 21.98° W), as well as the location of the O₃ sonde at Ny-Ålesund (Spitzbergen, 78.933° N, 11.883° E).

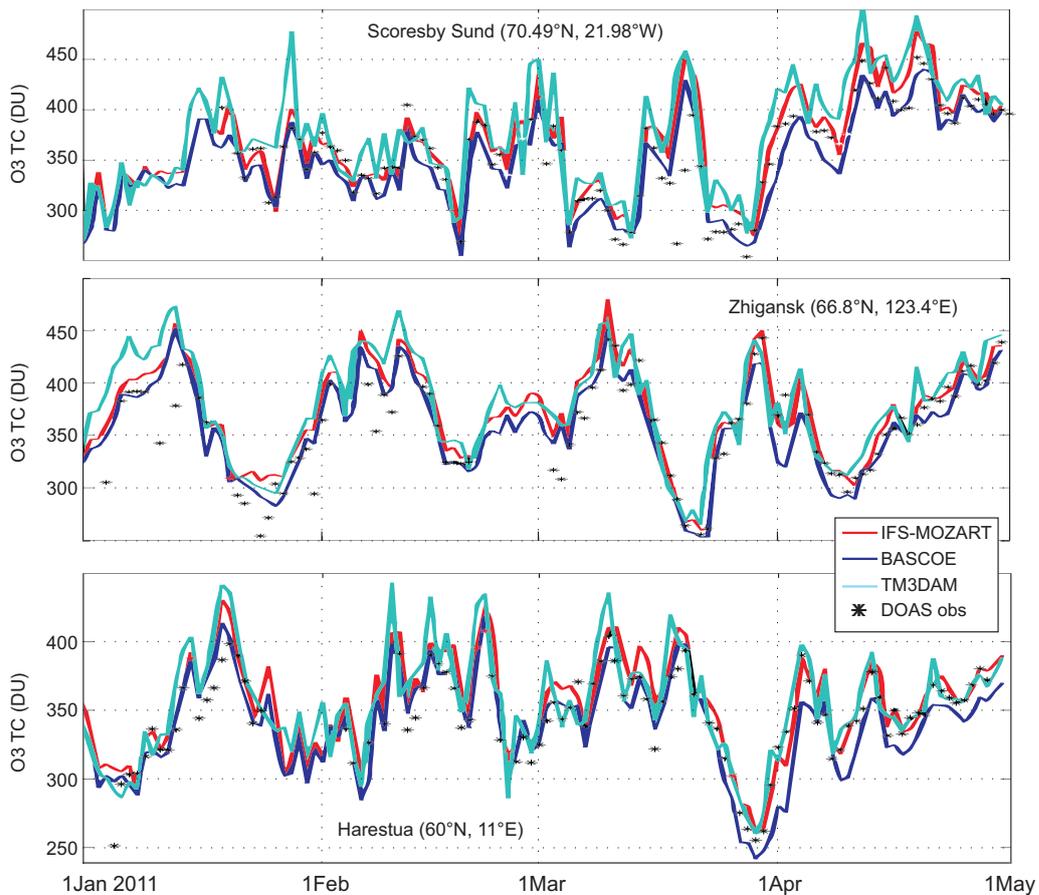


Fig. 9: Comparison of daily averaged total ozone columns (expressed in Dobson Units) for IFS-MOZART (red), BASCOE (blue), and TM3DAM (cyan) vs. ozone measurements from three SAOZ/DOAS stations in the Arctic: [Zhigansk-Scoresby Sund](#) ($66.870.49^\circ$ N, $123.421.98^\circ$ E, SAOZW), [Harestua-Zhigansk](#) ($60.66.8^\circ$ N, 1123.4° E, DOAS) and [Scoresby Sund-Harestua](#) (70.4960° N, 21.9811° W, SAOZE).

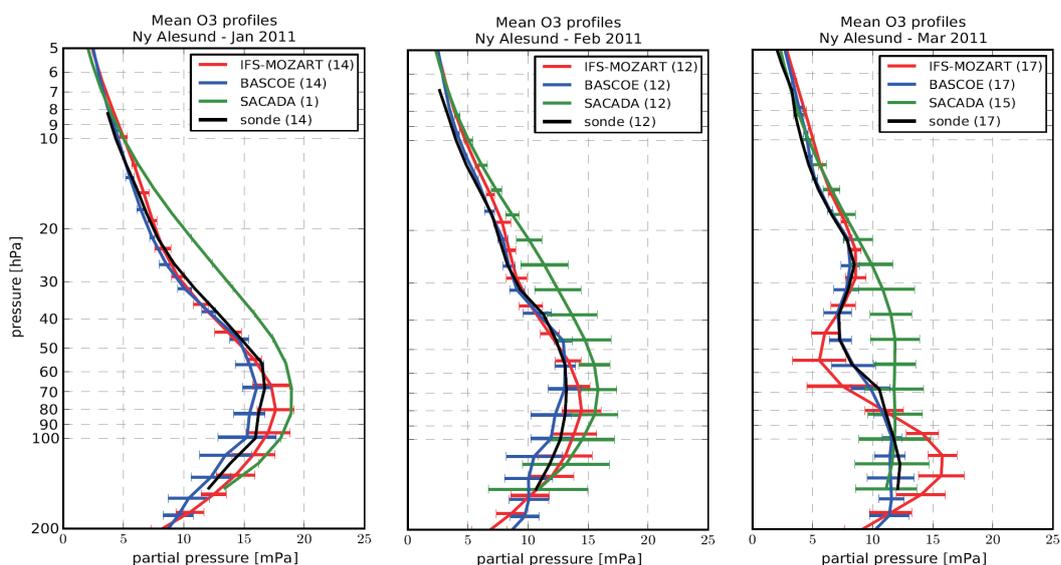


Fig. 10: Comparison of the monthly average O₃ profiles of IFS-MOZART (red), BASCOE (blue) and SACADA (green) with O₃ sonde profiles observed at Ny-Ålesund for January to April 2011. The number of available O₃ sonde profiles and the number of collocated system profiles are indicated in brackets.

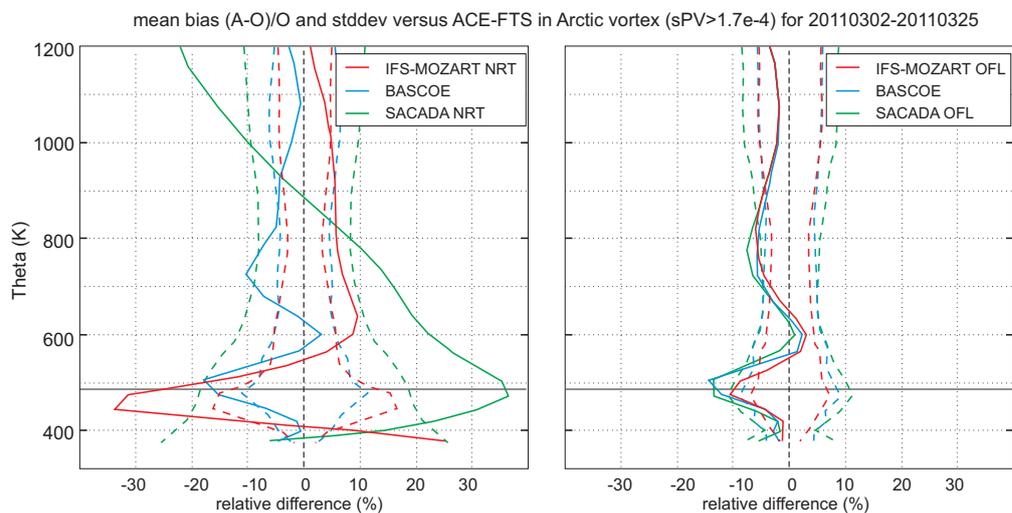


Fig. 11: Mean bias and standard deviations of the differences between the NRT analyses (left) and the offline experiments (right) of IFS-MOZART, BASCOE, and SACADA, on the one hand, and ACE-FTS observations, on the other hand, within the North Pole vortex (vortex edge calculated with an sPV of $> 1.7e^{-4} s^{-1}$) for March 2011. The ozone hole level used in Fig. 9.7 ($\theta \approx 485$ K) is indicated as the black horizontal line.

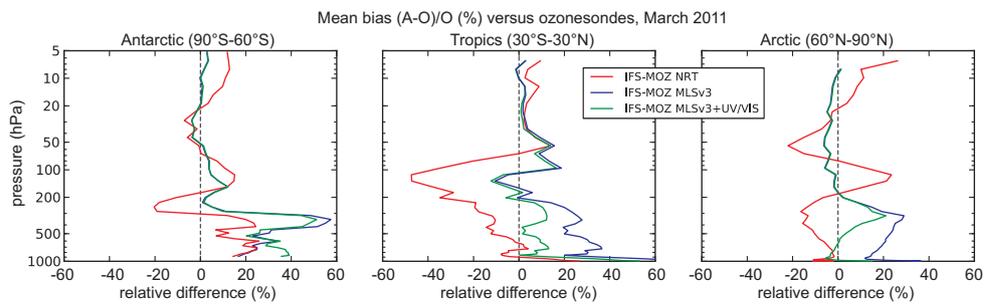


Fig. 12: Mean biases of three ozone analyses by IFS-MOZART using O_3 sonde profiles as reference, for March 2011. Results are shown for the Antarctic (left), Tropics (center) and Arctic (right) latitude bands using the IFS-MOZART NRT analyses (red lines), the offline experiment assimilating only MLS v3 (blue lines) and another offline experiment assimilating MLS v3 and the UV/VIS observations (green lines). See text for details.