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 adress 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 will be 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 will be 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 Reviwer 1), we will include 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 (NOx i.e. NO+NO_2), sulphur dioxide (SO_2) and ozone (O_3). 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. 12) but the analyses delivered in NRT by BASCOE indeed outperformed those delivered by SACADA due to different configuration choices. This will be 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 will be 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 will be explained 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:

(iv) **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 ESSD) 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 ressources to re-submit two papers. As mentioned by both reviewers, 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 will re-write 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 will even slightly modify the title to reflect the importance of the 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 whole 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 will be nearly completely re-written in the revised manuscript, with more focus on satellite data and CDA. Keeping only the beginning of the first paragraph (until P12464 L26), we will 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. Lorenc 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 will continue with an explanation of 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 finish 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) will be 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 reslution of SCIA nadir will be corrected. Since SCIA limb is not used, the mention of its vertical reslution will be 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 will be clarified (see similar specific comment by reviewer 1). In this section we prefer to discuss only the satellite instruments which were actually assimilated in thi study.

- *SBUV-2 – I would replace "larger precision" by "lower precision"*

  This will be 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, inclusing 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 will be 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 will be 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 will be 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 re-sized 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_METEOFR.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. Chabrillat, 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 NOy 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.