

Response to the reviewer comments on the manuscript

C-IFS-CB05-BASCOE: Stratospheric Chemistry in the Integrated Forecasting System of ECMWF

By V. Huijnen et al.

First, we would like to thank the reviewers for their critical, but useful comments. In view of their valuable suggestions in our revised manuscript we have:

- 1 included an additional model configuration containing full (tropospheric *and* stratospheric) chemistry within the whole atmosphere
- 2 revised our PSC-parameterization
- 3 extended our model evaluation with one additional year
- 4 revised some of our evaluations

The reviewer's comments are given in italic, and our responses in regular font. Textual modifications to the manuscript are highlighted in bold. Figure numbers refer to the revised manuscript.

Response to anonymous Referee #2

The purpose of the paper is to describe and benchmark a new version of the IFS model. This version has separate chemistry modules (mechanisms) for the troposphere and stratosphere, where the decision of which module to call is determined by the altitude of the grid box with respect to the tropopause. The stratospheric chemical mechanism comes from an assimilation system (BASCOE). The previous version of the IFS had tropospheric chemistry plus linearized stratospheric O₃. The paper concludes that a new simulation that uses both chemical mechanisms (called CIFS-TS) has good stratospheric O₃, NO₂ and other reactive trace gases compared to satellite data sets.

A goal of the paper is to demonstrate that their method of using the tropospheric mechanism/solver for tropospheric grid boxes and the stratospheric solver for stratospheric grid boxes is a computationally efficient way to calculate the full chemistry of the atmosphere. The biggest problem with this paper is that they did not actually test this. To demonstrate their method, they need to have run a simulation where tropospheric and stratospheric reactions were solved TOGETHER and NOT split into 2 mechanisms. Those results could then be compared with their 'split' method. Ideally, this would show that their method was faster (how much faster?) yet produced essentially the same results. I recommend they do this and then rewrite this manuscript. This experiment would not only satisfy the stated goal of the paper but it would also eliminate the confusion in the comparisons (see below) regarding transport/advection differences between BASCOE-CTM and the CIFS-TS.

We thank the reviewer for his/her valuable comments on our manuscript. Indeed a system with tropospheric and stratospheric chemistry resolved throughout the atmosphere (in the remainder referred to as 'C-IFS-Atmos'), as opposed to the reported more efficient approach in C-IFS-TS, had been already developed and briefly mentioned in the manuscript but its evaluation was missing. The main reason for not presenting this was that the stratospheric chemistry was treated very similar and hence small differences in model results between C-IFS-Atmos and C-IFS-TS could be expected in the

stratosphere. Discrepancies to the observations mainly raise from common stratospheric chemistry model assumptions (e.g. PSC treatment, photolysis) and differences in transport treatment between C-IFS and BASCOE-CTM. Larger differences can be only be expected when approaching the tropopause. We would like to point out that the original manuscript does mention that C-IFS-Atmos is 50% more expensive than C-IFS-TS (at the end of Sec. 2.3), essentially due to the larger chemical mechanism throughout the atmosphere that needs to be solved.

In response to the reviewer we acknowledge that an explicit evaluation of the differences between C-IFS-TS and C-IFS-Atmos does clarify our goal, which indeed also aims at presenting our methodology with separate tropospheric and stratospheric chemistry in C-IFS-TS. We therefore now include explicitly C-IFS-Atmos in our model evaluation and show that the differences with the more efficient approach in C-IFS-TS are as small as expected.

This expansion of the model evaluation also increases the usefulness of the comparison with BASCOE-CTM since it is now possible to compare two models with the same stratospheric chemistry but different transport schemes (BASCOE-CTM versus C-IFS-TS) and two models with the same transport but (slightly) different chemical schemes in the stratosphere (C-IFS-TS versus C-IFS-Atmos). A recent study by de Grandpré *et al.* (2016) is now cited to illustrate the type of issues raised in the stratosphere by the semi-lagrangian advection scheme.

I don't agree with the statement in the abstract that the new model configuration shows good performances of stratospheric O₃, NO₂, and other tracers. The figures chosen to demonstrate good representation of various stratospheric constituents in the CIFS-TS model generally show fair to poor agreement with observations. Stratospheric O₃, for example, often looks worse (or at least no better) than it did in the CB05 (trop only) or BASCOE (strat only) versions. This new model does not appear to be an improvement over previous model versions.

The comparisons between CIFS-TS and BASCOE-CTM are confusing. When stratospheric species such as NO₂, HNO₃, and O₃ are compared, the results are different. I thought the primary goal of the paper was to compare the chemical mechanisms, but since the results are rather different, there must be transport (or meteorological field) differences too. This is alluded to on page 7, lines 26-28. The transport/advection needs to be the same in the two simulations in order to compare the chemical mechanisms.

It should be made clearer in the text what the differences are between CIFS-TS and BASCOE-CTM. My overall recommendation is to test a combined (strat+trop) solver and compare the results to trop only, strat only, and the 'split' method presented here. The results will provide a good benchmark and will be easier to interpret if all experiments are performed with the same transport code and meteorological fields.

The reviewer appears confused by the selection of model setups chosen in our manuscript. In essence, the C-IFS is a Global Circulation Model (GCM) designed for meteorological analyses and forecasts where a module for chemistry has been included to extend its abilities in terms of atmospheric composition. On the other hand, the BASCOE system is a dedicated data-assimilation system for stratospheric composition, based on a Chemistry Transport Model (CTM) environment, i.e. a completely independent system to C-IFS.

The impact of various chemical mechanisms was evaluated through comparison of C-IFS-T (with linear ozone treatment in the stratosphere) and C-IFS-TS (which uses the identical chemical parameterization in the stratosphere as BASCOE-CTM).

The BASCOE-CTM is driven by meteorological fields from the C-IFS run, but still uses a different numerical scheme for the advection, and is running on a different grid as compared to the C-IFS runs. Note that in this setup of the BASCOE system the chemical data-assimilation is switched off, hence purely reflecting the forward model capabilities. Hence comparison between BASCOE-CTM and C-IFS-TS is a clean method to evaluate differences due to the representation of transport with identical meteorological fields, and not suited for the evaluation of differences in the chemical treatment since there are no such differences.

The additional model run, C-IFS-Atmos, where tropospheric chemistry is extended throughout the stratosphere, and vice versa, is now included to assess the impact of assumptions of the reduced chemistry in C-IFS-TS.

In response to the reviewer's concerns we now extend the table describing the model versions (see also the comment below), and extend the description between differences of the various setups. Notably in the introduction we now write:

“The CB05 tropospheric scheme has been combined with the stratospheric scheme from BASCOE-CTM to form a single chemistry mechanism that encompasses tropospheric and stratospheric chemistry throughout the atmosphere, here referred to as C-IFS-Atmos. However, this approach appears computationally expensive, due to the extended chemical mechanism. Therefore”

And also:

“In this optimized approach we developed a flexible setup where -within a single framework- either the tropospheric or stratospheric chemistry modules are addressed, referred to as C-IFS-TS. In this approach the parameterizations for the chemistry, including the respective chemistry mechanisms as optimized for troposphere and stratosphere separately, are retained.

In this paper we describe our **two merging approaches** and provide benchmark evaluations of the **C-IFS-Atmos** and C-IFS-TS systems with focus on the stratospheric composition. **The ancestor BASCOE-CTM is also included in the comparison through a forward model run (without chemical data assimilation) in order to provide insight in the differences caused by the treatment of transport between C-IFS and BASCOE.”**

The model evaluation has been extended to include results obtained with C-IFS-Atmos, as well as an evaluation of the stratospheric composition (including O₃, HNO₃ and NO₂) in C-IFS-T, to explicitly identify the impact of the newly implemented stratospheric chemistry within the C-IFS framework.

Other points

It would be helpful to add a table that lists the specifications of each of the models used and notes how dynamical fields are obtained (e.g., forecast, assimilation, ...?), chemical mechanism, resolution, etc. For example, BASCOE is an assimilation system, but it's only the BASCOE stratospheric chemical mechanism that used here, right? And BASCOE-CTM means the assimilated (reanalysis) fields have been saved and then are being used in an offline chemistry transport model? Presumably it is the same offline model that the C-IFS forecast fields are used in? If what I am asking does not make sense, please take this as an indication that I am confused by the descriptions of the models.

In response to the reviewer's concerns in Section 2.3 we expanded Table 2 which lists the specifics of the various model systems, as also given below. Further, we want to make clear that the C-IFS experiments have been run in 'nudged meteo' mode, by relaxation of the meteorology towards ERA-Interim, as we also write in Section 3. The BASCOE-CTM run is driven by the identical meteorology from the C-IFS experiment (and in turn from ERA-Interim), but applies its own advection algorithm which is clearly different from the one used in IFS. In Section 3 we also make more clear what are the differences between C-IFS and BASCOE-CTM. As discussed above, the BASCOE-CTM results are included as a reference of what can optimally be achieved with C-IFS-TS and C-IFS-Atmos in the stratosphere, using only simulations nudged with specified dynamics and unconstrained composition. Specifically we now write:

"Meteorology **in the C-IFS runs** is relaxed towards ERA-Interim (...) The performance of the C-IFS runs has further been compared against the BASCOE-CTM (without **chemical** data assimilation), using the same chemical mechanism and parameterizations **for photolysis and heterogeneous chemistry** as implemented in the C-IFS-TS. **This serves as a model reference for the C-IFS implementation of stratospheric chemistry. While C-IFS evaluates tracer transport on a reduced Gaussian grid, the BASCOE-CTM uses a regular latitude-longitude grid. It is run here with a resolution of 1.125° lon / lat similar to the resolution chosen for C-IFS, and on the same vertical grid of 60 levels. The BASCOE-CTM is driven by temperature, pressure and wind fields simulated by the C-IFS runs. However, while BASCOE adopts a flux-form advection scheme (Lin and Rood, 1996) the IFS uses the Semi-Lagrangian scheme for advection, accounts for vertical diffusion and includes a parameterization for convection (ECMWF, 2015).** Using essentially the same dynamical fields together with an identical implementation of the chemistry code should allow to identify differences **due to the different** transport schemes between C-IFS and the BASCOE-CTM. Common chemical biases between both systems also point at issues in the chemical parameterizations **such as reaction mechanism, photolysis, heterogeneous chemistry and sedimentation.**"

Table 2. Number of trace gases, the chemistry scheme in troposphere and stratosphere, and corresponding number of reactions (gas-phase/ heterogeneous and photolytic), as well as specification of the circulation model and computational expenses of a one-month run on T255L60 in terms of system billing units (SBU) for various C-IFS model versions. For completeness also the BASCOE-CTM system is indicated.

	C-IFS-T	C-IFS-S	C-IFS-Atmos	C-IFS-TS	BASCOE-CTM
No. trace gases	55	59	99	99	59
Chemistry scheme in troposphere	CB05	BASCOE (P<400hPa)	CB05+BASCOE	CB05	BASCOE (P<400hPa)
Chemistry scheme in stratosphere	CB05/ Cariolle	BASCOE	CB05+BASCOE	BASCOE	BASCOE
No. reactions (gas / het / photo)	93/3/18	142/9/52	211/11/60	93/3/18 or 142/9/52	142/9/52
Circulation model	GCM	GCM	GCM	GCM	CTM
SBU	2075	2500	4563	3076	- ^a

^aBASCOE does not run on the ECMWF supercomputing facility and hence cannot be compared directly to C-IFS in terms of computational resources.

Regarding ‘tracer species’ or similar expression found in many places, ‘tracer’ means a species that is unreactive and can be used to trace something, like transport. I think you mean ‘trace gas’ rather than tracer because that can be used in a general way to talk about any type of constituent in the model. Please search on ‘tracer’ in the document to identify where you mean trace gas or constituent.

We thank the reviewer for this comment, and changed the wording accordingly throughout the document.

p. 3, l.24. Are you saying the chemistry in the modules is parameterized? Or are you referring to the chemical mechanisms when you say ‘chemical parameterization’? A parameterization for chemistry is not the same thing as a chemical mechanism. Sometimes ‘chemical schemes’ is used, which is fine for referring to the mechanism. This confusion occurs throughout the paper. Please check each occurrence of ‘parameterization’ to verify the right words were chosen.

In this occasion the phrase ‘chemistry parameterization’ referred to all chemical conversion processes that require a parameterization, including aqueous phase and heterogeneous reactions as well as photolysis and parameterizations for sedimentation. Indeed this refers to more than just the definition of the chemical mechanism. To accommodate the concerns of the reviewer we had a critical look at our terminology for ‘parameterization’ in the complete manuscript, and changed it where appropriate (see also below). In this instance mentioned by the reviewer in Sec. 2.0 (p.3, l.24) we only wish to guide the

reader forward to the specific sections on stratospheric/tropospheric chemistry, but to prevent potential confusion we now write “tropospheric (CB05-based) chemistry parameterizations”.

p. 4, l. 15. The threshold temperature for NAT formation is pressure dependent. The manuscript indicates that 194 K was chosen as the threshold regardless of pressure. That would not be the correct way to calculate it.

While the BASCOE CTM was used some time ago for detailed studies of the processes leading to polar ozone depletion (Daerden et al., ACP, 2007), the corresponding microphysical module was removed (due to huge computational costs) and replaced by this very crude parameterization. Indeed the BASCOE CTM is now designed as a generic model which (until now) needs only to be good enough to allow the successful assimilation of satellite observations of stratospheric composition. Yet both reviewers indicated a simple improvement which could be implemented quickly enough for this revision of the manuscript.

Hence we have revised the PSC-parameterization, which is no longer purely temperature-dependent. We now remove H₂O and HNO₃ where their respective partial pressures exceed the equilibrium values, according to Murphy and Koop (2005), and Hanson and Mauersberger (1988). The time scale for irreversible removal of HNO₃ has been revised from 100 days in the original setup to 20 days, in accordance with the smaller regional and temporal extent where NAT particles are assumed to exist. This led to significant improvements in the H₂O and HNO₃ bias in the region where PSC formation is possible, and accordingly to a slight improvement in O₃ profile shapes in terms of a reduced positive bias at 100 hPa and reduced negative bias at 20hPa during August-September over the Neumayer and Syowa stations (see also below). Nevertheless, the HNO₃ timeseries for the BASCOE-CTM, CIFS-TS and C-IFS-Atmos models suggest that denitrification proceeds more slowly and ends one month later than observed by Aura MLS observations, which may be attributed to our crude modelling approach for the formation and sedimentation impact of NAT PSC.. We have modified the respective section as follows:

“Ice PSCs are presumed to exist at any grid point in the winter/spring polar regions where **water vapour partial pressure exceeds the vapour pressure of water ice (Murphy and Koop, 2005)**. Nitric Acid Trihydrate (NAT) PSCs are assumed **when the nitric acid (HNO₃) partial pressure exceeds the vapour pressure of condensed HNO₃ at the surface of NAT PSC particles (Hanson and Mauersberger, 1988)**.”

p. 5, l. 31. I don't understand what is meant by O1D and O3P being described implicitly, as opposed to being treated explicitly.

Within the troposphere the O¹D is produced from O₃ photolysis and assumed to react instantaneously, with only reaction products H₂O and again O₃. As the O¹D (and O) lifetime is much shorter than the integration time, while only reactions with N₂ and O₂ are assumed in the troposphere, the O¹D concentration can be considered in equilibrium over the integration time and hence does not need to be treated explicitly. The same argumentation holds for O³P, produced from O₂ photolysis in upper troposphere, and assumed to only react with O₃ to form O₂, and with O₂ to form O₃. This is different for

the stratosphere, where O¹D and O³P are involved in many more reactions. To clarify in the manuscript we reformulate this as follows:

“It is worth mentioning that the constituents O¹D and O³P, produced from O₃ and O₂ photolysis, are not explicitly computed in the troposphere, as O¹D and O³P are assumed to react with O₂, O₃ and N₂ only. This is different for the stratosphere, where O¹D and O³P are involved in many reactions.”

p. 6, l. 24. ‘solar radiation reaches the stratosphere earlier than the surface. . .’ as written this sounds like it is referring to delay caused by the speed of light! I doubt this was intended; it needs better wording.

The reviewer is clearly technically correct. We changed the formulation to a more compact formulation, leaving out the suggestion of a different timing:

“Also the presence of sunlight at solar zenith angles (SZA) larger than 90° at high altitudes needs to be accounted for in the stratosphere due to the Earth’s curvature, but may be neglected in the troposphere. This plays a role in the timing of springtime ozone depletion in the polar lower stratosphere. “

p. 7, Section 2.3.1. For JO3, the lack of a ‘jump’ in O3 may be because photolysis is unimportant (slow) near 100 hPa, so O3 is probably long-lived relative to the photochemical lifetime. JNO2 is much larger so I’m not sure why there isn’t a jump – can you explain this? It would be useful if you showed the simulated O3 and NO2 profiles in Fig. 1 to demonstrate the lack of a jump. What is the meaning of ‘JO3_TB’ in the title of one plot? No similar title for the other plot.

The reviewer is correct in that the presence or absence of jumps associated to the change in the reaction mechanism depends on the lifetime of the species, in combination with the magnitude of the change in the dominating reaction (or photolysis) rate with the different chemical mechanism. For O₃ the photolysis is a dominating loss term in this altitude range, but still the reaction rate is sufficiently low (i.e. the O₃ lifetime sufficiently long) such that jumps in the photolysis rate do not lead to jumps in O₃ concentrations. For NO₂ the photolysis rate is much larger, and resulting in a short (less than 1 hour) NO₂ lifetime. Jumps in photolysis rate potentially result in jumps in NO and NO₂ concentrations. Nevertheless, the jump is sufficiently small (for J-NO₂ we verified that the difference in photolysis rates around the tropopause is generally below 5%), such that the NO₂ concentrations do not show a significant jump. We now provide figures in the supplementary material where we present instantaneous profiles of a range of trace gases at the tropopause interface. We extended the discussion on this aspect with the following sentences:

“Even though such jumps are undesirable, no visible impact on local chemical composition was found, for any of the trace gases involved in both tropospheric and stratospheric chemistry, see also Figures S1-S3 in the Supplementary Material. This can be explained by the sufficiently small difference in the photolysis rates at the merging altitude of the photolysis and chemistry schemes, combined with the sufficiently long lifetime of the affected trace gases. ”

Section 2.3.2, l. 8. It’s unclear whether you’re saying NO, NO2, and have the mass fixed applied or whether they are the few species where the mass fixed isn’t applied. How badly is H2O not conserved in the stratosphere? This will conceivably cause problems for stratospheric chemistry. It would be useful to see a 1-year time series of the H2O mass above 100 hPa.

As explained in the manuscript the reason for switching off the mass fixer for the stratospheric H₂O tracer is because otherwise mass conservation errors originating from the troposphere lead to spurious redistribution of H₂O mass towards the stratosphere. Therefore, in fact due to switching off the mass fixer, the H₂O mass in the stratosphere remains very stable. We illustrate this by Figure R1 (below), which shows indeed absence of any trend in stratospheric H₂O columns over the years, indicating that H₂O mass conservation is sufficiently well ensured in the stratosphere. This figure also shows that H₂O total columns are essentially identical in C-IFS-Atmos and C-IFS-TS.

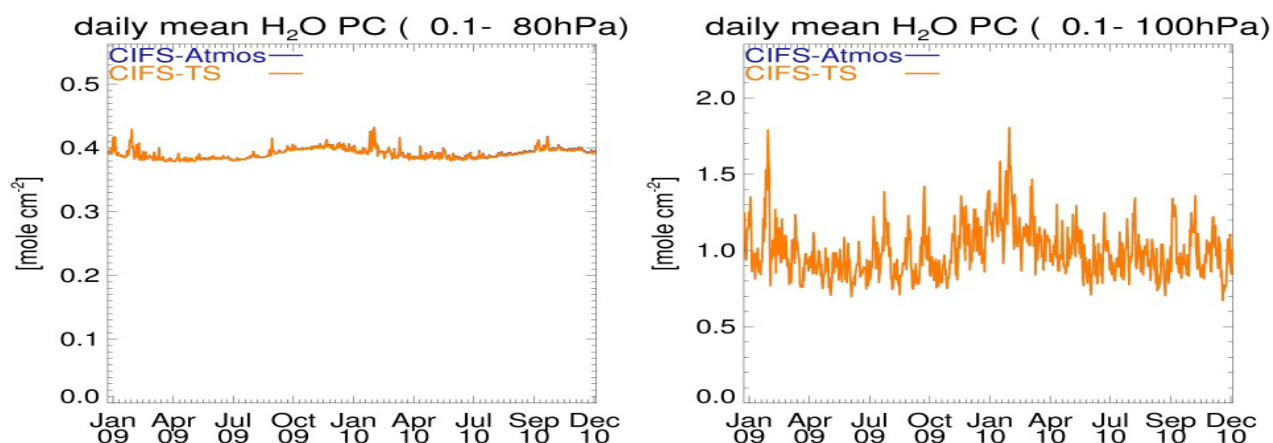


Figure R1. Evolution of global, daily mean H₂O partial columns (left: 0.1-80 hPa, right: 0.1 – 100 hPa) for the runs C-IFS-Atmos (blue) and C-IFS-TS (orange) for January 2009 to December 2010. C-IFS-TS is on top of C-IFS-Atmos.

In the manuscript we now write:

“The global advection errors in H₂O that essentially originate from the tropospheric part because by far most H₂O mass is located in the troposphere and the spatial gradients are much more pronounced. This should not affect the stratospheric H₂O mass budget, therefore the global mass fixer for the stratospheric H₂O tracer has been switched off. This prevents spurious trends in stratospheric H₂O columns over the years (not shown), indicating that H₂O mass conservation is well ensured in the stratosphere.”

p. 7, last 3 lines. This sentence says you are looking to identify differences in transport schemes. This confuses the issue of evaluating the chemical mechanisms (and their implementation). This evaluation should be performed using the same dynamical fields with the same model. If the advection schemes are also different, then we cannot actually test the impact of chemical mechanisms alone. And does ‘parameterization’ in line 28 refer to the different chemical mechanisms?

For a discussion on the selection of the model setups evaluated in our manuscript we refer to our response to the reviewer’s first general comment. We now extend the evaluation with results from run C-IFS-T, to explicitly identify the impact of the newly implemented stratospheric chemistry within the C-IFS framework. Indeed, the BASCOE-CTM run uses identical chemistry to C-IFS-TS and is not introduced to assess the chemical mechanisms, but rather differences due to the transport scheme while using the

same dynamical fields. Here, the ‘parameterizations’ refer to the reaction mechanism, photolysis, heterogeneous chemistry and sedimentation, as we now explicitly write.

p. 8, l. 27, ‘first Science Satellite’?

Indeed this is the meaning of the abbreviation ‘SCISAT-1’.

p. 8, l. 30-31. Suggest to change to ‘. . .between 6-30 km agree to within 15% of independent . . .’ For all the figures that are line plots (starting with Figure 2), the blue and black lines are hard to distinguish. Please do something with the line thickness and colors to improve readability.

We changed this according to the reviewer’s suggestions, thank you. We have improved color-coding and general figure quality, which unfortunately also had seen some degradation in the stage of pdf-generation from the word-document.

Section 4, Model Evaluation p. 9, lines 14-19. This paragraph would benefit by a general statement of the purpose of this comparison. It appears the purpose is to show that the TS mechanism looks more like the observed total column O₃ than does the trop-only code (with linearized strat O₃). One would expect the TS O₃ to be better than the linearized O₃ of CB05, but there should also be a comparison with the stratonly code. Comparing with the O₃ results in Fig. 6, I think the strat chem O₃ columns would be lower than the TS mechanism. I guess they aren’t the same because the BASCOE-CTM has different transport. Again, not having the same transport in all the simulations really interferes with a useful comparison.

We now replace this figure with an evaluation of the partial columns (10-100hPa) against Aura MLS observations, to emphasize the performance in the stratosphere. We now also include results from C-IFS-Atmos and C-IFS-T, as well as from BASCOE-CTM to assess the impact of different chemistry approaches, and different transport scheme. The new evaluation shows more clearly the benefits and limitations of the new approach in C-IFS-TS, as compared to C-IFS-T (with linearized O₃), as well as differences with BASCOE-CTM (which contains stratospheric chemistry only and uses the same dynamical fields as C-IFS but with a different transport algorithm). We have moved the assessment of O₃ total columns against the Multi-Sensor Reanalysis to the Supplementary Material . This does not include results from BASCOE-CTM, considering it’s missing tropospheric contribution. The manuscript text has been revised accordingly.

Furthermore, in Sec. 3 we now include a few general statements to clarify the purpose of the various model evaluations:

“Intercomparison of the runs C-IFS-TS and C-IFS-Atmos aims to provide a justification of our approach to split the chemistry into two regions, while intercomparison of C-IFS-TS with C-IFS-T can be used to identify the changes to stratospheric composition modelling between full stratospheric chemistry and the baseline approach with the linear ozone scheme.”

p. 9, discussion of Figs. 3-4. I do not agree that there are meaningful, reduced biases in the TS version. The linearized O₃ chemistry of the trop mechanism gives different results from the TS version, but not

really worse. These figures show that TS is not an improvement over trop-only. I think the use of mPa for the O₃ bias (lower panels) is misleading and probably minimizes the appearance of the disagreement in the middle stratosphere.

By evaluating O₃ profiles in terms of partial pressure biases in the original manuscript we intended to focus on the contribution of each pressure region to the O₃ TC, with larger weights in the lower stratosphere. This is now assessed in detail in the revised Figure 2 that presents the evolution of the O₃ partial columns (10-100hPa). Hence in accordance with the reviewer request we now present results of O₃ profiles in units ppmv, indeed giving more focus to the altitude with maximum O₃ concentration, at around 10 hPa. Also we now average over all profiles in 2009 and 2010, to improve the statistics, and include results from C-IFS-Atmos. We agree with the reviewer that we have been too positive when describing the C-IFS-TS results as compared to C-IFS-T (with linear chemistry). We have rewritten this section to provide a more balanced discussion.

p. 9, discussion of Fig. 5. I cannot tell the difference between obs and CIFS-T lines in the figure. There is no line color/style for the observations in each panel's legend. The TS O₃ agrees with one of the black lines (obs or CIFS-T??) near and below 100 hPa – sometimes – but the TS O₃ consistently has poor agreement above 50 hPa. Why? Since the TS (red) line often does not agree with either black line – I see no basis for claiming good agreement. Additionally, Syowa is often near the vortex and has large daily variability. Were the simulated profiles used in this figure calculated from the same days of the month as the Syowa data?

Figure 5 in the manuscript has been regenerated based on the revised model simulations. Color-coding has been updated, and error-bars denoting the model and observation variability are now included. Note that all comparisons with observations, including fig. 5, use three-hourly model output which has been collocated in time and space to the observations. This is now explained in Section 3. Please also note that in our section describing Figure 5 we do not claim general good agreement, as the reviewer suggests, but rather point at regions and months where C-IFS-TS performs well, and others where it shows biases compared to observations.

The revised simulations have seen some improvement in terms of vertical profile shape during ozone hole conditions, see also Figure R2, below, for an assessment of the differences to the C-IFS simulations presented in the GMDD manuscript. This is due to the improved PSC parameterization, especially above 50hPa in August and September where PSCs were allowed to exist in the C-IFS-TS run for the original GMDD version. The remaining discrepancies could still be caused by the limitations of the revised PSC parameterization. We now write:

“For the 2009 Antarctic ozone hole season (Fig. 5) the C-IFS-TS and C-IFS-Atmos shows a positive bias at ~100 hPa for August and September, and negative bias at higher altitudes (50-10 hPa), where C-IFS-T shows a positive bias.”

Additionally we now provide a closer analysis of the performance during polar ozone depletion, by presenting time series of HNO₃, ClO and O₃ during the 2009 ozone depletion over Antarctica (the new Figure 6). This new figure clearly shows the abilities and limitations of the different versions of C-IFS to describe this event. Specifically we now show that denitrification, which is clearly not modelled in C-IFS-T, starts at the correct time in the models with stratospheric chemistry, although it appears to last about one month too long as compared to the observations. We note that in the original manuscript, where this parameterization depended only on T, the denitrification started one month too late.

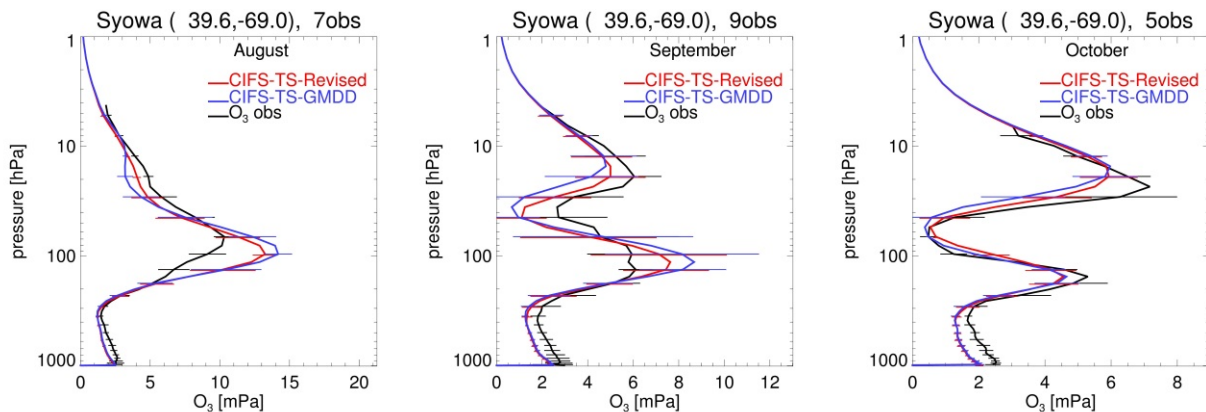


Figure R2. Evaluation of ozone in units mPa against WOUDC ozone sondes at Syowa station during August-October 2009. Black: ozone sonde, red: C-IFS-TS in the Revised model version, blue: C-IFS-TS in original GMDD version. Error bars denote the 1-sigma spread in the models and observations.

p. 9, lines 31-32. If you made a difference plot between MIPAS and the simulations, then you might be able to say whether there is good agreement. As presented, the conclusion can't be drawn that there are 'small biases'. Near the tropical maximum the TS looks slightly better than the BASCOE-CTM. Again, assuming that some of the differences are due to dynamical fields or advection scheme, this comparison isn't very useful.

As argued before, the inclusion of BASCOE-CTM is especially useful to diagnose if model biases arise due to different advection schemes or due to different chemistry schemes. To accommodate the reviewer's comments to better quantify the C-IFS versions, as compared to BASCOE-CTM we now include results from C-IFS-Atmos and C-IFS-T, and provide a more balanced discussion. Finally we have strengthened the evaluation of ozone with two new figures in the Supplementary Material: the quantitative comparison is strengthened by a new comparison of vertical profiles with Aura MLS (Fig. S7) and the discussion of Fig. 7 (top row) is confirmed with a corresponding evaluation also using Aura MLS (Fig. S8). The discussion of ozone on fig. 7 now reads:

"The evaluation of the zonal mean ozone mixing ratios against MIPAS observations shows good general agreement, Fig. 7, with all four modelling experiments providing similar features. The tropical maximum of O₃ mixing ratio at 10 hPa is under-estimated in all experiments but to a larger extent in those which model stratospheric photochemistry explicitly (BASCOE CTM, C-IFS-TS, C-IFS-Atmos) than in C-IFS-T, in line with the evaluation against O₃ sondes for June-July-August (figure 4). The same evaluation against MLS observations provides exactly the same conclusions (figure S8, supplementary material)."

p.10, lines 5-9. What is the message here? The CIFS has a terrible high bias in nighttime NO2 and a large low bias in HNO3. Why is the CIFS simulation worse than BASCOE? There is no clear explanation here.

We acknowledge that these results are not satisfying; unfortunately at current stage we do not have a clear explanation for this. Nevertheless, we want to highlight that the model performance has still improved compared to C-IFS-T, whose results we now include. Also we explicitly provide these figures to indicate current limitations of our model. We now write:

“Even though a clear improvement compared to run C-IFS-T is found, further investigation is necessary to diagnose origins of the biases in night-time NO₂ above 10 hPa and in HNO₃ between 10 and 70 hPa.”

Also in the conclusions section we include such a sentence.

p. 10, lines 10-20 (Fig. 8). N₂O and CH₄ profiles do NOT assess vertical transport. Their profiles below 10 hPa represent a balance between the vertical and horizontal components of the residual mean circulation. That balance depends on latitude, that is, whether the profile is from the tropical upwelling region or somewhere in the midlatitudes (horizontal and vertical motions matter and so does mixing), or isolated inside the polar vortex (descent). Above 10 hPa, profiles are more strongly influenced by chemical loss so the 2 simulations should look very similar there. The CIFS-TS simulation tends to look worse than the BASCOE CTM or the observations between 10-50 hPa. This suggests circulation and/or mixing problems in the tropics and SH.

O₃ at 20 hPa is strongly influence by chemistry, not just transport. These paragraphs indicate a lack of understanding of transport circulation and its diagnosis, as well as any understanding of what controls stratospheric ozone distributions.

These two paragraphs are indeed rather vague and mistakenly use Figure 8 as a diagnostic for “vertical” transport. We thank the reviewer for pointing this out and setting us on the right track.

The reviewer’s suggestion about circulation and/or mixing problems confirms that figure 8 is a preliminary yet valid diagnostic for transport processes in general, and that it indicates an unidentified issue for the representation of these processes in C-IFS. We do not think that circulation is the culprit because the BASCOE CTM is driven by meteorological fields which are the output of C-IFS. As noted above, the revised manuscript gives (at the beginning of section 3) a few more details about the modelling of transport in both models:

“The BASCOE-CTM is driven by temperature, pressure and wind fields simulated by the C-IFS runs. However, while BASCOE adopts a flux-form advection scheme (Lin and Rood, 1996) the IFS uses the Semi-Lagrangian scheme for advection, explicitly accounts for horizontal diffusion and includes a parameterization for convection (ECMWF, 2015).”

The revised manuscript also lists the relevant transport processes in the stratosphere along with a general reference on this topic (Shepherd, 2007) and states what specific pieces of C-IFS may be responsible for the problem(s). Since O₃ at 20 hPa is strongly influenced by both chemistry and transport, we stand with the statement that this transport issue “could also contribute” (directly) to the ozone biases noted below 10 hPa. But their attribution to an “excess of vertical transport” was clearly a mistake. No further statement can be made on this topic because further evaluation of stratospheric transport processes in C-IFS is beyond the scope of this paper. The two problematic paragraphs in section 4 have thus been re-written as follows:

“Fig. 9 shows an evaluation of N₂O and CH₄ profiles during September 2009 against observations by ACE-FTS. Owing to their long lifetimes these trace gases are good markers for the model ability to describe transport processes - i.e. not only the Brewer-Dobson circulation but also isentropic mixing, mixing barriers, descent in the polar vortex, and stratosphere-troposphere exchange (Shepherd, 2007). Moreover, N₂O is the main source of reactive nitrogen in the stratosphere while CH₄ is one of the main precursors for stratospheric water vapour. The figure suggests reasonable profile shapes for both CH₄ and N₂O in the upper stratosphere (10 hPa and higher) where their abundance is more strongly influenced by chemical loss but at lower altitudes (100-10 hPa) C-IFS-TS and C-IFS-Atmos show larger discrepancies to the observations than the BASCOE-CTM run, with weaker vertical gradients in the tropics and SH-mid latitudes and a sharper gradient in the extra-tropical Northern Hemisphere.

This discrepancy cannot be due to different wind fields because the BASCOE CTM experiment is driven by three-hourly output of the C-IFS-T experiment. We attribute it instead to the different numerical schemes for advection and/or to differences in the representation of sub-grid transport processes in the GCM and in the CTM. Convection and diffusion are indeed explicitly modelled in C-IFS but neglected in BASCOE CTM, which relies on the implicit diffusion properties of its flux-form advection scheme to represent sub-grid mixing (Lin and Rood, 1996; Jablonowski and Williamson, 2011). Since lower stratospheric ozone is strongly determined by both chemistry and transport, the transport issue indicated by fig. 9 could also contribute directly to the ozone biases seen below 10 hPa in Figures 3 and 4.”

The revised conclusions mention that this area deserves further investigation:

“For the long-lived tracers CH₄ and N₂O, larger errors with respect to limb-sounding retrievals were found between 10 hPa and 100 hPa than with the BASCOE-CTM, suggesting difficulties in representing slow transport processes. The BASCOE CTM experiment shown here was driven by three-hourly wind fields output of the C-IFS-T experiment. Hence this discrepancy is due to a difference in the representation of the transport processes between the GCM and the CTM, i.e. the numerical scheme used for advection (Semi-Lagrangian versus Flux-Form), the convection (parameterized in C-IFS but neglected in BASCOE CTM) or the diffusion (parameterized in C-IFS but not explicitly considered in the CTM). Hence, stratospheric transport in C-IFS will be an area for further evaluation and developments.”

Additional references

de Grandpré, J., Tanguay, M., Qaddouri, A., Zerroukat, M., and McLinden, C.A. (2016), Semi-Lagrangian Advection of Stratospheric Ozone on a Yin–Yang Grid System. *Monthly Weather Review* 144:3, 1035-1050. DOI: <http://dx.doi.org/10.1175/MWR-D-15-0142.1>

ECMWF: IFS documentation –Cy41r1. Operational implementation 12 May 2015. Available from <http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation> (last access on 5 July 2016), 2015.

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, *J. Geophys. Res.*, 113, doi:10.1029/2007JD008771, 2008a.

Hanson, D., and Mauersberger, K.: Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere, *GRL* 15 (8), pp 855–858, DOI: 10.1029/GL015i008p00855, 1988.

Hitchman, M. H., M. McKay, and C. R. Trepte (1994), A climatology of stratospheric aerosol, *J. Geophys. Res.*, 99(D10), 20689–20700, doi:10.1029/94JD01525.

Jablonowski, C., and Williamson, D. L.: The pros and cons of diffusion, filters and fixers in atmospheric general circulation models, in: *Numerical Techniques for Global Atmospheric Models* (pp. 381-493). Springer Berlin Heidelberg, doi:10.1007/978-3-642-11640-7, 2011.

Lin, S.-J. and Rood, R. B.: Multidimensional flux-form semi-Lagrangian transport schemes, *Mon. Weather Rev.*, 124, 2046–2070, 1996.

Murphy, D. M. and Koop, T. (2005), Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Q.J.R. Meteorol. Soc.*, 131: 1539–1565. doi: 10.1256/qj.04.94

Santee, M. L., A. Lambert, W. G. Read, N. J. Livesey, G. L. Manney, R. E. Cofeld, D. T. Cuddy, W. H. Daffer, B. J. Drouin, L. Froidevaux, R. A. Fuller, R. F. Jarnot, B. W. Knosp, V. S. Perun, W. V. Snyder, P. C. Stek, R. P. Thurstans, P. A. Wagner, J. W. Waters, B. Connor, J. Urban, D. Murtagh, P. Ricaud, B. Barret, A. Kleinboehl, J. Kuttippurath, H. Kuellmann, M. von Hobe, G. C. Toon, and R. A. Stachnik. Validation of the Aura Microwave Limb Sounder ClO measurements. *J. Geophys. Res.*, 113:D15S22, doi:10.1029/2007JD008762, 2008.

Shepherd, T. G.: Transport in the Middle Atmosphere, *J. Met. Soc. Jap.*, 85B, 165–191, doi: 10.2151/jmsj.85B.165, 2007.