

Replies to reviewer 2

for the manuscript submitted by Chabrillat et al. to EGU sphere
(doi:10.5194/egusphere-2025-1327)

We thank the reviewer for their constructive and helpful comments. Please find below the comments in **bold**, our replies in *italic*, and the manuscript modifications in **red**.

General comments

This manuscript presents a thorough evaluation of the stratospheric chemistry aspect of the IFS-COMPO Cy49R1 system. These updates will help the Copernicus Atmosphere Monitoring Service (CAMS) move towards a better stratosphere-troposphere chemistry scheme. The authors detail recent improvements, specifically focusing on advances in the heterogeneous chemistry part of the stratospheric chemistry. They attempt to evaluate the updated scheme using the Mt. Pinatubo-related perturbation as a test case.

These enhancements and detailed documentation will be highly beneficial to the large scientific community that utilizes CAMS data and will serve as a crucial guide for future model development. The manuscript's primary strength lies in the documentation of the updated scheme with some evaluation, rather than new scientific discoveries. This perfectly aligns with the scope and aims of Geoscientific Model Development. The authors have effectively detailed the upgrades and assessed their impact on stratospheric composition, making this a very suitable contribution to the journal. Overall, I recommend the manuscript for publication with minor modifications.

We thank the reviewer for these positive comments.

Comments

1. Decluttering Figures:

Many figures in the manuscript are currently difficult to read due to an excessive amount of text. To improve clarity and readability, I strongly recommend that the authors move explanatory text, detailed descriptions, and other information from the figures into the figure captions or the main body of the text. This will allow for larger tick labels and legends, making the figures much easier to interpret at a glance.

We guess that this comment addresses the figures in section 6 (time series in the polar lower stratosphere) because their legends were too long, and the figure in section 7 because it was formatted as a scorecard. The figures in section 6 have been re-plotted with smaller legends, and different proportions to enlarge the tick labels of the axes. The scorecard figure was simplified, and all explanatory text was moved either to the caption or to the main text.

2. Climatology:

I failed to understand the mean biases of modeled trace gases in a climatological context. Please consider adding an additional figure that compares key long-lived tracers/reservoir species (e.g., O₃, HNO₃, H₂O, N₂O, HCl, CFC-11, CFC-12) against well-established, publicly available observation-based datasets. There are many selective line plots but it is very difficult to understand basic strength and weakness in the chemistry scheme.

I suggest showing 5 to 10 mean latitude/height or pressure cross-sections (e.g., see Chapter 6 in the CCMVal 2010 report; tracer-tracer correlation plots in that report also provide a better guideline to understand issues with transport/chemistry). This approach would allow for a clearer assessment of mean biases across different regions of the stratosphere. For example, a 2x5 type figure could be used, with the first column displaying modeled climatological values from the lower stratosphere to the upper stratosphere, and the second column showing absolute or percent differences with respect to SPARC Data Initiative products (e.g., Hegglin et al., 2021). Alternatively, the authors could utilize TCOM (TOMCAT CTM and Occultation Measurement based data), as described in Dhomse et al., 2023 (data for all these species is available on Zenodo). While the current manuscript includes comparisons with daily satellite retrievals (ACE-FTS, Aura-MLS) and reanalyses (BRAM3), an additional comparison against observation-based climatologies would help readers understand mean biases in the updated chemistry scheme.

This comment makes a lot of sense for the evaluation of a Chemistry-Climate Model. Unfortunately, this approach is not feasible for IFS-COMPO for technical reasons. This limitation constrains our whole evaluation approach and should have been explained in detail. This is addressed in the revised manuscript, primarily by the following paragraph added to the Introduction:

In chemical forecast mode, IFS-COMPO operates similarly to a chemistry–climate model with meteorological fields nudged toward climate reanalysis (Ménard et al., 2020; Davis et al., 2022). Climate model evaluation primarily relies on mean bias estimates from comparisons between decadal simulations and observational climatologies (e.g., Froidevaux et al., 2019). Such long runs are impractical for IFS-COMPO because it is derived from an NWP system and designed for much higher resolutions than climate models: even at low resolution, a decadal simulation would take at least six months. Therefore, our evaluation focuses on the “chemical weather” context, where simulations of roughly one year are compared either with individual observations prior to the derivation of the statistics, or with a reanalysis of the observations over the same period. The next CAMS reanalysis of atmospheric composition, named EAC5, will span more than two decades and be accompanied with a control run with no chemical data assimilation (Flemming et al., 2025). These future large experiments will allow climatological evaluations of IFS-COMPO.

Section 4 (set-up of the modelling experiments) now starts with the following sentences:

As explained in section 1, IFS-COMPO Cy49 is too computationally expensive to run the long simulation required for climatological evaluation. We focus instead on three separate case studies with IFS-COMPO experiments lasting 6 to 24 months and summarized in Table 2. These case studies have been chosen to exemplify the performance of Cy49...

The lower resolution used by the Pinatubo experiment is better explained as it is also due to computing costs:

Lasting more than two years, this experiment would be too costly to run at the T511 resolution currently used for CAMS. Hence the T255 resolution, corresponding to an 80 km grid cell, was selected instead.