

Note: Reviewer's comments are presented in black font; authors' responses are presented in blue plain font; manuscript text quotations are presented in blue italics font.

Anonymous Referee #2

We would like to thank Reviewer #2 for his/her time devoted and the constructive and helpful comments.

Akritidis et al present a study which assesses a stratosphere-to-troposphere transport event (STT) that occurred over Europe during the cold winter of 2017. STT is a very important source of ozone into the troposphere but has and remains challenging to simulate given the laminar like structures that are associated with these events and their transient in time nature. Arkitidis use results from a range of models including the ECMWF Copernicus Atmosphere Monitoring Service to understand the drivers for this event and to use the observations of this event from aircraft and ozonesondes to evaluate the models. In general this is a well written manuscript and one I would recommend published after the following general and technical points are considered.

We thank the Reviewer for the comments, to which we will respond point by point.

General comments:

1. A table with the model acronyms and set ups used in the analysis would be very useful for the reader. Most of the information is already in the text but I feel a table would help the reader quickly appreciate the differences between RegEns, CAMS and IFS.

We agree with the reviewer's suggestion. The CAMS models and simulations used in the current study are now presented and described in Table 1 which is included in the revised manuscript (page 19).

Table 1. CAMS models and simulations used in the present study.

| CAMS Models and Simulations | Description |
|---------------------------------|--|
| CAMS | Copernicus Atmosphere Monitoring Service |
| IFS (CAMS global) | ECMWF Integrated Forecasting System |
| RegEns (CAMS regional ensemble) | Median ensemble of the seven CAMS regional model forecasts |
| IFS Forecast Day-1 | IFS forecast one day in advance |
| IFS Forecast Day-2 | IFS forecast two days in advance |
| IFS Forecast Day-3 | IFS forecast three days in advance |
| IFS Forecast Day-4 | IFS forecast four days in advance |
| IFS Forecast Day-5 | IFS forecast five days in advance |
| IFS no DA Forecast Day-1 | IFS forecast one day in advance without the use of data assimilation |
| RegEns Forecast Day-1 | Regional Ensemble forecast one day in advance |

The following phrase is also included in the revised manuscript, page 4, lines 30-31: *"Table 1 presents the CAMS models and simulations used in the present study."*

2. What more can we learn from this event? The RegEns models all differ in structure and I wonder what further analysis could be done to help understand (a) the role of resolution in the vertical (b) the role of horizontal resolution in biases that occur in the models. The use of the IFS with and without DA is instructive but I feel there is more to be teased out from RegEns and would like to see some more effort to that extent.

We thank the Reviewer for the comment. The role of resolution in vertical would be an interesting sensitivity study, which could provide a concrete answer to the question of how much can we improve a forecast, when we improve the vertical model resolution. If this was an in-house model, we could perform a series of sensitivity tests to investigate the hypothesis. Within CAMS84, however, we have access only to the operational forecast, and our task is to evaluate, as is. The same applies for the regional model forecasts.

In the same way, the impact of horizontal resolution on model performance would be another valuable exercise, which is usually performed with stand-alone models, however, is impossible to perform with operational datasets, with zero flexibility in the model data streams. As evaluation team within CAMS84, we cannot decide on the type of evaluation experiments there should be performed with the CAMS models. We should also note that despite the different regional model set-ups, we have access only to the post-processed datasets, which all have the same horizontal resolution as a result of re-gridding. Regional fields are only available in pre-defined vertical levels (0, 50, 250, 500, 1000, 2000, 3000 and 5000 m).

Technical comments:

Page 1, line 1: Technical point. I'm not sure I agree with the opening statement. STT tends to produce about 200 Tg (O₃)/yr. Lightning NO_x (natural) produces XXX..

We agree with Reviewer's comment. We have replaced the phrase *"is the dominant natural source of"* with the phrase *"is an important natural source of"* (page 1, line 1 in the revised manuscript).

Page 3, line 24: I find the phrase "weather-chemistry feedback" a bit puzzling.. I think you could be more specific here. What exact feedbacks are included and how are they represented?

More information on the included weather-chemistry feedbacks can be found in Inness et al., 2015, and references therein. In more detail Inness et al. state: *"It was therefore decided to implement the chemistry scheme and its solvers directly in the IFS, together with modules for photolysis, wet and dry deposition, as well as emission injection, to create a more efficient model system called the Composition-IFS (C-IFS, Flemming et al., 2015)"*. In the revised manuscript we have included the reference *"(Inness et al., 2015, and references therein)"* after the respective phrase (page 3, line 25).

Page 4, line 2: I guess when you mean data assimilation you mean chemical data

assimilation?

Yes. We have replaced the phrase "data assimilation" with the phrase "*chemical data assimilation*" (page 4, line 7 in the revised manuscript).

Page 4, line 8-9: Pernickety, I know, but you have used "seven" on line 8 and "7" on line 9 to refer to the number of models in the CAMS ensemble. Sticking with one or the other would be better.

Done. We have replaced "7" with "seven" (page 4, line 14 in the revised manuscript).

Page 6, line 10: Define FYROM please.

Done. We have included the definition "*(Former Yugoslav Republic of Macedonia)*" as suggested (page 6, lines 22-23 in the revised manuscript).

Page 6, line 13: Remove "the" before sea-level.

Done (page 6, line 25 in the revised manuscript).

Page 7, line 7: Can the authors confirm why pressure-interpolation of the chemical fields from model levels onto pressure levels could not be performed? I would think this is a fairly standard procedure. Could they elaborate on the errors introduced for example by not accounting for the exceptionally low real temperatures when using the US Standard Atmosphere for unit conversion? In addition, with respect to Figure 4, I would be intrigued to know what the spread is within RegEns or the standard deviation of the ensemble? In general, what can we learn more about the models from this event?

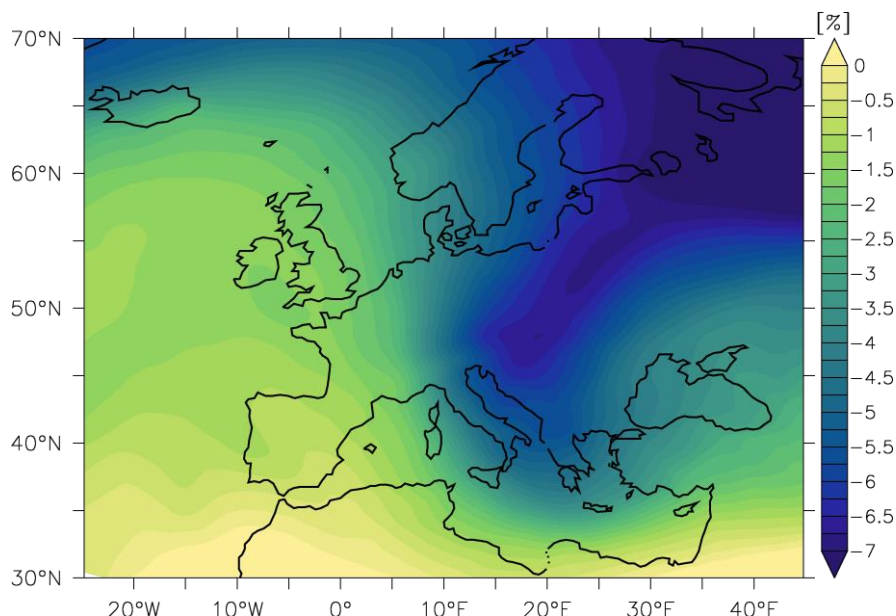
We agree with the Reviewer that the model-to-pressure levels interpolation is a standard procedure. Nevertheless, this study is performed within the framework of CAMS84 where the regional models are strictly provided in 8 height levels (up to 5 km), so actually data in model levels for the regional models are not available.

Regarding the error induced by the use of Standard Atmosphere temperature at 5000m (550.65K and 540.19 hPa) compared to the real temperatures during the period of interest, for a region of lon:10-30E and lat:32-58N and during the time span of Figure 4 there is a decrease of ozone concentrations of about 2.3 ppb when we consider the temperatures from the ERA-Interim dataset at 550 hPa. The spatial distribution of the percentage (%) differences between ozone concentrations calculated using the ERA-Interim temperatures and that of the Standard Atmosphere over the abovementioned time period are shown below. As expected the larger discrepancies in calculated ozone (up to 7%) are found over the regions exhibiting the lower temperatures. The respective text has been modified in the revised manuscript (page 7, lines 18-24) as follows:

"Although the spatio-temporal features of ozone in the RegEns agree well with that of the IFS, in quantitative terms there are discrepancies between the regional and the global product. This is likely due to the fact that (a) the RegEns is presented at 5000m level (the uppermost level available) and the IFS at 500 hPa, (b) different resolution and advection schemes are used in global and regional models and (c) pressure and temperature values from US Standard Atmosphere (USAF, 1976) were used for units conversion in RegEns. Considering the ERA-Interim (Dee et al., 2011) temperatures during the period of interest for the units conversion may result in even

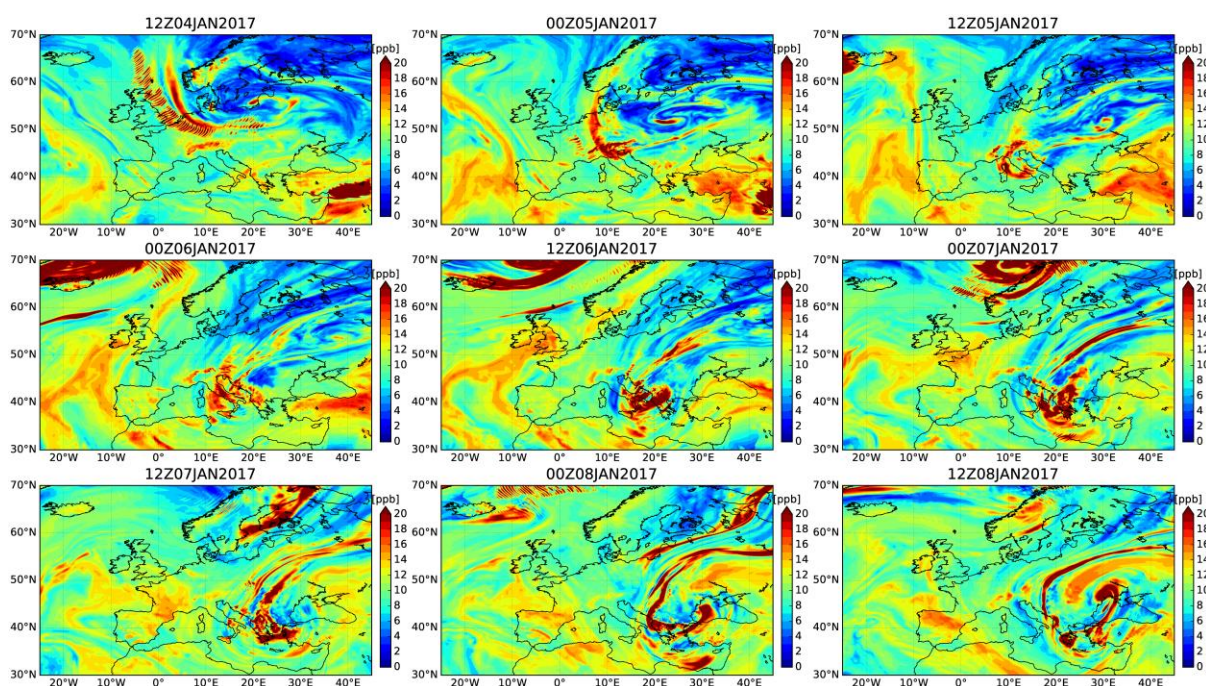
lower RegEnS ozone concentrations of up to ~7% in the regions exhibiting the lower temperatures (not shown).“

The reference Dee et al., 2011 is also included in the revised manuscript (page 14, lines 20-22).



Percentage (%) differences of ozone concentrations calculated using the ERA-Interim temperatures for the time period 12Z04JAN2017-12Z08JAN2017 and Standard Atmosphere temperatures at 5000m.

As concerns the spread among the regional models, the following figure shows the standard deviation of the regional ensemble ozone concentrations at 5000m for the same dates as that of Figure 4.



Standard deviation of the regional ensemble ozone concentrations at 5000m for the time period 12Z04JAN2017-12Z08JAN2017.

The “take-home message” from this study about the regional models is that the global forcing from the IFS results in capturing an extreme STT event of several stratospheric intrusions over Europe, consistent with the global model.

Page 8, line 16: I'm not sure what you mean by “mind the angle of view”?

We agree with the Reviewer that this is a bit confusing and thus has been removed in the revised manuscript. Our intent was to emphasize the orientation of the 3D structure.

Page 9, line 14: Can you speculate why the spread in RegEns increases in the vertical?

We can assume that the different dynamical cores, the schemes regulating vertical transport, and the methodology to relax top boundary concentrations could lead to the increased spread between the regional ensemble members.

Page 10, line 1: What constitutes “satisfactory”?

We understand that “satisfactory” gives a more qualitative interpretation. Still, a satisfactory forecast performance stands for a model forecast that captures the observed ozone variations in the vertical with a relatively small bias ($FGE < 0.3$).

Page 10, line 22: Insert “the” before CAMS.

Done (page 11, line 12 in revised manuscript).

Figure 2 caption, missing information about panels e and f.

We thank the Reviewer for the comment. The missing information are now included in Figure 2 caption in the revised manuscript (page 21).

*** Page 17, line 8 in the revised manuscript, “*LOTOS?EUROS*” is replaced with “*LOTOS-EUROS*”.

*** In the revised manuscript for the RegEns ozone vertical profiles the altitude of the sites was considered (if needed) for choosing Standard Atmosphere pressure and temperature.

*** In the revised manuscript in Figures 11a and c the RegEns ozone is plotted for time 12Z (in order to be consistent with IFS) instead of 13Z that was by mistake in the initial manuscript.