Author Responses to Reviewer Comments

We thank the reviewers and the Editor for their useful comments and feedback. We have reproduced the comments of Reviewer #2 and the Editor below in black text, followed by our responses in red text. Any additions to the manuscript are in blue text and our reference to line numbers is based on the revised submitted manuscript.

Reviewer #2's Comments:

I am still not convinced that a coarse resolution CTM is the right tool for this specific analysis (as explained in detail in the reviewer's comments). The author's responses to my questions/comments with regards to this, only stated all the reasons why a regional model would not be a suitable tool, but that does not answers my concerns. The author additionally pointed out 3 papers where the CTM in question was used for regional scale analysis. However, in the papers highlighted, the CTM is used to analyse ozone over broad regions (i.e. the Mediterranean, the North Atlantic and Eastern Amazon) and the model comparison were mainly to satellite data which samples ozone in the troposphere (as opposed to EMEP surface stations). Given the very short lifetime of ozone at the surface and the importance of model resolution to properly represent ozone precursor's emissions and deposition (which are important near the surface and near emission regions) I believe such a coarse resolution CTM (both horizontally and vertically) is not the right tool for this analysis. I would be happy to reconsider this manuscript if it can be revised to replace/remove the use of the CTM and focus more on the observations.

The reviewer is suggesting that we remove the modelling work from the paper, however, while the global model is not perfect, it acts an adequate and useful tool to help investigate the processes behind the signals seen in the observations. As outlined in our first response, a regional model would be a very useful tool for our study but there are many limitations (e.g. run time, lack of information outside the domain spatially and vertically). Therefore, the TOMCAT model provides more scientific benefits than it does limitations due to its coarser spatial resolution.

Reviewer #2 also states "*The author additionally pointed out 3 papers where the CTM in question was used for regional scale analysis. However, in the papers highlighted, the CTM is used to analyse ozone over broad regions (i.e. the Mediterranean, the North Atlantic and Eastern Amazon) and the model comparison were mainly to satellite data which samples ozone in the troposphere (as opposed to EMEP surface stations).*", which is not entirely true given that all three studies have used surface observations (i.e. Richards et., 2013; Monks et al., 2017; Pope et al., 2020) to evaluate the model surface ozone. The work by Richards et al., (2013) focussed on surface ozone in the Mediterranean (similar region to this work) and compared with 4 EMEP sites. Overall, the model was able to capture the seasonal cycle of the observations but did have a low bias ranging between approximately 3-24% depending on year (see Figure 1 of their manuscript). This low bias was predominantly in the winter, which still exists to a degree in our study, but our evaluation in this work actually shows better agreement in general given the model improvements since that study (e.g. see Monks et al., (2017)). Overall, TOMCAT has been used in multiple scientific studies, evaluated using surface observations and been a suitable tool to investigate regional surface ozone. Therefore, we are confident in the model's suitability for this work and should be kept in our study.

However, the Editor has made a useful suggestion to compare TOMCAT with other but higher resolution modelling data sets (i.e. CAMS) to evaluate the overall skill and suitability of TOMCAT in this study. Thus, please see our response to the Editor's comments below where we present the suitable skill of TOMCAT to simulate surface ozone.

Editor's Comments:

Anonymous referee #2 still has some serious concerns regarding the suitability of your modelling approach. A key issue here is the scale mismatch between the course grid resolution of your model and the spatial scale represented by the EMEP measurement sites as used in your study. This issue must be addressed before the paper can be accepted for publication. The referee has suggested that you remove the modelling analysis from the paper and focus on the observational analysis. I invite you to consider this option in a revised manuscript.

Clearly the referee is not convinced by the evaluation of your model against the EMEP network in your supplementary material. If you can show that your model performs sufficiently well in comparison with higher-resolution state of the art models for Europe, this may potentially help your manuscript. You could consider examining the data from the evaluation of the CAMS global reanalysis product (https://doi.org/10.1525/elementa.2020.00171) and the data available from the CAMS regional air quality products (https://atmosphere.copernicus.eu/regional-services). The opinion of the anonymous referee on your revised manuscript will still be considered for my final decision.

Firstly, we thank the Editor for the useful and constructive suggestion entailing the evaluation of TOMCAT against higher-resolution model out from the CAMS global reanalysis product and the CAMS regional reanalysis product. Here, we have focussed on May-August 2018 for the evaluation for several reasons: 1) the summer of 2018 is the key focus of this paper and 2) there was no regional CAMS data for 2017 (i.e. 2018 was the earliest year available). For the regional product, we have used the ensemble version of the data (i.e. the ensemble mean from the range of regional models used in CAMS).

The response Figure R1 below shows spatial maps of the TOMCAT and CAMS surface ozone (ppbv) averaged over Europe between 1st May and 31st August 2018. Overall, there is good spatial agreement between TOMCAT and the CAMS data. The spatial gradients in surface ozone are gradual and largely homogenous across most of the domain for all the model data sets. There is peak surface ozone in the Mediterranean (~50-60 ppbv), moderate ozone levels over continental Europe (~30-50 ppbv) and lower ozone over Scandinavia (~25-35 ppbv). Here, the largest discrepancy between TOMCAT and the other data sets occurs where TOMCAT minimum ozone is ~25 ppbv, while 30-35 ppbv for the CAMS data. Interestingly, there is no consistent offset between TOMCAT and the CAMS datasets. For instance, there is better agreement (spatially and in absolute terms) between CAMS global and TOMCAT over the Mediterranean than CAMS regional, while better agreement between TOMCAT and CAMS regional over central Europe in comparison to CAMS global.

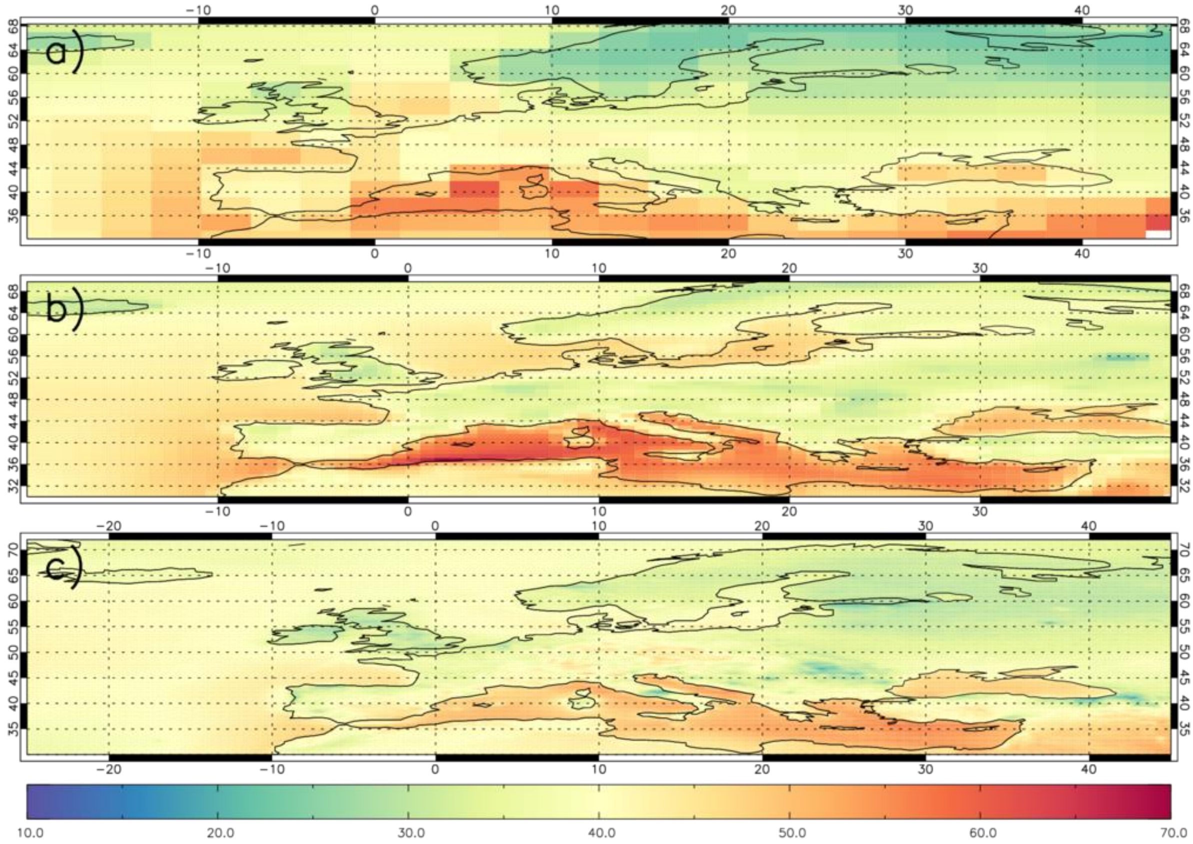

Figure R1: Surface ozone (ppbv) averaged between 1st May and 31st August 2018 for a) TOMCAT, b) CAMS global reanalyses and c) CAMS regional reanalyses on their original spatial resolutions.

When the CAMS data sets are interpolated onto the TOMCAT spatial resolution for statistical comparison (Figure R2), TOMCAT successfully reproduces the spatial pattern of CAMS surface ozone with good correlations (R) of 0.66 vs. CAMS global and 0.76 vs. CAMS regional. The percentage root-mean-square-error (RMSE%) between CAMS global and CAMS regional with TOMCAT is 16.8% and 14.5%, respectively. For reference, the difference between CAMS global and CAMS regional is 12.5% and the correlation is 0.77. Thus, the spatial domain metrics between TOMCAT and the CAMS data sets are consistent with that of CAMS global vs. CAMS regional.

Figure R2d shows the daily domain average time-series of the three data sets. As can be seen, TOMCAT typically simulates summer surface ozone values between that of CAMS global and CAMS regional (apart from the first half of May where it has a low bias of 3.0-5.0 ppbv). The median and 25th-75th percentile values of the time-series is also shown. The median values are very similar 38.1, 39.4 and 37.9 ppbv for TOMCAT, CAMS global and CAMS regional, respectively. The 25th-75th percentiles have similar surface ozone values, but the TOMCAT inter-quartile range (IQR) is slightly smaller (1.6 ppbv) than CAMS global (3.2 ppbv) and CAMS regional (5.1 ppbv). The time-series R (RMSE%) values between TOMCAT and CAMS global and CAMS regional are 0.76 (6.4%) and 0.75 (4.7%), respectively. The inter-CAMS R and RMSE% are 0.9 and 6.8%. Therefore, TOMCAT successfully reproduces the summer 2018 surface ozone values over Europe in comparison to higher resolution model datasets, which have assimilated observations. Thus, we are confidence in our TOMCAT simulations and hope the Editor agrees with our position that it is a suitable modelling framework for this work.

To make this clearer in the manuscript, we have added the following statement on Page 5 Line 174:

“TOMCAT surface ozone was also compared with higher resolution modelling (reanalysis) data from the Copernicus Atmosphere Monitoring Service (CAMS), which showed good spatial agreement between the modelling data sets and in the simulated surface ozone absolute values during the European summer 2018 pollution episode (**SM 5**).”.

In the supporting material (SM) section 5 (SM 5), we have added Figure R2 as a new Figure S9 and added the following text on Page 4 Line 143 of the SM:

“Given the relatively coarse resolution of TOMCAT, we compared the TOMCAT summer (1st May – 31st August 2018) simulation with higher resolution model data of surface ozone from the Copernicus Atmosphere Monitoring Service (CAMS). This involved two CAMS products including the global reanalysis product (available from https://ads.atmosphere.copernicus.eu/ and described by Wagner et al., (2021)) and the regional CAMS product for Europe (available from https://atmosphere.copernicus.eu/regional-services). Note, we used the ensemble reanalysis product which was an ensemble average of all the regional models involved in the CAMS regional programme. Here, CAMS global and CAMS regional had original spatial resolutions of 0.75°×0.75° and 0.1°×0.1° and were interpolated onto the TOMCAT spatial resolution for statistical comparison.

Figure S9a-c shows the summer-time average spatial surface ozone distribution over Europe for TOMCAT and both CAMS data sets. For all three models, there is peak surface ozone in the Mediterranean (~50.0-60.0 ppbv), moderate ozone levels over continental Europe (~30.0-50.0 ppbv) and lower ozone over Scandinavia (~25.0-35.0 ppbv). Here, the largest discrepancy occurs where TOMCAT has minimum surface ozone values of 25.0 ppbv while 30.0-35.0 ppbv for CAMS. However, there is generally no consistent systematic difference between TOMCAT and the CAMS datasets. For instance, there is better agreement (spatially and in absolute terms) between CAMS global and TOMCAT over the Mediterranean than CAMS regional, while the opposite is true for CAMS regional and TOMCAT in central Europe. TOMCAT surface ozone has a good spatial correlation with CAMS global (R=0.66) and CAMS regional (R=0.76) surface ozone. The percentage root-mean-square-error (RMSE%) between CAMS global and CAMS regional with TOMCAT is 16.8% and 14.5%, respectively. Comparisons of CAMS global and CAMS regional yield metrics of R=0.77 and RMSE%=12.5%. Therefore, the spatial domain metrics between all three data sets are consistent suggesting TOMCAT surface ozone is consistent with that of these higher resolution models.

Figure S9d shows the daily domain average time-series of the three data sets. TOMCAT typically simulates summer-time surface ozone values between that of CAMS global and CAMS regional (apart from the first half of May where it has a low bias of 3.0-5.0 ppbv). The median and 25th-75th percentile values of the time-series are also shown. The median values are very similar at 38.1, 39.4 and 37.9 ppbv for TOMCAT, CAMS global and CAMS regional, respectively. The 25th-75th percentiles have similar surface ozone values, but the TOMCAT inter-quartile range (IQR) is slightly smaller (1.6 ppbv) than CAMS global (3.2 ppbv) and CAMS regional (5.1 ppbv). The time-series R (RMSE%) values between TOMCAT and CAMS global and CAMS regional are 0.76 (6.4%) and 0.75 (4.7%), respectively. The inter-CAMS R and RMSE% are 0.9 and 6.8%. Therefore, TOMCAT successfully reproduces the summer 2018 surface ozone values over Europe in comparison to higher resolution model datasets, which have assimilated observations. Thus, providing further confidence in TOMCAT simulated ozone.”.

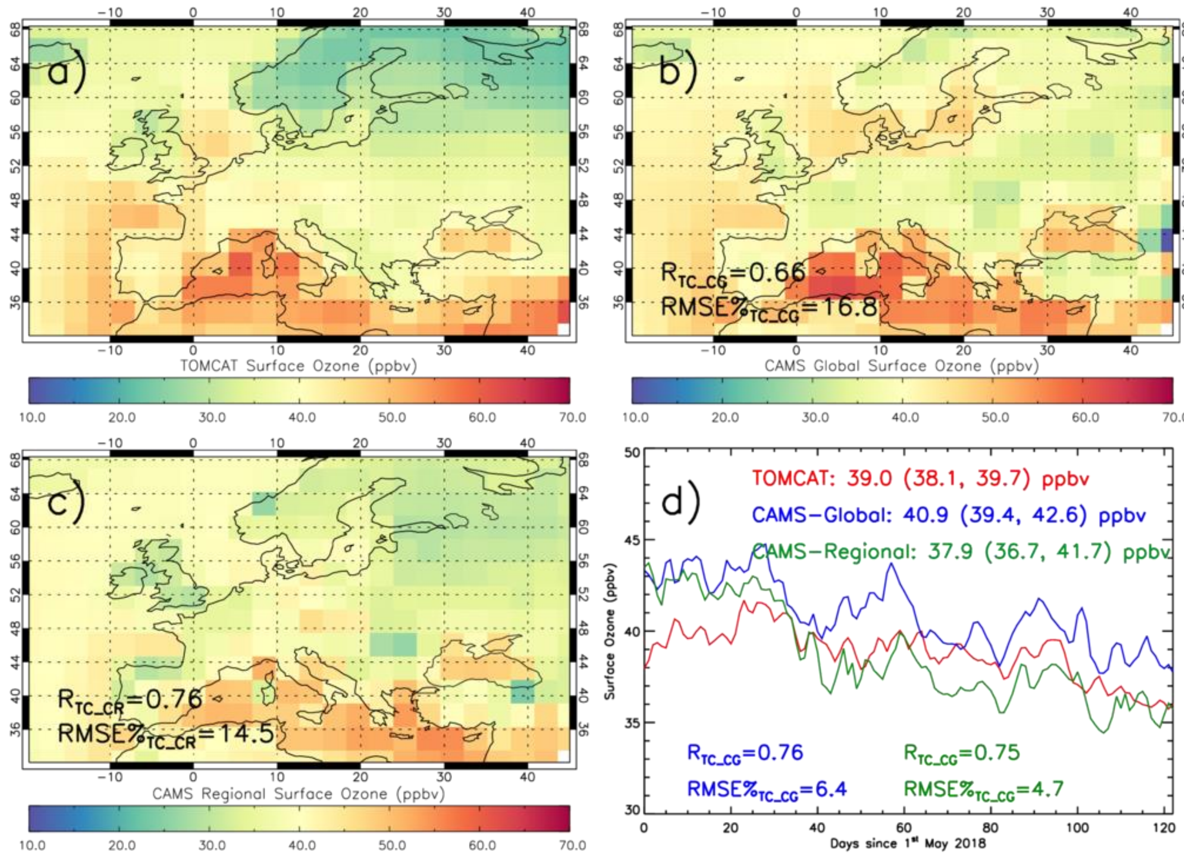

Figure R2: Surface ozone (ppbv) averaged between 1st May and 31st August 2018 for a) TOMCAT, b) CAMS global reanalyses and c) CAMS regional reanalyses. The correlation (R) and percentage root-mean-square-error (RMSE%) between TOMCAT and CAMS global (CG) and CAMS regional (CR) reanalyses are shown in the bottom of panels b) and c). Panel d) shows the daily domain average surface ozone (ppbv) time-series between 1st May and 31st August 2018 for TOMCAT (red), CAMS global reanalyses (blue) and CAMS regional reanalyses (green). The number after the model labels are the time-series median (25th percentile, 75th percentile) values. The R and RMSE% metrics show the same information as the maps but between the TOMCAT and CAMS time-series (coloured accordingly).

We have added the Wagner et al., (2021) reference to the SM reference list.

References:

Monks, S.A., Arnold, S.R., Hollaway, M. J., et al.: The TOMCAT global chemistry transport model v1.6: Description of chemical mechanism and model evaluation, *Geoscientific Model Development*, 10 (8), 3025–3057, doi:10.5194/gmd-10-3025-2017, 2017.

Pope, R.J., Arnold, S.R., Chipperfield, M.P., et al.: Substantial Increases in Eastern Amazon and Cerrado Biomass Burning-Sourced Tropospheric Ozone. *Geophysical Research Letters*, 47 (3), e2019GL084143, doi:10.1029/2019GL084143, 2020.

Richards, N.A.D, Arnold, S.R., Chipperfield, M.P., et al.: The Mediterranean summertime ozone maximum: global emission sensitivities and radiative impacts, *Atmospheric Chemistry and Physics*, 13, 2231-2345, doi:10.5194/acp-13-2331-2013, 2013.

Wagner, A., Bennouna, Y., Blechschmidt, A.M., et al.: Comprehensive evaluation of the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis against independent observations: Reactive gases, *Elementa: Science of the Anthropocene*, 9 (1), doi: 10.1525/elementa.2020.00171., 2021.