Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-79-AC1, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "The effects of intercontinental emission sources on European air pollution levels" *by* J. E. Jonson et al.

J. E. Jonson et al.

j.e.jonson@met.no

Received and published: 8 June 2018

We thank the reviewer for the effort to see into this multi-author paper. We apologise for oversights, partly due to the complex nature of the multi-model evaluation.

General comments

Lines 66-70: The structure in L66-70 seems not to match the actual content of the section (Section 5 became a very superficial discussion on model resolution, it seems that the authors forgot that they originally expected to suggest improvement in the experiment design in that section). It is unclear why comparison with measurements come back in Section 5, while it was introduced in Section 3. At the end of the introduction the reader is already sceptical to what extent the paper will address the problem at

Printer-friendly version

Discussion paper

hand.

We have changed this part of the paper, including a better motivation for Section 5 (in addition section 5 has also been improved). The motivation for section 5 is to sum up the results for the individual models and as such make the reader better prepared for the conclusions. These lines now reads: In this paper we first briefly discuss the model comparison to measurements in section 3. In section 4 we go on to describe the source receptor relationships for Europe, including a discussion on how and why the model results differ. Finally, in section 4 we sum up the results for the individual models. Based on model performance compared to measurements and where and when deviations in model results compared to the other models occur we try to indicate the origins of the differences in model behaviour. In the conclusions we then suggest some directions on how this information could be used to harmonize and improve future model calculations.

Specific comments

Line 23: define here the CTM acronym, which usually refers to chemistry transport models rather than chemical tracer models

Added CTM acronym

Line 25: TF-HTAP is organized under the EMEP programme of LRTAP

Added that HTAP2 reports to the EMEP stearing body

Line 50: add that the region targeted in that paper is Europe, but that (unlike in HTAP) the contribution is assessed by model tagging rather than sensitivity experiments.

We have added that tagging is used in his model.

Line 53-54: the sentence on additional papers is not relevant, suggest removing

We have removed the list and replaced in with a reference to the acp special issue.

Line 60-63 should be moved to the experiment description part (ex: L78)

This sentance is removed. This discussion is included elsewhere.

Line 63: rephrase “secondly we look at CO” to better introduce the actual chemical compound in opposition to the CO-like tracer.

We have changed to: Secondly we investigate CO as an interactive component of the atmosphere, participating in chemical reactions.

Line 87: For transparency and reproducibility concerns, but also with regards to the HTAP requirements, GFDM_AM3 should not be included if it is not part of the database.

We will now included GFDM_AM3 results in the database, note that these model results are present in a slightly different format. We have chosen to include these results as so few models have uploaded relevant results for this study.

Line 92: The Galmarini et al. article in the special issue is focused on the complementarity paper of global and regional models rather than model evaluation. In the version currently in discussion, only a Taylor diagram is given with models not labelled. Therefore it cannot be considered as a satisfactory reference regarding the capability of HTAP models in capturing surface ozone. Such an analysis should be included here if not covered elsewhere. The scatter plots in supplementary material is a good start, but further discussion is needed. The selection of Airbase sites is very questionable at this scale.

See also comments to reviewer 2 for this point. We have included ozone time series for several GAW sites in the supplementary material. Furthermore we have included a table with statistics similar to what is already included for CO. The mentioning of AirBASE as data source was wrong. The scatterplots in the supplementary material are based on European rural and remote sites from EMEP, as stored in the EBAS database.

Line 104: a reference is needed to conclude that GAW sites are affected by local

sources. Similarly, one could question why engaging in such analysis if global models are not capable to capture regional sources. If the author conclude that it is the case, it would be a major conclusion of the paper.

We agree with the reviewer that GAW sites are selected to represent regional background conditions. The argument was removed.

Line 116: what is the source of ozone profiles?

The original source for the ozone sondes is the "World Ozone and Ultraviolet Radiation Data Centre". This information is now included in the text. Data providers have been contacted and offered co-authorship.

Line 118: how “approximate” is the temporal matching between model and observations?

In HTAP2 model profiles are provided on an hourly basis. We have added "(to the nearest hour)" in the text.

Line 123: more quantitative results are needed to support the “tendency” for underestimation in tropospheric summertime ozone.

We have extended the section with the interpretation of the ozone profiles. This part now reads:

With a relatively inactive chemistry in the winter months the measured ozone profiles show little vertical variability, with ozone mixing ratios in the troposphere increasing gradually with height. Model calculated ozone profiles are close to the measurements. As the chemical activity increases in Spring and summer months the vertical variability increases, reflecting air masses of significantly different photochemical history at different levels. As was shown in Jonson (2010) the models are not capable of reproducing this vertical structure in ozone levels. Most of the models underestimate free tropospheric ozone in the summer months.

Line 154; why is GFDL_AM3 included in Figure 3 for CO_tracer but not in Figure 2?

Figure 2 requires data from the BASE, GlOALL and EURALL scenarios, whereas Figure 3 is based on the BASE and GLOALL scenarios only. For the GFDL_AM3 model we have no data for the CO tracer for the EURALL scenario. A comment has been added in the text.

Line 159: Why would EMEP have a too strong convection? If the comparison with measurement suggest that EMEP performs better than other models (L348), maybe the other models have a too weak convection ?

We have deleted the word too, subsequently changing the meaning so that we now say that the EMEP model has strong convention and not "too" strong convection.

Line 162: the larger vertical mixing seems to occur mainly in winter for EMEP, isn't that conflicting with the hypothesis about the role of convection? Maybe more discussion would be needed on the vertical diffusivity and resolution of the various models.

Running the global EMEP model with and without spinnup we see marked differences (in ozone) lasting all the way into Spring. We believe that this is caused by ozone lifted into the free troposphere increasing the free tropospheric reservoir of ozone in the following winter and spring. We believe this also the case for the CO tracer.

The text has been made more clear on this point: Differences in mixing ratios peak in the first part of the year when emissions are high and the exchange between the boundary layer and the free troposphere over Europe is weak. Differences in the free troposphere may reflect CO tracer advected from regions upwind with convective activity also in winter, or in the preceding autumn months increasing the free tropospheric reservoir in the following winter and spring.

Line 183: The difference between CO_tracer and CO seems larger for EMEP than for the other models. Would it also be the case in terms of relative increase, and if so, why?

The widening gap between EMEP end the IFS model is commented below in the same section. It is attributed to the possible lower OH levels in the EMEP model compared to the IFS model. (even though OH is not provided by the IFS model).

Line 193: the model differences for OH are very impressive (Fig 9 of the supplementary material). To the extent that one may wonder the relevance (and need) to produce a multi-model mean. Further discussion and external references are needed for that section. The sensitivity to upper boundary conditions, especially for EMEP that seems to behave differently.

A more detailed discussion on differences and of the effects of OH is now included in several places in the manuscript.

Line 213: a reference is needed to support the statement on the relative contribution of stratospheric/tropospheric ozone.

References to the HTAP 2010 report and Stevenson et al. 2006 now included.

Line 245: the discussion on aircraft emissions is interesting, but it seems that there are more important differences, such as the role of surface titration (why EMEP seems the less sensitive despite the higher resolution). Or the fact that the O3 response of Chaser is actually very close to that of CO_tracer (or is it a mistake in the Figure?)

Discussion on aircraft emissions strengthened following comments also from reviewer 2. Regarding the GFDL_AM3 (not CHASER), the figure is corrected. We believe that the reason why the EMEP model is less sensitive to titration must be sought in the chemistry schemes. This is now discussed later in the paper (see also comments from reviewer 2).

Line 256: the fact that CH4 is excluded from the experiments should appear before in the experiment description (Section 2).

The statement that CH4 is not included in the experiment is moved to section 2.

Line 259: The explanation of figure 5 needs to be re-worked in the text, in particular to explain that even if some sources are not isolated in some models, their contribution is still accounted for in the "remaining" fraction.

This section now reads: For each model the contribution from ROW (Rest Of the World) is calculated by subtracting the sum of the contributions from from available world regions from the BASE - GLOALL contribution. Thus the portion related to ROW includes a varying mixture of world region definitions depending on the model.

Line 266: more quantification is needed regarding the relative role of external/European sources. Figure 5 indicates that the external contribution seems indeed to exceed European contribution, but they are actually not that far. The percentage contribution (with error bar) should be given.

We have added an additional table listing the percentage contributions to annual ozone and summer ozone to Europe from Europe, North America and East Asia. In addition we also list the contributions to SOMO35 and POD forest calculated by the EMEP model. The numbers are a subset of those displayed in Figure 5.

Line 275: the comparison the HTAP1 is too weak. It is very frustrating not to better understand the added value of the new experiment and to what extent the earlier conclusions still hold. The benefit of having engaged in a complete new experiment should be better assessed. For instance by looking at a subset of models having participate to both and investigating clusters of regions for the reference/sensitivity simulations to conclude on the importance of (i) emission changes, (ii) region definition, (iii) participating models. Conclude on emission changes:

The paragraph has been rewritten and now reads: In comparison to HTAP1, HTAP2 regions are better defined. In addition emissions as well as models are up-to-date. To disentangle whether the changes from HTAP1 to HTAP2 are due to emissions, a changed model ensemble or changes in receptor regions is unfortunately not possible in a fully quantitative way. Source and receptor regions have been chosen in HTAP2

to cover the land-only politically connected regions accurately on a 0.1 degree grid. In HTAP1 the EUR region was a simple latitude - longitude box, also including parts of North Africa, the Middle East, Russia, Belarus, Ukraine and large sea areas, all of these identified as non European regions in HTAP2. In HTAP2 the European region is smaller, thus exporting larger fractions to nearby regions, but most major HTAP1 source regions are located within the smaller HTAP2 region, thus making this region more sensitive to titration effects. As a result the effects of emissions on ozone levels from the EUR region to itself is reduced.

The ensemble mean contribution to ozone levels from Europe to itself has decreased from 0.82 +- 0.29 ppb in HTAP1 to just 0.11 +- 0.32 ppb in HTAP2. Also - total and regional distribution of emissions for the base year changed from HTAP1 (2001) to HTAP2 (2010). Gaudelet al. (2018) have analysed the ozone trends between the years 2000 and 2014. Over Europe. They found a general ozone increase in the winter months (December, January, February) and a general decrease in the summer months (June, July, August). The emission trends in the HTAP1 world regions are given in Turnocket al. (2018) between 2001 (the base year for HTAP1) and 2010. The changes in measured ozone are consistent with the reductions in European (and North American) emissions of NOx (along with other ozone precursors) over the same period resulting in less titration and thus increased ozone levels in some areas mainly in the winter months, and simultaneously less net ozone production in summer. Likewise emissions in North America have decreased and may explain the 0.37 +- 0.10 to 0.22 +- 0.07 ppb decrease in the ensemble mean contributions from North America to European ozone levels. Over the same period emissions in other world regions as East Asia have increased. This increase may explain the 0.17 pm 0.05 to 0.22 +- 0.13 ppb ensemble mean increase from HTAP1 to HTAP2 in the East Asian contribution to European ozone levels. Contributions from South Asia are small in both HTAP1 and HTAP2 (0.07 versus 0.05).

Line 301: it is surprising to say that a comparison should not be made, when it is

made by the authors. There were no similar discussions on the realistic aspect of sensitivity experiments for non-CH4 species, so from an academic perspective regarding chemical sensitivity the comparison does hold.

Thanks. The point is that the effects of direct changes in concentrations is not the same thing as changing the emissions. This is now clarified and rewritten: However, comparing a 20% change in CH_4 concentrations, and the effects of the GLOALL emission scenario requires careful interpretation. Because of its relatively long lifetime of the order of 10 years in the atmosphere, a 20% change in concentration corresponds to an approximate 40 year long historic CH4 trend (Meinshausen 2011). The GLOALL scenario is not accounting for the full impact of a continued 20% reduction in emissions. With a continued emission reduction scenario, the overall ozone reductions would be larger, while the methane attributable fraction, relatively, would be smaller.

Line 330: the results related to ozone indicators are interesting and worth being highlighted in the abstract. It is frustrating that only one model can be used here, especially given the apparent different behaviour with regards to titration. More efforts should be given to investigate the HTAP database in order to include more models for a comparison of summertime mean of daily ozone maxima, or at least summertime mean ozone.

Unfortunately hourly data with attribution to source regions are only available for the EMEP and the CHASER model. We have thus instead included a table comparing annual average and summertime ozone based on the numbers in Figure 5. (see comments to line 266). SOMO35 and POD forest from the EMEP model are included in the same table. As SOMO35 and POD are mainly added up in the summer months the percentages are similar in the EMEP model and the table gives an indication on the differences between the models.

In addition to the table we have added some additional text: The regional contributions, expressed by these metrics, are also listed in table (new table). The figure and table

clearly shows that the choice of metric matters, in particular for the effects of European Emissions. POD forest is accumulated in the growing season in summer. A large portion of SOMO35 is also accumulated in the summer months. Table (new table) also lists the percentage contributions to summer ozone for all models. The similarities in the percentages for summer ozone and the ozone metrics in EMEP_rv48 is an indication that also for the other models these percentages are comparable.

Line 350: according to Section 3.2, EMEP is not the only model to display an overestimation of tropospheric ozone.

Yes, text is changed to: The model is one of the models with highest overestimation of ozone in the free troposphere in the winter and spring months.

Line 360: how can a lower diffusion can lead an underestimation of surface CO, the opposite would be expected.

Thanks - we agree. If the low RERER is caused by too much CO remaining in the PBL it should result in an overestimation of surface CO compared to measurements. We have deleted this statement.

Line 367: what is the rationale for a relaxation to zero in the GEOS-Chem adjoint?

The GEOS-Chem model has only ozone chemistry in the troposhere, and stratospheric levels should be disregarded.

Line 372: Section 5 is very descriptive and lacks a clear outcome

We have added more text, including text with additional motivation for this section: Here we try to point out if, and at what stage, the results from the individual models deviate from the other models. It should be stressed that such a deviation does not necessarily imply that the results from a particular model is wrong.

For the individual models we have tried to identify one ore more feature where the individual models differs from the other models.

Line 399: From the results provided in Section 5, it can not be concluded that the results are not sensitive to model resolution.

Yes, this is probably phrased too general - This part now reads: The model resolution differs between the individual models. Model results from the two CHASER models, differing in model resolution only, are qualitatively similar when compared to measured CO and O3} at background measurement sites and very similar in RERER for CO and ozone, suggesting that resolution differences at the scales investigated here, are not important to explain RERER differences between the global models. Still, it is difficult to conclude in general to what extent horizontal resolution affects the source receptor calculations at intercontinental scales.

Line 406: it is rather convection that is put forward rather than chemistry. The following sentence (L407) also goes in that direction. It is quite surprising to read that the conclusion and the content of the paper seem contradictory.

This part now reads: The joint and consistent analysis of a CO tracer, CO and O3 in this paper is a tool in understanding where and why (right or wrong) the models differ, however, it has a potential for wider use, enhancing our understanding of the result and also as a tool for model improvements, reducing the overall uncertainty in future model calculations. We believe that in order to close the gap in model results, and subsequently improving the reliability of the model output, possible future model inter-comparisons should be more process oriented (transport, depositions, chemistry etc). Our study shows that models differ already for CO and the inert CO tracer, where differences were established with 2 models, but that differences are amplified as more chemistry is added. Note that the CO RERER and O3 RERER values are not correlated taken the models as samples. The big additional spread in model results for ozone is clearly induced by differences in model chemistry and for instance treatment of titration in the winter boundary layer. However, differences in chemistry may well also be induced by differences in advection/convection as the level of exchange will inevitably affect the chemical regime in both the free trposphere and in the boundary

layer. We believe therefore that further process oriented evaluations (comparing advection/convection, chemistry, dry and wet deposition etc separately) should be made, making use of relevant meteorological and chemical measurements.

Line 419: from Figure 5, it seems that about half of ozone can be mitigated with European sources, isn't that "sizeable" already? More precise quantifications must be given in the conclusion and abstract with key figures and associated error bar across the multi-model ensemble.

We now include a comparison and discussion of RAIR in HTAP1 and HTAP2 in section 4.4. Furthermore RAIR is also included in the conclusions and the abstract.

The (almost) first parts of the conclusion section now reads: The model calculated relative contributions to surface ozone levels from non European sources is much larger in HTAP2 compared to HTAP1. Mainly because the contributions from Europe to it selves has decreased from 0.82 ppb to just 0.11 ppb. As a result RAIR has increased from 43 to 82\%. In parts differences could be explained by decreasing emissions in Europe and increased emissions in most other regions as East Asia from year 2001 to 2010. However, the results from the two HTAP phases can not easily be compared, ...

Table 1: what is given on the right part of the table? From the text, it appears to be correlation, but that should be stated clearly. Is CHASER_rel actually Chaser_t42 according to table 1 of the supp. Mat. A uniform model labelling would be appreciated.

We have added a line at the top of the table explaining what is calculated concentrations and what is correlations. Model labeling corrected.

Table 2: ibid about CHASER_rel. The labels of regions should be consistent with Fig 1.

Labels for regions now as in Figure 1.

Table 3: swap the first and second sentences. Use boldface rather than larger signs

for important deviations.

First and second sentence swapped. Large + and - replaced by bold face.

Figure 2. GFDM_AM3 missing for CO_Tracer

The calculation of RERER require BASE, GLOALL and EURALL. EURALL calculations are not available from GFDL_AM3.

Figure 3: panels g) and i) appear identical

g and i now corrected.

Figure 4 and 5: Truncation fixed

Figure 6: add that the results to 20% perturbation are plotted.

We have added that the the GLOALL scenario is calculated with with 20% reductions in anthropogenic emissions.

Supp. Table 1: Table 1: all models are in bold, not the first 7, since 7 models are displayed. The information about spin up should be given in the experiment description, not in Table 1. Footnote #2 is not references in the text.

We have removed bold face for the last model. The information about spinnup now in the experiment description. Footnote 2 deleted

Supp. figure 1: What happened in Heimaey in May in observations?

The peak is actually the IFS model. We have changed the colour scale for the models so that it is easier to tell models and measurements apart. In the IFS model it could be an artefact of the GFAS fire emissions (April/May 2010 is the time of EEyjafjallajökull eruption but it is not clear how this can effect CO).

Supp. figure 4: More details are needed in the legend about the source of data and the indicator displayed.

More information about data source added. Editing errors corrected.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-79, 2018.

Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-79-AC2, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "The effects of intercontinental emission sources on European air pollution levels" *by* J. E. Jonson et al.

J. E. Jonson et al.

j.e.jonson@met.no

Received and published: 8 June 2018

We thank the reviewer for the effort to see into this multi-author paper. We apologise for oversights, partly due to the complex nature of the multi-model evaluation.

Reply to general comments:

Most of the general comments are addressed in the detailed comments below, eg error in figure 3, fig 3i, AM reference etc.

At the time of the submission we believed that another paper would address the validation of ozone. As it turned out this was not the case, and as a result the ozone validation has been extended in this paper. We have reorganised the paper. The content of in particular section 5 has been expanded, and a motivation for this section is

Printer-friendly version

Discussion paper

included in the introduction.

The issue of ozone titration is discussed in more detail. The reviewer is right that as in particular European NOx emissions have decreased from 2001 to 2010, and as a result European controls have been offset by by removal of local suppression. These considerations are now discussed in section 4.4.

Detailed comments: ————————

Line 16: capitalization error.

Comma replaced by .

Line 25: Missing parenthesis

Added right)

Line 29: Verb agreement. Replaced is with are.

Line 46-54: The list of published papers should be used to provide context. Here it is simply a list.

We have deleted this list and replaced it with this: A large number of papers from HTAP2 have been published in the ACP (Atmospheric Physics and Chemistry) special issue: “Global and regional assessment of intercontinental transport of air pollution: results from HTAP, AQMEII and MICS”

Line 60: Details like model count would be better in the methods.

Model count deleted here.

Line 68-70: Differ should be differences?

Not applicable. The description of the sections later in the paper is changed.

Line 72-74: Poorly written.

Now corrected to: The HTAP2 model experiment was set up by the Task Force on

Hemispheric Transport of Air Pollution (TF∼HTAP). A project work plan, a description of the model experiments etc can be found on the TF∼HTAP web page (\url{ http://www.htap.org/}).

Line 82: "etc" seems particularly poor when later you will refer to advection schemes as a causal difference.

Advection added to list, and a reference to the supplementary material. These models have different resolutions, advection schemes, chemical mechanisms etc (see supplementary material and references therein).

Line 87: Space added

Line 92: Space added

Line 95: Replaces in by is

Line 95: How does evaluation of upwind sources affect conclusions about transport to Europe?

We have not included an evaluation of upwind sources here. Several other HTAP2 papers are addressing this.

Line 99: GAW (Global Atmospheric Watch) spelled out.

Line 101: How "high" correlations are expected given the resolutions of the models?

We have included some more text and references here: Correlations shown here are in the same range as correlations with MOPITT satellite measurements as reported by Naik et al.2013. However, as shown in Table 3, all models except IFS_v2 underestimate annual CO levels by 13% or more. Similar underestimations was also shown Strode et al. 2015.

Line 101-102: resolutions of all the models should be provided in the methods rather than the comparison to measurements.

The information about the resolution of the CHASER models is provided here as explanatory information for how the versions of the CHASER model differ. Information on model resolution for all models is given in a table in the supplementary, referenced in section 2.

Line 98-109 How is it that CO deserves a site-by-site comparison and ozone?

Unfortunately it was communicated to us until a few days before the manuscript had to be submitted that this would be included in the Galmarino et al. paper. Therefore is was not included in this ACPD submission. We are now including a site-by-site surface ozone evaluation based on GAW data as already included for CO.

Line 112. The authors should mention that they do have some surface evaluation in this paper. Currently, Table 3 in this manuscript is not referenced until Section 5.

We now say that there is additional surface evaluation in the Galmarini paper, and we refer to Table 3 also in this section and too all supplementary material on ozone evaluation.

Line 114: There is currently no discussion of ozone results except to say they exist somewhere in the supplement. Why is this sufficient?

We have added a more complete paragraph on the ozone comparisons made in chapter 3.

Lines 122 - 123: There must be more discussion of the basic results that will clearly affect transport.

This section has been extended and now reads: The profile comparison allows to identify differences between the models in vertical mixing of ozone useful for further interpretation in inter-hemispheric transport efficiency. Note that the GEOS-Chem model only simulates ozone in the troposphere and its ozone levels above 300 hPa should be disregarded. With a relatively inactive chemistry in the winter months the measured ozone profiles show little vertical variability, with ozone mixing ratios in the troposphere

increasing gradually with height. Model calculated ozone profiles are in general close to the measurements. As the chemical activity increases in Spring and summer months the vertical variability increases, reflecting air masses of significantly different photochemical history at different levels. As was shown in \cite{Jonson2010} the models are not capable of reproducing this vertical structure in ozone levels. Most of the models underestimate free tropospheric ozone in the summer months.

Line 138: Here and elsewhere the definition of regions is incorrect. Here you have NW, SW, SE, GR+TU. In the Figure, you have NW SW, E, GR+TU. Other places you have NW, SW, E, SE. Choose one, and be consistent.

The region notations are now consistent throughout the paper.

Lines 139 - 140: Is this source apportionment the same as contribution in sections 4.4 and 4.5?

Yes. We have now included references to the subsections.

Line 142: rate of decay is later explained, but here seems completely arbitrary.

We disagree. We think that the rate of decay is useful information/reminder here.

Line 182: Numbering of Figure 3 and 4 corrected.

Lines 185 - 189: The reasonableness of this should be discussed.

Differences between the individual models are very similar for CO and the CO tracer. Differences in the CO tracer can only be caused by advection as there is no chemistry for this species. The similarity between CO and the CO tracer for two models leads us to believe that the causes for the differences are the same.

This argument is included in the text.

Line 205: This gets discussed in several places and is really part of the methods.

OK, shortened here, but this information is also repeated here as part of the interpre-

tation of the results.

Lines 217 - 2019: Web citation is inappropriate. Further, the lifetime of ozone is expected to vary with respect to season and altitude (Wang et al. 1998; Brasseur, Orlando, and Tyndall 1999). Estimates of lifetime at 500hPa range from 15-160d and from 40-300d at 10km. Your upper bound of 18days is misleading. Table 1.1 of the HTAP 2010 report cites weeks to months in the free troposphere. The IPCC range of values do not acknowledge the complexity of ozone transport.

In acknowledgement of the complexity of ozone chemistry and transport, we now refer to the HTAP 2010 report for the lifetime of ozone. In addition we have replaced the web citation with a reference to the IPCC report.

Line 242: AM3?

The GFDL_AM3 model is added to the list of models not perturbing aircraft emissions.

Lines 246 - 247: Provide some reference or evidence.

We are now referring to a paper by Cameron et al. (2016} for the effects of aircraft emissions on surface ozone calculated by several global models. See updated discussion in the manuscript for details.

Line 247: here = PBL?

This part has been rewritten.

Lines 254 - 284: Is this contribution from a simple mean within seasons? What months were included in each season? Are the numbers in the text ensemble means? What about ensemble mean RBU? MDE? EU? 290-291: Did they "too" calculate smaller "than in this study" or did they "too" calculate "smaller as in this study"?

This section has been rewritten. See also comments from reviewer 1. In Figure caption 5 we now specify which months are included in WI, SP, SU and AU. 0.37 NA to EU and 0.17 EA to EU are from Table 4.2 in the HTAP1 report. The numbers are ensemble

means. This is now noted in the text. We have chosen not to compare the numbers for EU as the definition of the European domain is so different. We now also list list the numbers for the remaining regions. They are also shown in Figure 5.

Line 269: MDE appears to always be small.

We now say that contributions from the Middle East and North Africa are small.

Lines 290 - 291: Did they "too" calculate smaller "than in this study" or did they "too" calculate "smaller as in this study"?

The text is changed to make this clearer: They calculate a much smaller contribution from non European sources than in this study, similar to the contributions calculated in HTAP1.

Lines 272 - 273: Did these other studies use the same model?

We have added that in Jonson et al. (2015) the EMEP model was used. Brandt et al. (2013) used a different model.

Line 277: Right parenthesis added.

Lines 274 - 280: Methods?

This part is now rewritten.

Line 305: HTAP1?

We have added that the Fiore et al. paper was based on the HTAP1 model experiment.

Lines 306 - 335: There needs to be a clearer connection to the previous section. In fact, you could just add two bars to Figure 5a. That would help to connect the of POD and SOMO35 to the seasonality of titration.

We have added more material to this section following the recommendations also from reviewer 1

Line 361 - 371: Terse and uninformative

This section is rewritten bringing more information.

Line 390 - 392: See previous comments about ozone lifetime.

We now refer to discussions on lifetimes in previous section.

Line 400: Probably deleted. Improved text.

Line 405: added for ozone.

Comments regarding Table 1: If mountain sites are used at readers peril, consider making room for ozone evaluation by moving them from the first data result.

We have included a similar table as Table 1 with ozone. We have not included mountain sites in the ozone table as the "peril" is much larger for ozone as the dry deposition is faster and lifetime shorter.

Comments Table 2: Update region definitions to be consistent with figures and text.

Regional definitions updated.

Figure 1: update region names to be consistent. Also, too many extra colors so it is hard to tell what is included. Is the Baltic Sea part of Eastern Europe? Black Sea? Caspian Sea? Mediterranean?

Not changed. Difference in colour is visible both on the screen and on printout. The European seas are part of the OCN region.

Figure 2: Necessary?

We would definitely like to keep the figure. We think it illustrates very well the evolution in RERER going from a simple CO tracer to CO and finally ozone with a multi model ensemble.

Figure 3: lettering needs to be updated in the figure and in the text. What was the com-

mon grid and how was it treated when a grid cell at 1000hPa was below the surface?

Lettering is updated and is now the same as in the text. All model data has been interpolated to a common vertical grid. For gridcells below 1000hPa values at the lowest model level was used.

Figure 3: 3i is AM3 CO not ozone. Column 3: consider a scale that does not saturate in so much of the figure.

Panel 3i corrected.

Figure 4: North and south boundaries are unnecessarily different from figure 3. Further, this highlights that no meaningful discussion of the boundaries was made. In fact, 50E includes a lot of Russia and a lot of ocean. Column 3: consider a scale that does not saturate in so much of the figure.

North and south boundaries changed corresponding to Figure 3. This resulted virtually no visual changes in the figures. We have added a discussion on the boundaries: This area roughly corresponds to the European regions as shown in Figure 1, but also some additional land and sea areas. The main focus of the figure is in the free troposphere where horizontal gradients in concentrations are small. Liu et al. 2009} calculated the correlations between nearby pairs of sonde stations. They found low correlations near the surface indicating that local and regional effects are important here. From the surface correlations rose sharply to a local maximum in the lower troposphere. We therefore conclude that the selected area is a good representation of the atmosphere above Europe.

Figure 5: There is no discussion about the CHASER model being the only one without apparent titration, and this should be discussed somewhere. Region definitions should be consistent with the text or the text should be consistent with the figure. The units are cutoff on the first row.

We have commented the low level of ozone titration for the CHASER model in section

4.3: For all models, except the CHASER_re1 model, ozone titration dominates the overall European contributions when summed up over the three winter months. However, for all the models, including also the CHASER_re1 model, the net European contributions includes regions of net ozone production and net ozone destruction in winter.

Regional definitions now consistent with text.

Figure 6: Region definition nomenclature. I recommend showing as 3 stacked-bars (or adding to Figure 5). If I am interpreting this right, the RAIR is 84% compared to 43% from HTAP1. I suspect that all models provided annual and I think reporting RAIR would be useful (maybe in Figure 2).

Region definition nomenclature fixed.

Figure 6 is complemented by a table with results for summer ozone from the models following the recommendations from reviewer 1. For ozone this table lists the annual (and summer) percentage contributions to Europe from several regions, including Europe to it selves. We have also calculated average RAIR for the models in Figure 5. The HTAP2 RAIR of 82% compared to 43% in HTAP1 is discussed in section 4.4, and these numbers are also repeated in the conclusions. RAIR for the individual models proved difficult with European contributions to it selves was close to zero and even negative for some models.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-79, 2018.

Printer-friendly version

Discussion paper

site
Manuscript prepared for J. Name
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 25 June 2018

The effects of intercontinental emission sources on European air pollution levels

Jan Eiof Jonson[1], Michael Schulz[1], Louisa Emmons[2], Johannes Flemming[3], Daven Henze[4], Kengo Sudo[5], Marianne Tronstad Lund[6], Meiyun Lin[7], Anna Benedictow[1], Brigitte Koffi[8], Frank Dentener[8], Terry Keating[9], Rigel Kivi[10], and Yanko Davila[4]

[1]Norwegian Meteorological Institute, Oslo, Norway
[2]National Center for Atmospheric Research Boulder, Colorado, USA
[3]ECMWF (European Centre for Medium Range Forecast)
[4]University of Colorado Boulder, Colorado, USA
[5]NAGOYA-U,JAMSTEC,NIES, Japan
[6]Center for International Climate and Environmental Research (CICERO) - Oslo, Norway
[7]Program in Atmospheric and Oceanic Sciences of Princeton Univeristy and NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA
[8]European Commission, Joint Research Centre, Ispra, Italy
[9]U.S. Environmental Protection Agency
[10]Finnish Meteorological Institute, Sodankylä, Finland

Correspondence to: Jan Eiof Jonson (j.e.jonson@met.no)

Abstract. This study is based on model results from TF HTAP (Task Force on Hemispheric Transport of Air Pollution) phase II where a set of source receptor model experiments have been defined, reducing global (and regional) anthropogenic emissions by 20% in different source regions throughout the globe, with main focus on year 2010. All the participating models use the same set of emissions. Comparisons of model results to measurements are shown for selected European surface sites and for ozone sondes, but the main focus here is on the contributions to European ozone levels from different world regions, and how and why these contributions differ depending on model. We investigate the origins by use of a novel stepwise approach combining simple tracer calculations and calculations of CO and O_3. To highlight differences, we analyse the vertical transects of the mid latitude effects from the 20% emission reductions.

~~Based on the relative emission changes from different world regions the models agree that for~~ The spread in model results increase from the simple CO tracer to CO and then ozone as the complexity of the physical and chemical processes involved increase. As a result of non linear ozone chemistry the contributions from non European relative to European sources are larger for ozone compared to CO and the CO tracer. for annually averaged ozone the contributions from the rest of the world is larger than the effects from European emissions alone, with the largest contributions from North America and East Asia. ~~The contribution will however depend on the choice of ozone metric.~~ There are also considerable contributions from other nearby regions to the east and from international shipping~~,~~

. For ozone the European contributions to metrics reflecting human health and ecosystem damage, mostly accumulated in the summer months, are larger than for annual ozone. Whereas ozone from European sources peaks in the summer months, the largest contributions from non European sources are mostly calculated for the spring months when ozone production over the polluted continents starts to increase, while at the same time the lifetime of ozone in the free troposphere is relatively long. At the surface contributions from non European sources are of similar magnitude for all European sub regions considered, defined as TF HTAP receptor regions (north west, south west, east and south east Europe).

1 Introduction

This paper is based on the HTAP model experiment phase 2 (HTAP2), where CTMs (chemical tracer models) perform model sensitivity studies, perturbing the emissions in different world regions. TF HTAP (http://www.htap.org/) is organized under the auspices of the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP Convention) and reports to the Convention's EMEP Steering Body. The HTAP2 experiment is described in more detail in Galmarini et al. (2017) and in the HTAP2 work plan, posted on the HTAP2 web site www.htap.org. All models should use the same set of emissions, see Janssens-Maenhout et al. (2015).

In particular the experiments is set up to:

- Examine the transport of air pollution, including ozone and its precursors and particulate matter and its components (including black carbon), across the Northern Hemisphere.
- Assess potential emission mitigation options available inside and outside the UNECE region.
- Assess their impacts on regional and global air quality, public health, ecosystems, and near-term climate change.
- Promote collaboration both inside and outside the Convention.

HTAP2 is a follow up of the HTAP phase 1 model experiment (HTAP1). Results from HTAP1 has been described in a series of peer review papers, including (Casper-Anenberg et al., 2009; Fiore et al., 2009; Reidmiller et al., 2009; Jonson et al., 2010; Sanderson et al., 2008; Shindell et al., 2008), and in the the HTAP1 main report (TF HTAP, 2010). The HTAP1 model experiment showed that intercontinental transport of ozone and ozone precursors could explain a large portion of the ozone over Europe, but results differed substantially between the models.

A large number of CTMs ~~and GCMs~~ have uploaded their results to the HTAP2 database. This study is limited to those models that, in addition to the base run, as a minimum have uploaded their source receptor calculations for ozone reducing all anthropogenic global emissions and European emissions by 20%. Seven of the models fulfil these criteria.

Several A large number of papers from HTAP2 have been published. Janssens-Maenhout et al. (2015) describes the common set of emissions to be used by the models. Stjern et al. (2016) study the effects of changes in the atmospheric load of black carbon (BC), organic aerosols (AE) and sulphate on radiative forcing. Using a regional model with lateral boundaries from a global model, Huang et al. (2017) study the , in the ACP (Atmospheric Chemistry and Physics) Special issue: "Global and regional assessment of intercontinental transport of air pollution: results from HTAP, AQMEII and MICS"

The effects of intercontinental ozone on North America . With a regional model citeKaramchandan2017calculates the contributions from several emission sectors, including model boundary conditions, to tropospheric and . Turnock et al. (2018) use the HTAP2 source receptor relationships to parametrize future changes in ozone. Several additional transport of ozone to North America is discussed in Huang et al. (2017), but no such study has so far been made for Europe based on the HTAP2 papers have also been published or have been submitted for review.

data set. In this paper we aim to enhance our understanding of the contributions to European ozone levels from European and non-European sources. In order to better understand the transport patterns between the continents we use a novel stepwise approach, starting with a simple CO like tracer using the CO anthropogenic emissions and a fixed decay rate of 50 days. As all models use the same emissions, differences in model results can be ascribed to differences in transport (advection, including also convection and diffusion) only. Unfortunately only two models have provided such tracer data for at least the BASE, GLOALL and EURALL model run, while a third model has provided tracer data for BASE and GLOALL. Secondly we look at CO . Secondly we investigate CO as an interactive component of the atmosphere, participating in chemical reactions. The main sink for CO is the reaction with OH, and thus differences in OH is one of the main factors affecting CO. Finally we look at ozone. The causes of the differences in calculated ozone are hard to identify, but some clues can be identified based on the calculations of the CO like tracer and CO.

In this paper we first briefly discuss the model comparison to measurements in section 3. In section 4 we go on to describe the source receptor relationships for Europe, including a discussion on how and why the model results differ. Finallywe discuss how this information on how modelscompare to measurements , and in what way the model results differ , in section 5 we sum up the results for the individual models. Based on model performance compared to measurements and where and when deviations in model results compared to the other models occur we try to indicate the origins of the differences in model behaviour. In the conclusions we then suggest some directions on how this information could be used to harmonize and improve future model calculationsin section 5.

2 The HTAP2 model setup

The HTAP2 model model experiment was set up by the Task Force on Hemispheric Transport of Air Pollution (TF HTAP). A project work plan, a description of the model experiment etcexperiments

etc. can be found on the TF HTAP web page (http://www.htap.org/). The models were required to perform a 6 month spinnup for all model runs. A more detailed description of the requested model runs, emissions, requested model output and formats etc. is also included in (Galmarini et al., 2017) and references therein. A detailed description of the emissions can be found in Janssens-Maenhout et al. (2015). More documentation about the models can also be found in the supplementary material.

In this paper we focus on the effects on Europe. Even though a substantial number of models have uploaded their results to the database, model results for ozone (and CO) are only available from 7 of the models for the BASE model runs and for at least the two scenario runs reducing all anthropogenic emissions except CH_4 by 20% globally (GLOALL) and in Europe (EURALL). These models have different resolutions, advection schemes, chemical mechanisms etc (see supplementary material and references therein). Additional model runs reducing all anthropogenic emissions in North America (~~NAALL~~NAMALL), East Asia (~~EAALL~~EASALL), South Asia (~~SAALL~~SASALL), Middle East (~~MDALL~~MDEALL), Russia, Belarus, Ukraine (RBUALL) and ship emissions (OCNALL) are also discussed here. The ~~definition~~ definitions of these regions ~~is~~ are given in Koffi et al. (2016). The models are a subset of the HTAP2 models listed and described in Stjern et al. (2016). Since then additional model result have also been provided for ~~the~~GFDL~~the~~ GFDL_AM3 model~~(but not uploaded to the HTAP2 database)~~, raising the number of models to 8. (GFDL_AM3 model data are now included in the database, but in a different format than the other models). Additional information on the models are also listed in the supplementary material. Access to model data are available upon registration from http://aerocom.met.no.

3 Models vs measurements

In this section we ~~briefly~~ discuss the performance of the models compared to measurements. Wherever possible we have used the validation tools provided by AEROCOM: http://aerocom.met.no/cgi-bin/aerocom/surfobs_annualrs.pl?PROJECT=HTAP&MODELLIST=HTAP-phaseII. This enables the reader to explore the results on their own. For ozone a comprehensive model to ~~measurement~~comparison measurement comparison is published in Galmarini et al. (2017), including a comparison of both global and regional model results. However, this study focus mainly on the ensemble mean, and individual model results are anonymous. For surface ozone we ~~therefore~~ refer to this paper, but additional model validation is also included here. Comparisons of model calculated vertical profiles to ozone soundings are included in the supplementary material. As the focus of this paper ~~in~~ is on Europe, only European sites are shown. We have only included models with model output also for the GLOALL and the EURALL scenarios.

3.1 Surface

Monthly averaged timeseries of measured versus model calculated CO are shown in the supplementary material for a number of European GAW (Global Atmospheric Watch) sites. Some statistics for these sites are listed in Table **??**. At most sites CO has a clear winter maximum and a summer minimum. All models in general reproduce the seasonal cycle well at most sites ~~,~~ (see supplementary material), also reflected in their high correlations with the measurements. Correlations shown here are in the same range as correlations with MOPITT satellite measurements as reported by Naik et al. (2013). However, as shown in Table **??**, all models except IFS_v2 underestimate annual CO levels by 13% or more. Similar underestimations was also shown in Strode et al. (2015).

The results for the two CHASER model versions with high (1.1×1.1 degrees) versus low (2.8×2.8 degrees) resolutions differ, but they are qualitatively similar. ~~Some sites with very high concentrations (as Hegyhatsal) are clearly affected by local/regional sources not resolved by the global models.~~

This study also includes an evaluation of model results at several mountain sites. Results for these sites are shown, but should be interpreted with ~~great~~ caution. The elevation of mountain sites are poorly resolved in the models. Furthermore concentrations are likely to be affected by sub scale circulation patterns as mountain subsidence and upslope winds etc, that are not resolved by the models.

~~A~~ A more comprehensive comparison of the Base model calculations and ozone measurements from the EMEP and airbase measurement networks is presented in Galmarini et al. (2018) as part of HTAP2 and ~~ACMEII. Therefore we do not include an extensive comparison of measured and model calculated ozone here~~AQMEII (Air Quality Modelling Evaluation International Initiative). However, in the Galmarini et al. (2018) study the main focus is on the ensemble mean. An additional model validation of surface ozone is therefore also included here. Monthly averaged timeseries of measured versus model calculated O_3 are shown in the supplementary material for a number of European GAW sites. Some statistics for these sites are listed in Table **??**. The GAW sites are background sites relatively far from major sources. Scatter plots for the BASE model runs for ozone versus measurements are shown in the supplementary material. A summary of these results are also presented in Table **??**.

With coarse resolution, global models can not be expected to fully reproduce the measurements. The effects on model resolution on the validation of ozone measurements is demonstrated in Schaap et al. (2015) running the same set of models with variable horizontal resolutions. They show that for sites affected by local sources ozone is often over-predicted with coarse resolution as titration effects are watered out. Thus one may expect coarse global models to over-predict ozone levels at several sites classified as background sites. As shown in the scatter plots the OsloCTM3_v2 and the IFS_v2 model under-predicts the European annual ozone measurements by 22 and 18 percent, the other models overestimate ozone levels by 10 - 22%. This pattern of over and under-estimation is also apparent when comparing the

individual gaw sites. We only show results for one of the CHASER models as the two versions ~~again~~ are similar.

3.2 Vertical ozone profiles

Seasonal model calculated vertical profiles of ozone are compared to ozone sonde measurements downloaded from the World Ozone and Ultraviolet Radiation Data Centre (https://woudc.org/home.phpa) for several European sites in the supplementary material. Model calculated profiles are included in the calculations for the approximate same point in time (to the nearest hour) as the ozone sondes, and then averaged seasonally. The number of soundings included in the average for any site and season is listed in the individual panels. The figures have been produced by the AEROCOM tool: http://aerocom.met.no/cgi-bin/aerocom/surfobs_annualrs.pl?PROJECT=HTAP&MODELLIST=HTAP-phaseII.

The ~~vertical profile from the GEOSCHEMADJ model relax to zero in the upper part, and this is an artefact of the ajoint method. There is a tendency for most models to underestimate ozone in the free troposphere~~ profile comparison allows to identify differences between the models in vertical mixing of ozone useful for further interpretation in inter-hemispheric transport efficiency. Note that the GEOS-Chem model only simulates ozone in the troposphere and its ozone levels above 300 hPa should be disregarded. With a relatively inactive chemistry in the winter months the measured ozone profiles show little vertical variability, with ozone mixing ratios in the troposphere increasing gradually with height. Model calculated ozone profiles are in general close to the measurements. As the chemical activity increases in Spring and summer months the vertical variability increases, reflecting air masses of significantly different photochemical history at different levels. As was shown in Jonson et al. (2010) the models are not capable of reproducing this vertical structure in ozone levels. Most of the models underestimate free tropospheric ozone in the summer ~~(JJA) months.~~ months.

4 Source allocation, focusing on Europe

In this section we use the models to allocate the sources of ozone from different world regions, focusing on effects on European ozone levels. In order to better understand the differences between the models, we use a step-wise approach, starting the discussion with the CO like tracer in section 4.1, then we compare results for CO in section 4.2, where the treatment of the sources should be similar in all models, and the main sink is through the reaction with OH. Finally, in section 4.3 we compare the model results for O_3.

The calculations of the anthropogenic contributions from the different source regions are based on the difference between the base model runs and HTAP2 model scenario runs reducing all anthropogenic emissions globally (GLOALL), in addition to the reductions in the specific HTAP2 regions.

We first compare the model calculated effects of the GLOALL scenario for vertical trans-sections, and discuss the source allocation of domestic European anthropogenic sources versus external trans-continental anthropogenic sources expressed as ~~"response to extra-regional emission reductions" (RERER) (Galmarini et al., 2017)~~ RERER (Response to Extra-Regional Emission Reductions) as defined in Galmarini et al. (2017):

$$RERER = \frac{EURALL - GLOALL}{BASE - GLOALL}.$$

Again, BASE is the reference model run and EURALL the model runs reducing all European emissions by 20%. RERER is then a measure of the effects of external trans-continental versus domestic European emissions on the species in question. Given a fully linear chemistry, a RERER of one means that the concentrations in Europe are completely determined by sources outside Europe, whereas a RERER of 0 means that concentrations are determined by European sources alone. Unfortunately the chemistry is often far from linear. In particular for ozone, ozone titration, mainly in the winter months, can result in RERER values well above one, and in some cases even negative. In the section below annual RERER values are given for Europe as a whole and for four separate receptor regions, NW, SW, SE and GR+TU as shown in Figure **??**.

For ozone we also show the source attribution of European ozone further split into separate world regions for the the different models on a seasonal basis in subsection 4.4. Finally in subsection 4.5 we discuss to what extent the choice of ozone metrics will affect our findings.

4.1 CO tracer

The CO tracer is calculated with the same anthropogenic emissions as CO, and with a ~~set~~ fixed rate of decay giving a lifetime of 50 days. Any differences between the individual models can then be attributed to differences in transport processes. RERER for the CO tracers should be linear as there is no chemical interaction nor variability.

Table **??**, lists RERER calculated by the EMEP_rv48 and the IFS_v2 models (from the GFDL_AM3 model the CO tracer is calculated for BASE and GLOALL, but not EURALL. All three are needed for calculating RERER) for Europe and the four European sub regions. For Europe as a whole, RERER is also shown in Figure **??**. For the CO tracer RERER is ranging from 0.35 to 0.60, depending on model and European sub-region. There is a moderate difference in RERER between the two models. The highest RERER is calculated for the Gr+Tr region as this region is close to regions outside Europe as Russia, Belarus, the Ukraine, the Middle East and also the Mediterranean Sea and Black Sea.

Figure **??**a,d,g shows the annual mean difference in BASE - GLOALL of longitudinal CO tracer concentrations as an average between 30 and 60 degrees north. For all 3 models (EMEP_rv48, IFS_v2 and GFDL_AM3) the largest impacts of the 20% emission reduction on concentrations can be seen over the source continents in North America, Europe and in particular over ~~east~~ East Asia. There are

marked differences between the models as to what extent the CO tracer from the polluted boundary layer is lifted into the free troposphere. The EMEP_rv48 model (Figure **??**b), with high RERER (Table **??**) has ~~somewhat~~ higher tracer contributions in the free troposphere than the other two models (Figure **??**d,g). ~~This may indicate that material lifted~~ For the tracer the single factor that affects the concentrations is advection. Thus, the differences in the results are caused by various degrees of lifting into the free troposphere, possibly through ~~too~~ strong convection, ~~is transported~~ followed by rapid transport further from its sources~~and subsequently contributes~~ , subsequently contributing more to the tracer levels in distant regions before being decayed.

The seasonal cycle of the difference in BASE - GLOALL the over Europe, defined as the area bounded by 10°W to ~~50~~35°E and ~~25 to 65~~ 30 to 60 °N, is shown in Figure **??**a,d,g. ~~Differences in concentrations~~ This area roughly corresponds to the European regions as shown in Figure **??**, but also some additional land and sea areas. The main focus of the figure is in the free troposphere where horizontal gradients in concentrations are small. Liu et al. (2009) calculated the correlations between nearby pairs of sonde stations. They found low correlations near the surface indicating that local and regional effects are important here. From the surface correlations rose sharply to a local maximum in the lower troposphere. We therefore conclude that the selected area is a good representation of the atmosphere above Europe.

There are moderate differences in the seasonal behaviour of the CO tracer between the models, but tracer levels in the free troposphere are again highest in the EMEP_rv48 model. Differences in mixing ratios peak in the first part of the year when emissions are high and the exchange between the boundary layer and the free troposphere over Europe is weak. ~~There are moderate differences between the two models, but tracer levels in the free troposphere are again highest in the EMEP model~~Differences in the free troposphere may reflect CO tracer advected from regions upwind with convective activity also in winter, or in the preceding autumn months increasing the free tropospheric reservoir in the following winter and spring.

4.2 CO

Emissions of CO and the CO tracer are identical, and the results for CO resemble the results for the CO tracer in section 4.1. The dominant sink for CO in the atmosphere is the reaction with the OH radical, with a winter minimum and peaking in summer.

Table **??** lists RERER values for the seven models for Europe as a whole and for the four European sub regions shown in Figure **??**. RERER is ranging from 0.24 to 0.71, depending on model and European sub-region. Differences between the models are now caused by transport (as for the CO tracer) and chemistry. For the EMEP_rv48 and IFS_v2 models RERER is higher than for the CO tracer. Assuming that the CO chemistry is close to linear, this indicates a longer lifetime in the atmosphere than the 50 days for the CO tracer. IPCC Working group 1: the scientific basis (IPCC WG1, 2001),

https://www.ipcc.ch/ipccreports/tar/wg1/130.htm#tab41a) reports a lifetime of 0.08 to 0.25 years (about 30 to 90 days) depending on location and season, on average longer than 50 days.

As shown in Table **??** and Figure **??**, the spread in RERER between the models is again moderate. For the EMEP_rv48 and IFS_v2 models the difference in RERER is slightly larger than for the CO tracer. As for the CO tracer, the highest RERER is in general calculated for the GR+TR region as this region is close to the outer border of the European domain.

Figure **??**~~c~~b,e,h~~and **??**a,b,c~~ ,k,m,o,q shows the annual mean difference in BASE - GLOALL CO concentrations as an average between 30 and 60 degrees north. For all the models large differences in concentrations can be seen over the polluted continents North America, Europe and in particular over East Asia. As for RERER, there are differences between the models, in particular in the free troposphere. The EMEP_rv48 model (Figure **??**b), with high RERER, has higher CO contributions in the free troposphere than the other models. For the 3 models including results for the CO tracer, the results are similar between the CO tracer and CO.

As CO is lifted into the free troposphere transport between continents is rapid, and CO can be transported further before ~~decay, suggesting that the higher RERER is a result~~ decaying. This suggests that as for the CO tracer RERER to a large extent is controlled by the level of rapid lifting and subsequent efficient intercontinental transport in the free troposphere.

The seasonal cycle of the difference in BASE - GLOALL over Europe is shown in Figure **??**, middle panels. As for the CO tracer, differences in concentrations peak near the surface in the first part of the year when emissions are high and the exchange between the boundary layer and the free troposphere is weak. In addition the differences are magnified by the seasonal cycle in the OH sink.

We don't have access to the OH levels for all the models, but for those models providing OH (EMEP_rv4.8, CHASER_re1, OsloCTM3_v2 and CAMchem) annually averaged tropospheric levels are shown in the supplementary material along with the difference between the average and the four individual models. OH levels in the EMEP_rv4.8 model are low compared to the average, at least in the upper and middle troposphere. This may lead us to suspect that the widening gap in RERER from CO tracer to CO between the IFS_v2 and the EMEP_rv4.8 model is caused by differences in OH (however, this can not be confirmed, as OH is not available from the IFS_v2 model). Likewise, the higher than average OH levels in the OsloCTM3_v2 model may explain the lower than average CO RERER values for this model.

Furthermore the lifting of pollutants from the boundary level to the free troposphere is likely to affect the chemistry in the free troposphere ~~this causing (parts of)~~ causing parts of the differences in OH. The EMEP_rv48 model does not perturb aircraft emissions in the BASE-GLOALL scenario, and this could explain ~~large parts~~ some of the differences between this model and the 3 other models. See also discussion on ozone in section 4.3 below.

4.3 O_3

Tropospheric ~~Ozone~~ ozone differs from CO and the CO tracer as it is not emitted, but rather it is a secondary product involving combinations of chemical production and loss processes, exchange with the stratosphere, surface deposition and transport. Ozone in the troposphere is advected from the stratosphere mainly by stratospheric folding events, but its main sources (and sinks) are in the troposphere (TF HTAP, 2010; Stevenson et al., 2006). Net ozone production require ample sunlight and a sufficient supply (and mix) of mainly NMVOC (Non-Methane Volatile Organic Compounds), CH_4 CO and NO_x.

Table **??**, lists annual average RERER, for Europe and for the four European sub regions. RERER is ranging from 0.56 to 1.38, depending on model and European sub-region. As seen in Table **??** and Figure **??** O_3 RERER values are higher than for the CO tracer and for CO~~even though its lifetime in the atmosphere is ranging from 0.01 to 0.05 years (about 4 to 20 days) for ozone, see~~ . Lifetimes for ozone in the troposphere is highly variable, depending on season and altitude, ranging from hours to a few days in the boundary layer to weeks and even months in the free troposphere (TF HTAP, 2010). However, the overall lifetime in the troposphere is shorter than for CO, see also IPCC Working group 1: the scientific basis ~~,).~~ (IPCC WG1, 2001), Table 4.1a. The high RERER values are therefore caused by the non-linear chemistry that for some models can result in RERER values even exceeding one, and for seasonal RERER even negative values (not shown). The spread in RERER between the individual models is markedly larger than for CO and the CO tracer. Differences in transport, depositions and in particular a nonlinear chemistry, give substantial room for variability in ozone levels between the models. In NW Europe little sunlight throughout much of the year as a result of its northerly location and high cloud fractions, in combination with high NO_x emissions, result in ozone titration and calculated RERER around 1 for a majority of the models. The lowest RERER is calculated for the Gr+Tr (Greece + Turkey) and partially SW European regions. The EMEP_rv48 and the IFS_v2 are the only two models where RERER can be calculated for the CO tracer, CO and ozone. Whereas for the ~~IFS_v2 model RERER is lower than the EMEP_rv48 model for the~~ CO tracer and ~~for CO ,~~ CO IFS_v2 RERER jumps to well above one for ozone, well above any of the other models. To a less extent this is also applies to the CAMChem model. The GEOS-Chem and the OsloCTM3_v2 models have the lowest RERER for CO, but is well above and at the ensemble mean respectively for ozone. The CHASER models are close to the ensemble mean for CO, but has the lowest RERER for ozone. The EMEP model has the highest RERER for CO and the CO tracer, but is close to the ensemble mean for ozone. These changes in positions between CO and ozone are likely caused by differences mainly in model chemistry.

Based on the HTAP2 model calculations, Huang et al. (2017) have calculated RERER for the North American continent. In general these RERER values are markedly lower than for Europe. In addition to the effects of little sunlight discussed above, ~~less favourable for local ozone production,~~ Europe is also affected by nearby source regions as Russia, Belarus, Ukraine, the Middle East, North Africa and

shipping. These two factors are likely to explain the higher RERER values over Europe compared to North America.

Figure **??**c,f,i and **??**d,e,f shows the annual longitudinal mean difference in BASE - GLOALL O_3 concentrations as an average between 30 and 60 degrees north. The differences between the models are again markedly larger than for CO and the CO tracer. One notable difference stems from the interpretation of the scenario definition. The ~~Oslo CTM2~~ OsloCTM3_v2 model, CAMchem model and the CHASER models have included a 20% emission reduction also in aircraft emissions in the GLOALL scenario, whereas the EMEP_rv48 model, the IFS_v2~~and the GEOSchem adjoint~~ , the GFDL_AM3 and the GEOS-Chem models have not. As a result the additional ozone from BASE - GLOALL is much higher in the middle and upper troposphere for the first three models listed. For the ~~Oslo CTM3~~ OsloCTM3_v2 model the O_3 signal from aircraft emissions is located much lower in the troposphere than for the CAMchem and CHASER models. O_3 in the lower troposphere, and in particular in the boundary layer, ~~are~~ appears to be not so much affected by ~~the~~ aircraft emissions. ~~But also here the modelsdiffer substantially.~~ Based on several global models, run with and without aircraft emissions (as opposed to 20% perturbations in this study), Cameron et al. (2016) find that aircraft emissions increase near surface ozone by 0.3 to 1.9% globally, with the largest effects in the northern latitudes. In Europe and eastern North America, where population and aircraft emissions are particularly dense, the surface ozone perturbations are smaller than the zonal average.

As is the case for CO and the CO tracer, the EMEP_rv48 model (Figure **??**c), has higher O_3 contributions in the free troposphere than the IFS_v2~~and GEOSCHEMADJ~~ , GFDL_AM3 and GEOS-Chem models (the ~~two~~ three other models not perturbing aircraft emissions). This could be caused by lifting of ozone and ozone precursors from the boundary layer into the free troposphere and subsequent rapid transport between continents in the free troposphere.

The seasonal cycle of the difference in BASE - GLOALL over Europe is shown in Figure **??** right panels. Whereas the contributions from aircraft peaks in summer and autumn, the differences in BASE - GLOALL in general peaks in spring in the lower troposphere except for the CAMchem and GFDL_AM3 models peaking in mid summer. The CAMchem model has very high European net surface ozone contribution in summer compared to contributions from other regions, contributing to the shift in the seasonal maximum from spring into summer. See also discussion in sections 4.4 and 4.5 below.

4.4 European O_3 source ~~allocations~~ allocation by world region

Based on the difference between the BASE model runs and the 20% perturbations of global and European emissions ~~(not accounting for the effects of)~~ we attribute a major portion of ozone of anthropogenic origin in Europe to sources outside Europe. As part of the HTAP2 requests, model calculations have also been made reducing anthropogenic emissions by 20% in other major world regions. In Figure **??** the contributions to European ozone levels calculated by the different models are

shown with sources originating from these different world regions. None of the models have made the calculations for all the regions. For each model the contribution from ~~other regions~~ ROW (Rest Of the World) is calculated by subtracting the ~~added regional contributions from~~ sum of the contributions from from available world regions from the BASE - GLOALL contribution. Thus the portion related to ~~other regions~~ ROW includes a varying ~~number of~~ mixture of world region definitions depending on the model. In addition the percentage contributions to annual average ozone and summer ozone to Europe from the Europe, North America and East Asia, based on the numbers shown in Figure **??**, are shown in Table **??**. The percentage contributions to SOMO35 and POD_1 forest is also given in this table (see section 4.5 for definitions of SOMO35 and POD_1 forest).

There are large differences between the models, ~~but~~ in particular for the contributions of annual ozone from Europe, ranging from -48 to +37 percent. Still, there are some common features: For all models and all seasons except for the CHASER_re1 in summer, the contributions from regions outside Europe are larger than the contribution from European sources. The contributions from non European sources are largest in Spring. The largest non European contributions are from North America (NAMALL) and East Asia (EASALL). Contributions from Russia, Belarus, Ukraine (RBUALL) ~~and the Middle East (MDEALL)~~ are mixed, with significant calculated contributions calculated by two models (EMEP_rv48 and CHASER~~).~~ _re1). Contributions from the middle East (MDEALL) and North Africa (NAFALL) are small. There are also substantial contributions from ocean shipping (OCNALL), but this source has only been calculated by the EMEP_rv48 model. For Europe ~~substantial~~ contributions from shipping has also been shown in other studies as ~~Brandt et al. (2013); Jonson et al. (2015).~~ as Jonson et al. (2015) using the EMEP regional model and Brandt et al. (2013) using a different (non HTAP2) model. For all models, except the CHASER_re1 model, ozone titration dominates the overall European contributions when summed up over the three winter months. However, for all the models, including also the CHASER_re1 model, the net European contributions includes regions of net ozone production and net ozone destruction in winter.

~~Compared to the HTAP1 calculations (TF HTAP, 2010) the contribution to European ozone levels from the world regions differ. Part of the reason may be that the modelsincluded in the~~ The negative, or close to zero, net annual ozone production over Europe in the IFS_v2, GEOS-Chem and CAMChem models can explain the increase in RERER from CO to ozone in Figure **??** discussed in section 4.3. Likewise also the corresponding relative decrease in RERER for the CHASER models, and partially the EMEP_rv48 model can be explained by positive net ozone production over Europe.

In comparison to HTAP1~~and~~ , HTAP2 ~~ensemble are not the same. From 2001 to 2010 emissions of ozone precursors have decreased in USA (see and Europe (Mareckova et al., 2017), and increased in East Asia. Turnock et al. (2018) lists the percentage change in the HTAP1 world regions confirming these emission trends between 2001 (the base year for HTAP1) and 2010. This may explain the 0.37 to 0.22 ppb decrease in the contributions from North America, and the 0.17 to 0.22 ppb increase in the East Asian contribution~~ regions are better defined. In addition emissions as well as models

are up-to-date. To disentangle whether the changes from HTAP1 to HTAP2 are due to emissions, a changed model ensemble or changes in receptor regions is unfortunately not possible in a fully quantitative way. Source and receptor regions have been chosen in HTAP2 to cover the land-only politically connected regions accurately on a 0.1 degree grid. In HTAP1 the EUR region was a simple latitude - longitude box, also including parts of North Africa, the Middle East, Russia, Belarus, Ukraine and large sea areas~~included as separate,~~ , all of these identified as non European regions in HTAP2. In HTAP2 the European region is smaller, thus exporting larger fractions to nearby regions, but most major HTAP1 source regions are located within the smaller HTAP2 region, thus making this region more sensitive to titration effects. As a result ~~emissions~~ the effects of emissions on ozone levels from the EUR region ~~is no longer comparable~~to itself is reduced.

The ensemble mean contribution to ozone levels from Europe to itself has decreased from 0.82 $\pm$0.29 ppb in HTAP1 to just 0.11 $\pm$0.32 ppb in HTAP2. Also - total and regional distribution of emissions for the base year changed from HTAP1 (2001) to HTAP2 (2010). Gaudel et al. (2018) have analysed the ozone trends between the years 2000 and 2014. Over Europe. They found a general ozone increase in the winter months (December, January, February) and a general decrease in the summer months (June, July, August). The emission trends in the HTAP1 world regions are given in Turnock et al. (2018) between 2001 (the base year for HTAP1) and 2010. The changes in measured ozone are consistent with the reductions in European (and North American) emissions of NO_x (along with other ozone precursors) over the same period resulting in less titration and thus increased ozone levels in some areas mainly in the winter months, and simultaneously less net ozone production in summer. Likewise emissions in North America have decreased and may explain the 0.37 $\pm$0.10 to 0.22 $\pm$0.07 ppb decrease in the ensemble mean contributions from North America to European ozone levels. Over the same period emissions in other world regions as East Asia have increased. This increase may explain the 0.17 $\pm$0.05 to 0.22 $\pm$0.13 ppb ensemble mean increase from HTAP1 to HTAP2 in the East Asian contribution to European ozone levels. Contributions from South Asia are small in both HTAP1 and HTAP2 (0.07 versus 0.05).

A combined effect of the change in the definition of the European domain and the changes in emissions is that the relative model calculated contributions to surface ozone levels from non European sources is much larger in HTAP2 compared to HTAP1. In the HTAP1 final report (TF HTAP (2010), Table 4.2) the concept of RAIR (Relative Annual Intercontinental Response), defined as the ratio of the response in a particular region (Europe) due to the combined influence of sources in Europe and the three other regions (North America, East Asia and South Asia) to the response from all these four source regions. RAIR for the models in Figure **??** is 82% as opposed to 43% in the HTAP1 final report.

~~The~~ Using tagging in a regional model the calculated contributions from non European sources have also been calculated by Karamchandani et al. (2017)~~using a regional model. They too~~ . They calculate a much smaller contribution from non European sources than in this study, similar to the

contributions calculated in HTAP1. In the Karamchandani et al. (2017) study non European ozone is defined as the boundary influx to the model domain. As a result shipping, and nearby non Central European regions, are included in the domain, similar to the definition of the HTAP1 European domain.

4.4.1 Effects of a 20% CH_4 perturbation

As shown in Figure **??** four of the models have also calculated the effects of a 20% increase in CH_4 concentrations. Averaged over the four models the calculated effects for Europe of 20% changes in CH_4 levels is almost three quarters of the effects of the BASE - GLOALL model runs. However, ~~a direct comparison of~~ comparing a 20% change in CH_4 concentrations, and the effects of the GLOALL ~~scenario should not be made~~emission scenario requires careful interpretation. Because of its relatively long lifetime of the order of 10 years in the atmosphere, a 20% change in ~~concentrations~~ concentration corresponds to an approximate 40 ~~years of~~ year long historic CH_4 ~~trends~~ trend (Meinshausen et al., 2011). The GLOALL scenario is not accounting for the full impact of a continued 20% reduction in emissions. With a continued emission reduction scenario, the overall ozone reductions would be larger, while the methane attributable fraction, relatively, would be smaller. The effects of CH_4 is insensitive to the location of the emissions, and there are only moderate differences in the response in ozone levels by world region (Fiore et al., 2008). The agreement between the model estimates is a lot better for the CH_4 perturbation compared to the BASE - GLOALL estimates, and not too different for the HTAP1 estimate of about 1 ppb (Fiore et al., 2008). The sensitivity of ozone to CH_4 is discussed in more detail in Turnock et al. (2018).

4.5 Does the choice of ozone metric matter?

In Figure **??** the contributions to European ozone levels are shown as seasonal and annually averaged ozone and in Table **??** the percentage contributions to annual and summer ozone from European, North American and East Asian sources are listed based on the numbers from Figure **??**. In Europe several other metrics are also used calculating the effects of ground level ozone. The two metrics listed below are designed to capture the effects of ground level ozone on human health (SOMO35) and on the environment (POD_1 forest):

- SOMO35: Sum of Ozone Means Over 35 ppb is the indicator for health impact assessment recommended by WHO. It is defined as the yearly sum of the daily maximum of the running 8-hour running average of ozone above 35 ppb.

- POD_1 (deciduous) forest: Phyto-toxic Ozone Dose for forests is the accumulated stomatal ozone flux over a threshold Y integrated from the start to the end of the growing season. For deciduous forests, discussed here, the critical level of 4 mmol m^{-2} is exceeded in most of

Europe, indicating a risk of ozone damage to forests. See Mills et al. (2011a, b) for further description of this metric.

~~Unfortunately the two latter metrics have only been provided by the EMEP model. The annual effects of the 20% reductions in anthropogenic emissions from different world regions are shown for mean ozone, SOMO35 and POD_1 (deciduous) forest in Figure ?? as percentage contributions where 100% refers to the difference betwee the BASE and GLOALL scenario. The figure clearly shows that the choice of metric matters, in particular for the effects of European Emissions.~~

POD_1 forest is only accumulated over the growing season ~~, in particular excluding the winter month with frequent ozone titration from local sources. This metric is mostly accumulated in the summer ,~~ in summer when the contributions from local European sources are high. Likewise SOMO35, with a cutoff value at 35 ppb, is accumulated mainly in the summer months~~when local ozone production peaks over the European continent~~. Thus both metrics these netrics largely exclude the effects of ozone titration mainly taking place in other seasons.

Contributions to annual mean ozone are accumulated regardless of season and ambient ozone levels. In the EMEP_rv48 model contributions from ~~NAM and EAS~~ NAMALL and EASALL have already been shown to be little affected by ozone titration and a major source mainly in the spring months before the local European sources gathers momentum. Contributions from ~~RBU and OCN~~ RBUALL and OCNALL are a mixture of nearby and more distant sources, and effects on annual mean ozone, SOMO35 and POD_1 forest are similar. It is likely that the difference between the ozone metrics would be considerably larger if calculated with the other models, and in particular those models with substantial titration effects from European Emissions as already shown in Figure **??**.

Unfortunately the two latter metrics have only been provided by the EMEP_rv48 model. The annual effects of the 20% reductions in anthropogenic emissions from different world regions are shown for mean ozone, SOMO35 and POD_1 forest in Figure **??** as percentage contributions where 100% refers to the difference between the BASE and GLOALL scenario. The regional contributions, expressed by these metrics, are also listed in table **??**. The figure and table clearly shows that the choice of metric matters, in particular for the effects of European Emissions. POD_1 forest is accumulated in the growing season in summer. A large portion of SOMO35 is also accumulated in the summer months. Table **??** also lists the percentage contributions to summer ozone for all models. The similarities in the percentages for summer ozone and the ozone metrics in EMEP_rv48 is an indication that also for the other models these percentages are comparable.

5 Discussion on individual models

As shown above ~~there are large~~ differences between the models ~~. Furthermore, these differences~~ amplify going from the simple CO tracer, via CO, to ozone. This stepwise amplification provides an opportunity to pinpoint probable causes. At the same time we also use the comparisons to

measurements as a guidance. Some of the results from the individual model calculations are summed up in Table **??**. Below we discuss the characteristics and the results for the individual models. Here we try to point out if, and at what stage, the results from the individual models deviate from the other models. It should be stressed that such a deviation does not necessarily imply that the results from a particular model is wrong.

The horizontal resolution of the EMEP_rv48 model is 0.5 $\times$ 0.5 degrees, higher than any of the other models. Compared to the other models, the difference between BASE and GLOALL is among the highest compared to the other models for CO and the CO tracer. Much of this may be caused by a larger rate of exchange (possibly by convection) between the boundary layer and the free troposphere. On the other hand this model performs among the best both for CO and ozone compared to measurements. Calculated CO levels at remote sites are not high, see Table **??**) and supplementary material, compared to the other models. ~~Ozone sondes show~~ The model is one of the models with highest overestimation of ozone in the free troposphere in the winter and spring months. The EMEP_rv48 model differs from the other models by having a larger rate of exchange between the boundary layer and the free troposphere.

The horizontal resolution of the IFS_v2 model is 0.7 $\times$ 0.7 degrees. The RERER results for CO are close to the ensemble mean and CO levels close to observations. For ozone RERER is higher than the other models, and above 1 in all European regions except Greece and Turkey. European net Ozone production is strongly affected by ozone titration resulting in net ozone loss from European sources ~~for~~ in all seasons except summer. Calculated ozone levels in Europe are low compared to measurements, in particular for low ozone sites. The IFS_v2 model differs from the other models by having the highest level of ozone titration. The underestimation of ozone at low ozone sites is most likely caused by the high level of titration.

The horizontal resolution of the OsloCTM3_v2 model is 2.8 $\times$ 2.8 degrees. ~~For CO RERER is well below the model ensemble mean. A possible reason for this is that the~~ The advection is solved using the Prather scheme, giving very little numerical diffusion~~, possibly restraining the vertical exchange more than for the other models~~. For CO RERER is well below the model ensemble mean. The model ~~Underestimates~~ underestimates CO, and overestimates O_3 compared to measurements. For CO the low RERER and the underestimation of surface CO compared to measurements could be affected by higher OH values compared to the other models.

The two models CHASER_re1 (resolution 2.8$\times$ 2.8 degrees) and CHASER_t106 (resolution 1.1$\times$ 1.1 degrees) differ only in resolution, and results from the two models are very similar. RERER for CO is close to ensemble mean. RERER for ozone almost 30% lower than ensemble mean. The CHASER models differs from the other models by having lower RERER for ozone and little or no ozone titration over Europe even in winter. The lack of ozone titration may be the cause of the overestimation of ozone at low ozone sites seen in the ozone scatter plot shown in the supplement.

The horizontal resolution of the ~~GEOSCHEMADJ~~ GEOS-Chem model is 2.0 × 2.5 degrees. CO concentrations ~~underestimated by~~ on acerage underestimated by more than 20 ~~+~~ percent. O_3 concentrations overestimated by 14%. ~~The vertical profiles of ozone relax to zero~~ O_3 is only simulated in the troposphere and ozone levels above the tropopause are based on boundary concentrations (see supplementary material) and should be disregarded here. Like most models the GEOS-Chem model underestimates CO and overestimates O_3 in EU. The GEOS-Chem model has the lowest RERER value for CO, but at the same time a high RERER for ozone. It has high ozone titration in winter and high European ozone production in summer. As for the IFS_v2 model the underestimation of ozone at low ozone sites is most likely caused by the high level of titration.

RERER calculated by the GFDL_AM3 model is close to the ensemble mean for both CO and O_3. RERER for CO 20% below ensemble mean. RERER o3 17% higher than the ensemble mean.

The horizontal resolution of the CAMchem model is 1.9 × 2.5 degrees. CO concentrations are on average underestimated by 25% and O_3 concentrations are overestimated by 22%. RERER is close to ensemble mean for both CO and ~~O3~~O_3. Similar to the GEOS-Chem model the CAMchem model has high RERER for ozone in combination with high ozone titration in winter and high European ozone production in summer. The high net ozone production in summer is the likely cause for the shift in the O_3 maximum for BASE - GLOALL from Spring to Summer in the lower troposhere above Europe.

6 Conclusions

The HTAP1 experiment showed a very large spread in model results. (TF HTAP, 2010). Part of this spread may have been caused by differences in the 2001 emissions, as each modelling group used their own set of emissions. In HTAP2 all models are required to use a common set of emissions. Even so, the spread in model results remains large. The model calculated relative contributions to surface ozone levels from non European sources is much larger in HTAP2 compared to HTAP1. Mainly because the contributions from Europe to it selves has decreased from 0.82 ppb to just 0.11 ppb. As a result RAIR has increased from 43 to 82%. In parts differences could be explained by decreasing emissions in ~~North America and~~ Europe and increased emissions in most other regions as East Asia from year 2001 to 2010. However, the results from the two HTAP phases can not easily be compared, partially because the model ensemble has changed, but mainly because the definition of the European area has changed considerably from HTAP1 to HTAP2. The HTAP2 source and receptor regions are better designed for characterising export and import of air pollution to and from the individual regions. For HTAP2 additional diagnostics were defined which allow better understanding of transport efficiencies, such as the utilisation of idealized CO tracer and more information on the vertical distribution of tracers in the output requirements.

Not surprisingly, our study reveals that the magnitude of the inter-model spread in hemispheric transport, characterised by RERER, increases with the complexity of the processes involved. We demonstrate that the spread in European RERER increases from the ~~idealized~~ idealised CO tracer to fully prognostic CO and ozone. Atmospheric transport alone can not be made responsible for the ~~large RERER difference between CO and ozone, as~~ larger spread between the models in RERER going from CO to ozone. As the residence time in the troposphere is ~~of the order of 0.01 to 0.05 years for ozone, and for CO considerably longer, 0.08 to 0.25 years, see IPCC Working group 1: the scientific basis,). For comparison the 50 days lifetime of the CO tracer translates to about 0.14 years. The~~ longer for CO compared to ozone (see discussion in sections 4.2 and 4.3). the increase in RERER from CO to O_3 ~~is likely~~ must be caused by more complex non-linear chemistry forming and destroying ozone and not by a longer atmospheric lifetime of O_3 compared to CO.

The model resolution differs between the individual models. Model results from the two CHASER models, differing in model resolution only, are qualitatively similar when compared to measured CO and O_3 at background measurement sites ~~. Horizontal resolution does not affect~~ and very similar in RERER for CO and O_3, suggesting that resolution differences at the scales investigated here, are not important to explain RERER differences between the global models. Still, it is difficult to conclude in general to what extent horizontal resolution affects the source receptor calculations ~~much at intercontinental sales~~at intercontinental scales.

The joint and consistent analysis of a CO tracer, CO and O_3 in this paper is a tool in understanding where and why (right or wrong) the models differ, however, it ~~could probably be used more. and as a result~~ has a potential for wider use, enhancing our understanding of the result ~~an~~ and also as a tool for model improvements, reducing the overall uncertainty in future model calculations. We believe that in order to close the gap in model results, and subsequently improving the reliability of the model output, possible future model inter-comparisons should be more process oriented (transport, depositions, chemistry etc). ~~The largest spread in model results is clearly induced by differences in the model chemistry.~~ Our study shows ~~that the models differ markedly~~ that models differ already for CO ~~(and CO tracer~~ and the inert CO tracer, where differences were established with 2 models~~). We believe that a comparison of the chemical mechanisms used in the models , in combination with an extensive evaluation with atmospheric measurements, may be a first step improving the models~~, but that differences are amplified as more chemistry is added. Note that the CO RERER and O_3 RERER values are not correlated taken the models as samples. The big additional spread in model results for ozone is clearly induced by differences in model chemistry exemplified by the treatment of titration in the winter boundary layer. However, differences in chemistry may well also be induced by differences in advection/convection as the level of exchange will inevitably affect the chemical regime in both the free troposphere and in the boundary layer. We therefore believe that further process oriented evaluations (comparing advection/convection, chemistry, dry and wet deposition etc separately) should be made, making use of relevant meteorological and chemical measurements.

The HTAP2 results, using state of the art global models, reflecting updated emission estimates and refined receptor region definitions, confirm the importance of hemispheric transport of air pollution. Based on seasonal and annual averaged ozone, all the models agree that the contribution from non European sources to European surface ozone levels is considerable. However, calculations with the EMEP_rv4.8 model shows that this conclusion to some extent will depend on the choice of ozone metrics. Alternative metrics, such as SOMO35 and POD_1 forest, will to a larger extent accumulate in the summer months when ozone production peaks over the European continent. The dependence on ozone metrics seen in the EMEP_rv4.8 model is corroborated by the other HTAP2 models all showing the effects of summer ozone pointing in the same direction. As a result the potential for reducing the detrimental effects from ozone caused by European emissions alone is higher when applying these metrics.

The model results suggest that ~~it will be difficult to achieve~~ sizeable reductions in ~~ozone levels with European emission reductions alone, and that reductions in the emissions of ozone precursors should be made in~~ European ozone levels can best be achieved through a combined global effort (or at least throughout the northern hemisphere) ~~. Emissions~~ to reduce the emissions of ozone precursors. Efforts to curb regional pollution in other non European regions, exemplified by the reductions in North American emissions of ozone precursors~~have already been reduced~~ , have most likely reduced the ozone burden also in Europe. Further reductions in the Emissions of ozone precursors are expected in Europe and North America~~and are expected to decrease further here~~. However, ~~this decrease~~ decreases here has so far been partially counteracted by increases elsewhere. Other regions, such as East Asia, are currently facing severe air pollution problems. Part of the remedy for the elevated European ozone levels may well be local and regional air pollution control to curb air pollution also in these regions.

Acknowledgements. This work has been partially funded by EMEP under UNECE. Computer time for EMEP model runs was supported by the Research Council of Norway through the NOTUR project EMEP (NN2890K) for CPU, and NorStore project European Monitoring and Evaluation Programme (NS9005K) for storage of data. The AeroCom database at Met Norway received support from the CLRTAP under the EMEP programme, through the service contract to the European commission no. 07.0307/2011/605671/SER/C3, and benefited from the Research Council of Norway project no. 229796 (AeroCom-P3) and no 235548 (SLCF). The National Center for Atmospheric Research is funded by the National Science foundation. The University of Colorado has received support from NASA under grant NNX16AQ26G. We would also like to thank WOUDC for making the ozonesonde measurements available. Some data used in this publication were obtained as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available through http://www.ndacc.org. EMEP surface measurements have been made available through the EBAS web site, http://ebas.nilu.no/Default.aspx.

Please add after "and benefited from the Research Council of Norway project no. 229796 (AeroCom-P3)" another project code:

"and benefited from the Research Council of Norway project no. 229796 (AeroCom-P3) and no 235548 (SLCF).1000

References

Brandt, J., Silver, J. D., Christensen, J. H., Andersen, M. S., Bønløkke, J. H., Sigsgaard, T., Geels, C., Gross, A., Hansen, A. B., Hansen, K. M., Hedegaard, G. B., Kaas, E., and Frohn, L. M.: Assessment of past, present and future health-cost externalities of air pollution in Europe and the contribution from international ship traffic using the EVA model system, Atmospheric Chemistry and Physics, 13, 7747–7764, doi:10.5194/acp-13-7747-2013, https://www.atmos-chem-phys.net/13/7747/2013/, 2013.

Cameron, M. A., Jacobson, M. Z., Barrett, S. R. H., Bian, H., Chen, C. C., Eastham, S. D., Gettelman, A., Khodayari, A., Liang, Q., Selkirk, H. B., Unger, N., Wuebbles, D. J., and Yue, X.: An intercomparative study of the effects of aircraft emissions on surface air quality, Journal of Geophysical Research: Atmospheres, 122, 8325–8344, doi:10.1002/2016JD025594, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025594, 2016.

Casper-Anenberg, S., West, J., Fiore, A., Jaffe, D., Prather, M., Bergman, D., Cuvalier, K., Dentener, F., Duncan, B., Gauss, M., Hess, P., Jonson, J., Lupu, A., MacKenzie, I., Marmer, E., Park, R., Sanderson, M., Schultz, M., Shindell, D., Szopa, S., Vivanco, M., Wild, O., and Zeng, G.: Intercontinental impacts of ozone pollution on human mortality, Environ. Sci. Tech., 43, doi:10.1021/es900518, 2009.

Fiore, A., West, J., Horowitz, L., and Schwarzkopf, M.: Characterizing the tropospheric ozone response to methane emission controls and the benefits to climate and air quality, J. Geophys. Res., 113, doi:10.1029/2007JD009162, 2008.

Fiore, A., Dentener, F., Wild, O., Cuvelier, C., Schultz, M., Textor, C., Schulz, M., Atherton, C., Bergmann, D., Bey, I., Carmichael, G., Doherty, R., Duncan, B., Faluvegi, G., Folberth, G., Garcia Vivanco, M., Gauss, M., Gong, S., Hauglustaine, D., Hess, P., Holloway, T., Horowitz, L., Isaksen, I., Jacob, D., Jonson, J., Kaminski, J., keating, T., Lupu, A., MacKenzie, I., Marmer, E., Montanaro, V., Park, R., Pringle, K., Pyle, J., Sanderson, M., Schroeder, S., Shindell, D., Stevenson, D., Szopa, S., Van Dingenen, R., Wind, P., Wojcik, G., Wu, S., Zeng, G., and Zuber, A.: Multi-model estimates of intercontinental source-receptor relationships for ozone pollution, J. Geophys. Res., 114, doi:10.1029/2008JD010816, 2009.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmospheric Chemistry and Physics, 17, 1543–1555, doi:10.5194/acp-17-1543-2017, https://www.atmos-chem-phys.net/17/1543/2017/, 2017.

Galmarini, S., Kioutsioukis, I., Solazzo, E., Alyuz, U., Balzarini, A., Bellasio, R., Benedictow, A. M. K., Bianconi, R., Bieser, J., Brandt, J., Christensen, J. H., Colette, A., Curci, G., Davila, Y., Dong, X., Flemming, J., Francis, X., Fraser, A., Fu, J., Henze, D., Hogrefe, C., Im, U., Garcia Vivanco, M., Jiménez-Guerrero, P., Jonson, J. E., Kitwiroon, N., Manders, A., Mathur, R., Palacios-Peña, L., Pirovano, G., Pozzoli, L., Prank, M., Schultz, M., Sokhi, R. S., Sudo, K., Tuccella, P., Takemura, T., Sekiya, T., and Unal, A.: Two-scale multi-model ensemble: Is a hybrid ensemble of opportunity telling us more?, Atmospheric Chemistry and Physics Discussions, 2018, 1–44, doi:10.5194/acp-2018-86, https://www.atmos-chem-phys-discuss.net/acp-2018-86/, 2018.

Gaudel, A., Cooper, O., Ancellet, G., Barret, B., Boynard, A., Burrows, J., Clerbaux, C., Coheur, P.-F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garcia, O., Granados Muños, M., Hannigan, J., Hase, F., Huang, G., Hassler, B., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik,

S., Latter, B., Leblanc, T., Le Flochmoën, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., , Querel, R., Rahpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V., , Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M., Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, Elem. Sci. Anth., 6, 39, doi:10.1525/elementa.291, 2018.

Huang, M., Carmichael, G. R., Pierce, R. B., Jo, D. S., Park, R. J., Flemming, J., Emmons, L. K., Bowman, K. W., Henze, D. K., Davila, Y., Sudo, K., Jonson, J. E., Tronstad Lund, M., Janssens-Maenhout, G., Dentener, F. J., Keating, T. J., Oetjen, H., and Payne, V. H.: Impact of intercontinental pollution transport on North American ozone air pollution: an HTAP phase 2 multi-model study, Atmospheric Chemistry and Physics, 17, 5721–5750, doi:10.5194/acp-17-5721-2017, https://www.atmos-chem-phys.net/17/5721/2017/, 2017.

IPCC WG1: IPCC Third Assessment Report: Climate Change 2001. Working Group I: The Scientific Basis, https://www.ipcc.ch/ipccreports/tar/wg1/130.htm#tab41a/, 2001.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmospheric Chemistry and Physics, 15, 11 411–11 432, doi:10.5194/acp-15-11411-2015, http://www.atmos-chem-phys.net/15/11411/2015/, 2015.

Jonson, J., Stohl, A., Fiore, A., Hess, P., Szopa, S., Wild, O., Zeng, G., Dentener, F., Lupu, A., Schultz, M., Duncan, B., Sudo, K., Wind, P., Schulz, M., Marmer, E., Cuvelier, C., Keating, T., Zuber, A., Valdebenito, A., Dorokhov, V., De Backer, H., Davies, J., Chen, G., Johnson, B., Tarasick, D., Stübi, R., Newchurch, M., von der Gathen, P., Steinbrecht, W., and Claude, H.: A multi-model analysis of vertical ozone profiles, Atmos. Chem. Phys., 10, 5759–5783, doi:10.5194/acp-10-5759-2010, 2010.

Jonson, J. E., Jalkanen, J. P., Johansson, L., Gauss, M., and Denier van der Gon, H. A. C.: Model calculations of the effects of present and future emissions of air pollutants from shipping in the Baltic Sea and the North Sea, Atmospheric Chemistry and Physics, 15, 783–798, doi:10.5194/acp-15-783-2015, https://www.atmos-chem-phys.net/15/783/2015/, 2015.

Karamchandani, P., Long, Y., Pirovano, G., Balzarini, A., and Yarwood, G.: Source-sector contributions to European ozone and fine PM in 2010 using AQMEII modeling data, Atmospheric Chemistry and Physics, 17, 5643–5664, doi:10.5194/acp-17-5643-2017, https://www.atmos-chem-phys.net/17/5643/2017/, 2017.

Koffi, B., Dentener, F., Janssens-Maenhout, G., Guizzardi, D., Crippa, M., Diehl, T., Galmarini, S., and Solazzo, E.: Hemispheric Transport Air Pollution (HTAP): Specification of the HTAP2 experiments – Ensuring harmonized modelling Establishing ozone critical levels II. UNECE Workshop Report, Eur 28255 en – scientific and technical research reports, doi:10.2788/725244, 2016.

Liu, G., Tarasick, D., Fioletov, V., Sioris, C., and Rochon, Y.: Ozone correlations lengths and measurement uncertainties from analysis of historical ozonesonde data in North America and Europe, J. Geophys. Res., 114, doi:10.1029/2008JD010576, 2009.

Mareckova, K., Pinterits, M., Tista, M., Ullrich, B., Wankmüller, R., and Nyíri, A.: Emissions for 2015, In *Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components.* EMEP/MSC-W Status Report 1/2017, The Norwegian Meteorological Institute, Oslo, Norway, 2017.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213, doi:10.1007/s10584-011-0156-z, https://doi.org/10.1007/s10584-011-0156-z, 2011.

Mills, G., Hayes, F., Simpson, D., Emberson, L., Norris, D., Harmens, H., and Büker, P.: Evidence of widespread effects of ozone on crops and (semi-)natural vegetation in Europe (1990-2006) in relation to AOT40- and flux-based risk maps, Global Change Biology, 17, 592–613, doi:10.1111/j.1365-2486.2010.02217.x, 2011a.

Mills, G., Pleijel, H., Braun, S., Büker, P., Bermejo, V., Calvo, E., Danielsson, H., Emberson, L., Grünhage, L., Fernández, I. G., Harmens, H., Hayes, F., Karlsson, P.-E., and Simpson, D.: New stomatal flux-based critical levels for ozone effects on vegetation, Atmos. Envir., 45, 5064 – 5068, doi:10.1016/j.atmosenv.2011.06.009, 2011b.

Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J.-F., Lin, M., Prather, M. J., Young, P. J., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R., Eyring, V., Faluvegi, G., Folberth, G. A., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., van Noije, T. P. C., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R., Shindell, D. T., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., and Zeng, G.: Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmospheric Chemistry and Physics, 13, 5277–5298, doi:10.5194/acp-13-5277-2013, https://www.atmos-chem-phys.net/13/5277/2013/, 2013.

Reidmiller, D., Fiore, A., Jaffe, D., Bergmann, D., Cuvelier, C., Dentener, F., Duncan, B., Folberth, G., Gauss, M., Gong, S., Hess, P., Jonson, J., Keating, T., Lupu, A., Marmer, E., Park, R., Schultz, M., Shindell, D., Szopa, S., Vivanco, M., Wild, O., and Zubera, A.: The influence of foreign vs. North American emissions on surface ozone in the US2, Atmos. Chem. Phys.Discuss., 9, 7.927–7.969, 2009.

Sanderson, M., Dentener, F., Fiore, A., Cuvelier, K., Keating, T., Zuber, A., Atherton, C., Bergmann, D., Diehl, T., Doherty, R., Duncan, B., Hess, P., Horowitz, L., Jacob, D., Jonson, J., Kaminski, J., Lupu, A., Mackenzie, I., Mancini, E., Marmer, E., Park, R., Pitari, G., Prather, M., Pringle, K., Schroeder, S., Schultz, M., Shindell, D., Szopa, S., Wild, O., and Wind, P.: A multi-model study of the hemispheric transport and deposition of oxidised nitrogen, Geophys. Res. Lett., 35, doi:10.1029/2008GL035389, 2008.

Schaap, M., Cuvelier, C., Hendriks, C., Bessagnet, B., Baldasano, J., Colette, A., Thunis, P., Karam, D., Fagerli, H., Graff, A., Kranenburg, R., Nyíri, A., Pay, M., Rouil, L., Schulz, M., Simpson, D., Stern, R., Terrenoire, E., and Wind, P.: Performance of European chemistry transport models as function of horizontal resolution, Atmos. Envir., 112, doi:10.1016/j.atmosenv.2015.04.003, 2015.

Shindell, D., Chin, M., Dentener, F., Doherty, R., Faluvegi, G., Fiore, A., Hess, P., Koch, M., MacKenzie, I., Sanderson, M., Schultz, M., Schulz, M., Stevenson, D., Teich, H., Textor, C., Wild, O., Bergmann, D., Bey, I., Bian1, H., Cuvelier, C., Duncan1, B., Folberth, G., Horowitz, L., Jonson, J., Kaminski1, J., Marmer, E., Park, R., Pringle, K., Schroeder, S., Szopa, S., Takemura, T., Zeng, G., Keating, T., and Zuber, A.: A multi-model assessment of pollution transport to the Arctic, Atmos. Chem. Phys., pp. 5.353–5.372, 2008.

Stevenson, D., Dentener, F., Schultz, M., Ellingsen, K., van Noije, T., Wild, O., O.Zeng, Amann, M., Atherton, C., bell, N., Bergmann, D., Bey, I., Butler, T., Cofala, J., Collins, W., Doherty, R. D. R., Drevet, J., Askes, H., Fiore, A., Hauglustaine, M. G. D., Horowitz, L., Isaksen, I., Lamarque, M. K. J., Lawrence, M., Monanaro, V., Müller, J., Pyle, G. P. M. P. J., Rast, S., Rodriguez, J. M., Sanderson, M., Savage, N., Shindell, D., Strahan, S., Sudo, K., and Szopa, S.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone, J. Geophys. Res., 111, doi:10.1029/2005JD006338, 2006.

Stjern, C. W., Samset, B. H., Myhre, G., Bian, H., Chin, M., Davila, Y., Dentener, F., Emmons, L., Flemming, J., Haslerud, A. S., Henze, D., Jonson, J. E., Kucsera, T., Lund, M. T., Schulz, M., Sudo, K., Takemura, T., and Tilmes, S.: Global and regional radiative forcing from 20 % reductions in BC, OC and SO_4 – an HTAP2 multi-model study, Atmospheric Chemistry and Physics, 16, 13 579–13 599, doi:10.5194/acp-16-13579-2016, http://www.atmos-chem-phys.net/16/13579/2016/, 2016.

Strode, S. A., Duncan, B. N., Yegorova, E. A., Kouatchou, J., Ziemke, J. R., and Douglass, A. R.: Implications of carbon monoxide bias for methane lifetime and atmospheric composition in chemistry climate models, Atmospheric Chemistry and Physics, 15, 11 789–11 805, doi:10.5194/acp-15-11789-2015, https://www.atmos-chem-phys.net/15/11789/2015/, 2015.

TF HTAP: Hemispheric transport of air pollution. Part A: Ozone and partuculate matter edited by: Frank Dentener, Terry Keating, and Hajime Akimoto, http://www.htap.org/publications/2010_report/2010_Final_Report/HTAP%202010%20Part%20A%20110407.pdf, 2010.

Turnock, S., Wild, O., Dentener, F., Davila, Y., Emmons, L., Flemming, J., Folberth, G., Henze, D., Jonson, J., Keating, T., Kengo, S., Lin, M., Lund, M., Tilmes, S., and O'Connor, F.: The Impact of Future Emission Policies on Tropospheric Ozone using a Parameterised Approach, Atmospheric Chemistry and Physics Discussions, 2018, 1–41, doi:10.5194/acp-2017-1220, https://www.atmos-chem-phys-discuss.net/acp-2017-1220/, 2018.