

Dear referee#2

Thank you very much for your detailed review of our manuscript acp-2019-715. Please find our replies to your comments below. In the following, referee comments are given in *italics*, our replies in normal font, and text passages which we included in the text are in **bold**.

Mertens et al. perform a source attribution study examining the contribution of different emission sectors to air pollution over Europe, with a focus on ozone as a pollutant, a special focus on emissions from the road transport sector, and a regional focus on Europe and Germany. They employ a uniform methodology for "tagging" the emissions of ozone precursors in a system of coupled models, allowing a consistent downscaling to be made from the global scale to the national scale. Furthermore, they compare two simulations performed with different emission inventories, showing the sensitivity of the sectoral contributions to the way in which the emissions from each sector are represented in the emission inventory. This combination of sensitivity and source attribution reveals some interesting information about the behaviour of tropospheric ozone in the model system used, for example the particularly strong differences in the contribution of land transport emissions to the higher percentiles of the ozone distribution when a spatially more explicit inventory is used.

Reply: We thank referee#2 for this elaborate summary and the detailed review which helped to improve the manuscript considerably.

The manuscript is clearly within the scope of ACP, and the method clearly has a lot of potential to inform international air quality policy. Unfortunately the manuscript in its current form suffers from a number of serious flaws, which must be corrected before it can be accepted for publication.

Reply: We thank you for your overall positive comment. In accordance with your comments (see below for details) and the comments from referee#1 we strongly revised parts of the manuscript. Currently, a final proofreading is performed after which we will upload the revised manuscript. We hope it can then be accepted for publication.

Firstly, the quality of the written English is terrible. The manuscript is littered with grammatical and spelling errors, and written in a generally inaccessible style. I do not feel that it is my job as a reviewer to provide an exhaustive list of these errors. The authors should seek additional help to get the language up to an acceptable standard. I will give one example though: the very title of the manuscript contains a jarring error. The current title basically implies that ozone causes land transport emissions. Clearly this is the other way around. Land transport emissions happen first, and this leads to ozone production. A grammatically correct title could be "Attributing ozone and its precursors to land transport emissions in Europe and Germany".

Reply: We are really sorry for the errors in the first draft of the manuscript. Of course it is not the work of the referees to perform a language editing. We checked the manuscript in detail and corrected many errors. In addition we revised the title according to the suggestion of referee#2.

In the abstract, the authors state that tagging is "required" and that their method is the "only possible" way to examine global to regional scale effects. This language is way too strong and should be toned down before publication. This is especially true given that the authors themselves state on line 28 of page 25 that their results are "consistent" with a perturbation study, and also given the fact that the experiment design doesn't actually make a distinction between land transport emissions in Europe and the rest of the world.

Reply: We think that the scientific community agrees that tagging methods are the only correct way to calculate contributions (for non-linear species). Impacts and contributions can be similar, but they answer completely different questions. Therefore we don't agree that the language of the sentence '[...]the contribution of land transport emissions to tropospheric ozone cannot be calculated or measured directly, instead atmospheric-chemistry models equipped with specific source apportionment methods (called tagging) are required' is too strong. Yet, we rephrased the other sentence to:

We investigate the combined effect of long range transported ozone and ozone which is produced by European emissions by applying a tagging method simultaneously and consistently on the global and regional scale.

Page 2, lines 26-27: while this is generally true on very small scales (eg. urban areas), the response of ozone to perturbation of precursor emissions in remote regions has been shown to be approximately linear. See for example Wild et al. (2012) and Turnock et al. (2018). Since the authors are also discussing long-range transport, some additional discussion of this here would be relevant.

Reply: We agree with referee#2 that in remote regions (i.e. with low NO_x mixing ratios) the ozone chemistry is (almost) linear w.r.t. NO_x and VOC perturbations. However, with increasing NO_x mixing ratios the chemistry of ozone cannot be considered any more as linear (see for example Fig. 1 in Grewe et al., 2012). Concerning your comment we think it is essential to further discuss the differences of the tagging and the perturbation method to clarify this point. The perturbation approach is based on a Taylor approximation around a base state w.r.t. the chemical regime (called x_0). The goal of this approach is to estimate a sensitivity (e.g. $d\text{O}_3/dE$, where E are the emissions) of the ozone chemistry around x_0 by a Taylor approximation. This sensitivity can be used to estimate a response of ozone on emission changes (as done by Wild et al., 2012). Clearly, this approximation is only valid around x_0 , but not for a different base state $\tilde{x}_0$. Further, only for small perturbations non-linear effects can

be neglected (first elements of Taylor series). This means that in regions with large NO_x emissions non-linear effects can be neglected only for very small perturbations (e.g. 5 %) around x_0 . As the approximation is only valid around x_0 an extrapolation to larger perturbations leads to larger errors. Therefore, Wild et al., 2012 introduces a non-linearity factor (equation 6 therein) to account for non-linearities for much larger perturbations as the original 20 %. Further they state: 'For emission reductions greater than 60 % this correction remains insufficient, and we do not expect the parameterization to work as well under these conditions.' Clearly, for such large perturbations the Taylor approximation is not valid anymore. This has been discussed in great detail also by Grewe et al., 2010. The tagging approach, however, works in a completely different way. It does not consider the sensitivity of the ozone chemistry to an emission change. Instead, it attributes ozone at any base state w.r.t. the chemistry x_y to the corresponding emissions. Thus, the non-linearities are implicitly taken into account. However, the tagging approach gives no information about the sensitivity of an emission change (e.g. $d\text{O}_3/dE$). In addition, Wild et al. (2012) clearly state:

'The 20 % emission perturbations applied in the HTAP studies were chosen to be small enough to give an approximately linear response while being sufficiently large to provide robust signals in all models. However, the response of O_3 to its precursor emissions is known to be non-linear (e.g. Lin et al., 1988), and it is important to characterize where these non-linearities become significant. Scaling a 20 % emission reduction by a factor of five has been shown to underestimate the response to a 100 % reduction (Wu et al., 2009), and while this underestimation is relatively small for VOC emissions, generally less than 10 %, it can exceed a factor of two for NO_x emissions (Wu et al., 2009; Grewe et al., 2010), and shows a strong seasonal dependence (Wu et al., 2009). For this reason the sensitivity approach used in the HTAP studies is unsuitable for deriving a full source apportionment for O_3 . However, it does not preclude its use in estimating the impact of less severe emission changes.'

This is clearly in line with our argumentation. Yet, to clarify this point in the revised manuscript we changed the paragraph accordingly:

For a chemical specie that is controlled by linear processes, the perturbation and the tagging approaches lead to identical results, however, the ozone chemistry is strongly non-linear. Accordingly, the response of ozone on small emission changes can be considered as almost linear, but it does not allow a complete ozone source apportionment (e.g. Wild et al., 2012). As an example, Emmons et al. (2012) have reported that tagged ozone is 2–4 times larger than the contribution calculated by the perturbation approach. As has been outlined in numerous publications, this difference is due to different questions these methods answer. The perturbation approach investigates the impact of an emission change on the mixing ratios of ozone and is therefore well suited to evaluate for example mitigation op-

tions. The tagging approach quantifies the contribution of specific emission sources onto the ozone budget for a given state of the atmosphere (Wang et al., 2009; Emmons et al., 2012; Grewe et al., 2017; Clappier et al., 2017; Mertens et al., 2018). These contributions do, however, not necessarily change linearly with potential changes in emissions. The difference between the results from the perturbation and tagging approaches can actually be used as an indicator for the degree of non-linearity of the chemistry as pointed out Mertens et al. (2018) in their equation 6. In the following we use the terms 'impact' to indicate results from perturbation approaches and 'contribution' to refer to results of tagging methods. In this study, we are interested in the contribution of land transport emissions to ozone in Europe. Therefore, we chose a tagging method for source apportionment.

Page 4, lines 3-4: Aren't the last two points in this list in fact exactly the same thing?

Reply: We are sorry for the confusion the sentence caused in the original manuscript. Point 3 is dedicated to year to year variability (e.g. years with large biomass burning emissions or summer heatwaves). Point 4 is dedicated to the seasonal variability (e.g. strong biogenic emissions in summer). To clarify this we rephrased the part in the revised manuscript (see our reply to the next point).

Page 4, line 5: I can see how using two different inventories can somewhat account for uncertainties in the emissions, but three years is way to short a period to account for interannual variability. I also do not see how the model uncertainty or the uncertainty in the choice of source apportionment method is accounted for at all in this experiment design. It's fine to mention that there can be a lot of uncertainty, but the authors should not claim to be doing more to address these uncertainties than they actually are.

Reply: Of course three years are not enough to catch the full range of inter-annual variability. Referee#2 is completely right that we do not account for model and/or methodological uncertainties (e.g. different source apportionment methods). During the writing process of the manuscript we changed the order of the four points, but forgot to change the sentence on p4 l5. We clarified this accordingly. The changed paragraph reads:

Typically, the uncertainties of such source apportionment studies are large. Reasons are:

- **uncertainties in the models (e.g. chemical/physical parametrizations) and through the choice of source apportionment methods;**
- **uncertainties of the emissions inventories;**

- seasonal variability of the contributions caused by meteorological conditions and seasonal cycles of emissions (e.g. stronger biogenic emissions and more active photochemistry during summer than winter);
- year to year variability of the contributions caused by meteorological conditions or large emissions of specific sources in specific years (for example yearly differences of biomass burning emissions);

To account for the uncertainties due to different emission inventories we performed simulations with two different anthropogenic emission inventories. To further account for the seasonal variability we investigate the contributions for winter and summer seasons. In addition, we consider always three simulation years to gain insights in the variability of the contribution in different years. The investigation of uncertainties caused by models and/or source apportionment methods is beyond the scope of this study.

Page 5, lines 32-35: These are the only lines in the paper where the authors discuss model evaluation. I understand that the model has been evaluated elsewhere, and the model is basically as good, bad as other models, but I would appreciate some more discussion about how the model performance could be expected to influence the conclusions of the manuscript. Since the authors also want to use their model to examine extreme ozone events (in Section 4.2), there must be at least some analysis of how well the model is capable of representing these events in comparison with observations.

Reply: We added section (Sect. 2.5) including a short evaluation of simulated ozone concentrations in comparison to Airbase data. This Section reads:

A model set-up very similar to the one used for the present study has been evaluated with observational data by Mertens et al. (2016). Generally, the comparison showed a good agreement with observations. The biases are similar to comparable model systems and exhibit a positive ozone bias and negative biases for NO₂ and CO. One important reason for these biases is the too efficient vertical mixing within the COSMO-CLM model. Further, an evaluation of the ozone mixing ratios simulated by *REF* and *EVEU* have already been presented by Mertens et al. (2020), however, mainly focusing on JJA mean values. To investigate the models ability to represent extreme values, we present a brief evaluation of the simulated ozone concentrations in comparison to the Airbase v8 observational dataset (available at, <https://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-8>, last access 14.2.2020). As the model resolution of 50 km is too coarse to resolve hot-spots of individual cities we restrict the comparison to those stations which

are classified as area types 'suburban' and characterised as 'background'. We focus on JJA 2008 to 2010 and compare the results to overall 350 measurement stations. The measurements are subsampled at the same temporal resolution (3 hourly) as the model data. The comparison with the observational data shows a positive ozone bias of the model, which has been discussed in previous studies (Mertens et al., 2016, 2020). The JJA 2008–2010 average ozone concentrations, however, are overestimated by the model. The average root-mean-square-error (RMSE) over all 350 stations is $29.2 \mu\text{g m}^{-3}$ for *REF*, and $24.3 \mu\text{g m}^{-3}$ for *EVEU*, respectively. The corresponding mean biases (MBs) are 26.6 % and 20.5 %, respectively (see Table S1). In addition, we calculated also the RMSE and MB for the *REF* simulation considering only measurements and model data at 12 and 15 UTC. For this subsample, both, RMSE and MB decrease considerably. Accordingly, the largest ozone values during daylight are captured very well by the model. As a more detailed comparison between measurements and model result shows, the overestimation of ozone is particularly strong during night. This can partly be attributed to a too unstable boundary layer during night, which is a common difficulty in many models (Travis and Jacob, 2019). In addition, the vertical mixing in COSMO-CLM/MESSy is very efficient, which also leads to positive ozone biases at noon and during the night (see also Mertens et al., 2020, 2016). Currently, further investigations are undertaken, about how this bias could be reduced in the future. Besides the too efficient vertical mixing, also too less ozone deposition during night, too low NO or VOC emissions, and successively underestimated ozone depletion during nights could also partly contribute to this bias. For analysing extreme ozone values, we also compare the 95th percentiles of ozone with measurements (see Fig. S1). Overall, the model is able to capture most of the regional variability of the extreme values over Europe. Near the densely populated regions in Benelux, Germany and Italy, however, the model is not able to reproduce the extremes. In these areas the model resolutions (i.e. also for the 12 km domain, which is not shown here) are too coarse to allow for a representation of extreme ozone values in urban areas. As has been shown by prior studies (e.g. Tie et al., 2010) resolutions below 10 km are required to capture high ozone values near cities. Terrenoire et al. (2015) have noted that even with 8 km resolution the performance of the applied CHIMERE model is better at rural than at urban sites. This underestimation can also be quantified using the RMSEs and MBs for the 95th percentile which are listed in Table S1. These results have important implications for the analyses, which are presented in this manuscript. First of all, the too strong vertical mixing in COSMO-CLM/MESSy leads to a positive bias of the contribution of stratospheric ozone at ground level. Further, also contributions of lightning and aviation at ground-level are likely larger due to this overestimated

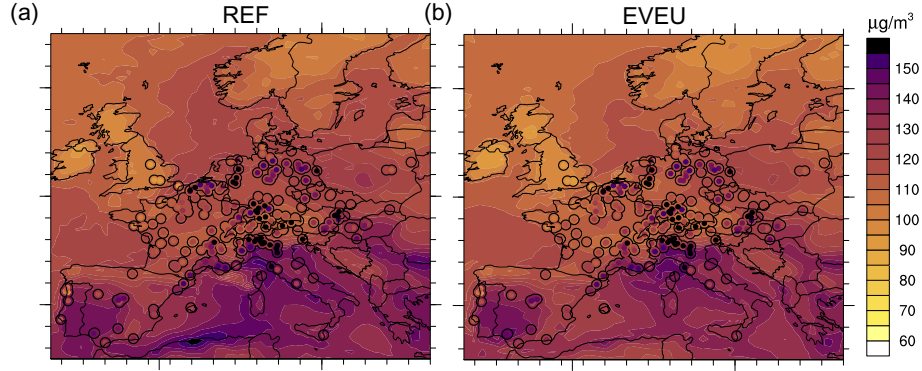


Figure S1: 95th percentile of ozone (in $\mu\text{g}/\text{m}^3$) for the period JJA 2008 to 2010. The background colors show the ozone concentrations as simulated by CM50, the circles represent the location of stations of the Airbase observation data. The inner point represents the measured concentrations, the outer point the concentrations in the respective grid box, where the station is located. All values are based on data every 3 hours.

vertical mixing. This leads to a around 1 percentage point lower contribution of anthropogenic emissions in COSMO-CLM/MESSy compared to EMAC (see Mertens et al., 2020). Due to the coarse model resolution our results are representative for the regional scale, but not for specific urban areas. In these urban areas local emissions and local ozone production/destruction might be more important such that contributions of local sources can be much larger than the values we present. On the regional scale, however, Mertens et al. (2020) showed that the results are quite robust w.r.t. the model resolution (down to 11 km). Because of the stronger ozone bias during night, we further compared the contributions at 12 and 15 UTC with the contributions considering all times of the day. The relative contributions show only small differences, i.e. a slightly larger contribution of anthropogenic emission sources during day (not shown). Therefore, we present always results for all times of the day.

Section 2.1: the authors need to do a lot more here to compare their source apportionment method with other methods in the literature. This is especially important, since the authors themselves have stated on Page 4 (line 1) that differences between source apportionment methods are an important source of uncertainty. Kwok et al. (2015) is already mentioned in Section 2.1, and Dunker et al. (2002) is mentioned in the introduction. Both of these studies use a regime-dependent attribution methodology, which is actually correctly acknowledged by the authors on page 26 in the Discussion section, but a discussion of how these methodologies differ from the methodology employed by the

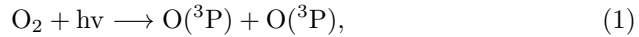
Table S1: Root-mean-square error (RMSE, in μ/gm^3) and mean bias (MB, in percent) of the *REF* and *EVEU* simulations compared to Airbase observation data. Given are the scores for the mean values during JJA and DJF, as well as values for the 95th percentile for JJA. For *REF* listed additionally also the scores considering only the values at 12 and 15 UTC.

	RMSE	MB
REF JJA mean	29.2	26.6
REF JJA 12 and 15 UTC	18.7	13.4
EVEU JJA mean	24.3	20.5
REF JJA 95th percentile	26.9	-10.0
EVEU JJA 95th percentile	28.7	-14.2
REF DJF mean	35.1	32.8
EVEU DJF mean	32.8	30.1

authors, and how this could be expected to influence the results of the study is required already in Section 2.1. Similarly, since the authors are also considering the global scale, they should also put their methodology into the context of the existing techniques for source attribution at the global scale. The authors already cite Emmons et al. (2012) elsewhere in the paper, but do not mention this work in Section 2.1, where it would be appropriate to have some discussion of how these methods differ, and how this might influence the results of the study. One very important difference is that Emmons et al. (2012) only consider NO_x as a precursor of ozone, while the technique employed by the authors combines the effects of both NO_x and VOC precursors. Similarly, the study of Butler et al. (2018) is also missing from the discussion. Butler et al. (2018) account for effects of both NO_x and VOC as ozone precursors, but they make some very different design decisions to the technique employed by the authors. The authors must do more to put their method in the context of the previous work, and discuss the relative strengths and weaknesses of the approach they have chosen.

Reply: The method we apply has been discussed in detail by Grewe et al. (2017). Therefore, the intention of this section was only to recap the general idea of the applied tagging method to the reader. It was never intended as a full discussion of our tagging approach compared to other approaches. However, we agree with referee#2 that a short discussion about the different approaches is helpful here, as these differences will also be discussed in the discussion. The newly added part reads:

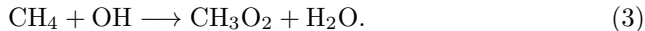
Some of the categories listed in Table 3 are not directly associated with emission sectors. These categories are stratosphere, CH₄ and N₂O. All ozone which is formed by the photolysis of oxygen, i.e



is labelled as stratospheric ozone. The degradation of N_2O is a source for NO_y (and loss of ozone) by the reaction:



The degradation of CH_4 is considered as source of $\text{NMHC}^{\text{CH}_4}$. This refers to the reaction:



As discussed recently in detail by Butler et al. (2018) all tagging methods are based on specific assumptions and have specific limitations. The scheme of Grewe et al. (2017), which we apply in the current study, is based on specific assumptions, which differ from other tagging schemes used in regional and global models. One important difference is the question whether ozone formation is attributed to NO_x , VOC. The schemes which are available in the regional models CMAQ (called CMAQ-ISM, Kwok et al., 2015) and CAMx (Dunker et al., 2002, called CAMx OSAT,) use threshold conditions to check, whether ozone formation is NO_x or VOC limited. Depending on this the production is attributed to NO_x or VOC precursors only. The scheme of Emmons et al. (2012), applied on the global scale, tags only NO_x and therefore ozone production is only attributed to NO_x precursors. Based on the work of Emmons et al. (2012), Butler et al. (2018) presents a scheme, which attributes ozone formation either to NO_x or VOCs (implying that usually 2 simulations, one with NO_x and one with VOC tagging, are performed). This scheme has also been applied by Lupascu and Butler (2019) in a regional model simulation over Europe, using the NO_x tagging only. Compared to these schemes the scheme of Grewe et al. (2017) attributes ozone production always to all associated precursors (i.e. NO_x , HO_2 and VOCs). In VOC limited regions, this approach leads to the effect that a NO_x emission reduction of an emission sector reduces the contribution of that sector, and increases the contribution of the other sectors. In contrast, a reduction of VOC emissions decreases the contribution of the respective sector only. The latter is similar to the approaches integrated in CMAQ-ISM or CAMx, which attribute ozone production in the case of a VOC limit to VOC precursors only. Compared to a NO_x only tagging, our approach leads to lower contributions of NO_x sources, since they compete, not only with other NO_x sources, but also with VOC sources. Because of the family concept, which is necessary to keep the memory consumption and the computational costs low, the tagging method applied in our study can lead to some unphysical artefacts. As an example, Grewe et al. (2017) discuss the production of PAN by NMHCs from CH_4 degradation. Due to the combinatorial approach unphysical artefacts may occur, as for instance NMHCs from stratospheric origin. The main reason for this

is the definition of the PAN family, which transfers tags from NO_y to NMHCs. Other tagging schemes have specific issues as well. As an example, the scheme of Emmons et al. (2012) does not neglect the O_3 - NO_x null cycle, which leads to an overestimation of local sources compared to long range transport sources (see also Kwok et al., 2015). Overall, the impacts of the underlying assumptions on the results are difficult to quantify. Therefore, it is important to study effects of different emission sources with different methods (at best in the same model framework), in order to better understand the strengths and weaknesses of the different approaches and their impact on the source apportionment results.

Also in Section 2.1, the authors could briefly mention how stratospheric ozone is tagged in their approach, since this does not fit into the framework of their Equation 2.

Reply: We added a note about tagging of stratospheric ozone in the revised manuscript (see reply above).

Section 2.2: The authors should make it clear that the tags are applied globally, with no distinction between emissions in Europe and the rest of the world. This is acknowledged later in the manuscript, but the reader would benefit from having this made clear already in this section.

Reply: We added a note in the Section which reads:

In the configuration of the tagging method applied for the present study we use only one global tag for every source category. While this allows to investigate the contributions of all global emissions of a specific emission source to ozone mixing ratios, we are not able to separate contributions from local and long range transport (e.g., we cannot separate contributions from European and Asian land transport emissions to European ozone levels, but we can quantify the contribution of global land transport emissions to European ozone levels).

Page 10, lines 17-24: For some additional context here, it would be nice to know how the proportional contributions of land transport to ambient modelled NO_y compare to the proportional contribution of land transport to total NO_x in the inventories. Is the contribution as would be expected from simply looking at the emissions, or is it disproportionately higher or lower?

Reply: This is indeed a good question. We calculated the relative share of land transport emissions to all anthropogenic + soil NO_x emissions for June (see Fig. S2, which we add also to the revised Supplement). The contributions of land transport emissions are in the range of 50 % to 70 %. The contribution

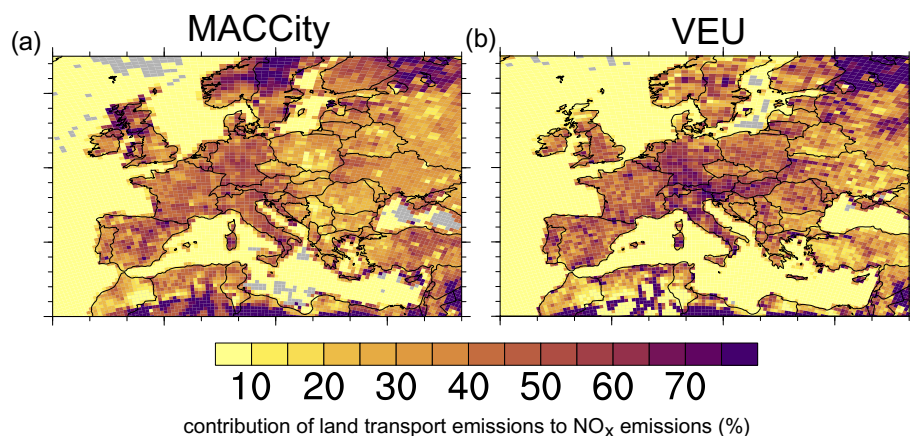


Figure S2: Relative contribution of land transport NO_x emissions to all other emissions (considering soil- NO_x , shipping, anthropogenic, AWB and biomass burning emissions; in %) for July 2009; (a) for the MACCity emission inventory and (b) for the VEU emission inventory.

is larger in the VEU emission inventory compared to the MACCity emission inventory. The contributions of the emissions are in a similar range as the contribution of land transport emissions to NO_y , however, the regional distribution differs slightly. Near the hot-spots (e.g. Paris) we found smaller relative contributions of land transport emissions to NO_y , while the values are larger in rural areas. We added a note about this in the revised manuscript:

The relative contribution of land transport emissions to ground level NO_y is in the range of 40 % to 70 % in most parts of Europe (see Fig 4). These relative contributions are similar as the share of land transport NO_x emissions to all NO_x emissions (see Fig. S9 in the Supplement), but compared to the share of the emissions the contributions to NO_y are slightly lower near hot-spots, and larger in rural areas.

Section 4.1, page 15: The authors rightly interpret the ozone due to land transport in DJF as coming from long-range transport. I also understand that the limits of the experimental design (one global tag for land transport) make it hard to say anything about long-range transport in JJA, when local photochemistry is more active. But could it be possible to try? For example, could they look at the land transport contribution at the western boundary of the refined grid in JJA, and use this as a rough estimate of the contribution of land transport (and other sectors) in remote regions to baseline ozone in Europe? This could add a lot of value to the study and would be highly relevant for international policymaking.

Reply: This is indeed a good point. As mentioned by referee#2 we cannot directly estimate the relative importance of 'global land transport emissions' com-

pared to 'European land transport emissions', as we consider only one global tag. To answer this question in detail more tags would be necessary. Based on your suggestion we added a Figure to this Section (see Fig S3) in which we show area averaged contributions for different categories and for different regions. One region we defined here is called inflow and spreads over a large area of the western boundary of the finer domain. We added a paragraph describing this in the manuscript. As you see, we added also a separation between Northern Alps and the Po Basin (see our answer below).

To quantify the contributions of land transport emissions and other emission sources in different regions in more detail, Fig. S3 shows area-averaged relative contributions for JJA and DJF for the *REF* and *EVEU* simulations (absolute contributions are given in Table S1 to Table S8 in the Supplement). The geographical regions were defined according to the definitions of the PRUDENCE project (Christensen et al., 2007), but slightly modified. The region Alps was split up in two separate regions called 'Northern Alps', defined as rectangular box ($46^{\circ} : 48^{\circ}$ N and $9^{\circ} : 13^{\circ}$ E), and 'Po Valley' ($44^{\circ} : 46^{\circ}$ N and $5^{\circ} : 15^{\circ}$ E). Note, however, that the region Northern Alps contains parts of Switzerland and Southern Germany, which are still rather flat and subject to large land transport emissions. In addition, we defined a region called 'inflow', which is defined from $40^{\circ} : 60^{\circ}$ N and $-13^{\circ} : -11^{\circ}$ E, and is used to quantify contributions in the air coming towards Europe. A figure summarizing the definition of all regions is part of the Supplement (Fig. S12). The relative contribution of land transport emissions in the 'inflow' region is about 9 % and very similar in both seasons and for both European emission inventories. During DJF the contributions of all categories and regions are very similar. During JJA the contribution of land transport emissions increases in most regions compared to the 'inflow' (≈ 9 %). In the Po Valley the contribution reaches up to 16 %. Unfortunately, the difference between the contribution in a specific region compared to the contribution in the region 'inflow' cannot be used to calculate O_3^{tra} from European emissions. Such a calculation requires different tags for global and European land transport emissions. The relative contribution of other anthropogenic emissions in the 'inflow' region (≈ 34 %) is also very similar in both seasons. During DJF the contributions in the different regions remain very similar to the contributions in the 'inflow' region. During summer, in contrast, a West-East gradient of the contribution of anthropogenic emissions is present over Europe with a decrease of the contribution of up to ≈ 27 % in Eastern Europe. This decrease is mainly caused by the seasonality of the different emissions (discussed further below). The biogenic emission category shows different relative contributions in the 'inflow' region during DJF (≈ 11 %) compared to JJA (≈ 14 %), which is mainly caused by the strong increase of biogenic emissions during summer

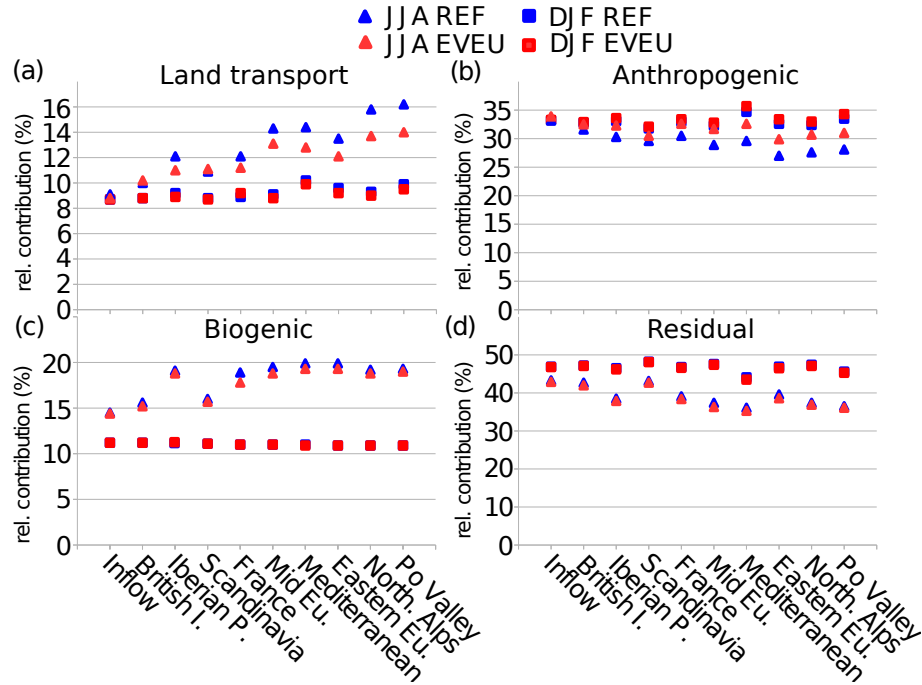


Figure S3: Relative contributions to ground-level ozone (in percent) area averaged in different geographical regions for DJF 2008 to 2010 (triangles) and JJA 2008 to 2010 (squares). Shown are the results of the *REF* (blue) and the *EVEU* simulations (red) for (a) the land transport category, (b) the anthropogenic categories, (c) the biogenic category, and (d) all other categories. For simplicity the anthropogenic category contains the categories anth. non-traffic, aviation and shipping. The residual contains all other categories. The vertical-axis scale differs for (a) to (d).

compared to winter. In the different regions the relative contributions increase during JJA compared to DJF, and, compared to the 'inflow' of up to ≈ 20 %. The contribution of all other tagging categories during DJF is around ≈ 47 % in most regions, and ranges between 41 % and 36 % during JJA.

Page 16, line 3: the seasonal cycle of photochemical activity also plays a role here.

Reply: Indeed. We changed this sentence to:

This seasonal cycle is caused by a complex interplay of the seasonal cycles of different emission sources, meteorology and photochemical activity.

Page 16, line 8: is there any way in this study to separate the influence of soil NOx and biogenic VOC? Or are these two different sources inextricably joined together into the "biogenic" sector?

Reply: As they are emitted in the same category there is no possibility to separate them anymore from the simulation results. However, we see where this can deliver important insights, and we are currently revising the tagging method in such a way that these two emissions could be handled separately.

Section 4.2: As mentioned earlier, it would be nice to know how well the model is capable of reproducing the extreme values of ozone as measured. If the model is doing a good job at this, then the results reported here could help to understand these extreme ozone measurements. If the model is not doing well at this, then the results reported here could potentially provide information about systematic model biases, and point the way towards improving the model. As it currently stands, it is not clear at all how these results should be interpreted.

Reply: As discussed above we added a short section with a model evaluation to the revised manuscript. This evaluation shows that the model is able to reproduce the measured 95th percentile of ozone values quite well on the rural scale, but for strong local ozone enhancements the resolution of our model is too coarse (e.g. Tie et al., 2010). The evaluation clearly shows that the model results are not well suited for analysis of the contributions during extreme ozone events on the levels of individual cities. In our analysis, however, we focused on larger geographical regions. We think that on the basis of these larger geographical regions the model results are well suited to investigate the general trends of ozone contributions. Further, our finding that the relative contribution of land transport emissions increase during extreme ozone events compared to the mean conditions is in line of Valverde et al. (2016). They reported a large importance of land transport emissions during high ozone events for Barcelona and Madrid surroundings.

Page 19, line 11: The region "Alps" includes the Po Valley. Does this mean that high mountains are in the same region as a polluted valley? The influences on air quality would be expected to be very different in these regions. High mountains will be more influenced by the free troposphere (and long-range transport), while the valley will be more influenced by local sources. Furthermore, "Alps" and "Po Valley" are used individually in this section and elsewhere in the manuscript. It is not always clear which region is meant. The authors could consider disaggregating this region into two sub- regions for their analysis (which could be quite informative), or at least being clearer about exactly which region they are referring to throughout the text.

Reply: Yes indeed, the region called 'Alps' includes the Po Valley and the Alps. The main intention for this was that we wanted to stick to the geographical regions defined in the PRUDENCE project. However, we agree that from the

point of view of air quality these regions strongly differ. To take this into account, we split the region 'Alps' in two subregions called 'Northern Alps' (defined as $44^{\circ} : 48^{\circ}$ N and $5^{\circ} : 15^{\circ}$ E) and 'Po Valley' (defined as $44^{\circ} : 46^{\circ}$ N and $9^{\circ} : 13^{\circ}$ E). However, the results for both regions are still very similar. The main reasons are:

- The region Northern Alps contains parts of Southern Germany and also Switzerland, where the mountains are not very high, and much traffic is present.
- Even in the 'higher alps' there are some very important roads with large land transport emissions (e.g. Brenner and Inn valley) which can be clearly seen in the emission inventory. On the 50 km resolution these emissions are mixed over quite large regions.

To better represent the sharp contrast between Alps and the Po Valley a much finer resolution (and fine resolved emission inventories) are necessary, which pose challenging tasks for the future.

Page 19, lines 15-16: the discussion about "uncertainties" in the inventory is very vague here. Could the large range in the contribution of land transport to extreme ozone when using EVEU emissions be related to the higher spatial heterogeneity and existence of more "hot spots" in this inventory compared with REF? There could potentially be some important information here about the need to get the distribution of NO_x right in order to capture the high ozone events. A comparison of the REF and EVEU ozone timeseries with some measurements from urban background stations during extreme events could potentially add a lot of value here.

Reply: It is indeed interesting to investigate how different geographical distributions of NO_x emissions could influence the ability of the model to simulate high ozone events. This issue has partly been investigated in previous publications (Tie et al., 2010; Markakis et al., 2015). Compared to these previous studies, the resolution applied here is rather coarse (50 km). The 95th percentiles of ozone for *REF* and *EVEU* are rather similar (see also the newly added evaluation section). When comparing individual stations during specific periods, we noticed that maximum ozone values are not better represented by *EVEU* compared to *REF*.

As the 95th percentiles of the ozone values are very similar, we think that the differences of the contributions between the two emission inventories are only caused by the different geographical and sectoral distributions. To clarify this we rephrased and extended the discussion. However, for follow up work we agree that this is still an interesting question and should be further investigated using an improved model set-up at finer resolution. The modified text reads:

The ozone values at the 95th percentile (see Sect. 2.3) and at the other percentiles (see Figs. S1 and S2 in the Supplement), however,

are similar for *REF* and *EVEU* (i.e., none of the emission inventories leads to a better representation of extreme ozone events in the model). Accordingly, the discussed differences of the relative contributions are not caused by a different representation of the ozone values themselves, but only due to the different geographical and sectoral distributions of the emissions in *REF* and *EVEU*. This demonstrates the large uncertainty, especially for contributions during high ozone events, of the source apportionment analyses which is caused by the uncertainties of emissions inventories (e.g. geographical distribution of emissions, total emissions per sector). These uncertainties must be taken into account in source attribution studies targeting at contributors during high ozone events.

Page 23, line 4: the results are not "rather similar", but actually have some important differences, which are subsequently discussed. I think what the authors are trying to say here is that the contribution of land transport is similar in each case, but this is not the meaning which comes across.

Reply: We agree with referee#2 that our original intention of the discussion does not come across. Also the comment from referee#1 shows that this part of the discussion caused confusion. Therefore we revised this part of the discussion completely, taking also into account some more (recently published) work to discuss potential reasons for the differences between the results of the different source apportionment results (taking also into account one of the next comments from referee#2). The new part of the discussion reads:

A detailed comparison of our results with previous studies is complicated: First, we apply one global tag for the land transport sector and do not differentiate between local produced ozone and long range transported ozone. In comparison to our approach similar regional studies usually attribute ozone only to the emissions within the regional domain and attribute long-range transported ozone to the boundary conditions. Second, the tagging methods applied in various studies differ. Third, the applied emission inventories differ, so do ozone metrics and simulated periods. Tagaris et al. (2015), who calculated the impact of different emission sectors on ozone using a 100 % perturbation of the respective emission sectors reported an impact of European road transport emissions of 7 % on average for the maximum 8 hr ozone values in July 2006. In most regions impacts above 10 % have been reported, with maximum local impacts (Southern Germany, Northern Italy) of above 20%. While their largest impacts occur in similar regions as our largest contributions (Southern Germany, Northern Italy), our mean contributions are larger than their impacts, but the maximum contributions are lower than their maximum impacts. Further, around London and in parts of Northern England their impacts (see Fig. 3 therein) are

around 2 to 4 %, while our contributions are in the range of 8 to 10 %. Hence, impact and contribution differ largely in these regions. This is in line with previous work, stating that the contributions to ozone are more robust, i.e. less dependent on the background, as the perturbations or impacts (Grewe et al., 2012, 2019). All the studies that we are aware of and which reported contributions of land transport emissions to ozone over Europe using a tagging method either use the CAMx model (CAMx OSAT method, Karamchandani et al., 2017) or the CMAQ model (CMAQ-ISM method, Valverde et al., 2016; Pay et al., 2019). As discussed, these two methods apply a sensitivity approach to check, whether ozone production is NO_x or VOC limited. These previous studies considered only European emissions, while we consider the combined effect of European emissions and long range transport. Therefore, one would expect that our contribution analysis shows larger contributions as previous studies. However, our contributions in general are lower compared to previously reported values. As an example, Karamchandani et al. (2017) reported contributions around larger European cities in the range of 24 to 11 %, in Budapest even up to 35 %. Valverde et al. (2016) reported contributions of road transport emissions from Madrid and Barcelona of up to 24 % and 8 %, respectively. Similarly, Pay et al. (2019) diagnosed contributions of road transport emissions on ozone of 9 % over the Mediterranean Sea and up to 18 % over the Iberian Peninsula, however for a specific summer episode only (July 2012). To discuss potential reasons why our contributions are lower compared to previous estimates, we analysed our results for July 2010, to compare these contributions directly with the findings of Karamchandani et al. (2017). As an example, Karamchandani et al. (2017) reported contributions of 17 % around Berlin, while our contributions are in the range of 12–14 %. Further they diagnosed contributions from the biogenic sector of around 11 % around Berlin, while we find contributions of the biogenic sector of around 18 %. Generally, the contributions reported by Karamchandani et al. (2017) seem to be much more variable over Europe compared to our results. A reason for this might be the different treatment in the apportionment of NO_x and VOC precursors. Land transport emissions contribute mainly to NO_x emissions, while biogenic emissions are an important source of VOCs. As shown by Butler et al. (2018), anthropogenic emissions contribute most to ozone over Europe, if a NO_x tagging is applied, while biogenic emissions are the most important contributor, when a VOC tagging is applied (Figs. 3 and 4 therein). Accordingly, those approaches which use a threshold to perform either a VOC or NO_x tagging, attribute ozone production under VOC limitation mainly to biogenic sources, while under a NO_x limitation ozone is attributed mainly to anthropogenic sources (including land transport emissions). Most likely this leads to a much stronger variability between anthro-

pogenic and biogenic contributions compared to our approach, where ozone is always attributed to NO_x and VOC or HO_x precursors. Similar effects can also be observed when comparing our results to the results of Lupășcu and Butler (2019), who applied a NO_x -only tagging for the period April to September 2010 and considered regional as well as global sources similar to our approach. They reported contributions of biogenic emissions in Europe for the period July - September between 5 and 13 % over Europe. Our results show contributions of biogenic emissions which are much larger (15 to 26 % for the same period). In their approach, ozone is only attributed to biogenic NO_x emissions, while we attribute ozone to biogenic NO_x and VOC emissions. Further, our estimated stratospheric contribution to ground-level ozone is also larger than the contributions reported by Lupășcu and Butler (2019). In this case, our results indicate contributions for July to September in the range of 5 to 10 % compared to their 2 to 4 %. Similarly, for lightning- NO_x our model shows larger contributions (6–12 %) compared to the 3–6 % diagnosed by Lupășcu and Butler (2019). These differences of the contributions for the stratospheric and the lightning category can partly be attributed to the more efficient vertical mixing in COSMO-CLM. Mertens et al. (2020) reported a maximum difference of the contributions from the stratosphere and lightning to ozone between EMAC and COSMO-CLM/MESSy of 30 % at maximum. As the difference between our results and the results of Lupășcu and Butler (2019) are much larger as these 30 %, the difference can most likely not entirely be attributed to differences in vertical mixing. Rather, the differences can probably be explained by the different contributions of the biogenic category (due to different tagging methods) and by the different contributions of lightning and stratospheric sources. However, the different studies provide not enough insights about the applied emissions (e.g. for lightning- NO_x , soil NO_x and biogenic VOCs) to fully analyse these differences. The discrepancies between the results of the different source apportionment methods show clearly the importance of a co-ordinated comparison of different source apportionment approaches, as has already been suggested by Butler et al. (2018).

Page 25, lines 25-26: This sentence basically conveys no meaning and could be easily deleted with no loss to the manuscript. Alternatively the authors could try to be clearer about what they mean here.

Reply: We rephrased the sentence to:

The result that regions are hot-spots for NO_y from land transport emissions, but not for O_3 from land transport is counter intuitive. The reasons for this is that large amounts of NO_x emissions alone are not sufficient for large ozone production. This is caused by the

non-linearity of the ozone chemistry and the strong interdependence between ozone production and meteorological conditions (e.g. Monks et al., 2015).

Page 25, last paragraph: if the previous work only accounts for the contribution of European land transport emissions to European ozone, and the current study also includes global emissions, then shouldn't the current study result in a higher contribution than the previous work? The opposite appears to be the case. Can the authors explain this apparent discrepancy?

Reply: We agree. As already discussed above, we tried to clarify this in the new discussion. The differences between the studies, however, are so large that we cannot fully explain the discrepancies, but the discussion hopefully provides some insights. What is really needed to understand the differences between the tagging methods is a detailed inter-comparison of them.

Page 26, line 23: the authors appear to be concluding from the strong influence of the "biogenic" sector that soil NO_x emissions are strongly influencing ozone. But couldn't this also be biogenic VOC? How do they separate the influence of these two different sources? A comparison with Butler et al. (2018) could be instructive here, since in that study the separate roles of NO_x and VOC as ozone precursors were examined. Comparison of their Figure 3 and Figure 4 indicates that biogenic VOC make a larger contribution to European ozone in summer than biogenic NO_x. The authors should discuss this here.

Reply: Of course also biogenic VOCs are very important for the ozone production. As discussed above, we cannot differentiate between ozone produced by biogenic VOCs and soil-NO_x, as we join them together in one category. The soil-NO_x emissions are an important contributor to the NO_x emissions in Europe in summer (see Fig. S4 in the Supplement). Uncertainties of these emissions cause uncertainties of the simulated contributions. However, also the biogenic VOC emissions are uncertain. Therefore, we rephrased this part to clarify that we do not want to say that biogenic VOCs are not important. The discussion of the importance of biogenic VOCs and NO_x with the reference to Butler et al. (2018) was already introduced in the revised discussion above.

Page 28, lines 14-15: the future work proposed by the authors would indeed be extremely interesting from a policymaking perspective. If possible, they should also include as many other sectors as possible. This could help to inform decisions about where emission reductions would be most effective.

Reply: Thanks for this positive comment. Actually, the main focus is on land transport, but we intent to investigate also other categories.

Page 28, lines 28-29: again, it appears that the authors are over-interpreting their results when they conclude that soil NO_x has a strong influence on Euro-

pean ozone levels.

Reply: We guess referee#2 meant lines 18-19. Here we write: 'Here, the focus should not only be on the land transport emissions, but also on other important emissions, including especially biogenic and soil-NO_x emissions, which have large uncertainties and contribute strongly to European ozone levels.' We do not see where we over-interpret the results. However, to make the sentence more clear we add 'VOC' to biogenic:

'Here, the focus should not only be on the land transport emissions, but also on other important emissions, including especially biogenic **VOCs** and soil-NO_x emissions, which **are subject to** large uncertainties and contribute strongly to European ozone levels. '

We are looking forward to your reply,
Mariano Mertens
(on behalf of all co-authors)

References

- Butler, T., Lupascu, A., Coates, J., and Zhu, S.: TOAST 1.0: Tropospheric Ozone Attribution of Sources with Tagging for CESM 1.2.2, *Geosci. Model Dev.*, 11, 2825–2840, <https://doi.org/10.5194/gmd-11-2825-2018>, URL <https://www.geosci-model-dev.net/11/2825/2018/>, 2018.
- Christensen, J. H., Carter, T. R., Rummukainen, M., and Amanatidis, G.: Evaluating the performance and utility of regional climate models: the PRUDENCE project, *Climatic Change*, 81, 1–6, <https://doi.org/10.1007/s10584-006-9211-6>, URL <http://dx.doi.org/10.1007/s10584-006-9211-6>, 2007.
- Clappier, A., Belis, C. A., Pernigotti, D., and Thunis, P.: Source apportionment and sensitivity analysis: two methodologies with two different purposes, *Geosci. Model Dev.*, 10, 4245–4256, <https://doi.org/10.5194/gmd-10-4245-2017>, URL <https://www.geosci-model-dev.net/10/4245/2017/>, 2017.
- Dunker, A. M., Yarwood, G., Ortman, J. P., and Wilson, G. M.: Comparison of Source Apportionment and Source Sensitivity of Ozone in a Three-Dimensional Air Quality Model, *Environmental Science & Technology*, 36, 2953–2964, <https://doi.org/10.1021/es011418f>, URL <http://dx.doi.org/10.1021/es011418f>, pMID: 12144273, 2002.
- Emmons, L. K., Hess, P. G., Lamarque, J.-F., and Pfister, G. G.: Tagged ozone mechanism for MOZART-4, CAM-chem and other chemical transport models, *Geosci. Model Dev.*, 5, 1531–1542, <https://doi.org/10.5194/gmd-5-1531-2012>, URL <http://www.geosci-model-dev.net/5/1531/2012/>, 2012.
- Grewe, V., Dahlmann, K., Matthes, S., and Steinbrecht, W.: Attributing ozone to NO_x emissions: Implications for climate mitigation measures, *Atmos. Environ.*, 59, 102–107, <https://doi.org/10.1016/j.atmosenv.2012.05.002>, URL <http://www.sciencedirect.com/science/article/pii/S1352231012004335>, 2012.
- Grewe, V., Tsati, E., Mertens, M., Frömming, C., and Jöckel, P.: Contribution of emissions to concentrations: the TAGGING 1.0 submodel based on the Modular Earth Submodel System (MESSy 2.52), *Geosci. Model Dev.*, 10, 2615–2633, <https://doi.org/10.5194/gmd-10-2615-2017>, URL <https://www.geosci-model-dev.net/10/2615/2017/>, 2017.
- Grewe, V., Matthes, S., and Dahlmann, K.: The contribution of aviation NO_x emissions to climate change: are we ignoring methodological flaws?, *Environmental Research Letters*, 14, 121 003, <https://doi.org/10.1088/1748-9326/ab5dd7>, 2019.
- 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, *Atmos. Chem.*

- Phys., 17, 5643–5664, <https://doi.org/10.5194/acp-17-5643-2017>, URL <https://www.atmos-chem-phys.net/17/5643/2017/>, 2017.
- Kwok, R. H. F., Baker, K. R., Napelenok, S. L., and Tonnesen, G. S.: Photochemical grid model implementation and application of VOC, NO_x, and O₃ source apportionment, *Geosci. Model Dev.*, 8, 99–114, <https://doi.org/10.5194/gmd-8-99-2015>, URL <http://www.geosci-model-dev.net/8/99/2015/>, 2015.
- Lupaşcu, A. and Butler, T.: Source attribution of European surface O₃ using a tagged O₃ mechanism, *Atmos. Chem. Phys.*, 19, 14 535–14 558, <https://doi.org/10.5194/acp-19-14535-2019>, URL <https://www.atmos-chem-phys.net/19/14535/2019/>, 2019.
- Markakis, K., Valari, M., Perrussel, O., Sanchez, O., and Honore, C.: Climate-forced air-quality modeling at the urban scale: sensitivity to model resolution, emissions and meteorology, *Atmos. Chem. Phys.*, 15, 7703–7723, <https://doi.org/10.5194/acp-15-7703-2015>, URL <https://www.atmos-chem-phys.net/15/7703/2015/>, 2015.
- Mertens, M., Kerkweg, A., Jöckel, P., Tost, H., and Hofmann, C.: The 1-way on-line coupled model system MECO(n) – Part 4: Chemical evaluation (based on MESSy v2.52), *Geosci. Model Dev.*, 9, 3545–3567, <https://doi.org/10.5194/gmd-9-3545-2016>, URL <http://www.geosci-model-dev.net/9/3545/2016/>, 2016.
- Mertens, M., Grewe, V., Rieger, V. S., and Jöckel, P.: Revisiting the contribution of land transport and shipping emissions to tropospheric ozone, *Atmos. Chem. Phys.*, 18, 5567–5588, <https://doi.org/10.5194/acp-18-5567-2018>, URL <https://www.atmos-chem-phys.net/18/5567/2018/>, 2018.
- Mertens, M., Kerkweg, A., Grewe, V., Jöckel, P., and Sausen, R.: Are contributions of emissions to ozone a matter of scale?, *Geosci. Model Dev.*, 13, 363–383, <https://doi.org/10.5194/gmd-13-363-2020>, URL <https://www.geosci-model-dev.net/13/363/2020/>, 2020.
- Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., and Williams, M. L.: Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer, *Atmos. Chem. Phys.*, 15, 8889–8973, <https://doi.org/10.5194/acp-15-8889-2015>, URL <http://www.atmos-chem-phys.net/15/8889/2015/>, 2015.
- Pay, M. T., Gangoiti, G., Guevara, M., Napelenok, S., Querol, X., Jorba, O., and Pérez García-Pando, C.: Ozone source apportionment during peak summer events over southwestern Europe, *Atmos. Chem. Phys.*, 19, 5467–5494, <https://doi.org/10.5194/acp-19-5467-2019>, URL <https://www.atmos-chem-phys.net/19/5467/2019/>, 2019.

- Tagaris, E., Sotiropoulou, R. E. P., Gounaris, N., Andronopoulos, S., and Vlachogiannis, D.: Effect of the Standard Nomenclature for Air Pollution (SNAP) Categories on Air Quality over Europe, *Atmosphere*, 6, 1119, <https://doi.org/10.3390/atmos6081119>, URL <http://www.mdpi.com/2073-4433/6/8/1119>, 2015.
- Terrenoire, E., Bessagnet, B., Rouil, L., Tognet, F., Pirovano, G., L  tinois, L., Beauchamp, M., Colette, A., Thunis, P., Amann, M., and Menut, L.: High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE, *Geosci. Model Dev.*, 8, 21–42, <https://doi.org/10.5194/gmd-8-21-2015>, URL <https://www.geosci-model-dev.net/8/21/2015/>, 2015.
- Tie, X., Brasseur, G., and Ying, Z.: Impact of model resolution on chemical ozone formation in Mexico City: application of the WRF-Chem model, *Atmos. Chem. Phys.*, 10, 8983–8995, <https://doi.org/10.5194/acp-10-8983-2010>, URL <http://www.atmos-chem-phys.net/10/8983/2010/>, 2010.
- Travis, K. R. and Jacob, D. J.: Systematic bias in evaluating chemical transport models with maximum daily 8 h average (MDA8) surface ozone for air quality applications: a case study with GEOS-Chem v9.02, *Geosci. Model Dev.*, 12, 3641–3648, <https://doi.org/10.5194/gmd-12-3641-2019>, URL <https://www.geosci-model-dev.net/12/3641/2019/>, 2019.
- Valverde, V., Pay, M. T., and Baldasano, J. M.: Ozone attributed to Madrid and Barcelona on-road transport emissions: Characterization of plume dynamics over the Iberian Peninsula, *Science of The Total Environment*, 543, Part A, 670 – 682, <https://doi.org/http://dx.doi.org/10.1016/j.scitotenv.2015.11.070>, URL <http://www.sciencedirect.com/science/article/pii/S0048969715310500>, 2016.
- Wang, Z. S., Chien, C.-J., and Tonnesen, G. S.: Development of a tagged species source apportionment algorithm to characterize three-dimensional transport and transformation of precursors and secondary pollutants, *J. Geophys. Res. Atmos.*, 114, n/a–n/a, <https://doi.org/10.1029/2008JD010846>, URL <http://dx.doi.org/10.1029/2008JD010846>, d21206, 2009.
- Wild, O., Fiore, A. M., Shindell, D. T., Doherty, R. M., Collins, W. J., Dentener, F. J., Schultz, M. G., Gong, S., MacKenzie, I. A., Zeng, G., Hess, P., Duncan, B. N., Bergmann, D. J., Szopa, S., Jonson, J. E., Keating, T. J., and Zuber, A.: Modelling future changes in surface ozone: a parameterized approach, *Atmos. Chem. Phys.*, 12, 2037–2054, <https://doi.org/10.5194/acp-12-2037-2012>, URL <https://www.atmos-chem-phys.net/12/2037/2012/>, 2012.