Dear editor,

in this document we listed the comments (in bold), our replies and the changes applied to the manuscript (in blue color). Beside the text changes, a new co-author has been added to the author list (Matthias Kohl), and table 2 (now table 3) was corrected as a error was found in the calculation algorithm.

Best regards,

Andrea Pozzer

referee #1

1. line 16: use "secondary organic aerosols" instead of just "aerosols", because it is already said that the aerosols discussed here are from VOCs.

We have applied the change.

2. line 37: Can the authors elaborate in the manuscript why explicit chemical identities would be useful in atmospheric simulations here?

Explicit chemistry schemes allow a representation of Secondary Organic Aerosols formation from VOCs without the use of lumped species and experimentally derived parameters (e.g., reaction rates, aerosol yields). These tuning parameters can increase model uncertainties and result in large differences between Atmospheric Chemistry models. In addition, empirical chemical schemes are not mass conserving (e.g., for carbon) and the higher-generation reaction products are lumped or ignored, even if, for instance, they play a pivotal role for OH recycling and ozone chemistry (Taraborrelli et al., 2012), or are a major component of atmospheric brown carbon (Laskin et al., 2015). In general, explicit identities of model species are essential for making further progress in quantifying the atmospheric budget of Secondary Organic Aerosols. By relying on high-quality (experimental and theoretical) data of the physico-chemical properties of precursors and intermediates, an ever more realistic modelling of removal, ageing and formation pathways would be possible.

In the revised text at line 37 we add the following lines: Instead, explicit chemistry schemes allow a representation of Secondary Organic Aerosols formation from VOCs without the use of lumped species and experimentally derived parameters (e.g., reaction rates, aerosol yields). These tuning parameters can increase model uncertainties and result in large differences between Atmospheric Chemistry models. In addition, empirical chemical schemes are not mass conserving (e.g., for carbon) and the higher-generation reaction products are lumped or ignored, even if, for instance, they play a pivotal role for OH recycling and ozone chemistry (Taraborrelli et al., 2012), or are a major component of atmospheric brown carbon (Laskin et al., 2015). In general, explicit identities of model species are essential for making further progress in quantifying the atmospheric budget of Secondary Organic Aerosols. By relying on high-quality (experimental and theoretical) data of the physico-chemical properties of precursors and intermediates, an ever more realistic modelling of removal, ageing and formation pathways would be possible.

3. For the OA treatment (starting from line 103), has the MOM + ORACLE

approach been tested against chamber SOA experiments of, for example, alpha-pinene, to see if the model correctly captures SOA mass production in the experiments? This reviewer trusts that the two submodels have been well tested in previous studies but just wonders if this can be done for closure.

This is a very important aspect to look at in order to ensure an ever increasing realism of our model as experimental capabilities and knowledge progress. Beside the global model, MOM is normally run in the photochemical box model CAABA/MECCA (Sander et al., 2019), which allows for multi-phase kinetics simulations with deliquescent aerosols and cloud droplets, and it has been evaluated with chamber studies (e.g., Novelli et al., 2020). Nevertheless, the MOM + ORACLE approach has so far not been evaluated against chamber experiments, but this is planned for modelling the experiments routinely done at the SAPHIR and SAPHIR* chambers. However, the MOM + ORACLE approach has been evaluated against observation with the mixed layer model by Janssen et al. (2017).

Added in line 112: An earlier version of MOM has been compared against chamber measurements (Nölscher et al., 2014) and improved by Novelli et al. (2020), while ORACLE was derived empirically from chamber experiments (Donahue et al., 2011) and has been evaluated against observations from a field campaign by Janssen et al. (2017). Nevertheless, the MOM + ORACLE combination still has to be fully evaluated.

4. For section 4.2.2, the authors do not seem to offer a potential explanation for the underestimation. Can this be added to the text?

It is indeed difficult to assess a potential explanation for the strong and homogeneous underestimation. Due to their lower solubility, these tracers are only removed by reactions with OH, NO_3, and O_3, with the first being at least two orders of magnitude faster than the other. There are therefore strong indications that the reaction with OH of these tracers are too fast. Analogously to Pozzer et al. (2006), we again point to a revision of such reaction rates. Furthermore, beside possibly too fast decomposition of these tracers, a substantial lack of emissions could be present, for instance from natural sources as shown by J.-L. Li et al. (2021).

This text has been added in line 355: As alkenes are mostly removed by reaction with OH, there are strong indications that these reactions are too fast, in addition, a substantial lack of emissions could be present, even from natural sources as suggested by (J.-L. Li et al., 2021).

5. In the outlook section, the authors discuss potential improvements to the model. Since SOA lifecycle in the atmosphere can also be affected by aerosol phase state, heterogeneous reaction with oxidant, etc., which are still highly uncertain, could the authors include these processes in the discussion as well, and talk about how they can be captured in the MOM+ORACLE framework?

The referee pointed out correctly some further issues that could be implemented in the future. Some of them have already been studied or implemented in the EMAC model, although not explicitly described or used here.

Initial work on the state of the secondary organic aerosols with the ORACLE model has already been performed by Shiraiwa et al. (2017), although more work is needed to estimate its impact on the global secondary organic aerosol budget. The MOM+ORACLE framework calculates the phase partitioning of organic compounds by assuming a bulk equilibrium. However, the phase

state of the organic aerosol can affect the mixing time of the condensed organic compounds within the aerosol. In general, equilibrium partitioning to the particle phase is a reasonable assumption if the aerosol is liquid, however, if the phase state is solid, non-equilibrium partitioning should be considered. The MOM+ORACLE framework does not consider any kinetic limitations in the bulk. Shiraiwa et al. (2017) suggests that kinetic limitations in the bulk may not significantly affect SOA partitioning in the boundary layer, justifying the use of equilibrium partitioning in this part of the atmosphere.

Finally, MOM comprises an explicit gas-phase VOC oxidation scheme but does not include any heterogeneous reactions. However, the MOM+ORACLE framework can be used as a basis to link ORACLE with an aqueous phase mechanism which is planned for a future study.

The following text has been added to the conclusions (line 757): Finally, the MOM+ORACLE framework calculates the phase partitioning of organic compounds by assuming a bulk equilibrium, without any kinetic limitation. However, the phase state of the organic aerosol can affect the mixing time of the condensed organic compounds within the aerosol. In general, equilibrium partitioning to the particle phase is a reasonable assumption if the aerosol is liquid, however, if the phase state is solid, non-equilibrium partitioning should be considered. Shiraiwa et al. (2017) suggests that kinetic limitations in the bulk may not significantly affect SOA partitioning in the boundary layer, justifying the use of equilibrium partitioning in this part of the atmosphere, but kinetic limitations should be implemented and investigated for the free and upper troposphere.

referee #2

1. No comparisons are made to previous model versions. The authors have made no arguments for which conditions such a complex mechanism is needed. This work ought to show what difference, and particularly improvement, is achieved with this complex chemistry scheme over standard, more reduced schemes.

We thank the reviewer for pointing this out. We indeed have discussed this issue in the manuscript (line 46-54) : "A comparison for VOCs with the MIM chemical mechanism [...] that was used previously [...] is not shown here as in such mechanism (i) most of organics are either lumped [...] or missing; [...] (ii) primary species common to MIM and MOM would be influenced only by the different sinks (mainly OH), and a detailed description of OH budget is presented here; (iii) the model bias with respect to secondary species, e.g. oxygenated VOCs, has been linked to a mis-representation or lack thereof of processes like in-cloud chemistry." Therefore we believe that most of the comparison would not yield any additional information. It must be also stressed that we do not claim that this chemical mechanism is better than those previously used; rather the MOM mechanism is more complete, which allows studies which are not possible with simplified chemistry. In the few cases of specific trace gases (e.g. ozone or formaldehyde), which are present in both, MIM and MOM, a comparison could indeed be useful. This, however, would require a different model set-up, specifically aimed at such a comparison, and therefore is beyond the scope of this manuscript.

No changes have been applied to the manuscript, as the information is already contained in the text (lines 46-54)

2. The VOC measurements used for evaluation of the model are really not adequate and numerous additional observations could have been used to evaluate the VOCs. The Emmons compilation of aircraft observations only includes data through 2000, and the quality and quantity of in-situ (aircraft) measurements of VOCs have increased dramatically since then. For example, the ATom missions provide global coverage of the compounds presented here (HCs, CH3CHO, CH3COCH3, etc.). Numerous other aircraft campaigns led by NASA, NOAA and NSF in the U.S. provide measurements over polluted regions (e.g., INTEX-NA, DISCOVER-AQ, SEAC4RS, SENEX, KORUS-AQ, FIREX-AQ, etc., etc.). These datasets are all freely available and in standard, similar formats which making them fairly easy to use, though it does require a bit more effort than taking the single profiles provided by the Emmons climatology.

We believe that an evaluation of a global model of such complexity, as the one presented in this study, should show the capability of the model to represent background/climatological conditions, rather than polluted/episodic ones. For this reason we adopted the so called "Emmons database", as each profile is based on multiple flights in the same regions, mostly for background conditions. As mentioned in the manuscript (line 138), we consider the used aircraft campaigns as representative for the period and the regions, i.e., equivalent to an observational-based climatology, although we acknowledge its limitations by using additional satellite observations. In addition, as these data have been used in many other evaluations, we can compare our results with previous studies.

A comparison with single flight measurements would give us a more "episodic" comparison, and should be accompanied with detailed process studies, as done before with this mechanism (e. g. Lelieveld et al., 2018; Tadic et al., 2021; Wang et al., 2020). Therefore, we restrain from using the campaign data suggested by the referee in such a general evaluation as the one presented here, as we would miss an overall overview.

Nevertheless, we believe that the Atmospheric Tomography Mission (ATom campaign), does fit perfectly our needs, being over background regions, for different seasons, and with an extensive set of observed trace gases, and we thank the referee for pointing that out. [. . . , see replied to referee #2 in the open discussion]

In lines 141-147 we clarified what we consider to be an evaluation of a model of such complexity as EMAC. Furthermore, the comparison with the ATom(s) field campaigns has been added, with additions in many part of the manuscript:

- **The description of the campaign is now in Sect.3.1.1, line 149-165,**

- **Fig.2 with the ATom(s) campaign flight paths and the regions defined for the comparison,**

- **Tab.2 with the statistics of the observations-model results comparison based on the AToms database,**

- **Additional general remarks from the comparison in line 270-279,**

- **Fig.3 scatter plots for CH_3OOH and $HCHO$,**

- **Additional text in line 318-319, for the comparison with alkanes,**

- **Additional comparison (line 365 and Fig.9) for methanol,**

- **Additional investigation of North Hemisphere model underestimation for acetone, against ATom observations (lines 289-294 and Fig.10),**

- **Comparison with $HCOOH$ in lines 416-422,**

- **confirmation of previous results for PAN in line 455.**

The IASI retrievals are used extensively, but much more should be included about their accuracy and the results of published validation results. Perhaps they have much greater uncertainty at high latitudes, or remote regions - this should be discussed here.

Random and systematic components affect the uncertainties of the VOC measurements from IASI. Every IASI product comes with its own estimate of the random uncertainty associated with an individual retrieved column (see, e.g., Franco et al., 2018). For a non-background abundance of the considered species, the relative uncertainty on an individually retrieved column ranges typically between 20% and 50%, with the highest uncertainties found for the low columns. This single-pixel uncertainty increases for lower and background columns as the lower abundance of the target species approaches the IASI detection threshold. Nevertheless, these random uncertainties become negligible for the column averages calculated here for the comparisons with EMAC, because of the total number of measurements used per model grid cell. Indeed, after data filtering, 17 IASI measurements per day on average fall into each model grid box close to the Equator, i.e., more than 6,000 over the year 2010. Owing to the satellite polar orbits, this number increases with latitude and with the higher spatial sampling of IASI.

With respect to the systematic uncertainties, comparisons with independent measurements were performed to identify any potential bias in the IASI VOC columns. For the species that are also retrieved from ground-based FTIR measurements, namely CH3OH, HCOOH and PAN, column comparisons at various latitudes and environments were performed (Franco et al., 2020; Mahieu et al., 2021). Since no column measurements of acetone and acetic acid are currently available, the IASI data were compared to independent columns derived from tropospheric aircraft profiles taken from various campaigns, complemented in the lower stratosphere by model data (Franco et al., 2019, 2020). These comparisons confirmed the absence of any large systematic biases of the IASI data, and that there is no noticeable latitudinal impact on the discrepancies. However, an underestimation (locally up to 30%) of the highest columns over tropical source regions (e.g., the Amazon Basin) during the dry season has been identified for all the species, except for PAN. Such underestimation of the elevated VOC columns affects the nadir-viewing sounders in general. For instance, it has also been observed in an extensive comparison of the TROPOMI formaldehyde product with over 20 FTIR measurement sites (Vigouroux et al., 2020).

Despite the underestimation of the highest columns locally, in the context of our study, the accuracy of the IASI products is relatively stable and sufficient to provide a global evaluation of the EMAC performance, considering the large uncertainties that still affect the emissions and atmospheric modelling of these VOCs.

In the manuscript, we have updated the section 3.3.3 dedicated to the IASI measurements as follows: The uncertainties on the individual retrieved column can be large, but are considerably reduced by averaging numerous observations in space and time, as done in this study by comparing annual averages on the model grid. A full description of the ANNI framework, the characterization of the VOC products, and comparisons of the satellite data with independent measurements can be found in Franco et al. (2019, 2020, 2018) and Mahieu et al. (2021), and references therein. In

these studies, comparisons with ground-based column measurements of VOCs, as well as with columns derived from aircraft profiles, indicate no large systematic biases of the IASI data and no dependence on the latitude, although an underestimation (locally up to 30%) of the highest columns over tropical source regions (e.g., the Amazon Basin) has been identified for all the IASI VOC products, except PAN. The accuracy of the IASI measurements is therefore sufficient to provide a global evaluation of the EMAC performance, considering the large uncertainties that still affect the emissions and atmospheric modelling of the VOCs. For both, IASI and the model, daily gridded averages were constructed at the spatial resolution of the model grid. To have similar temporal coverage over the year between model and satellite observations, the EMAC daily averages were masked when the corresponding IASA data were missing for the same day and location.

referee #3

(1) Line 81: Are those emissions reasonable compared with other global model inputs or estimates? It is better to add a few sentences to discuss it. It looks like the work did not apply regional emission inventories (e.g., NEI for US, EMEP for Europe, MIX for East Asia or MEIC for China). Would this affect the simulations in anthropogenic-influenced regions? In particular, for aerosol composition, I think the regional inventories may matter significantly.

The emissions database used in this work has been extensively compared with other global emission databases by Crippa et al. (2018), and therefore it is not repeated here. It can be mentioned that differences can be up to " 43% for BC " (Crippa et al., 2018). As we needed to have a global emissions dataset, we adopted the one we considered most realistic at the time of the simulation, although different global emissions dataset are possible within the EMAC model (see for example Jöckel et al., 2016; Reifenberg et al., 2021). As mentioned by the referee, we did not apply any regional emission inventories: as we were not focusing on specific locations, we preferred to adopt a homogeneous global emission dataset rather than over-imposing regional emissions datasets. It must, however, be stressed that the uncertainties in these global emissions are often similar to the regional ones. For example Saikawa et al. (2017) compared 2 global and 3 regional emissions dataset for China, showing that "large disagreements exist among the five inventories [. . .]". As a comparison between different emissions datasets is outside the scope of this study, we will simply refer to the work of Crippa et al. (2018) for the details.

This text has been added to line 77 when describing the anthropogenic emissions: [. . .] and has been compared to other global emission databases by (Crippa et al., 2018).

(2) Section 3.2.5: I think the authors only used the EANET data for 2010. Please specify.

Yes, we used the data for the year 2010. Thanks for pointing this out. We will correct this in the revised version.

This has been amended in line 210

(3) Line 238-239: What kind of biogenic CO precursors?

An example would be methanol CH$_3$OH, as presented by Rosanka et al. (2021b) and in Sect.4.2.3 in our manuscript. Also CH$_3$COCH$_3$ (see Sect.4.2.4) is overestimated in these regions. We will refer to these sections in the revised version of the manuscript.

We added to line 301 the following text: (see Rosanka et al., 2021a,b), as also shown in Sect.4.2.3 and Sect.4.2.4.

(4) Line 280-281: How high? I am surprised that temperature difference can be significant between the model parameters and the observations. If the overestimation is due to excess of biogenic VOCs, it is better to compare the surface concentrations of isoprene etc. Some measurements are available in those tropical areas and the data are published (e.g., from AMAZE-08).

The referee is correct. The bias in temperature is indeed very limited, as shown by Hagemann and Stacke (2015). Nevertheless, it must be stressed that even a small overestimation can have a strong impact on the emissions. As shown by Guenther et al. (2006, Fig.4), the MEGAN model shows a strong response of the emissions on changed temperature. Once plants are exposed for several days to temperatures above 300 K, MEGAN predicts strong dependencies of isoprene emissions on actual temperature, so that even a difference of a few degrees implies large emission differences. Following the referee's suggestion, we compared the simulated isoprene mixing ratios in the Amazon rainforest with observations from the AMAZE-08 campaign (Martin et al., 2016) and from the ATTO tower (Yáñez-Serrano et al., 2015). In comparison to the AMAZE-08 campaign, the simulated isoprene measurements are overestimated by more than a factor of 3 (6.1 ± 1.2 ppb simulated and 1.9 ± 1.4 ppb observed), while the overestimation factor for the ATTO tower is on average approximately 1.6, being higher in February/March and lower in October/November.

In line 348-353, the following lines have been added: [. . .] due to its high temperature sensitivity (Guenther et al., 2006). This is confirmed by a comparison of the simulated isoprene mixing ratios in the Amazon rainforest with observations from the AMAZE-08 campaign (Martin et al., 2016) and from the ATTO tower (Yáñez-Serrano et al., 2015). In comparison to the AMAZE-08 campaign, the simulated isoprene measurements are overestimated by more than a factor of 3 (6.1 ± 1.2 ppb simulated and 1.9 ± 1.4 ppb observed), while the overestimation factor for the ATTO tower is on average approximately 1.6, being higher in February/March and lower in October/November.

(5) Consistent model underestimation present in mid- and high latitudes of Northern Hemisphere for CH3OH, CH3COCH3, HCOOH, CH3COOH, and PAN. The authors made some explanations. For example, Biogenic or biomass burning VOCs emissions might be too low during summertime in those regions. In-cloud chemistry may happen. In what regions, such in-cloud chemistry can be a significant source (e.g., HCOOH)? This model bias is too consistent, which made me wonder how this is related to underrepresented emission inventories of VOCs in EDGAR (e.g., contributed by outflow of polluted areas). If biomass burning is not well represented, what does the model-satellite comparison look like for different seasons? Or if it is biogenic, would the simulation become better in winter? I think the current discussion about the potential bias is a bit too general.

Indeed the model bias is very large for those species at high latitude . We agree with the referee that the anthropogenic emissions could be underestimated, but this would unlikely explain such a

homogeneous underestimation. As example, Jacob et al. (2002) showed that anthropogenic emissions would roughly contribute only to 25% of the total CH_3COCH_3 budget. The underestimation in our model at high latitude is roughly of 50%, i.e. larger than the total anthropogenic contribution. Furthermore, if anthropogenic emissions would be the main source of error, we would find a constant underestimation also near source regions, which is, again, not the case for CH_3COCH_3. On contrary, for CH_3OH, the anthropogenic emissions can easily be distinguished in the model results, both, in North America and in South Asia (see Fig.6), and they are clearly higher than the satellite observations suggest.

[...] In summary, due to their different budget, a more detailed analysis is necessary to understand the reason of the discrepancies between model results and observations for the mentioned trace gases, which is well beyond the scope of this study. This work should, however, present the weakness in representing such gases, so that future work could use this study as a starting point.

No changes have been applied here, as more detailed analysis is required to obtain meaningful and scientifically sound conclusions

(6) Line 389-390: What are those concentrations? The magnitudes seem not matching with annual mean surface concentrations of OH.

We are indebted to the Reviewer for spotting this typo (units factor exponent should read "5" instead of "15") which occurred during the typesetting of the manuscript. Indeed, typical average tropospheric concentrations of OH are in the order of 10^5–10^6 molecules cm^{-3}.

The units have been changed from 10^{15} to 10^5.

(7) Section 4.3: What about the model performance of surface OH in different regions? This would affect VOC chemistry and SOA formation significantly. Surface OH concentrations can be quite high in polluted areas (e.g., in China). A recent paper pointed the model underestimation of OH because of the lack of HONO source [Miao et al., 2021]. Would that be a problem in the updated model presented here?

Here we intend to show that the model is reproducing OH similar to that in previous simulations with EMAC. A detailed study on OH and its reactivity on a global scale from this model results is planned in the near future. It has been also shown that the model HO_x budget agrees well when compared to observation in the free troposphere (Tadic et al., 2021). Nevertheless, we agree with the referee that a comparison in polluted regions at the surface could be problematic, also because of the apparent missing HONO source (see for example Elshorbany et al. (2014) for the approach mimicking HONO sources in EMAC, not adopted in this study). We, however, refer to the answer to referee#2: in an evaluation, of a global model of such complexity as the one presented here, the background conditions should be evaluated at first, showing that the model is able to reproduce such conditions. More specific studies on polluted areas should be conducted with dedicated simulation(s) and with more focused analyses.

No changes were made for this point.

(8) Section 4.5: Is the model performance lower in winter than in summer? Residential coal burning in China seems to be a big source of PM2.5 associated with large uncertainty in the inventory. Would that affect the comparison in Asia?

Sadly this cannot be investigated with the dataset of Hammer et al. (2020), as these data are provided as annual average (albeit at high spatial resolution). On the other side, we believe the referee to be completely right: as shown by Pozzer et al. (2012, 2017), and more recently by Miao

et al. (2020), BC and OC are very important for the $PM_{2.5}$ budget in East and South Asia. We think that the large uncertainties associated with such emissions do strongly affect our results in these regions.

In lines 525-2527, we aknowledge the importance of BC and OC emissions for $PM_{2.5}$ over East and South Asia, as well as their large uncertainties that could affect our results.

(9) Line 470-473: What do the model values in higher vertical layers compare with the observations?

We have compared SO_4^{2-} concentrations with observations from aircraft campaigns, as compiled by Heald et al. (2011), similarly to Fig. 26 of the manuscript. [...]. As mentioned in the manuscript, the stations present at high altitude were not representative because of the difficulties in reproducing the correct orography by a model of such resolution. Nevertheless, the simulated sulfate concentrations agree well with the aircraft observations in the free and lower troposphere, with an underestimation in a few cases ([...], ITCT-2K4, ADIENT or IMPEX campaigns), but always within the measurement standard deviations.

We have added a new figure (Fig.22) with comparison of sulfate in the troposphere. In addition, the text was augmented in line 565-570 with the following text: We therefore compared SO_4^{2-} concentrations with observations from aircraft campaigns, as compiled by Heald et al. (2011). The results are presented in Fig.31. The simulated sulfate concentrations agree well with the aircraft observations in the free and lower troposphere, with an underestimation in a few cases (see Fig.31, ITCT-2K4, ADIENT or IMPEX campaigns), but always within the measurement standard deviations, confirming that the vertical profile of sulfate is generally well reproduced in the lower troposphere.

(10) Line 487-488: Nitrate overestimation seems a general problem in chemical transport models. A main reason could be the inappropriate deposition parameters [Miao et al., 2020; Zhai et al., 2021]. But more pronounced bias in winter than in summer is inconsistent with the other studies.

Indeed, the referee is correct that a more pronounced bias in winter is inconsistent with other studies, and this was acknowledged in the manuscript (see lines 482-484). Due to the intricate chemistry involved in the aerosols budget, we prefer not to draw any conclusions or hypothesis for the behaviour of the simulated nitrate.

No changes applied for this point.

(11) Section 4.6.5: It would be good to make some connections for model biases on species. For example, how would the significant biases for the oxygenated organics discussed in the paper affect the OA simulations? So does OH as I mentioned in 7.

Following the referee's previous comments, it is clear that the possible underestimation of OH over polluted regions will results in an underestimation of Organic Aerosols (OA) in the same region. This is visible in Fig. 25 of the manuscript, where the few locations in highly polluted regions over east Asia are strongly underestimated. Nevertheless, it is hardly possible to comment on what effect could have the underestimation of the oxygenated organics mentioned in the paper to the OA. First of all, the discussed trace gases are too light to contribute significantly to OA directly, even for products of their oxidation chain. Although some of these might contribute to OA

mass through aqueous phase chemistry, this process is not (yet) included in the model. Secondly, no OA observations are present for the regions of underestimation, so it cannot be confirmed that the simulated OA deviates from the observations in these regions. Any discussion on this point would be extremely speculative and we prefer to avoid it.

No changes applied for this point.

(12) Line 553- 562: The observations in Asia used here for aerosol composition are quite limited. Studies have suggested the emission inventories for semivolatile and intermediate volatility organic compounds are insufficient and lead to the majority of the model biases in simulating OA [Miao et al., 2021; Zhao et al., 2016] in regions that are influenced significantly by anthropogenic emissions. In Europe, residential wood burning also contributes a lot to the primary OA in winter. Would that be a possible bias? For the rural areas, fragmentation of OA is not included in the model, which may contribute to the overestimation over North America. OA includes both primary OA and SOA. It might be more clear if the two cases are separated discussed.

We thank the reviewer for pointing out these potential sources of discrepancy between our model and observations. In contrast to other observational networks (e.g., EMEP, IMPROVE), EANET does not perform measurements of OA. Therefore, the observations presented here are based on individual field campaigns, which usually last for approximately one month, and not from a network of observational sites that routinely collect samples throughout the year. Therefore, given that the model simulation lasts for only one year, the number of observations used are limited compared to other components (e.g., sulfates from EANET) or other regions (e.g., OA from EMEP in Europe or from IMPROVE in USA). This is emphasized now in the revised text.

ORACLE simulates the contribution of SVOC and IVOC emissions from fuel combustion and biomass burning to OA formation, including additional IVOC emissions that were not accounted for by the traditional emission inventories. However, we do agree with the reviewer that the SVOC/IVOC emission factors used are highly uncertain and can add to the model bias.

Residential wood burning has been identified as the main source of discrepancy in model bias over Europe during winter (A. P. Tsimpidi et al., 2016). This is already mentioned in the text (i.e., as biomass burning). In the revised text, this is clarified by replacing the "biomass burning emissions" in line 557 with "biomass combustion emissions from residential heating".

The effect of fragmentation is included in the aging scheme of ORACLE. The oxidation of organic compounds can result in products with different volatilities due to functionalization (reducing the volatility) or fragmentation (increasing the volatility). To minimize the computational cost, ORACLE uses a simple photochemical aging scheme that efficiently simulates the net effects of fragmentation and functionalization of the organic compounds. However, even the conservative aging scheme of ORACLE assumes that the net volatility transformations point to only one direction (functionalization) which can lead to increasingly higher OA concentrations in lower-volatility bins. This may result in an overestimation of OA at long aging timescales downwind of the source regions. This overestimation is partially offset by the use of MOM for simulating explicitly the degradation of the traditional VOCs.

Our manuscript focuses primarily on the evaluation of the model performance for the tropospheric gas phase composition and secondarily on the general properties of the aerosol composition. We believe that an in-depth analysis of the simulated organic aerosol composition along with AMS factor analysis data can provide valuable insights on the benefits of the proposed modeling approach and we have left this analysis for a planned future study.

In line 655-657, the following lines have been added: As pointed out by Miao et al. (2021) and B. Zhao et al. (2016), the emission inventories for semivolatile and intermediate volatility organic compounds are insufficient and lead to the majority of the model biases in simulating OA in these regions that are influenced significantly by anthropogenic emissions.

Technical remarks [...]
All the technical remarks were be taken into account in the revised version of the manuscript, with the exception of Fig.16, which was not changed.

References

Crippa, M., D. Guizzardi, M. Muntean, E. Schaaf, F. Dentener, J. A. van Aardenne, S. Monni, U. Doering, J. G. J. Olivier, V. Pagliari, and G. Janssens-Maenhout (2018). "Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2". In: *Earth Syst. Sci. Data* 10.4, pp. 1987–2013. ISSN: 1866-3516. DOI: 10.5194/essd-10-1987-2018.

Donahue, N. M., S. Epstein, S. N. Pandis, and A. L. Robinson (2011). "A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics". In: *Atmospheric Chemistry and Physics* 11.7, pp. 3303–3318.

Elshorbany, Y. F., P. J. Crutzen, B. Steil, A. Pozzer, H. Tost, and J. Lelieveld (2014). "Global and regional impacts of HONO on the chemical composition of clouds and aerosols". In: *Atmospheric Chemistry and Physics* 14.3, pp. 1167–1184. DOI: 10.5194/acp-14-1167-2014.

Franco, B., L. Clarisse, T. Stavrakou, J.-F. Müller, A. Pozzer, J. Hadji-Lazaro, D. Hurtmans, C. Clerbaux, and P.-F. Coheur (2019). "Acetone Atmospheric Distribution Retrieved From Space". In: *Geophysical Research Letters* 46.5, pp. 2884–2893. DOI: 10.1029/2019gl082052.

Franco, B., L. Clarisse, T. Stavrakou, J.-F. Müller, D. Taraborrelli, J. Hadji-Lazaro, J. W. Hannigan, F. Hase, D. Hurtmans, N. Jones, E. Lutsch, E. Mahieu, I. Ortega, M. Schneider, K. Strong, C. Vigouroux, C. Clerbaux, and P.-F. Coheur (2020). "Spaceborne Measurements of Formic and Acetic Acids: A Global View of the Regional Sources". In: *Geophysical Research Letters* 47.4, e2019GL086239. DOI: 10.1029/2019gl086239.

Franco, B., L. Clarisse, T. Stavrakou, J.-F. Müller, M. Van Damme, S. Whitburn, J. Hadji-Lazaro, D. Hurtmans, D. Taraborrelli, C. Clerbaux, and P.-F. Coheur (2018). "A General Framework for Global Retrievals of Trace Gases From IASI: Application to Methanol, Formic Acid, and PAN". In: *Journal of Geophysical Research: Atmospheres* 123.24, pp. 13, 963–13, 984. DOI: 10.1029/2018jd029633.

Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P. I. Palmer, and C. Geron (2006). "Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature)". In: *Atmospheric Chemistry and Physics* 6.11, pp. 3181–3210. DOI: 10.5194/acp-6-3181-2006.

Hagemann, S. and T. Stacke (2015). "Impact of the soil hydrology scheme on simulated soil moisture memory". In: *Climate Dynamics* 44.7, pp. 1731–1750. DOI: 10.1007/s00382-014-2221-6.

Hammer, M. S., A. van Donkelaar, C. Li, A. Lyapustin, A. M. Sayer, N. C. Hsu, R. C. Levy, M. Garay, O. Kalashnikova, R. A. Kahn, et al. (2020). "Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998-2018)". In: *Environmental Science & Technology* 13.54, pp. 7879–7890.

Heald, C., H. Coe, J. Jimenez, R. Weber, R. Bahreini, A. Middlebrook, L. Russell, M. Jolleys, T.-M. Fu, J. Allan, et al. (2011). "Exploring the vertical profile of atmospheric organic aerosol: comparing 17 aircraft field campaigns with a global model". In: *Atmospheric Chemistry and Physics* 11.24, pp. 12673–12696.

Jacob, D., B. Field, E. Jin, I. Bey, Q. Li, J. Logan, and R. Yantosca (2002). "Atmospheric budget of acetone". In: *Journal of Geophysical Research: Atmospheres* 107.D10, 4100, doi:10.1029/2001JD000694.

Janssen, R., A. Tsimpidi, V. Karydis, A. Pozzer, J. Lelieveld, M. Crippa, A. Prévôt, W. Ait-Helal, A. Borbon, S. Sauvage, et al. (2017). "Influence of local production and vertical transport on the organic aerosol budget over Paris". In: *Journal of Geophysical Research: Atmospheres* 122.15, pp. 8276–8296.

Jöckel, P., H. Tost, A. Pozzer, M. Kunze, O. Kirner, C. A. M. Brenninkmeijer, S. Brinkop, D. S. Cai, C. Dyroff, J. Eckstein, F. Frank, H. Garny, K.-D. Gottschaldt, P. Graf, V. Grewe, A. Kerkweg, B. Kern, S. Matthes, M. Mertens, S. Meul, M. Neumaier, M. Nützel, S. Oberländer-Hayn, R. Ruhnke, T. Runde, R. Sander, D. Scharffe, and A. Zahn (2016). "Earth System Chemistry integrated Modelling (ESCiMo) with the Modular Earth Submodel System (MESSy) version 2.51". In: *Geoscientific Model Development* 9.3, pp. 1153–1200. DOI: 10.5194/gmd-9-1153-2016.

Laskin, A., J. Laskin, and S. A. Nizkorodov (May 2015). "Chemistry of Atmospheric Brown Carbon". In: *Chemical Reviews* 115.10. Publisher: American Chemical Society, pp. 4335–4382. ISSN: 0009-2665. DOI: 10.1021/cr5006167.

Lelieveld, J., E. Bourtsoukidis, C. Brühl, H. Fischer, H. Fuchs, H. Harder, A. Hofzumahaus, F. Holland, D. Marno, M. Neumaier, et al. (2018). "The South Asian monsoon—pollution pump and purifier". In: *Science* 361.6399, pp. 270–273.

Li, J.-L., X. Zhai, Y.-C. Wu, J. Wang, H.-H. Zhang, and G.-P. Yang (2021). "Emissions and potential controls of light alkenes from the marginal seas of China". In: *Science of The Total Environment* 758, p. 143655.

Mahieu, E., E. V. Fischer, B. Franco, M. Palm, T. Wizenberg, D. Smale, L. Clarisse, C. Clerbaux, P.-F. Coheur, J. W. Hannigan, et al. (2021). "First retrievals of peroxyacetyl nitrate (PAN) from ground-based FTIR solar spectra recorded at remote sites, comparison with model and satellite data". In: *Elem Sci Anth* 9.1, p. 00027.

Martin, S., P. Artaxo, Q. Chen, A. Guenther, S. Gunthe, J. Jimenez, A. Manzi, K. Prenni, U. Pöschl, J. Schneider, and E. Swietlicki (2016). *AMAZE-08 Aerosol Characterization and Meteorological Data, Central Amazon Basin: 2008*. en. DOI: 10.3334/ORNLDAAC/1308.

Miao, R., Q. Chen, Y. Zheng, X. Cheng, Y. Sun, P. I. Palmer, M. Shrivastava, J. Guo, Q. Zhang, Y. Liu, Z. Tan, X. Ma, S. Chen, L. Zeng, K. Lu, and Y. Zhang (2020). "Model bias in simulating major chemical components of $PM_{2.5}$ in China". In: *Atmospheric Chemistry and Physics* 20.20, pp. 12265–12284. DOI: 10.5194/acp-20-12265-2020.

Miao, R., Q. Chen, M. Shrivastava, Y. Chen, L. Zhang, J. Hu, Y. Zheng, and K. Liao (2021). "Process-based and observation-constrained SOA simulations in China: the role of semivolatile and intermediate-volatility organic compounds and OH levels". In: *Atmospheric Chemistry and Physics* 21.21, pp. 16183–16201.

Nölscher, A., T. Butler, J. Auld, P. Veres, A. Muñoz, D. Taraborrelli, L. Vereecken, J. Lelieveld, and J. Williams (2014). "Using total OH reactivity to assess isoprene photooxidation via measurement and model". In: *Atmospheric Environment* 89, pp. 453–463. DOI: 10.1016/j.atmosenv.2014.02.024.

Novelli, A., L. Vereecken, B. Bohn, H.-P. Dorn, G. I. Gkatzelis, A. Hofzumahaus, F. Holland, D. Reimer, F. Rohrer, S. Rosanka, D. Taraborrelli, R. Tillmann, R. Wegener, Z. Yu, A. Kiendler-Scharr, A. Wahner, and H. Fuchs (2020). "Importance of isomerization reactions for OH radical regeneration from the photo-oxidation of isoprene investigated in the atmospheric simulation chamber SAPHIR". In: *Atmospheric Chemistry and Physics* 20.6, pp. 3333–3355. DOI: `10.5194/acp-20-3333-2020`.

Pozzer, A., P. Jöckel, R. Sander, J. Williams, L. Ganzeveld, and J. Lelieveld (2006). "Technical Note: The MESSy-submodel AIRSEA calculating the air-sea exchange of chemical species". In: *Atmos. Chem. Phys.* 6.12, pp. 5435–5444. DOI: `10.5194/acp-6-5435-2006`.

Pozzer, A., A. de Meij, K. J. Pringle, H. Tost, U. M. Doering, J. van Aardenne, and J. Lelieveld (2012). "Distributions and regional budgets of aerosols and their precursors simulated with the EMAC chemistry-climate model". In: *Atmos. Chem. Phys.* 12.2, pp. 961–987. DOI: `10.5194/acp-12-961-2012`.

Pozzer, A., A. P. Tsimpidi, V. A. Karydis, A. de Meij, and J. Lelieveld (2017). "Impact of agricultural emission reductions on fine-particulate matter and public health". In: *Atmospheric Chemistry and Physics* 17.20, pp. 12813–12826. DOI: `10.5194/acp-17-12813-2017`.

Reifenberg, S. F., A. Martin, M. Kohl, Z. Hamryszczak, I. Tadic, L. Röder, D. J. Crowley, H. Fischer, K. Kaiser, J. Schneider, R. Dörich, J. N. Crowley, L. Tomsche, A. Marsing, C. Voigt, A. Zahn, C. Pöhlker, B. Holanda, O. O. Krüger, U. Pöschl, M. Pöhlker, P. Jöckel, M. Dorf, U. Schumann, J. Williams, J. Curtius, H. Harder, H. Schlager, J. Lelieveld, and A. Pozzer (2021). "Impact of reduced emissions on direct and indirect aerosol radiative forcing during COVID–19 lockdown in Europe". In: *Atmospheric Chemistry and Physics Discussions* 2021, pp. 1–23. DOI: `10.5194/acp-2021-1005`.

Rosanka, S., B. Franco, L. Clarisse, P.-F. Coheur, A. Pozzer, A. Wahner, and D. Taraborrelli (2021a). "The impact of organic pollutants from Indonesian peatland fires on the tropospheric and lower stratospheric composition". In: *Atmospheric Chemistry and Physics* 21.14, pp. 11257–11288. DOI: `10.5194/acp-21-11257-2021`.

Rosanka, S., R. Sander, B. Franco, C. Wespes, A. Wahner, and D. Taraborrelli (2021b). "Oxidation of low-molecular-weight organic compounds in cloud droplets: global impact on tropospheric oxidants". In: *Atmospheric Chemistry and Physics* 21.12, pp. 9909–9930. DOI: `10.5194/acp-21-9909-2021`.

Saikawa, E., H. Kim, M. Zhong, A. Avramov, Y. Zhao, G. Janssens-Maenhout, J.-i. Kurokawa, Z. Klimont, F. Wagner, V. Naik, et al. (2017). "Comparison of emissions inventories of anthropogenic air pollutants and greenhouse gases in China". In: *Atmospheric Chemistry and Physics* 17.10, pp. 6393–6421.

Sander, R., A. Baumgaertner, D. Cabrera-Perez, F. Frank, S. Gromov, J.-U. Grooß, H. Harder, V. Huijnen, P. Jöckel, V. A. Karydis, K. E. Niemeyer, A. Pozzer, H. Riede, M. G. Schultz, D. Taraborrelli, and S. Tauer (2019). "The community atmospheric chemistry box model CAABA/MECCA-4.0". In: *Geoscientific Model Development* 12.4, pp. 1365–1385. DOI: `10.5194/gmd-12-1365-2019`.

Shiraiwa, M., Y. Li, A. Tsimpidi, V. Karydis, T. Berkemeier, S. Pandis, J. Lelieveld, T. Koop, and U. Pöschl (2017). *Global distribution of particle phase state in atmospheric secondary organic aerosols, Nat. Commun., 8, 15002.*

Tadic, I., C. M. Nussbaumer, B. Bohn, H. Harder, D. Marno, M. Martinez, F. Obersteiner, U. Parchatka, A. Pozzer, R. Rohloff, M. Zöger, J. Lelieveld, and H. Fischer (2021). "Central role of nitric oxide in ozone production in the upper tropical troposphere over the Atlantic

Ocean and western Africa". In: *Atmospheric Chemistry and Physics* 21.10, pp. 8195–8211. DOI: 10.5194/acp-21-8195-2021.

Taraborrelli, D., M. G. Lawrence, J. N. Crowley, T. J. Dillon, S. Gromov, C. B. M. Groß, L. Vereecken, and J. Lelieveld (2012). "Hydroxyl radical buffered by isoprene oxidation over tropical forests". In: *Nature Geoscience* 5.3, pp. 190–193. ISSN: 1752-0908.

Tsimpidi, A. P., V. A. Karydis, S. N. Pandis, and J. Lelieveld (2016). "Global combustion sources of organic aerosols: model comparison with 84 AMS factor-analysis data sets". In: *Atmospheric Chemistry and Physics* 16.14, pp. 8939–8962. DOI: 10.5194/acp-16-8939-2016.

Vigouroux, C., B. Langerock, C. A. Bauer Aquino, T. Blumenstock, Z. Cheng, M. De Mazière, I. De Smedt, M. Grutter, J. W. Hannigan, N. Jones, et al. (2020). "TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations". In: *Atmospheric Measurement Techniques* 13.7, pp. 3751–3767.

Wang, N., A. Edtbauer, C. Stönner, A. Pozzer, E. Bourtsoukidis, L. Ernle, D. Dienhart, B. Hottmann, H. Fischer, J. Schuladen, J. N. Crowley, J.-D. Paris, J. Lelieveld, and J. Williams (2020). "Measurements of carbonyl compounds around the Arabian Peninsula: overview and model comparison". In: *Atmospheric Chemistry and Physics* 20.18, pp. 10807–10829. DOI: 10.5194/acp-20-10807-2020.

Yáñez-Serrano, A. M., A. C. Nölscher, J. Williams, S. Wolff, E. Alves, G. A. Martins, E. Bourtsoukidis, J. Brito, K. Jardine, P. Artaxo, and J. Kesselmeier (2015). "Diel and seasonal changes of biogenic volatile organic compounds within and above an Amazonian rainforest". In: *Atmospheric Chemistry and Physics* 15.6, pp. 3359–3378. DOI: 10.5194/acp-15-3359-2015.

Zhao, B., S. Wang, N. M. Donahue, S. H. Jathar, X. Huang, W. Wu, J. Hao, and A. L. Robinson (2016). "Quantifying the effect of organic aerosol aging and intermediate-volatility emissions on regional-scale aerosol pollution in China". In: *Scientific reports* 6.1, pp. 1–10.