We thank referee#1 for the very helpful comments which helped to improve the manuscript. Here are our replies:

- *The paper is the fourth part of a paper series about MECO(n), an online coupled model version of COSMO-CLM and EMAC. In this paper tropospheric chemistry is discussed in detail for the first time. The model results are compared to different observations. Potential problems in the model system are discussed in a sufficient way. The paper is well written and I only suggest some minor corrections. Therefore the paper should be published in GMD*

  Reply: We thank referee#1 for these very positive and encouraging comments.

- *As this is the fourth part of a paper series about MECO(n), can you please add a few words about part 1 to 3 in the introduction?*

  Reply: This is indeed a very good point. The corresponding publications of part 1–3 were mentioned in the introduction, but they are not highlighted in detail. We rephrased the introduction slightly to highlight the contributions of part 1–3 in more detail.

- *p.3, line 30: Please tell the reader where this emission data set is published or described. Published elsewhere is not enough.*

  Reply: The mentioned study is not finished yet, therefore we can't give a reference here. This dataset, however, is not used in the present study. Here the 'MACCIty' emission scenario (Granier et al., 2011) is used. We rephrased the sentence slightly to make this more clear.

- *In the introduction you write that one of the advantages of MECO(n) is that for standard CCMs "current computational resources pose an upper limit". Here you write that you have to exchange data between the different instances (which also costs time) and that there are additional waiting times for data exchange. Can you give an estimation how much time you save in total compared to doing a high resolution EMAC simulation?*

  Reply: That is indeed a good question. Actually the computational time (and especially the time needed for exchange of the date) heavily depends on the network of the computing system. In addition also the amount of core per node (which defines the possibilities to distribute the different tasks on the nodes) influence the computational resources needed for MECO(n). For this reasons we hesitated to discuss this in more detail in the manuscript.

  As a rough estimate EMAC at T42 resolution (with 31 vertical layers) needs around 130 node-h per year on 'mistral' at the Deutsches Klimarechenzentrum. The resolution of COSMO is around 6 times higher, as the resolution of EMAC, which ends up in a multiplication of the computation time with a factor of 36 (assuming perfect scaling with the increased

amount of gridboxes). In addition the time-step must be decreased by a factor of 3–4 (for further calculations we assume 3). This gives a computational demand of EMAC at COSMO resolution of roughly 14000 node-h per year ($130 \cdot 36 \cdot 3$). The nested set-up as described in this study with one instance over Europe needs roughly 3200 node-h per year on Mistral. Please keep further in mind, that COSMO features a finer vertical resolution (40 instead of 31 vertical levels) and that one timestep of COSMO is 'more expensive' than one timestep of EMAC, as for example a much more detailed land-model is used by COSMO in comparison to EMAC. Thus there is a benefit of a factor of four based on this estimates.

- *p.6, line 24: You say that you don't consider Averaging Kernels (AK) in your comparisons and therefore you focus on horizontal patterns. This is only possible if AKs do not change in horizontal direction. Probably that is not a problem but can you please check?*

  Reply: Referee#1 is totally right with this remark. For SCIAMACHY Blond et al. (2007) compared model results with and without averaging kernel with satellite measurements and did not find huge differences (compare Fig. 5 b and d in Blond et al., 2007). For OMI differences between the diagnosed tropopause are likely more problematic than not considering AKs (see also discussion by Righi et al., 2015; Jöckel et al., 2016). This differences can also change horizontal patterns, which we overlooked. We therefore decided to rephrase the sentence:

  Therefore, only a qualitative comparison of the data is possible. A quantification of biases is rather based on the comparison with the ground-level observations.

- *p.7, line 22: Here it may sound as the cold bias is due to the coupling but I guess it is the same known problem in COSMO-CLM you mention on page 14. Please also add a short remark here.*

  Reply: We thank referee#1 for this remark. A short note, similar as on page 14, is added in the revised manuscript.

- *p.8, section 4.2: Please clarify where height corrected values are used and where not. Especially in the beginning of the section I don't know if height corrected values are used or not.*

  Reply: We added a note that all metrics are based on the height corrected values.

- *In figure 6 the highest values for the height corrected values seem to appear in Belgium/Netherlands and near Nantes (France). Both areas seem to be rather flat. Can you explain why you have the highest corrections there?*

  Reply: We do not see a large difference between the 'height corrected' and the uncorrected values at these stations. Nevertheless we think that referee#1 might refer to the large differences between the values displayed by the inner and the outer circle. These differences show that much higher

values are measured (inner circle) than simulated by the model (outer circle, 'height corrected'). Likely these differences correspond to large local sources which are not well represented in the used emission database. We improved the paragraph about height correction (see below).

- *Height correction: Can you please give a short description (maybe in the appendix) how the height corrected values are calculated? Is it just scaling with pressure or is it something more complicated?*

  Reply: We rephrased the description of Section 3 regarding the height correction to explain the procedure. As no inter- or extrapolation of the model results is performed we think that this description is sufficient:

  To allow for a fair comparison between EMAC, COSMO/MESSy and the observations a 'height correction' of the model results from EMAC and COSMO/MESSy is applied. For the EMAC data the geometric height of each station is compared with the geopotential height of the individual model levels at the corresponding gridbox in which the station is located. For the COSMO data the procedure is analogue to EMAC, but the height of the model level instead of the geopotential height is chosen. We pick the model results at the vertical level, where the geopotential height (EMAC)/model level height (COSMO) is nearest to the geometric height of the station. No interpolation of the model results between different levels is performed. However, this option only works, if the station is located higher than the ground of the lowest model layer. In the opposite case, the values of the lowest model layer are chosen and no extrapolation of the simulated data is performed. This height correction is very important, especially over mountainous terrain, as the topography is much finer resolved by COSMO/MESSy. In other words, if the observations would always be compared to the model values at the lowest model level, COSMO/MESSy would outperform EMAC solely because of the finer resolved topography. The usage of these height corrected values is indicated in the corresponding sections.

- *p. 19: MECCA: Can you please specify which version of the recommendations from JPL you used? Meanwhile the newest version is from 2015, which (according to your supplement) is not used.*

  Reply: We used the version of the evaluation cycle 17 (Sander et al., 2011), not the newest from cycle 18. A corresponding note is added in the revised manuscript.

- *p.1, line 7: "... and a one ..."; change to "... and one ..."*

  Reply: Done.

- *p.3, line 25; p.8, line 33: line to long (happens several times)*

Reply: We changed the long terms COSMO(12km)/MESSy and COSMO(50km)/MESSy to shorter abbreviations CM12 and CM50 in order to overcome some of the problems.

- *p.8, line 5: "we are do not present" → we do not present*

  Reply: Done.

- *p.8, line 15: Atlantic sea → Atlantic ocean*

  Reply: Done.

- *p.9, line 28: located in?*

  Reply: Yes indeed. Fixed.

- *p.13, line 13: mixing ratios*

  Reply: Done.

References

Blond, N., Boersma, K. F., Eskes, H. J., van der A, R. J., Van Roozendael, M., De Smedt, I., Bergametti, G., and Vautard, R.: Intercomparison of SCIAMACHY nitrogen dioxide observations, in situ measurements and air quality modeling results over Western Europe, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, 10.1029/2006JD007277, http://dx.doi.org/10.1029/2006JD007277, d10311, 2007.

Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., van der Gon, H. D., Frost, G., Heil, A., Kaiser, J., Kinne, S., Klimont, Z., Kloster, S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M., Smith, S., Thompson, A., Aardenne, J., Werf, G., and Vuuren, D.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980 - 2010 period, Climatic Change, 109, 163–190, 2011.

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

Sander, S. P., Abbatt, J., Barker, J. R., Burkholder, J. B., Friedl, R. R., Golden, D. M., Huie, R. E., Kolb, C. E., Kurylo, M. J., Moortgat, G. K., Orkin, V. L., and Wine, P. H.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, JPL Publication 10-6, Jet Propulsion Laboratory, Pasadena, available at: http://jpldataeval.jpl.nasa.gov (last access: 15. July 2016), 2011.

Righi, M., Eyring, V., Gottschaldt, K.-D., Klinger, C., Frank, F., Jöckel, P., and Cionni, I.: Quantitative evaluation of ozone and selected climate parameters in a set of EMAC simulations, Geosci. Model Dev., 8, 733–768, 10.5194/gmd-8-733-2015, http://www.geosci-model-dev.net/8/733/2015/, 2015.