We thank referee#2 for the very helpful comments which helped to improve the manuscript. Here are our replies:

- *This paper describes a new nested, coupled system for chemistry climate modelling, which has online coupling of the nested grids. This paper evaluates the gas phase tropospheric chemistry in this new model in comparison to the coarser resolution model and to observations. I think this paper is successful in its aims of describing and evaluating the new model. Subject to my minor comments below, I recommend this paper is accepted*

  Reply: We thank referee#2 for these very positive and encouraging comments.

- *A point with respect to rainbow colour scales – there is a good argument for not using them, as they artificially distort the field that they are visualizing, and they also cause problems for people with colour blindness. I recommend changing the rainbow colour scales to ones that do not suffer from these flaws. See here for more about this: http://www.climate-labbook.ac.uk/2014/end-of-the-rainbow/*

  Reply: We agree with referee#2 that the rainbow colour scale has some problems. For the revised manuscript we changed the colour scale of the figures which used the rainbow scale before.

- *P1 line 21: This sentence is awkward and difficult to read: Especially, as some of the relevant processes (for example tropospheric ozone chemistry) are non-linear, it is desirable to resolve smaller scales, since with finer resolution the capabilities of chemistry-climate models in simulating species like ozone or nitrogen dioxide can be enhanced. Possible alternative: It is desirable to resolve smaller scales because finer resolution chemistry-climate models can simulate species like ozone or nitrogen dioxide better, as some of the relevant processes are non-linear (for example tropospheric ozone chemistry).*

  Reply: We thank referee#2 for this comment. We adopted the suggested change.

- *P7L4: This sentence is awkward and hard to understand: To allow for a fair comparison between EMAC and COSMO/MESSy always the value of the model layer which is nearest to the elevation of the station is selected. Suggested alternative: To allow for a fair comparison between EMAC and COSMO/MESSy, the value of the model layer which is nearest to the elevation of the station is always selected.*

  Reply: We thank referee#2 for this comment. According the comments from referee#1 the whole paragraph was rephrased in order to explain the 'height correction' in more detail.

- *P7L32: "The maximum over the Mediterranean sea is underestimated in COSMO(50km)/MESSy. EMAC simulates a higher extend of this maximum, which better corresponds to the satellite measurements." The word*

"extend" should be "extent". I also don't agree with this when I look at fig 3. What is meant by a "higher extent"? Higher in altitude – which we cannot see in this figure? Higher latitudes – which doesn't agree with what I can see from the figure? Larger figures with a better colour scale may help to show the reader what you are trying to convey here. The same applies to the following statement about the Alps, as it's hard to see details over such a small region.

Reply: Referee#2 is right with the fact, that our sentence was unclear. We rephrased this part for the revised manuscript:

The overall patterns of all three ozone columns look very similar with a strong north-south gradient. Investigating into more detail, some differences are apparent. COSMO simulates the maximum ozone column mainly along the coastline of Turkey. Compared to this the maximum in EMAC extents further to the West and South. This corresponds better to the satellite measurements, which show the largest values in the whole south-eastern part of the Mediterranean Sea.

Further the colour scale was changed, as mentioned above and the panelling of the figure was changed.

- *Sec 4.1: How good are the satellite retrievals? Is there any bias that may account for some of the differences – a difference between the land and sea? I am not an expert in satellite retrievals, so I think a sentence or two about whether there are any biases in the satellite data would be helpful here.*

  Reply: We add a note about the problems with the satellite retrievals in Sec. 3 where the data are described. We felt that the discussion of these biases is more suitable in this section than in Sect. 4.1. A reference to this discussion was added in Sect. 4.1.

- *Sec 4.2.1 on Taylor diagrams. I don't think you explain anywhere what the x axis is on the Taylor diagram.*

  Reply: We rephrased the first part of the description to better explain the meaning of the x-axis.

  For a more quantitative comparison, Taylor diagrams (details are given by Taylor, 2001) are calculated. These diagrams combine the (normalised) standard deviation (as radius) and the correlation between the observed and the simulated time series (as angle). The observational reference point is marked with REF on the x-axis. The calculations are based on hourly averaged model output and observations, respectively. The bias in percent between the simulated and observed ozone concentration is displayed by the size of the symbols. The dashed circles indicate the root mean square error. Again, only the height corrected values are used, which improve the results of EMAC

considerably. The Taylor diagrams for the uncorrected cases are part of the Supplement (Sect. S1.4).

- *Have I also understood correctly that the EMAC model does better according to the metrics described in this section than the COSMO/MESSy model? If this is the case then maybe some commentary to explain why this is would be good here. Or to refer forward to a section where you discuss this*

  Reply: Yes indeed, EMAC performs better according to the described metrics. This is likely due to the problems with the diurnal cycle in COSMO. In the revised manuscript we added a short note with references to the next sections.

  The overall better results for EMAC compared to COSMO are likely caused by the deficits in the representation of the diurnal cycle in COSMO as discussed in Sect. 4.2.2. A more detailed discussion about potential reasons for this is provided in Sect. 5.

- *Sec 4.5: I assume the unit "a" means "annum". It took me a few seconds to work this out, and it isn't explicitly stated anywhere. I would have uses "years" or abbreviated to "y" or "yr", as this seems to be the convention in the literature. I don't know if GMD have a policy on this.*

  Reply: Yes, 'a' stands for the latin 'annum'. We know that in many other publications the abbreviation 'y' or 'yr' is used. In the informative annex C, the International Standard ISO 80000-3 proposes the symbol 'a ' to represent a year of either 365 or 366 days. Also the IUPC recommend the usage of 'a' (http://media.iupac.org/publications/books/gbook/IUPAC-GB3-2ndPrinting-Online-22apr2011.pdf). As GMD points to this document in the guidelines we think that 'a' is the right choice.

- *Unless the reader has a good feel for what numbers to expect for the methane lifetime, the numbers in this section are not very helpful, in my opinion. Most non-specialists will simply know that the methane lifetime is approx 10 years globally, however the numbers in this section are very different to that. Some context or literature values would help here.*

  Reply: We agree with referee#2 on this point. The values we are showing are only for a part of the globe and are not comparable to 'typical' numbers in the literature. We use the lifetime simply as a proxy for the tropospheric oxidation capacity to check if this capacity changes between both models. In the revised manuscript we added a more detailed note on this.

  As shown by Jöckel et al. (2016) the methane lifetime against OH of the *RC1SD-base-10a* simulation, which has a very similar set-up as used in the present study (see Sect. 2.3), is around 7.7 a for the year 2008. As analysed in detail by Jöckel et al. (2016) this is at the lower end compared to results from other models which are mainly in the range from 8–9 a. The values we present

here are not directly comparable to these global estimates of the methane lifetime, as we calculate the lifetime only for a part of the globe. Here, for a more detailed comparison of the results from EMAC and the two COSMO/MESSy instances we further calculate the lifetime separately for three different vertical layers of the atmosphere: From the ground to 850 hPa, from 850 hPa to 500 hPa and finally from 500 to 200 hPa. For this we sum up all grid boxes within the respective area.

- *Abstract, Line 15: Change to: "In comparison with observations, both EMAC and COSMO/MESSy show strengths and weaknesses."*

  Reply: Done.

- *P2L18 – "consistence" should be "consistency"*

  Reply: Done.

- *P4L15 remove second comma:"Thus, it is desirable that all..."*

  Reply: Done.

- *P7L5: Rearrange to: This is very important, especially in mountainous terrain, as COSMO/MESSy resolves the topography much better.*

  Reply: Done.

- *B7L8: suggest adding a comma before the word measurements to make this sentence a bit clearer*

  Reply: Done.

- *Caption fig 5: please specify that the middle row is for June 2008*

  Reply: A note is added

- *P9L28: missing "in"*

  Reply: Done.

- *P12L10: Remove first comma: "In order to check if the vertical distribution of ozone is well simulated,..."*

  Reply: Done.

- *P16L8: Remove first comma: "To investigate if we can..."*

  Reply: Added.

- *P16L16: "In addition, also" – you don't need both of these terms in this sentence*

  Reply: Indeed - wee agree with referee#2.

- *P16L17: "of" should be "off"*

  Reply: Done.

- *P17L15: The word "especially“ seems a bit out of place to me here. Suggest "Particularly“ instead.*

  Reply: Done.

- *P17L23: Again, "especially“ seems out of place here. Suggest "particularly,“ or "in particular,“. Later in this sentence "as by the coarser“ should be "than by the coarser“*

  Reply: Done.

- *P17L26: "good“ should be "well“*

  Reply: Done.

- *P17L27: Another sentence beginning with especially – perhaps you wish to keep this one, however I'd remove the word as the sentence works as well without it.*

  Reply: Removed.

- *P17L28: Sentence starting with "The comparison“ – remove the first comma. The final "to“ should be "too“.*

  Reply: Done.

- *P17L32: remove first comma*

  Reply: Removed.

References

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.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmos., 106, 7183–7192, 10.1029/2000JD900719, http://dx.doi.org/10.1029/2000JD900719, 2001.