Review of "The high-resolution version of TM5-MP for optimised satellite retrievals: Description and Validation"

The work described in this paper is certainly interesting and the motivation for the work is clearly explained. The general structure and flow is good. However, in the current state, it is difficult to assess the reproducibility and scientific quality of this work for two main reasons: the experimental setup and differences between the two model resolutions are not clearly explained; the analysis and presentation of the results is not well structured and leaves a number of unanswered questions.

Major Comments

1) A much more rigorous description of the two model setups and their differences is required. There is currently no mention of the vertical resolution although this is later reported as being important for e.g. STE fluxes. In its current format the model description in Section 2 is lacking information and is written in a way that makes it unclear which modifications are applied to the new TM5 model version and which to the higher resolution.

2) A paragraph should be added in Section 2 to clearly describe all details for the two model integrations used for this work. In particular, the authors should specify:
- start date of integrations and run length
- chemical initial conditions and spin up periods
- details of the analysis used (horizontal, vertical and temporal resolution).
Are these the same for both model resolutions or are they different? Are the simulated model years the same used for the emissions and observational datasets?

3) The use of EMEP observations in it's current state is confusing.
In line 242-243 the authors state that sites in "Norway, Finland, The Netherlands, Belgium, Poland, Germany, Spain, Italy and Portugal" are used for comparison. Why just use EMEP stations from the above and not the ones in other countries? Or is this a mistake (see later)?
The same uncertainty regarding only selected sites being used and lack of explanation on why they are used is present throughout the manuscript and affects the interpretation of the results.

In Fig 3, EMEP sites from " Finland, The Netherlands, Belgium, Poland, Slovakia and Italy" are used, aggregated by nation. The authors should explain why just these six countries? Why aggregate the sites? Poland and Slovakia should be mentioned in the list of sites in Section 2!

In Fig 4, four EMEP sites are selected for comparison. Why these 4 sites? Again the sites in the Czech Republic and Great Britain used in this figure are not listed in Section 2.

In Fig 5, two sites are selected for comparison and again no explanation as to why those specific sites are used.

Similarly, Table 6 and 7 use yet two different subsets of EMEP stations for comparison without explaining the reason for their choice.

In Fig 8, four selected EMEP sites are shown (from Norway, Germany, Austria and Slovakia). Same issues as above.

If the purpose of the comparison with EMEP is to evaluate the model performance in the new configuration, as well as addressing the differences in model resolution, the current analysis is not convincing. Comparison of model data with tropospheric ozone column from satellite would help better evaluate model performance on the global scale. This could also lead to better evaluate the model ozone profiles which currently show significant discrepancies with MOZAIC data. Further comparison with EMEP surface sites (and other campaign data) would then add to the analysis, so far as the comparison is done across all suitable sites and a clear explanation is given if only a subset of sites is selected.

4) In Section 5 the authors provide an analysis of budget terms for tropospheric ozone and compare these at the two different model resolutions. They state: "the chemical tropopause calculated for 3x2 is applied for the analysis of 1x1 budget terms to ensure that a valid comparison is performed". However, if convection and convective transport is significantly different in the two model resolutions (as the authors suggest) the position of the chemical tropopause at 1x1 should be at a higher altitude in the tropics compared to 3x2. Using the 3x2 chemical tropopause to analyse 1x1 budget terms is in my view inconsistent and the reduction in STE term with increasing resolution is likely to change if the 1x1 chemical tropopause is used.

Minor Comments

line 16: change 'coastal' to 'coastal regions'

line 66: 'and the extent of mixing by convective upwelling (i.e. land type)'; it is not clear why land type is mentioned here

line 75: change 'where' to 'in which'

line 103: change 'regional' to 'regional domain'

line 105: 'massivelyntersta' ? Correct.

line 218-219: Explain which method is used for coarsening of emissions (area averaged? Linear interpolation?) and comment on its suitability

line 220-222: not clear from the text how the lightning Nox parameterisation works. Does it use convective precipitation (line 220) or convective flux (line 221)? And if it uses convective flux (from ERA?) why does it need rescaling?

line 235: Fisher et al., 2015 is not present in the References section

line 292-293: 'more accurate temporal distribution of regional 222Rn emissions at 1x1'; is the timing of the emission different at 1x1 compared to 3x2? If so it should be clearly stated and it should be explained why this would be desirable or necessary.

line 295: 'coasted' to 'coastal'

line 308: '(i.e.) orography and land type' to '(i.e. orography and land type)'

line 323: 'that the significant differences exist' remove 'the'

line 362: 'the also the efficiency' remove 'the'

line 409: 'the surface deposition flux to e.g. should...' correct

line 514: 'related to lower [OH] (i.e.) chemical production' remove '(i.e.)'

line 596-597: 'that the chemical production term is too.' ? Not clear, rephrase

line 665: 'across diverse global regions'? Either 'globally' or 'across various regions'