General comments

This manuscript presents an evaluation and sensitivity analysis of the WRF-CHIMERE model, with a focus on biogenic secondary organic aerosol and the Hyytiälä monitoring site in Finland. While the topic is within the scope of GMD, the current version of the manuscript is not suitable for publication in the journal. The evaluation of the model is too limited to be of general interest, as it only covers a single site for two and a half months, which is not sufficient for a regional CTM model evaluation paper in GMD.

The paper does not present any substantial novel concepts, ideas, tools, or data, and does not represent a significant advance in modeling science. The short simulation period and apparently poor emission data for isoprene make it difficult to draw firm conclusions from the study, and there are few interpretations or conclusions presented in the manuscript.

If the model evaluation were extended to include more sites across Europe, it could be of sufficient interest to warrant publication in GMD. Organic carbon (OC) measurement data from 2019 are available from ebas.nilu.no for about 30 different regional sites in Europe, which would be a valuable addition to the evaluation. Given the poor agreement with observations for isoprene at Hyytiälä, it would be interesting to include isoprene measurements from other European sites (data from almost 20 sites are available for 2019 in ebas – including data from Pallas in Finland). It would also be useful to investigate the effect of isoprene emissions on ozone across Europe, using observations from the many regional background sites available in ebas. For some unclear reason, the comparison with observed PM_{2.5} measurements was restricted to sites in Spain and Italy (and this comparison is not discussed in detail in the manuscript). It would be interesting to evaluate the model's performance for PM_{2.5} across all of Europe (including Hyytiälä) and discuss these results in the manuscript.

In its current form, I recommend that the manuscript be rejected for publication in GMD. However, if the authors are willing to substantially extend their comparison with observations across Europe, their work may be reconsidered for publication.

Specific comments

- Section 2.2 only describes BSOA how did you treat SOA from anthropogenic VOC? Did you include a VBS-treatment with (or without) aging also for ASOA?
- You base your VBS scheme on Hodzic and Jimenez (2011) their scheme only included SOA from OH-reactions. Did you include SOA formation from oxidation by ozone and/or NO₃ radicals? Please provide details about how the BVOC + O₃ and BVOC + NO₃ are treated in the model.
- How did you treat SOA from SQT?
- Hodzic and Jimenez (2011) only had a single monoterpene species (TERP) you split the MT into four different species; please provide details of the differences in SOA yields (and reactivities) for the different MTs.
- Hodzic and Jimenez (2011) also included SOA production from biomass burning, POA ageing, acidenhanced BSOA production and anthropogenic pollution-enhanced SOA production. Did you include all (or any) of these SOA formation reactions?
- Lines 146–147: You use ΔH_{vap} of 36 kJ mol⁻¹ Hodzic & Jimenez 2011 used 88 kJ mol⁻¹ please explain why you chose the lower value.

- Did you only include the organic mass (including or excluding particulate water; or the full particle mass) when calculating the gas-particle partitioning of the SVOCs?
- How was the deposition of gas-phase SVOCs treated in the different model simulations?
- Line 165: How were the annual anthropogenic emissions from CAMS "hourly distributed" for the simulation period? Please provide some details (and/or reference) regarding the temporal distribution of the emissions.
- Did you include any emissions from biomass burning (wildfires) in the simulations?
- Lines 179–181: Were the initial and boundary concentrations of aerosols and gases taken from LMDz-INCA3 simulations for 2019 or for some other year? Were the boundary concentrations constant or varying in time?
- Lines 198–204: The selection of PM_{2.5} data for the model evaluation is very odd for a study focussed on the "Finnish boreal forest". Also, there is no discussion of the results of the PM_{2.5} evaluation in the manuscript (only a figure and a table in the Supplement, with no accompanying text). As mentioned in my General comments, I think that the comparisons of modelled and measured PM_{2.5} should be made for all of Europe and definitely include data from Hyytiälä. The Spanish and Italian sites included here are probably among the least interesting sites for the boreal Finnish forest so unless the PM_{2.5} evaluation is extended I think it should be removed completely.
- Lines 252–254: In what sense are your modelled BVOC emissions "generally in line with" the data presented by Hellén et al. (2018)? I do not think they are very much in line at all. For the summer 2016 they give (in their Table 1) mean total MT concentrations of 427 ppt and only 11 ppt isoprene and 13 ppt SQT; this seems very far from your ratios between the three BVOC types.
- Figure 8 and Table 4. How does the statistics for the BVOCs change if you exclude the sawmillemission related time periods from the analysis? Does this improve the statistics for MT significantly?
- Could you please provide some data on how well the SQT are modelled compared to measurements? If no measurements are available from 2019 it would be good to show the mean modelled concentration compared to the observations from some other year.
- Lines 269–276, regarding the isoprene evaluation there are isoprene measurements from about 20 sites in Europe during 2019 in ebas; a comparison of your model results to these data should be included in order to determine if the problems of overestimated isoprene emissions in MEGAN really are severe across Europe, or if it is more of a local problem in the forests around Hyytiälä.
- Lines 286–287: The sentence about the PM_{2.5} comparison gives no useful information regarding the contents of the section (Analysis and source apportionment of OA). A proper model evaluation of PM_{2.5} results should be included elsewhere in the manuscript (or not at all) and include sites all over Europe (not only in Spain and Italy, which makes no sense at all). The Supplemental

information Figure S1 and Table S2 both lack results from the C_5H_8 -emissions-off simulation; this scenario should also be included, in case a more extended PM_{2.5} model evaluation is included.

- Lines 290–293: "Extremely low OA concentrations are missed by the model, and there is a tendency of zeroing out such concentrations throughout the entire simulations (Figure 9 and Figure 10). The latest might suggest uncertainties in the background OA fields used in the model and/or in the concentrations injected at the very boundaries of the coarser domain (i.e., long-range transport)." These two sentences are unclear. As far as I can see from Fig. 9 and 10, the model produces lower OA concentrations than the measurements? What do you mean by "zeroing out such concentrations"? Regarding the suggested "uncertainties in the background OA fields" how do you set the background OA concentrations in the model?
- Figure 13. It would be interesting to see the same type of plots for the large-scale model domain. Do the isoprene emissions have similar effects also in continental and southern Europe?
- Figure 15 and lines 361–362: You suggest that the high night-time concentration of NOx in the model could be due to a "too shallow planetary boundary layer in the model". Is this a general problem in WRF-Chimere or a local problem in the region around Hyytiälä? Please provide evaluation against NOx measurements at other background stations in Europe, to make this manuscript of more general interest as a model evaluation study.
- Considering the overestimated NOx concentrations in the model (by a factor of two according to Table 4) how does this influence the SOA-production? Will this lead to an overestimation of the "high-NOx"-path for SOA formation and underestimation of the "low-NOx"-path?

Technical corrections

There are a large number of minor language errors in the text and the manuscript would benefit from thorough language editing. I will not go through all language mistakes (I think that is the job of the author-team and possibly the language editing of the journal). I only list some of the mistakes I spotted here. In case the manuscript is revised, please make sure to have the language checked and corrected before resubmission.

- Line 29: "We attributed the latest" what do you mean by that? Please reformulate.
- Line 72: "of the air mass" → "of the organic aerosol"?
- Line 88: "for the latest" what do you mean? Please reformulate.
- Line 121, regarding the CAMS operational ensemble: I guess it is not WRF-Chimere that is part of the CAMS ensemble but rather the Chimere CTM using meteorology from the IFS model?
- Line 127 (and at many places in the manuscript): Your definition of the "astronomical summer" is not correct – please remove the word "astronomical" from the period description (it should be removed everywhere in the manuscript) – the astronomical summer of 2019 was 21 June – 23 September.
- Line 140. Oxidization \rightarrow Oxidation
- Figure 2 caption, lines 763–765; the text "The text font size represents the tendency of both particles and gas-phase organic material (OM) to transition in the one or the other phase (i.e. larger font size indicated a better attitude towards that phase, and vice versa)" is awkward and needs rephrasing; you probably mean something like "stronger affinity" for the phases rather than "better attitude".

- Captions of Figures 5,6,7,11 and 12: change "the astronomical summer of 2019" to the actual time period included (15 Jun–30 Aug, 2019)
- Line 289: "flat diurnal of $OA'' \rightarrow$ "flat diurnal variation of OA''
- Line 319: DMSP \rightarrow DMPS
- Table 4 is not referred to anywhere in the text
- Table 4 and Table 5. SEMAR II \rightarrow SMEAR II
- Lines 381 and 427: could \rightarrow cloud
- Line 404: respond \rightarrow response