

**Review of the manuscript “Atmosphere-Ocean-Aerosol-Chemistry-Climate Model SOCOLv4.0: description and evaluation” submitted by Sukhodolov et al. for publication in GMD**

The authors describe and evaluate the new version v4.0 of the SOCOL CCM. It seems that the model performs comparatively well and I would like to congratulate the authors to this model development. I find that overall both model description and evaluation are done in an excellent way, the paper is well written, and it will be a very useful reference for future studies using this model. I have only minor suggestions that I would like to see considered before publication. I start my list with a few general issues and then proceed in the order of appearance in the text.

As a major novelty in comparison to the predecessor SOCOLv3 it is mentioned that the model is now coupled interactively to an ocean and interactive aerosol at the same time. It would be good to say clearly if the authors think that any of the evaluated features are influenced by this novelty. I wouldn't consider it a problem if not but also if that is the case it should be clearly spelled out. In this context I'm a bit at odds with the statement in the introduction that “the interaction of Earth system components is required for reasonable model performance in most cases”. What are these cases? And if it really were “most cases”, it should be possible to clearly identify the benefits of the full coupling in the model evaluation.

Model tuning: A major component of climate model development is the model tuning which should be adequately represented in the model description. My understanding is that the authors haven't changed anything in the model tuning in comparison to the reference model (without chemistry and aerosols) as presented by Mauritsen et al. (2019). If that is true it should be mentioned explicitly, if not, the tuning strategy would need to be described.

Computational performance: I would find some information useful on the increase of computing time due to the interactive treatment of chemistry and aerosols in comparison to the base model. I know that this may be machine and configuration dependent but some numbers for typical configurations would be informative.

Figures: Most of the figures appear fairly blurred in the pdf accessible to me. This should be taken care of before publication.

L95: “an upgraded dynamical core” Has there really been a change of the dynamical core in the family of ECHAM atmosphere GCMs since version 4? If yes, please specify.

L153 My understanding is that the sea ice model has plenty more than just two parameters that could be used for tuning. But it is also not clear to me what the

authors mean when they speak of “parameterizations of changes in ice”. As said above, if the authors did any tuning it should be mentioned.

L193 “tropospheric GW sources prescribed as a function of latitude”: Which GWs? Was this option used in these experiments? If yes, using what function.

L194 “RRTMG”: If SOCOL follows Mauritsen et al. (2019), then the PSRAD variant of RRTMG is used, which should be mentioned.

L198 “radiation calculation of SCOCOLv4”: I guess this is true only for the radiation calculation interacting with the model dynamics, not photochemistry, right? It would be useful to mention that these things are, to my understanding treated inconsistently as it is done in many other CCMs.

L198 “uses prognostic tracer concentrations ... except CO<sub>2</sub>” This seems to clash with L305 where it is claimed that concentrations of GHGs from scenarios are used. Or do you mean emissions for the latter? What about CH<sub>4</sub> and N<sub>2</sub>O, which are not mentioned at line 305?

L202 MACv2-SP: I thought that SOCOL calculates at least sulfate aerosols interactively. Please explain where the climatology is used and where interactive aerosol.

3.1.2: In particular here it would be interesting to know if SOCOL performs in any way different than the base model without interactive chemistry and aerosols.

ENSO evaluation: I don’t think a Nino3.4 time series is very useful for evaluating the model performance. A power spectrum and teleconnection patterns would be much more helpful. But again, if SOCOL performs very similar to the base model it would be sufficient to just state that.

L398 “conclude that forcing and model response are adequately represented”: Well in principle it could also be that the forcing is too strong and the response too weak.

L422 It would be good to explain the QBO nudging approach or to provide a reference.

L615: For the ozone bias in the lowermost stratosphere only chemistry is discussed as a potential reason. Could it be that the tropopause is at a slightly wrong altitude which may have large effects in this regions of large vertical gradients?

L656 “southern (polar) night jet”

Figure in the appendix: I don’t know the journal policy concerning this, but if the appendix contains only three figures I would consider it more convenient to not have an appendix but include the figures in the main part.

Figure 15: I wouldn’t speak of “normalized”, but that an anomaly with respect to some reference is presented.

L769 “whereas for SOCOLv4”: I find the whereas confusing because at the first read I related it to the tropopause definition.

L807 “underestimating the aerosol lifetime” Well, if the observations are flawed, as mentioned in the following sentence, it’s not sure that this is an underestimation.

L822 "SOCOL includes an interactive ocean" As said in the beginning: If this is put so prominently as an achievement, effects of that should be discussed.