

# Answers to Reviewer 1

We would like to thank the anonymous reviewer #1 for his/her great review of our publication and provide here the answers to his/her questions.

*This study presents results from the SOCOL-MPIOM coupled chemistry-atmosphere-ocean model. Rather than focusing on validating the overall performance of the meteorology and chemistry compared to observations, the study focuses on the impacts of interactive chemistry on the atmosphere by comparing experiments with and without the chemistry module implemented. They go on to present historical all forcing experiments and focus on the effects of solar variability on surface temperatures. The paper is generally well written and the examination of the effects of coupling the chemistry will be of interest to the wider community, particularly in the context of current Earth System Model development activities. I therefore deem the topic of sufficient interest for the GMD community. However, I think a number of aspects of the paper could be improved before I would recommend publication. In particular, I think the discussion could be more focused onto some of the fundamental aspects of the model behavior, such as the climate sensitivity, rather than on a diverse range of topics that don't necessarily fit together well. I therefore think removing some of the topics, such as the lengthy discussion of the role of solar variability during the Maunder and Dalton Minima, would shorten the paper and help to increase the overall impact. I have made some suggestions for ways to do this below.*

## **Recommendation: Major revisions.**

Thank you for your comments and questions. We revised the manuscript and focus now on three aspects only:

1. The differences between CHEM and NOCHEM (effect of chemistry-climate interactions)
2. Climate sensitivity
3. Temperature increase since 1850

The discussion of the solar forcing has been almost completely removed. Furthermore, we show results from the former M1 and M2 simulations only and removed the simulations with the larger solar forcing (L1 and L2). Consequently, we substantially shortened the discussion of temperature variability during the MM and DM.

For the evaluation of the climate sensitivity of SOCOL-MPIOM, we decided to re-perform the climate sensitivity experiments under a pre-industrial climate state. These experiments are initialized using restart conditions from CHEM. Consequently we also updated the description of the climate sensitivity experiment. With the different climate state, the TCR/ECS estimates changed as well and we extended the feedback analysis, as suggested.

Please note, that the results of the sensitivity experiments also changed slightly. We found an error in the detrending of the experiments, which lead to an overestimation of the response in the order of 0.05 K. The temperature increases for the different experiment are therefore slightly reduced.

## **Major points**

*Section 4.1: This part comes across more as an evaluation of whether the Shapiro et al. (2011) solar forcing dataset is plausible rather than anything specific to do with the SOCOL-MPIOM model. The conclusions are mixed, with poor model-proxy data agreement during the MM, but better agreement during the DM. There is also a strong solar induced warming during the early 20th century, which contributes to an overestimation of the temperature trend compared to observations. These results raise doubts as to whether the Shapiro dataset is plausible for use in climate models.*

We substantially reduced the solar forcing discussion and the model-proxy comparison in the revised manuscript. Section 4.1 has been condensed into a single paragraph and the focus of this part of the manuscript is now on the temperature increase in the industrial period.

*This brings me to another point, which is that the authors offer little justification for why they have used the Shapiro et al. (2011) dataset (L9 3025) rather than another more moderate construction (e.g. Wang et al. (2005)). The authors state that the Shapiro construction is outside of the uncertainty range given in the IPCC and their results appear to confirm that this does not produce results that can be squared with observations (L9-14 3052). The reasons for including this particular solar forcing dataset in the model therefore needs more justification.*

The experiments presented have been performed in a research project focusing on the role of solar variations on the climate system and the evaluation of the Shapiro et al. spectral solar forcing reconstructions. The Shapiro forcing is therefore not only used to evaluate the model SOCOL-MPIOM, but also in a number of sensitivity studies presented in Anet et al. (2013a, 2013b, 2014).

For the revised manuscript we decided to consider only the experiment with the 'medium' amplitude solar forcing, which has a Maunder Minimum to present day amplitude of  $3 \text{ Wm}^{-2}$  and agrees better with other so-called 'strong solar forcing reconstructions' (e.g. Lean 1995, Bard 2000).

*More generally, I am not convinced that the detailed discussion around the role of solar forcing during the Dalton and Maunder Minima really fits into this study. There are other more relevant aspects of the model evaluation that could be expanded upon (see below) and the solar specific aspects might be better suited in a separate paper. I would therefore recommend taking out most of Section 4.1. This would also help to shorten the paper, which in its current state feels a bit too long.*

We followed your advice and removed most of the former Section 4.1 from the manuscript.

*Section 3.3: The model is shown to have a too high Equilibrium Climate Sensitivity and Transient Climate Sensitivity. However, little attempt is made to explore the reasons for this. The Gregory et al. (2004) method could be used to separate out longwave and shortwave clear and cloudy components (see e.g. Andrews et al. (GRL, 2012)) and this could help to elucidate where the model's feedbacks come from. Some further analysis of this would help to strengthen discussion on L15-26 P3051.*

For the revised manuscript, we re-performed the climate sensitivity experiment in a pre-industrial climate state. With this change the climate sensitivity of the model is no longer 'too high' (see below). In the updated result section we furthermore extended the feedback analysis and explore reasons for differences between ECHAM5 and SOCOL and SOCOL with chemistry feedbacks vs. without chemistry feedbacks.

*I also suggest doing an ECS experiment for the NOCHEM run. The effect of interactive chemistry found here is smaller than that of Dietmuller et al. (2014) and much smaller than the 20% effect found by Nowack et al. (A large ozone-circulation feedback and its implications for global warming assessments, Nature Climate Change, submitted). This is an emerging area, and if this effect is as large as other models suggest it has the potential to be important for the wider climate modeling community and therefore dependencies on model/experimental design need to be understood. In the discussion (L15 3051 – L6 3052), the authors suggest that the apparently weaker effect of chemistry in SOCOL-MPIOM may be due to the small decrease in ozone in the tropical lower stratosphere, but this would suggest a weak Brewer Dobson circulation response. I think this effect needs to be better diagnosed to establish why the results shown here differ from other recent studies on the role of interactive chemistry in climate sensitivity.*

The new experiments include TCR and ECS simulations for both model versions (with and without interactive chemistry) as well as similar simulations for ECHAM5/MPIOM. The comparison is therefore

now complete. We extended the analysis of chemical feedbacks (see below) and can now explain the ECS differences between CHEM and NOCHEM.

We contacted the main author of the mentioned publication and if he can provide us the submitted manuscript, we will also compare our results to this publication.

The new results show that the different climate state has a strong effect on the estimated TCR and ECS, and that the TCR and ECS of SOCOL/MPIOM agree now much better with the estimates for the CMIP5 models.

Model	TCR [K]	ECS [K]
SOCOL_chem/MPIOM	1.8	3.8
SOCOL_nochem/MPIOM	1.8	4.0
ECHAM5/MPIOM	1.8	5.4

Given the differences to the former estimates, we hypothesize that the warm climate state leads to some amplified positive feedbacks in the former present day climate sensitivity experiments (compare Meraner et al., 2013). Furthermore, the transient experiment are characterized by a very strong positive surface air temperature drift around 1990, which might also affect the TCR/ECS estimates.

Since we are anyway planning another publication focusing on the differences in the model response to CO<sub>2</sub> and solar forcing, we will also analyse this difference in the response between the two climate states at a later point. In the revised manuscript we will include the pre-industrial climate sensitivity experiments only.

*Section 3 P3023 L6-20: The explanation that there is a model surface temperature drift and how it is corrected seems rather disconnected and it is not until P3026 L7-9 that we learn the reason for this is related to the choice of solar forcing dataset, which includes very different irradiances in the visible part of the spectrum. I think the discussion on P3026 needs to be moved to the point at which the model drift is discussed to make this whole issue clearer. Furthermore, on L20 the fact that the adjusted TSI ends up being comparable to Kopp and Lean (2011) is probably more due to luck than judgment, so I think this statement about the comparison with observed TSI needs to be toned down or removed.*

Thank you for this comment. We shortened the description of the spectral irradiance differences and included it in the description of the control experiments as suggested.

We decided to keep the comparison of the 1990 TSI value to Kopp and Lean, since we think that the TSI tuning did not lead to a completely unrealistic low value of the TSI is relevant.

## Minor comments

**Introduction** Please note, we restructured and rewrote parts of the introduction. It is now substantially shorter.

*L28 3015 ‘Very strong’ – this is vague and since we don’t really know how stratospheric wind anomalies impact on the troposphere I suggest removing this and just saying ‘Wind anomalies. . .’*

Thank you.

*L5 3016 ‘unusual’ – I suggest changing this to ‘anomalously high’ and adding a reference to e.g. L. M. Polvani and D. W. Waugh: Upward wave activity flux as precursor to extreme stratospheric events and subsequent anomalous surface weather regimes, J. Climate, 17, 3548-3554 (2004)*

Thank you, the citation was added to the manuscript.

*L15-17 3016 This sentence is unclear and confusing.*

The sentence was replaced by: „The surface equivalent of the NAM is the Arctic Oscillation (AO). For the North Atlantic and European region the AO is closely related to the North Atlantic Oscillation (NAO) .“

*L25 3016 Add a reference e.g. Kolstad, E. W., Breiteig, T. and Scaife, A. A. (2010), The association between stratospheric weak polar vortex events and cold air outbreaks in the Northern Hemisphere. Q.J.R. Meteorol. Soc., 136: 886–893. doi: 10.1002/qj.620*

The reference is included in the new version of the manuscript.

*L18 3016 Replace ‘both’ with ‘the tropospheric annular modes’*

Changed as suggested.

*L21 3016 Add a reference e.g. Baldwin and Dunkerton, 2001.*

Done.

*L5 3017 I’m not sure Meehl et al. (2009) is the best reference for the top-down pathway. Rather e.g. Kuroda and Kodera (2002), Matthes et al., (2006).*

Thank you for this comment. However, we shortened the introduction and removed the discussion of the top-down and bottom-up mechanism.

*L10 3017 ‘Differently’ → ‘In contrast’*

Thank you.

*L29 3017 insert ‘temperature’ before gradient*

Thank you.

*L8 3018 ‘proven’ → ‘shown’*

Proven has been replaced.

*L8 3018 ‘essential’ – they are not always essential, it depends very much on what you are interested in. I suggest changing this to ‘important tool’*

Ok, we replaced essential by important.

*L10 3018 between THE ocean and atmosphere*

Thank you.

*L27 3018 The effect of atmospheric chemistry → remove ‘the’*

Thank you

## **Model description**

*L7 3020 ‘the QBO input data’*

Thank you.

*L13 3020 ‘forcing’ → ‘effect’*

We changed forcing to effect.

*L19 3020 What PSC scheme is used? Please give more details.*

Details of SOCOL version 3 and also the PSC scheme used are described in great detail in Stenke et al. (2013b). Therefore, our description of the model is rather brief (as the description of MPIOM is). Nevertheless, the parametrization and schemes used are important and we added another explicit reference to Stenke et al (2013b) to the beginning of the SOCOL model description.

„An in-depth description of the model and the parametrizations used in the chemical module is given in \cite{Stenke2013}. In the following we refer only to the most important fact that are needed to understand the characteristic of the coupled model SOCOL-MPIOM.“

*L14-15 3021 This discussion of vertical interpolation of tracers comes from nowhere and it is unclear to the reader as to the potential importance of this – can you clarify?*

We found in earlier simulations that the vertical interpolation from the model levels to pressure levels in the post-processing and back from pressure levels to model levels when the data is read in, can lead to substantial differences in the ozone concentrations. Furthermore, it depends on the number of pressure levels chosen in the post-processing. To avoid this we keep the ozone data on the original model grid.

To clarify our statement, we write now:

„By forcing the model with ozone concentrations directly on the model grid, differences between CHEM and NOCHEM, related to the vertical interpolation between pressure levels and model levels can be avoided.“

*L21-23 3021: The parameterization of absorption in the Lyman-alpha, Schumann– Runge, Hartley, and Higgins bands in the CHEM run alone seems rather arbitrary and unphysical. This is shown to have impacts on the stratospheric climatology, but since it's unphysical to neglect this effect in the first place these changes seem rather spurious.*

We agree that omitting the effect of Lyman-alpha, Schumann- Runge, Hartley, and Higgins bands in the NOCHEM experiment is unphysical. However, omitting the effect of the atmospheric chemistry is unphysical as well and since this parametrization is closely connected to the atmospheric chemistry in SOCOL it is not enabled in NOCHEM. Since this effect is large, as the results show, we propose to add the parametrization to the NOCHEM as well, as stated in the discussion.

*L21 3021 ‘including a’ → ‘which includes a’*

Thank you.

## experiments

*L24 3022 What do you mean by ‘scratch’?*

From scratch refers to present day conditions. Clearly, this state of the atmosphere differs from the 1600 state, but given the short adjustment time of the atmosphere we think that this effect is negligible.

We clarified this in the manuscript:

„The atmospheric and chemistry components are initialized by present day conditions, which adjust to the pre-industrial climate state within a few years.“

*L26-29 3023 How are other chemical species (CH<sub>4</sub>, N<sub>2</sub>O etc.) represented in the NOCHEM run? Do they follow the same treatment as for ozone? Please clarify.*

CH<sub>4</sub> and N<sub>2</sub>O are considered as uniformly mixed gases in NOCHEM. The global average concentrations, however, are identical to CHEM.

We state in the revised version of the manuscript:

„CH<sub>4</sub> and N<sub>2</sub>O are considered as uniformly mixed gases with same global average concentration as in CHEM.“

*L17 3024 ‘radiative flux imbalance’*

Thank you

*L18 3024 'global mean surface temperature change'*

Thank you

*L21 3024 'without changing' → with fixed*

Thank you.

*L19-20 3024 Experiment M1 has not been introduced by this point in the manuscript, so it is not clear what you mean.*

Oh yes, thank you for this comment.

Given the comments of reviewer 2 we decided to replace the climate sensitivity experiment performed in a 1990th climate state with experiments performed under a pre-industrial climate state. Therefore the description of the experiments has changed and M1 is no longer needed to initialize the simulations.

*L13-14 3025 'by a larger amplitude' what? 11 year solar cycle?*

We refer here to the TSI difference between Maunder Minimum and present day, as stated in the following sentence. We rewrote this part to make this more clear:

„In comparison to many other state-of-the-art solar forcing reconstructions, this reconstruction is characterized by a larger amplitude (compare Schmidt et al., 2012), with a TSI difference between the Maunder Minimum (end of the 17 century) and present day of  $6 \pm 3 \text{ Wm}^{-2}$ .“

*L3-5 3026 The lower UV irradiance in the Shapiro dataset must mean that stratospheric temperatures are lower in this version of the model? This issue is not mentioned at all, but if it is the case it seems that it would be important and should be discussed.*

Differences in the stratospheric temperatures can indeed be expected between a simulation with ECHAM5 and SOCOL. Furthermore, even larger differences in the stratospheric temperatures between both models can be expected due to large differences in the ozone concentrations. ECHAM5 typically uses the ozone climatology of Fortuin and Kelder (1997). This climatology differs strongly from the values simulated by SOCOL.

However, besides for the climate sensitivity, the aim of our publication is not the comparison of SOCOL and ECHAM5. Therefore, we prefer to not include a comparison of the stratospheric temperatures and dynamics between ECHAM5 and SOCOL in the manuscript.

*L12 3028 You should be consistent here about the use of M and L that you introduced earlier for the solar forcing sensitivity experiments.*

Yes, thank for this comment. We removed the L forcing completely from the manuscript, therefore the separation between L and M forcing is no longer needed.

*L20 3028 'used as the forcing'*

Thank you

*L9 3029 'simulated global surface temperature increase'*

Thank you.

## results

*L3 3030 'a' → the*

*L3 3030 'is' → are*

*L3 3030 'development' → evolution*

*L17 3030 'at a depth of'*

*L22 3030 'However, the oceanic temperatures are still not'*

*L27 3030 'not yet reached'*

*L28 3030 delete 'so far'*

Thank you for the corrections.

*L21 3031 I think it is important to stress here that because the QBO is nudged there is limited potential for the ozone response to feedback onto the circulation.*

This is an important remark, thank you. We included the following in the manuscript:  
„Note that the QBO nudging applied to the model may weaken feedbacks between ozone and circulation changes. “

*L7 3032 'in austral spring, during the break-up of the polar vortex.'*

Thank you.

*L14/L15 Do you mean statistically significant? If so, please state at what confidence level and how this is calculated.*

We stated in the caption of Fig 4 that a Student's t-test was used and a threshold of  $p \leq 0.05$  was used in the significance test. We included this information now also in the text.

*L14 3032 'on' → in*

*L18 3032 'are the result of a number of different processes'*

*L24 3032 'in summer (not shown)'*

*L26 3032 negative signal → cooling effect*

Thank you.

*L5 3033 undergoes → exhibits. Also add reference.*

We added a reference to the textbook from Brasseur and Solomon (2005) compare, e.g., their Fig. 5.4.

- Brasseur, G. P. and Solomon, S. (2005). Aeronomy of the Middle Atmosphere (pp. 443–531). Dordrecht, Netherlands: Springer.

*L8 3033 'day reach' → 'day can reach'*

Thanks you.

*L20 3033: The findings are not really contrary to the results of Maycock et al. (2011), you have just done a different experiment altogether. I suggest rephrasing to: 'Maycock et al. (2011) reported a maximum cooling in the lower stratosphere after a uniform increase of the stratospheric water vapour; however, the cooling effect in SOCOL-MPIOM is strongest in the upper stratosphere and mesosphere. This is probably because the water vapour difference between CHEM and NOCHEM is not uniformly distributed and the largest differences are found in the higher stratosphere.'*

We agree and changed the paragraph accordingly.

*L2 3034 'the differences in the zonal mean zonal wind reflect'*

*L12 3035 reflected in the NAM → 'reflected as a negative NAM index.'*

Thanks you.

*L26 3035 Give numbers for the total SSW frequency in SOCOL-MPIOM. The error bars on the seasonal distribution in reanalyses are large, so I suggest removing the part about the seasonality of SSWs in the model being too uniform unless a more robust statistical comparison is made between the model and reanalyses.*

Thank you for this comment. We removed the statement on the differences in the seasonality from the manuscript.

*L2-4 3036 Is this difference in SSW frequency statistically significant? You can use the t-test in the Appendix of Charlton et al. (2007; A new look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations. J. Climate, 10, 470-488, doi:10.1175/JCLI3994.1) to test this.*

We tested the significance of the differences for the number of events per winter season and found no significant differences between the data sets in any case. We mention this in the revised version of the manuscript.

„However, the differences between the data set are in no case statistically significant (statistical test following Charlton et al., 2007b).“

*L20-21 3038 'is obviously' → are*

We changed this in the manuscript.

*L21-22 3039 This is not the formal definition of climate sensitivity.*

We rewrote the section about the climate sensitivity analysis and no longer use this definition of the climate sensitivity.

*L23 3039 'transient climate simulations of past and future climates.'*

*L28 3039 With 2.2 K the TCR of → With a TCR of 2.2K,*

Thank you. We completely rewrote this subsection, therefore this suggestion could not be applied to the revised manuscript.

*L8-10 3040 'In comparison to the MPI-ESM based on ECHAM5-MPIOM, the TCR is the same but the ECS is considerably higher.' – why does only the ECS change between the model versions, but not the TCR? It is not clear to me why the effect should be so sensitive to the particular idealized climate change experiment used. This needs more explanation.*

In the results for the new experiment (see above) we find again very similar TCR estimates between the models and larger differences for the ECS. In the results we see that the rate of change, e.g., sea ice loss or changes in the cloud cover, are almost identical between SOCOL/MPIOM and ECHAM5/MPIOM. What is different is the equilibrium response, which is reflected in the ECS experiments and this difference might, to some extent, be related to differences in the initial state of the experiments. The ECHAM5/MPIOM experiments, for instance, include more sea ice in the NH, while the positive temperature drift in CHEM already lead to some melting at the sea ice edge in the Arctic. When the 80 year long TCR experiment would be continued for another 50 years, we would probably also see differences in the response in these experiments.

Furthermore, when comparing the ECS and TCR from other CMIP5 models (Flato et al., 2013), it seems to be a common feature of models, that models with the same TCR do not necessarily own the same ECS.



*L21 3040 I suggest adding some discussion here about the comparison with Dietmuller et al. (2014). This is currently in the discussion, but should be moved here.*

We extended the analysis of ozone changes and their role for the climate sensitivity in the revised version of the manuscript and moved the discussion from the 'Discussion' section to the results. Dietmueller et al (2014) explain the negative feedback of the atmospheric chemistry on the climate sensitivity by a combination of ozone changes and changes in the stratospheric water vapour. Our new results show that the pattern of ozone anomalies is very similar to Dietmuller et al. (2014), but the anomalies are weaker, suggesting a smaller effect in the climate sensitivity. The changes in the stratospheric water vapour, however, and the differences in the stratospheric water vapour changes between a simulation with and without chemistry-climate feedbacks are larger than in Dietmuller et al. (2014). Therefore, only the relative importance of these changes has shifted, but the net effect is very similar.

*L27 3044 sufficient → larger*

Done as suggested.

*L1 3050 do you mean higher stratopause?*

We meant higher stratosphere, similar to middle and lower stratosphere, but changed this to 'upper stratosphere' in the revised manuscript.

*L4-5 3051 This sentence has been erroneously pasted in: With a transient climate response (TCR) of 2.2 K and an equilibrium climate sensitivity (ECS) of 3.7 K. Please remove.*

Sentence is removed, thank you.

*Table 1 caption: → In column chemistry the usage of the interactive chemistry module is indicated.*

*Table 2 caption: → 'winter (DJF) zonal mean zonal wind at 50 hPa'*

Thank you.

*Figure 11 caption: What method have you used to account for the autocorrelation?*

We used the approach by Zwiers and von Storch (1995) as implemented in the `ncl` function `equiv_sample_size`.

- Zwiers, Francis W., Hans von Storch, 1995: Taking Serial Correlation into Account in Tests of the Mean. *J. Climate*, 8, 336–351.

We clarified this in the caption:

„... and taking auto-correlation into account following Zwiers and von Storch (1995).“