Anonymous Referee #1
Received and published: 29 February 2016

Summary

The paper compares the solar cycle signal in several stratospheric ozone data sets. It is obviously the first part of a two-part paper that forms the basis of a model assessment study (to follow). It is also related to the upcoming CMIP6 ozone forcing paper. The paper goes relatively deep into the comparison of data products, and I must stress that I am not an expert on these products. In my view, the comparison of different product or product versions with respect to solar cycle signals is a worthwhile task and the conclusions are strong but interesting. One focus of the paper is on the seasonality, which (according to the authors) should be captured by any potential climate model forcing data set, but also the general altitude-latitude structure. The authors conclude that there are considerable differences between data sets in this respect. The paper is well written, albeit quite technical at times. It is scientifically sound. In my view the paper is publishable after minor revisions.

We thank the reviewer for reading the manuscript and providing their helpful comments and suggestions. We address their specific issues in turn below.

Specific Comments

1. My main comment concerns how the outline of the paper and its position amongst the other paper (Part 2, CMIP6): If the assessment of data sets for use as forcings in CMIP6 is the main topic, then perhaps it should be made more clear at the beginning, which properties such a data set should have and which not and how such a data set is (or might be) generated. Will some smoothed satellite-based data set be used and then extended forward and backward? Will an existing model simulation or ensemble be used? Or will an ozone data set be generated purely statistically? Perhaps a few words on that would help, otherwise the paper is in danger of being misunderstood. The starting point (and recommendation) of the paper is that in any case a realistic solar-cycle imprint should be in, and the conclusion (in the abstract) is that satellite-based ozone data sets alone will not be good enough to get that signal. I find that interesting and well demonstrated in the analysis. However, I would like to know a little bit more about other effects, although this is not the topic of the paper: Obviously ozone depleting substances should be in such a data set. It is also clear that climatic influences such as SSTs should be excluded because the coupled models will generate their own SSTs. What about the QBO, should it be in or out? How will volcanic eruptions be specified in CMIP6? This sounds a bit off topic, but it might help the reader to position the paper amongst the other two upcoming papers before the focus then goes entirely towards the many data sets and the solar cycle imprint. Also, it would help to assess the relevance of the uncertainties found against other uncertainties. Once the position of the paper is made clear, I can agree with most of the paper.

The reviewer raises many important points regarding the ozone database for CMIP6. As part of the CMIP6 special issue in Geoscientific Model Development Discussions (GMDD) there will be a paper that describes the CMIP6 ozone dataset in full (cited as Hegglin et al. (in prep.) in the manuscript) and a paper that describes the solar forcing for CMIP6

(Matthes et al (2016)). The ozone database will be based on chemistry-climate model simulations from the WCRP/SPARC Chemistry Climate Model Initiative (CCMI), and will include effects of ozone depleting substances, greenhouse gases and the solar cycle. The authors are not sure whether it will also include the effects of QBO and volcanic eruptions, since these are represented differently across models.

Because the CMIP6 ozone dataset has not yet been finalized and will be fully documented in the GMDD special issue, we are cautious of adding additional information to the current manuscript. However, the revised manuscript actually places much less emphasis on the assessment of satellite ozone datasets for CMIP6, and instead focuses more on the interpretation of the datasets themselves. Part II will provide more detail about a solar-ozone response recommendation for CMIP6. However, we have added some more detailed text at the end of Section 1 that describes the main properties that an ozone dataset for climate models must possess which helps to put the analysis into a modeling context.

2. Temperature used for conversion: It is often not clear how temperatures were used to convert number densities to mixing ratio. In which cases were daily profiles used, in which monthly, or even just a climatology? In which cases were zonally averaged temperatures used, in which the full 3D fields? Perhaps add a table.

All of the conversions presented in Section 3.1.1 use zonal and monthly mean temperature profiles to convert the zonal and monthly mean SAGE II profiles. We have expanded the text in this section to give a more detailed description of the temperature fields used to conduct the conversions to mixing ratios. We hope that the reviewer finds this explanation clearer.

3. Regression models: Perhaps it is common practice to write the model in this way. I am still surprised that no lags are used. Also, the volcanic term is basically Pinatubo, so perhaps it might be better just to cut that period out.

Near identical multiple linear regression models have been recently applied to satellite ozone datasets (see e.g. SI^{2}N papers by Tummon et al., 2015; Harris et al., 2015 in ACP). The main term in the regression model for which a lag might be appropriate is ENSO. We have tested the sensitivity of the results to lags in the ENSO index in the range 0-12 months, but find no significant effects on the diagnosed solar-ozone response. We therefore do not include any lags in the regression model. Text has been added to the Methods section of the revised manuscript to explain this.

At the suggestion of the reviewer, we now exclude the periods following major volcanic eruptions from the analysis and no longer include the volcanic regressor in the MLR model. However, these changes have little impact on the results.