

Anonymous Referee #2

This paper presents an analysis of global stratospheric ozone trends from data reevaluated by a number of groups as part of the SI2N initiative. It is well-written and generally well-presented. There are no major issues, but I find a number of minor ones.

Thank you. The comments given below were helpful and have improved the quality of the manuscript.

Detailed comments:

Page 8568, line 18: A “hint of a trend”?

We spent some time on the phrasing of this sentence. Quantifying the uncertainty is a major part of this study, and this sentence would be written differently depending on which uncertainty estimate is used. We would prefer to stay with this wording in the absence of a better suggestion.

Page 8571, Section 2.1: It would be helpful if the merged datasets (GOZCARDS, SWOOSH) were defined here. GOZCARDS is defined in Section 3.1.2, but only for the pre-1998 period, and I cannot find SWOOSH described at all.

We have revised section 3 and now have a third section describing the merged data sets.

Page 8576: Equation 1: The large number of fitting parameters (~30) in the model is alarming. What checks have the authors made to be sure that their fits are robust, with all those degrees of freedom? QBO_orthog is not defined.

First it should be noted that the apparently large number of fitting parameters results from the expansion of some of the regression model fit coefficients in the Fourier series to account for seasonality. So, it is not the case that there are ~30 independent basis functions in the regression model fit. Standard practice in regression modeling is that there should be at least twice as many data points as fit coefficients in the regression model and this requirement has been applied throughout our analysis. When fitting the model to monthly mean data from 1979 to 2012 there are 408 data points which significantly exceeds this criterion. Previous studies that have applied the same regression model to diagnose trends in ozone have demonstrated that the resultant fits are robust and that there is no danger of over-fitting. In this study, for shorter periods with fewer data, the number of basis functions is reduced with a resultant reduction in the number of fit coefficients. Visual inspection of selected fits to the data ensures that there is no over-fitting.

We have added a sentence to the manuscript that explains the QBO orthogonal function.

Page 8576, line 18: What is a “simple linear trend”? The term seems to suggest a simple linear regression to the data without the 30-odd adjustable coefficients of Equation 1. However, I expect you mean a linear trend derived using Equation 1, but with free endpoints. In that case I suggest “single linear trend”, or “unconstrained linear trend”.

We agree and have changed it to ‘single linear trend’.

Equation 2: Why is an AR(2) model used? How exactly is Equation 2 applied?

In making the decision to use an AR(2) model we did test the effects of an AR(1) model in determining the uncertainty on the fit coefficients. We found that the regression model fit coefficient uncertainties could be under-estimated when using an AR(1) model and so, to be as robust as possible in our estimates of the trend uncertainties, we felt it prudent to use an AR(2) model consistent with what has been done in other previous applications of regression analyses to derive trends (Reinsel et al., 1994; Bodeker et al., 1998; Fioletov et al., 2003).

Page 8578, line 27: The phrase “care needs to be taken when comparing them”, used here and elsewhere, is a vague phrase that doesn’t really say anything. The issue is one of representativeness: these geographically sparse data may (no matter how much care is taken) yield trends that are different from the zonal mean.

We have made this point clearer.

Page 8581, lines 11-13: The error bars here on the ozonesonde data are quite a bit larger than one would expect from weekly data at just a single site, so presumably something other than simple atmospheric variability (representativeness error?) has been included. This statement should be expanded to explain what the large uncertainties really reflect. Moreover, in Figures 2, 4 and 6, the uncertainties for ozonesondes appear to be for the most part larger than the Umkehr, FTIR, microwave or lidar uncertainties, which seems quite inconsistent with the fact that there are far more ozonesonde sites. Something is funny here...

The reviewer is correct. We have removed that statement and added a couple of sentences explaining the difficulties associated with comparing the trend uncertainties calculated for the ground-based records.

Figure 6: On the other hand, this figure is so small I could be reading those error bars incorrectly. Please make the little tiny plots at least as big as the little plots in Figures 2, 4, 7 and 8.

We agree and will make sure this happens in the final version.

Page 8581, Section 3.2: The discussion of Figure 5 is unsatisfying; perhaps “...is hard to be confident about the significance of this feature...” is meant to be typical British understatement? GOZCARDS and SWOOSH have (apparently statistically significant) trends in opposite directions below 40 hPa, and above 10 hPa the trends show little coherence between datasets. SAGE-GOMOS trends at the poles look seriously wonky.

We are not trying to hide the obvious disagreements between the various records – indeed dealing with them is fundamental to the main thrust of the paper and to SI2N as a whole. This particular phrasing was chosen because (a) it is true, and (b) the main conclusion we want to draw is that whichever record you look at, there has been a clear change from the negative period in the earlier period. We discuss the differences in some detail in the second paragraph, but in general there are other papers which do a more thorough comparison (e.g., Tummon et al., and the more recent intercomparison paper by Hubert et al., both in the special issue). For these reasons we would prefer

not to change the text.

We are sorry the reviewer finds this example of British understatement is unsatisfying but we note that it is not in the class of Watson and Crick's "This structure has novel features which are of considerable biological interest".

Page 8582, line 12: Again, please define "simple linear trend".

We agree and have changed it to 'single linear trend'.

Page 8582, line 13: Is there a rationale for using a PWLT in the troposphere? Do you include QBO, solar cycle, etc., there too?

We are not trying to present an analysis of the troposphere, and so the reviewer is entirely correct to say the PWLT model is not appropriate there. We choose 10 km as a lower bound so that the trends presented are largely representative of the stratosphere. This is reasonable in mid-latitudes, but not in the tropics. Little emphasis is put on the lower altitude trends, and we have now removed any confusing references to the troposphere.