Response to Anonymous Referee #2

General comments:

The expected stratospheric ozone recovery from the effect of halogenated ozone depleting substances (ODSs) has received much attention in recent years. Yet detecting the recovery of the ozone layer is complex due to a number of factors, including internal and external variability, that obscure the emerging signal associated with the slow decline in ODSs levels. The manuscript addresses this issue by investigating three stages of ozone recovery. To this end, the authors use total column ozone (TCO) changes based on experiments of the UM-UKCA and multiple linear regression (MLR) analysis. Although models are not perfect (e.g. often show significant disagreement compared to observations), they are a valuable mean to explore ozone changes due to specific factors (i.e. ODSs levels). Overall, the manuscript addresses relevant issues with regard to the evolution of the stratospheric ozone layer and uses appropriate data and methods. The text is technically well written. I have minor specific comments (detailed below), which I hope will help the authors improve the paper. In general, I suggest more detailed description and evaluation, additional comparison with ozone measurements, and further discussion on existing literature. Therefore, the manuscript is recommended for publication after the specific and technical comments are addressed.

Specific comments:

a. In the Introduction section, the authors clearly set out the stratospheric ozone depletion in the last decades associated with man-made emissions of ODSs. Due to international efforts banning the use of these substances, the ozone layer is expected to recover and the study aims to explore different stages. However, significant work has been done on detection and attribution of ozone recovery, hence it would be appropriate (and helpful for the broader audience) to briefly introduce key findings, remaining issues, and link it with the novelty of this work. Moreover, this will help relate and put into context the main findings here later in the manuscript.

Reference to the recent findings of Chehade et al. (2012), Pawson et al. (2014) and Weber et al. (2018) has been added to the introduction section in order to further establish the novelty of this work.

b. In the Model configuration and simulations section (page 3, lines 25–26), the authors explain that the simulations used were performed in support of the CCMI activity, and that are described in more detail in Bednarz et al. (2016) and Keeble et al. (2017). Bednarz et al. (2016) described that the simulations included a future climatological solar cycle since 2009 based on the observed cycle 23, which is not consistent with the description given in the manuscript (page 3, lines 16–17). Please clarify.

This is a mistake here – the details of the solar cycle should be those provided by Bednarz et al. (2016) – historic observed solar forcings are applied until 2009, after which cycle 23 is repeated until the end of the simulation. The manuscript has been corrected to account for this. The MLR and results presented in this study are not affected by this error as we use the correct top of atmosphere solar flux prescribed in the model for this analysis.

c. In the Removing natural cycles section, the text describes a MLR analysis to identify the impacts of natural variability on TCO. Since the results of this study heavily rely on the MLR analysis, I think this section requires more detailed description of the statistical method. In particular, the TO3i and Ne,l,t terms need better description (i.e. "… some constant value" and "Any noise…"). Also, an evaluation of the MLR analysis is important – i.e. How good is it? How much of the model raw data is captured by the MLR and how much "noise" is left? –. The manuscript already includes some references on MLR analysis that may help.

Further detail has been added to section 3 describing the MLR model and its terms. The TO3i term (now changes to I following advice from reviewer 1) corresponds to the intercept term of the MLR. N corresponds to the month to month variations not accounted for by the other explanatory variables. We feel that the MLR does a good job accounting for the natural cycles we seek to remove, as shown by comparing the red and blue lines in figure 1 and discussed in the manuscript in section 4.

d. For the Modelled global column ozone and minimum values section, it may be appropriate a statement about the choice of not including the polar regions in this analysis (Figure 1), since, in other sections and figures these regions are included and also discussed in the last paragraph here (page 5, lines 18–25). In fact, the latter paragraph argues that minimum column ozone values are a poor indicator of ozone recovery by giving examples based on polar regions. Is there any particular reason for not using the latest version (3.3) of the Bodeker Scientific database? The latest version, in addition to include some improvements on the methodology, could be expanded until 2016 in the inset of Fig. 1. Also it would be nice to include the Bodeker Scientific database in the acknowledgements, as recommended on the website.

For figure 1 we present only data from 60S-60N as i) the interannual variability at high latitudes, particularly the Arctic, is very large and identification of longterm changes is more difficult, and ii) high latitude polar ozone depletion is strongly seasonal, and this feature

dominates monthly mean time series as the seasonal cycle changes with changing stratospheric ozone depletion. We have clarified in the discussion about minimum column values that we only consider values between 60S-60N.

An earlier version of the Bodeker Scientific database (v2.8) was used as this dataset includes monthly mean values, as are presented from our model results. We are reticent to use the latest version as it does not, at present, include monthly mean data, and these would have to be calculated from the daily data provided. This requires a number of decisions which would need to be made (e.g. what spatial and temporal coverage is required for a monthly mean datapoint) which are not trivial and may not match those reached by Bodeker Scientific themselves. In this case, any representation of monthly mean time series we produce may not match the final monthly mean dataset provided by Bodeker Scientific in the future, which may lead to confusion. Bodeker Scientific has been added to the acknowledgements.

e. Regional trends section. This section includes very interesting results. However, modelled results in Fig. 2, both "raw" and residual data, could be compared to observed trends. In turn, this may lead to some evaluation/discussion and to put into context these results with existing literature. Nevertheless, there is some discussion (outlook) on TCO trends between 2000–2017 in the Discussion and Conclusions section (page 8, lines 15–19). Error bars representing the 95% confidence interval may need a line or two detailing how these are estimated and whether they account for autocorrelation. Are these confidence intervals calculated in the same way for all analyses?

Additional information on how trend uncertainties are calculated has been added to the manuscript, alongside comparison of our modelled trends with observed trends calculated by Chehade et al. (2012), Pawson et al. (2014) and Weber et al. (2018). Please see our response to review 1 for further information.

f. Return to historic values section. Figure 4 shows that TCO values in the tropics (<30º) reach the "1980 last recovery" between ~2060s–2070s. However, the main text (page 7, lines 12–14) explains that "…, it is the only region in which total column ozone abundances are not greater than their 1980s values by the end of the simulation,…". Please clarify. Also, is there any particular reason for not showing(using) "ozone residuals" on Fig. 4 as in previous analyses? I understand the study aims to explore ozone recovery addressing natural cycles.

The aim of the final section of the paper is to highlight that at some latitudes the absolute ozone column abundance may not recover to its 1980 values, due in part to the fact that the 1980s were years during a solar maximum, and in part to other factors which overwhelm the recovery

trend expected by decreasing CFCs (e.g. increasing BDC speeds resulting from increased GHG concentrations). Further, due to the high interannual variability in the Arctic, it may be possible for years late in the 21st century to have very low column ozone abundancies due to the high natural variability in these regions (see e.g. Bednarz et al., 2016). For these reasons it was felt that discussion of the raw modelled column ozone abundancies was more pertinent than the ozone residuals calculated elsewhere in the manuscript. Extra text has been added to the figure captions for figure 2-4 which provide information on how the error bars were calculated.

Technical comments:

Page 1, lines 9–10. "This approach…". The approach or method has not really been introduced. I suggest rephrasing this sentence (e.g. Here internal atmospheric variability… is accounted for by…).

This sentence has been reworded to read "The impacts of modelled internal atmospheric variability are accounted for by applying a multiple linear regression model to modelled total column ozone values, and ozone trend analysis is performed on the resulting ozone residuals."

Page 1, line 28. Substitute "ODSs" for "ODS" for consistency throughout the text.

Substituted

Page 1, line 34. Randel and Wu (1995) did not explore the effects of Mt Pinatubo eruption on stratospheric ozone.

This reference should be Randel et al., 1995 (Randel, W. J., Wu, F., Russell, J. M., Waters, J. W., and Froidevaux, L.: Ozone and temperature changes in the stratosphere following the eruption of Mount Pinatubo, J. Geophys. Res., 100, 16753–16764, doi:10.1029/95JD01001, 1995.) and has been corrected in the text and reference list.

Page 1, lines 37–38. References order.

References have been reordered.

Page 1, line 38. Delete "," between "other" and "non-chlorinated".

Deleted

Page 2, line 2. References order.

References have been reordered.

Page 2, lines 13–15. “Good agreement…” This sentence is a bit confusing, rephrasing maybe?

This sentence has been rewritten to read:

“If good agreement is found between the model and observations when all processes are included, then evidence of ozone recovery due to decreasing stratospheric halogen loadings can be identified by excluding other processes. For example, Solomon et al. (2016) found evidence for healing of the Antarctic ozone layer in September when polar halogen chemistry is included but interannual dynamical variability and volcanic factors are excluded.”

Page 2, lines 22–23. “…, data from fully coupled chemistry-climate model…” is a bit misleading since you use imposed SSTs. I would clarify “fully coupled” (chemistry and radiation schemes?).

To avoid confusion, the word fully has been removed so that the sentence now reads “To explore future ozone trends and recovery, data from coupled chemistry-climate model (CCM) simulations are required.”

Page 2, line 26. Spell out “SSTs”.

SSTs have been defined here before the first use of SST and later in section 2 just “SSTs” is used.

Page 2, line 37. Substitute “-” from “–”.

Corrected

Page3, line 23. Could use just “SSTs”, as it was introduced before.

See comment above

Page 5, line 33. Substitute “… DU year-1…” for “… TCO (DU year-1)…”? I am aware that “for the column ozone” is mentioned later in the sentence, though it is somehow confusing.

This sentence has been reworded to read "Figure 2 shows total column ozone trends (in DU year^{-1}) obtained from the raw data from the UM-UKCA simulation and the ozone residuals for the decline (1980-1997) and recovery (2000-2017) phases, averaged over 10° latitude bands."

Page 7, lines 9–10. "… 1980s values for the first time (light red)…" should be "blue".

We apologise for the confusion – this figure was replotted a number of times. The text now matches the colours in the figure.

Page 7, line 21. Typo: "airmasses".

Corrected

Page 7, line 26. I would substitute "expected" for "projected" (e.g. acknowledging these are modelled results, which are model and scenario dependent).

This change has been made

Page 8, line 15. Typo: "… Unlike a recent analyses…"

This has been corrected to "In contrast to a recent analysis of total ozone measurements…"

Page 8, line 26. Typo: "… The tropics have too small a trend…"

This sentence has been altered to read:
"The magnitude of the column ozone recovery trend in the tropics is too small in comparison with the natural variability resulting from the solar cycle and the QBO to identify significant trends."

Page 15, line 2; and Figure 2, legend. Please follow same consistency in the naming, both for the figure and the main text.

All figures have bee reproduced so that figure legends, figure caption and the main text use consistent naming conventions.