This paper investigates changes in lower stratospheric ozone over the recent past in model simulations and satellite observations, an important topic when assessing the effect of the Montreal protocol or effects of climate change on stratospheric dynamics. It is found that lower stratospheric ozone changes scale linearly with tropical upwelling velocity across different model simulations with different nudgings, suggesting a primary role of tropical upwelling for controlling decadal ozone changes in that region. However, none of the model simulations reproduces the observed ozone trend and, if the linear ozone-upwelling relation holds, a large upwelling trend would be needed to explain the observed ozone trend. Furthermore, nudging the model dynamics towards reanalysis turns out to be tricky, such that tropical upwelling trends in nudged simulations are often very different from the original reanalysis trends, such that the usefulness of nudged simulations to investigate observed ozone variability appears questionable.

The topic of lower stratospheric ozone trends is of much interest to the stratospheric community, and this paper makes an important contribution to further our understanding on these trends. The paper is concise, well structured and well written and the results are presented clearly, and I enjoyed reading. I do strongly recommend publication and only have a few minor and specific comments, which could help to further improve the presentation and discussion.

Thank you for the review! We have addressed the minor comments below in bold.

Minor comment:

I don't fully agree with the statement that ERA-Interim shows "inconsistency of its long-term upwelling trend against ... observations..." (L315), or "ERA-Interim being a particular outlier" (L375). Indeed, there is an inconsistency, but only if one considers residual circulation upwelling velocity calculated using the standard TEM residual circulation definition, and also only for a particular period (e.g. Seviour et al., 2011). On the other hand, the ERA-Interim upwelling calculated from momentum or thermodynamic balances shows a long-term increase in the lower stratosphere (Abalos et al., 2015) which is (at least qualitatively) consistent with observational estimates (e.g., Ray et al., 2014). Also, mean age trends based on upwelling from thermodynamic balance estimate (diabatic heating rates) from ERA-Interim appear to agree better with mean age observations than other reanalyses (e.g., Ploeger et al., 2019, 2021). I'd suggest to include a more careful discussion (e.g. L315ff) and also reconsider the statement that "upwelling trends explain roughly half of the discrepancy between modeled and observed ozone changes" (e.g., L297ff, L395), as this is related to the former one. There are a few more specific comments below related to that.

We appreciate the reviewer's comment on this, and have updated the text to include the information provided by the reviewer, including discussion of the fact that ERA-I's "outlier" status is related to the standard TEM definition (near line 315 in the original text). We have noted that the ERA-Interim upwelling trend (coupled with WACCM's upwelling-LSCO trend relationship) is consistent with observations, and have removed the value judgements regarding ERA-Interim.

Regarding upwelling trends explaining half of the discrepancy, we have addressed this issue in the specific comments below.

Specific comments:

L166: Remarkably, the QBO-related upwelling increase during 2015-2016 in AMIPQBO, which is likely responsible for the positive upwelling trend (at least partly), is not seen in the original reanalysis data. This could be worth a note.

We added a note on this.

L180, Fig. 2: Why is the correlation between the SD-simulations and their respective reanalysis decreasing below about 70hPa? Is the nudging strength varying with level?

We don't have a good explanation for this decrease in correlation below 70 hPa. The nudging strength is uniform below 0.8 hPa, which we have noted in Section 2 (also in response to RC1). We can speculate here that differences in the TTL structure between the model and reanalyses may lead to inconsistencies in the conversion of wave variability to upwelling.

L201, Fig. 2: Any idea why the nudging of the climatology shifts trends to be less negative / more positive in the lower stratosphere (i.e. nudging only anomalies results in more positive upwelling trends)? The same happens for nudging T, in particular for the "ca" (green) case.

We speculate that nudging absolute values of winds and temperatures (as opposed to anomalies) can act as (potentially unwanted) diabatic heating and impact gravity wave momentum forcing, as discussed in Davis et al. 2020. We have added a note on this in Section 2 where we motivate the alternative SD configurations.

Fig. 2 and 3: Related to the last comment, I'd find it noteworthy that nudging temperature climatological anomalies (T-ca) changes the lower stratospheric upwelling trend from negative to positive and the lower stratospheric ozone trend from positive to negative. For the zonal anomaly nudging simulations this is not the case.

We agree that this is noteworthy, and have added a few sentences mentioning this.

Fig. 3: Another interesting detail is that the extent of negative ozone trends into the NH middle latitudes is not reproduced by any simulation. This might point to mixing effects which are perhaps not well represented in the model (as also suggested by Wargan et al., 2018; Orbe et al., 2020). Maybe also worth mentioning.

Agreed. Done.

L239, Fig 4: I think it could make sense to include the figures for tropical latitudes also in Fig. 4, just to show how clear the relation is for the region where we expect it to be clearest.

The tropical latitude version of this was already included as Fig. S1.

L297: I don't understand this remark ("...the negative trend in upwelling in that simulation appears to explain roughly half of..."). My problem is that we don't know the true upwelling. If ERA-Interim would be the truth (and not MERRA-2) its positive upwelling trend would be in the range where the linear relation in Fig. 6 is consistent with the observed ozone trend.,

so that the entire ozone trend difference could be explained by the upwelling trend difference. (This is related to my minor comment above).

The key point in this statement is that for these simulations MERRA-2 is the input to the reanalysis, so it is the "truth" to which we are comparing (even though we of course don't know the true upwelling). The sentence after this one clarifies that the upwelling-LCSO relationship in Fig. 6 can account for around half of the difference in LSCO trend between the UVT L88 simulation (~ +1.2 DU/decade) and SWOOSH (~ -1 DU/decade). We've modified the second sentence to try and make this clearer.

L375: Only trends in ERA-Interim upwelling calculated using the standard definition of TEM residual circulation velocities are an outlier (Abalos et al., 2015). (This is related to my minor comment above).

We've removed the reference to ERA-Interim being an outlier here.

L394: Also here it is not entirely clear to me what is exactly meant. Here, I understand that 50% of the trend difference can be attributed to the spurious upwelling trend due to nudging - and with this statement I would agree. Above (L297), it was not so clear to me what was meant. (Also related to my minor comment above).

We don't understand what the reviewer doesn't understand here. The reviewer is interpreting this line correctly, i.e., "that 50% of the trend difference can be attributed to the spurious upwelling trend due to nudging".

L395: Given the linear relation in Fig. 6, isn't it most likely that the simulations underestimate the true upwelling trend? If the true upwelling trend would be positive - similar to ERA-Interim - this would explain the difference. Couldn't this be hypothesized here? (Also related to my minor comment above).

We don't know what is most likely, but the model upwelling-LCSO trend relationship suggests that the negative ozone trends in observations are consistent with a positive upwelling trend. We've changed the text here to note that the Fig. 6 relationship is consistent with the ERAI upwelling trend and observations.

Technical corrections:

L325: I can't find eqn. 1.

This was added back in. See also response to RC1

L352: There is one "the" too much.

Fixed