Dear Editor and Reviewers,

We thank the reviewers for their second round of constructive comments, which we believe improve the manuscript. We address these comments point-by-point below.

Best wishes,

Tereza Jarníková, on behalf of the authors

Reviewer 1

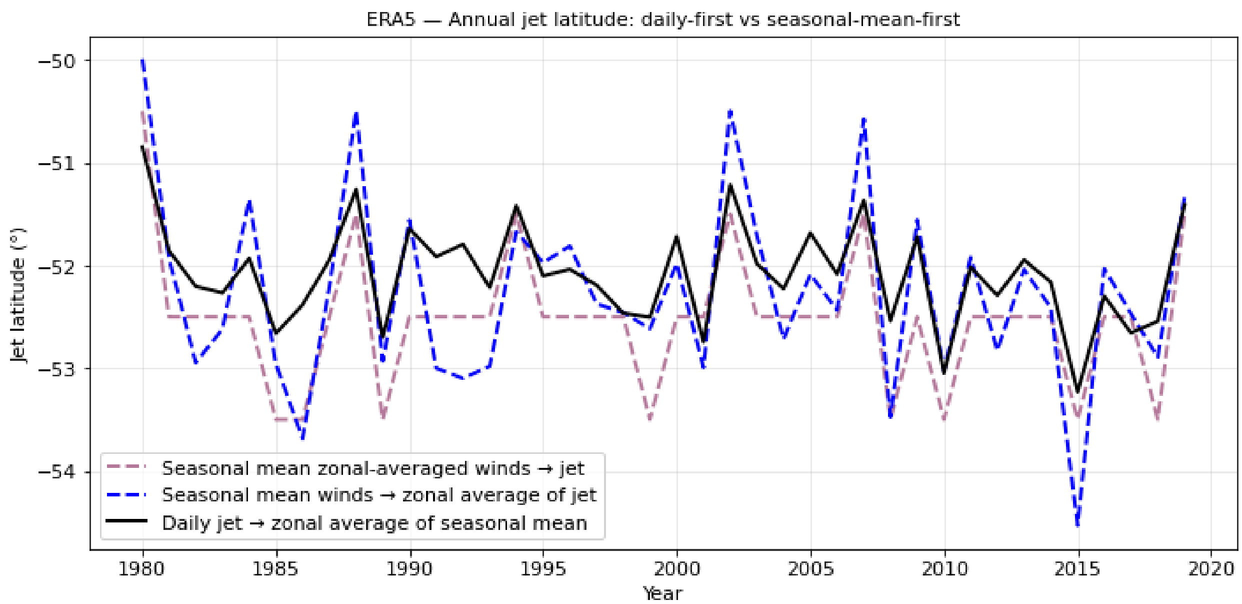
Minor comment:

The method in which the authors are calculating the zonal mean jet position is somewhat problematic. Finding the jet position at each longitude on a daily timescale and then zonally and seasonally averaging adds a lot of noise into the calculation, as the jet location on daily timescales at individual longitudes is quite variable and subject to how you are interpolating to find the location of the maximum. For example, see Section 4.1 of Adam et al. (2018) for an explanation of why this isn't a good idea:

https://gmd.copernicus.org/articles/11/4339/2018/gmd-11-4339-2018.html

I think the authors should double check to make sure that they would get the same results in Figure 4 (and associated tables) if they found the jet maximum using the seasonally-averaged zonal-mean zonal wind. In other words, average the data as much as you can before trying to find the location of the jet maximum, and compare with the existing calculations in the paper.

> *Thank you for this suggestion; we had tried calculating the jet multiple ways. However, using a seasonally averaged zonal-mean zonal wind (ie, finding the zonal wind, then taking the seasonal average, then finding the jet, or, equivalently, taking the seasonal average, and then finding the zonal wind and then finding the jet) reduces the resolution to 1 degree and does not allow us to find trends (purple dashed line below). We thus tried two options: in the first, we find a daily jet at each longitude and then average seasonally and across longitude (original calculation, in black). In the second, we average the winds seasonally, find a jet at each longitude, then take the mean jet as the average across longitudes (blue dashed line below). While the two calculations yield similar means and trends, the second introduces more variability, so we prefer to keep our original calculation in black.*



Typos/Technical corrections:

Figures 2a, 2b, 5: Consider making blue line for zero on y-axis a different color. It is easy to confuse with ERA5 in its present form.

> *I've now made the y=0 line grey in these plots.*

Figure 1 caption: Should refer to Table 4, not Table 2

> *Thank you, I had missed this when adding the two extra tables!*

Figure 4 caption: Correct supplementary table numbers here

> *I believe these are correct: ST3 and ST4 refer to the climatological jet and the jet position trend, respectively.*

Figure 6 caption: needs to provide latitude range for wind speed and wind speed trend averages\

> *Thank you for flagging, done.*

Figure 7 caption, typo: Delete Table 4

> *Done.*

Table 1: The NCEP-NCAR reanalysis is being discontinued. You may wish to denote an end date of 2026:
https://psl.noaa.gov/news/2026/r1datanotice.html

> *Because of this and following similar comments from Reviewer 2, we've ultimately decided to remove NCEP-NCAR from this analysis.*

Reviewer 2
I had reviewed the previous version of the manuscript and had commented on the lack of novelty of the manuscript. The authors have generally responded well to my comments. In particular, they now expand the discussion to CMIP6 models other than UKESM1. This enhances the scientific value of the paper. I have no further major comments and thus now recommend publication of the paper subject to addressing the minor (though possibly laborious) comments listed below:

1. The authors now discuss several CMIP6 chemistry-climate model. I recommend dropping the GISS-E2-1-G model from this discussion. Morgenstern et al. (2020, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020GL088295) discuss that this model incorrectly handles volcanic eruptions, affecting simulated ozone trends during the period of increasing ODSs. Their author team includes the GISS model PIs, lending credence to this assertion. It is thus unsuitable for inclusion here. On the other hand, a further chemistry-climate model (ECEarth3-AerChem) is available at ESGF. This model was late to contribute to CMIP6 and thus has not received the same level of scrutiny as the other models. It would be good to see it included here.

With GISS-E2-1-G dropping out, UKESM1 would be simulating the most extreme ozone depletion of the remaining models, making it more important to balance out the paper with the other CMIP6 models. Doing so would decrease the multi-model mean trend biases versus the observational biases.

> *We now add ECEarth3-AerChem to the comparison in Figs 6 and S1-S3.*
>
> *Because GISS is still commonly included in CMIP6 intermodel comparisons, including ozone comparisons (e.g. Griffiths et al, 2021, https://doi.org/10.5194/acp-21-4187-2021), and remains available for download on the ESGF, we believe it is best to include it, for completeness, but to add a caveat regarding its spuriously large ozone trends in both the text and caption.*

Text: Note that GISS-E2-1-G is known to have biases in simulating the recovery of volcanic eruptions, which amplifies the influence of ODS on TCO, so while we include it in the comparison of models, we keep this bias in mind when interpreting results.

Caption: Note the caveats regarding TCO trends in GISS-E2-I-G discussed in section 3.4.

2. Also displaying wind trends over the period 1980-2019 is maybe not ideal as the reversal of ozone depletion suggests that the wind trends change sign in this period. it may be more instructive to display trends e.g. from 1980-2000 (as you have done) and from e.g. 1998-near present. This way you perhaps more clearly work out the contrast between these two periods.

Though we understand the value of looking in the two trend periods, we wanted to provide an overview of how the winds behave over the entire 40 year period, which we deem as more relevant for climate timescales, and also to be consistent with, e.g. Figures 1 and 3. The interested reader can get a sense of how the winds behave in the second part of the timeseries (2000-2020) from Fig. 2 - the general trend remains positive, though less steep than in 1980-2000.

3. I remain unconvinced about the value of retaining NCEP-NCAR in this analysis. As noted by the other reviewer, this reanalysis is superseded, very low resolution, and has been shown to exhibit some spurious behaviour (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD034161). It's not outrightly wrong to use it, of course; it just does not add to the discussion. I am particularly concerned about the change of sign of the bias versus ERA5 exhibited by NCEP-NCAR, with underestimated wind speeds at high latitudes and overestimated wind speeds at midlatitudes (figure 1). This suggests erroneous pressure patterns; this behaviour differs from JRA3Q and MERRA2 that exhibit consistent biases likely reflecting properties of their boundary layer schemes. The inconsistent bias also explains the relatively small mean bias of R1 (table 4) that is however not a sign of better quality than the other products.

We originally included NCEP-NCAR because it is still used in the ocean modelling community, despite its age. However, because of this spurious behaviour, and because NCEP-NCAR is also being discontinued as of March 2026, we have now decided to remove it from the analysis.