Overall comments

In this study the authors explore the impacts of various combinations of specified-dynamics (SD) schemes and its impacts on convective and wave-mean flow dynamics and the associated tracer transport in the most recent version of the high top CESM2-WACCM version. The novelty of this study lies in providing a thorough test of the specified dynamics scheme by nudging the model circulation to its own free-running output, unlike previous studies in the literature where the model was nudged towards a reanalysis meteorology. This is indeed a very important topic as recent studies have highlighted that specified dynamics schemes, as implemented in various model frameworks, do not constrain the stratospheric mean meridional circulation or the underlying trends in the reanalysis products, ultimately adversely affecting transport processes. The study provides an in-depth analysis of the effects of combining different parameters including the meteorology frequency and the nudging timescale within the specified dynamics scheme; all by nudging the zonal mean winds and the temperature (u,v and T). The authors analysed these parameters with 18 different experiments that span one year with WACCM to understand how nudging affects the error structure and propagation in the dynamics and the transport of species in the troposphere and the stratosphere. The study points out several deficiencies in the current application of the nudging scheme in WACCM which may be generalised for other models as well. Moreover, one of the main findings is that even for just one model year, the residual circulation and eddy mixing processes accumulate tracer errors at the end of their characteristic transport pathways in the upper troposphere/lower stratosphere as well as in the polar stratosphere. However, the authors identify nudging parameter combination scenarios within WACCM's configuration that often minimise errors in the circulation and tracer transport. Overall, this study is a cautionary tale of the degree of impact of nudging schemes on various aspects of the coupled chemistry-climate system identifying the artificial limitations and implications that arise from the use of such a scheme and it greatly improves our understanding of such processes. It fits naturally within the scope of Atmospheric Chemistry and Physics journal and I recommend this study to be accepted and published with minor revisions. I don't have any major reservations related to specific parts of the text, but some minor specific comments and recommendations that may enhance the read of the paper follow just below.

Specific comments

Paragraph starting at line 31: I find that the CO discussion is short, and it could be enhanced a bit more in this paragraph.

Page 3, line 63: Please specify which meteorological products were used in Kinnison et al. (2007) or at least state that one of them was from an older version of WACCM.

Page 3, line 65: Specify that these results are based on the Chemistry-Climate Model Initiative (CCMI) output and provide the time scales (multi-decadal, climatological) in which the SD model spread in those studies was found to be as large or larger than in the free-running (FR) simulations. Overall, I feel you can mention some of the most relevant details of these multi-modelling studies that first investigated the impacts of the implementation of SD schemes in a multi-modelling framework. I would also add the Orbe et al., (2020) paper as a reference.

Page 3, lines 82-83: Apart from O_3 being the obvious choice for looking at transport processes especially in the stratosphere, what was the defining criteria you selected CO in the study and not another tracer? The reader could find a short comment around the choice of the tracers interesting.

Page 4, lines 96-98: Perhaps I'm not familiar enough with this but would you expect the met list nudging implementation to produce different results to the ones you are showing with the "alternative" dynamical-core-independent nudging technique?

Page 4, lines 107-108: Please clarify that these refer to the the absolute values.

Page 4, lines 132-136: Although the text is quite clear, perhaps a table with all the combinations including the ones that require an increase in advection sub-cycling crossed out might be helpful to the reader.

Page 5, lines 142-144: Very interesting indeed. Perhaps you could at least make a note of this (or even caveat) in the discussion section of the document.

Page 8, lines 213-214: How do you estimate the percentage change relative to the total variability? Is it ±3.x stddev assuming a normal distribution? Please clarify in the text or perhaps add a short note of this in the methodology section.

Page 8, start of line 217: Please clarify here that longer timescales refer to the nudging timescale.

Page 8, end of line 232: clarify that in this case the minimal global mean error occurs at the longer timescale.

Page 8, lines 235-236: However, they do not seem to scale. In Figure 3d the 2hr line is also missing. Is it out of bounds in this plot?

Page 8, lines 238-239: This is more of a remark than something to address; you can add a note of this perhaps. It's quite interesting to see in Figure 5b that the 24h nudging performs very well throughout the depth of the stratosphere.

Page 9, start of line 264: You could maybe combine Figures 9 and 10 in a single figure with 5 panels? Currently Figure 9c and 10c are duplicates.

Page 10, lines 278-279: This is unclear to me. Do you mean that you will focus on the mean error negative regression values or the regression values with decreasing meteorology frequency and/or nudging timescale?

Page 10, line 300: Are these EP flux vector anomalies wrt to the reference meteorology or do they represent the spatial errors? It seems the latter but please clarify that in the caption of Fig. 11 or in a relevant part of the text.

Page 10, lines 307-308: The way you have written the brackets denoting Fig. 3c and d prompt the reader to find the tropospheric and stratospheric Psi and EPF Div mean errors. This seems to be not the case as the quantities shown are the global mean vertically averaged mean errors. Please rephrase this sentence so it conveys what you meant originally. "Global average mean errors in the circulation (Fig. 3c,d)... in the stratosphere."

Page 11, lines 329-331: Perhaps a lat/height cross section figure of the e-folding O_3 and CO timescales could be added in the supplement. I wouldn't consider that a must, but it might be helpful to the reader. I don't feel that this is necessary to be shown in the paper; you could just include it in the responses.

Page 11, end of line 332: "…Northern Hemisphere" – compared to?

Page 12, line 362: "…are only substantially…" – why only?

Page 13, line 378: Please note that you still talk about the UTLS as in the mid to upper stratosphere it seems to me that the fluxes do not have the same structure. Also "…over the pole" – clarify which pole?

Page 13, line 382: Clarify compared to what CO is increased? Reference climatology or that it is the mean error that increases?

Page 14, line 427: Please provide an average estimated range of the stratospheric AoA for the reader.

Typos

Page 3, line 60: "…one *of* none…" – correct to or

Page 3, line 72: "strength" *of the* "meridional circulation,"

Page 15, line 444: "asking sjust" – just

References

Orbe, C., Plummer, D. A., Waugh, D. W., Yang, H., Jöckel, P., Kinnison, D. E., Josse, B., Marecal, V., Deushi, M., Abraham, N. L., Archibald, A. T., Chipperfield, M. P., Dhomse, S., Feng, W. and Bekki, S.: Description and Evaluation of the specified-dynamics experiment in the Chemistry-Climate Model Initiative, Atmospheric Chemistry and Physics, 20(6), 3809–3840, doi:10.5194/acp-20-3809-2020, 2020.