Response to comments by Reviewer #3

Thank you very much for your valuable comments and suggestions. Please see below for our answers to yours.

In this paper, a comparison is made on estimates of the transformed Eulerian mean momentum and thermodynamic budget terms from 4 reanalysis datasets: MERRA-2, JRA-55, ERA-Interim, and CSFR. Their results clearly show differences amongst these datasets. My main concerns on this manuscript are:

1. Presenting the difference magnitudes and large uncertainties shouldn't be the highlight. The manuscript needs to explain what these differences mean in terms of the physics of the atmosphere for the highlights to be suitable for publication in a journal like Atmospheric Chemistry and Physics. Are the differences enough to suggest that some of the reanalysis suggests a significantly different form of dynamics is occurring? There may also be instances when the difference values may simply be just noise. The noise-signals and physical signals need to be clearly pointed out. Once the authors focus more on what differences actually signify crucial Physics differences amongst the reanalysis datasets, they may be able to improve the organization of the manuscript.

We believe that showing the uncertainty range for TEM variables and terms from multiple reanalyses is very important, and furthermore that these estimates are fundamental to both the current prevailing practice in analyzing these data products and our ability to interpret such analyses. We explain in further detail below.

The reanalysis system consists of a forecast model, assimilation scheme, and assimilated observational data. For the dynamical core of the forecast model, we believe that all four reanalyses use good, reasonable models. Choices of particular sub-gridscale parameterizations differs among different reanalyses, as summarized in Chapter 2 of the S-RIP Final Report (SPARC, 2022) for these four and other reanalyses (e.g. radiative transfer schemes in Table 2.4, convective parameterizations in Table 2.6, and gravity wave drag parameterizations in Table 2.7). We believe that the four reanalyses analyzed in this manuscript all make reasonable choices for these parameterizations (please refer to the response letter to Reviewer #1 where we show zonal acceleration data for these four reanalyses.) These systems are complex, and it is sometimes but not always possible to attribute particular anomalies to the use of a particular parameterization. A good example is provided by MERRA-2's treatment of gravity wave drag. In MERRA-2, an increased latitudinal profile of the gravity wave drag background source at tropical latitudes and increased intermittency are applied to ensure that the forecast model can produce a spontaneous QBO. However, this treatment resulted in unrealistic tropical zonal winds in the stratosphere in the 1980s (Figure 2 of Kawatani et al., 2016), when data assimilation constraints were weaker than in the 2000s.

Furthermore, the final reanalysis data products are largely determined not by particular choices in the forecast models, but rather by the observational data assimilation. In some cases, particular parameter settings in the assimilation scheme can result in obvious biases in the reanalysis products. Good examples include the CFSR QBO issue (Saha et al., 2010, see the section on "QBO problem in the GSI"), the near-zero and sometimes even negative values of water vapour at and above the tropopause in CFSR (Davis et al. 2017, Wright et al. 2020), and the ERA5 vs. ERA5.1 issue (Simmons et al., 2020). Some of these issues have been noticed and solved by re-running the reanalysis systems for particular periods, but smaller issues are often not corrected or even noticed by the reanalysis centres. Some of the differences shown in the manuscript may emerge from these kinds of issues, but as data users it is practically impossible to attribute these issues unequivocally without conducting parameter-perturbation experiments, which is naturally the province of the data producers. Our role is instead to identify and highlight the issues so that they are more likely to be attributed and addressed in future development.

In addition, we do not have "reference" observations for each of the TEM terms and variables, and we therefore must rely on reanalyses for these terms and variables. Uncertainty ranges obtained from multiple recent reanalyses are thus important for evaluating and especially quantifying our current understanding of the atmosphere from the TEM point of view.

References:

Davis, S. M., Hegglin, M. I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G. L., Nash, E. R., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP, Atmos. Chem. Phys., 17, 12743–12778, https://doi.org/10.5194/acp-17-12743-2017, 2017.

Kawatani, Y., Hamilton, K., Miyazaki, K., Fujiwara, M., and Anstey, J. A.: Representation of the tropical stratospheric zonal wind in global atmospheric reanalyses, Atmos. Chem. Phys., 16, 6681–6699, https://doi.org/10.5194/acp-16-6681-2016, 2016.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Hou, Y.-T., Chuang, H., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P. V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, B. Am. Meteorol. Soc., 91, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1, 2010.

Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., Flemming, J., Haimberger, L., Healy, S., Hersbach, H., Horányi, A., Inness, A., Muñoz-Sabater, J., Radu, R., and Schepers, D.: Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1, ECMWF Technical Memoranda, 859, 38 pp., https://doi.org/10.21957/rcxqfmg0, 2020.

SPARC: SPARC Reanalysis Intercomparison Project (S-RIP) Final Report, edited by Fujiwara, M., Manney, G. L., Gray, L. J., and Wright, J. S., SPARC Report No. 10, WCRP-6/2021, 612 pp., https://doi.org/10.17874/800dee57d13, available also at https://www.sparc-climate.org/sparc-report-no-10/ (last access: 16 February 2023), 2022.

Wright, J. S., Sun, X., Konopka, P., Krüger, K., Legras, B., Molod, A. M., Tegtmeier, S., Zhang, G. J., and Zhao, X.: Differences in tropical high clouds among reanalyses: origins and radiative impacts, Atmos. Chem. Phys., 20, 8989–9030, https://doi.org/10.5194/acp-20-8989-2020, 2020.

2. In describing the reanalysis datasets somewhere in the methodology, more needs to be said on the differences in the physics that each model is known to already exhibit. These need to be described in a way that would help readers already get an idea of potential differences amongst the model.

We will add an overview of the reanalysis systems written above to Section 1.

3. The results spend too much time describing dynamics that are already well-known. For example, the first sub-section describing REM means may be reduced to solely focus on the issues regarding the calculation of v* or w*. The results need to be re-written in a way that immediately focuses on the physical and/or unphysical differences amongst the reanalysis with one another and/or with REM.

We will shorten the description of the REM results, but would like to retain the relevant figure panels (i.e. for temperature and zonal wind) and thus some text. This is because the REM results are needed for part of the interpretation later. Also, we believe that these panels provide a helpful reference for future studies, as we have already found them useful in this context in our own work.

4. Showing DJF dynamics without mentioning sudden stratospheric warming dynamics isn't a good idea. Your methodology indicates the calculations uses monthly-mean. To mention SSWs requires the use of, at least, daily-mean datasets. You can choose to do this, or instead show one equinox season.

First, in this manuscript, we intend to show and discuss the seasonal background states against which various short-term fluctuations and events occur, including SSWs. We believe that the DJF climatology will be important and useful for further studies including those on SSWs.

Regarding SSWs, Chapter 6 of the S-RIP Final Report (SPARC, 2022; see also references therein) extensively investigated SSWs across multiple reanalyses. In particular, an intercomparison of the momentum budget in reanalysis products during SSW events was conducted by Martineau et al. (2018a), who also compared temperature and meridional heat flux. Also, Martineau et al. (2018b) showed the thermodynamic budget for a particular SSW event.

Please note that results for MAM and SON are shown in the Supplement. We chose to show DJF and JJA in the main text because DJF and JJA are the two contrasting seasons often discussed in the literature.

References:

Martineau, P., Son, S.-W., Taguchi, M., and Butler, A. H.: A comparison of the momentum budget in reanalysis datasets during sudden stratospheric warming events, Atmos. Chem. Phys., 18, 7169–7187, https://doi.org/10.5194/acp-18-7169-2018, 2018a.

Martineau, P., Wright, J. S., Zhu, N., and Fujiwara, M.: Zonal-mean data set of global atmospheric reanalyses on pressure levels, Earth Syst. Sci. Data, 10, 1925–1941, https://doi.org/10.5194/essd-10-1925-2018, 2018b.

I recommend major revisions for this manuscript.

We hope that the above explanation has clarified the key points of our manuscript.