Interactive comment on "Climatology and Interannual Variability of Dynamic Variables in Multiple Reanalyses Evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP)" *by* Craig S. Long et al.

Anonymous Referee #2

The manuscript describes the results of intercomparisons of the zonal mean temperature and zonal winds obtained by assimilation at different centers as a part of the SPARC-Reanalysis Intercomparison Project (S-RIP). The focus of the comparison is the middle atmosphere below 1 hPa during the period of satellite observations (1979–2014). The comparison is mainly of the reanalyses produced by the different centers. A large discontinuity is found corresponding to the change in the NOAA observing system from TOVS to ATOVS. During the ATOVS period (1999–2014), the agreement among recent reanalyses largely improves. Comparisons against independent satellite observations such as HIRDLS temperature have also been conducted as supplementary

analysis. Reanalysis zonal winds and temperature are widely used in different areas of atmospheric science research. Therefore, the paper merits publication in Atmospheric Chemistry and Physics once the general remarks below have been addressed.

General remarks. The introduction of the present paper states that the goal of this project is to better understand the differences between current reanalysis products and their underlying causes. However, the present paper mainly describes the differences but there is little discussion of their causes. To understand the differences in temperature and zonal winds, it is also necessary to analyze the difference in forcings such as radiative heating rate, and resolved and unresolved momentum forcings. For instance, MERRA-2 is a unique assimilation system making use of non-orographic gravity wave parameterization. Such an effect on the equatorial zonal winds should be detectable by analyzing momentum forcing.

Also, in the comparison using the reanalysis ensemble of just three members, it is not possible to detect errors or deficiencies that are common to the three members. Temperatures are compared with HIRDLS, but this is not sufficient. The assimilation increment is an important measure of the quality of assimilated products, and should therefore be analyzed and compared.

The large differences between the periods of observation by TOVS and ATOVS are repeatedly mentioned, but the reason for the improvement is not discussed. Is this due to increased vertical resolution resulting from the increased number of channels for ATOVS, or something else?

It would more convenient to display and discuss the results such as in Fig. 1 grouped by family similarly to the companion paper by Fujiwara et al. (2017): "Introduction to the SPARC Reanalysis Intercomparison Project". ECMWF reanalyses: (a) ERA-I, (b) ERA40, JMA reanalyses: (c) JRA55, (d) JRA25, NASA GMAO reanalyses: (e) MERRA-2, (f) MERRA, NOAA/NCEP and related reanalyses: (g) CFSR, (h) R1, (i) 20CR.

Minor comments. 1) Page 2, line 23 Typo: Fiorina -> Fiorino

2) Page 7, lines 25–27: "The easterly SAO phase is believed to result from ...." Please add the reference.

3) Page 8, line 1 Agreement does not necessarily mean good performance, if they have common errors.

4) Page 8, line 13: "the 0.5 K contour occurs moves upward from between 20 and 10 hPa to between 7 and 5 hPa." After 2001, there is no reduction of differences in the upper stratosphere. Is there some explanation for this?

5) Page 8, line 15: "The disagreement between the three reanalyses is greater in June-August" Why? Is this due to the dynamical heating difference?

6) Page 8, line 20: "disagreement between the three reanalyses in determining this temperature during the TOVS period than during the ATOVS period" Is this because ATOVS has better vertical resolution around the tropopause region?

7) Page 10, lines 4–5: "The left hand column shows the gross monthly mean differences ...." Because there are large differences between the TOVS to ATOVS periods, it would be better to plot the monthly difference for two different periods before and after 1998.

8) Page 10, lines 10–12: "In general, the earlier reanalyses (JRA-25, ERA-40, and R-1) show greater differences from the REM than the more recent reanalyses (CFSR, ERA-I, JRA-55, MERRA, and MERRA-2)" This may be true, but because the reanalysis ensemble mean (REM) is calculated from ERA-I, JRA55 and MERRA, the difference from REM should be smaller for these three reanalysis products.

9) Page 13, line 24: "westerlies in MERRA-2 are more than 10 m/s stronger than those in the REM westerlies." Is this because of the non-orographic gravity wave parametrization introduced in MERRA-2? Comparison of the momentum forcing may clarify the cause.

10) Page 14, line 18: "linear slope of their matched monthly winds" This phrase is not easy to understand. Is this the linear regression line between observed and assimilated QBO winds?

11) Page 16, line 1 : " the fact that the CFSR did not bias correct the SSU Channel 3" Please give information about the bias corrections introduced in ERA, JRA and MERRA. JRA55 has an apparently smaller gap between the TOVS and ATOVS periods. Is this due to bias corrections applied for JRA55?

12) Page 16, lines 19–20: " differences are generally similar to those of the reanalyses from the REM (Figure 6)." Because the comparison so far has been made against REM, it would be useful first to show the difference between REM and HIRDLS.