Review of “Trends in Some Characteristics of the Warm-Season Tropopause-Level Jet Streams in Both Hemispheres” by Orr & Martin
Reviewed by Gloria Manney
(My apologies for the delay in returning this review. I beg the usual excuses of time pressures and assorted disruptions, but also offer my delay in part as a compliment to the paper, since it led me down many interesting avenues that took some time from completing the review itself.)

Recommendation: Publish in WCD after some important clarifications and added discussion.

General Comments: This paper uses a jet characterisation framework based on PV gradients at the dynamical tropopause to detail basic characteristics (waviness, latitude, altitude, windspeed) of the upper tropospheric jet streams in the warm season (extended summer season) and trends therein, focusing primarily on the NH but including comparison with the SH. Results are compared for three reanalyses. The work is new and interesting, the analysis is sound, and I believe the results will be of value to many readers of WCD; I enjoyed reading this paper. That said, in addition to a few small corrections, there are a couple of areas where (in most cases with only addition of discussion, though one would require some additional analysis) this paper could and should be substantially improved, and I recommend that these improvements be made before publication in WCD. These areas are:

(1) Discussion of relationships to and context of previous work: The discussion of previous work neglects numerous references that either analyse warm-season jet stream characteristics, or use methods for jet characterisation similar to those used herein, or both. Rather than detract from the motivation for this work, that previous work (which, as for all good research, highlights additional questions and areas that have not been studied) enhances the motivation for this work, and in some cases will be helpful to explain results herein or put them in the context of other phenomena. Places where I see the opportunity to do this are noted in the specific comments below.
(2) Discussion of reanalysis comparisons and relation to previous such comparisons: The use of multiple reanalyses in this work is commendable (and particularly important for trend studies, e.g., Manney & Hegglin 2018, S-RIP Final Report (2022, https://aparc-climate.org/publications/sparc-report-no-10/). This analysis would be greatly strengthened by putting the results from multiple reanalyses in the context of the findings of the S-RIP (now A-RIP) SPARC (now APARC) activity. S-RIP results are documented in the S-RIP Final Report mentioned just above, and (amongst many other papers) the S-RIP ACP/ESSD Special Issue (https://acp.copernicus.org/articles/special_issue829.html) and the S-RIP Phase 2 (now A-RIP) ACP/WCD Special Issue (https://wcd.copernicus.org/articles/special_issue10_1242.html). (https://aparc-climate.org/publications/sparc-report-no-10/). Chapters 2 (reanalysis systems description, which should be cited herein for characteristics of all the reanalyses used), 7 (Extra-Tropical UTLS), and 12 (Synthesis & Summary of the report) in particular are especially relevant; Chapters 4 (Overview of winds and temperatures) and Chapter 8 (Tropical Tropopause Layer; includes Asian Summer Monsoon analyses). Some of the key messages from the S-RIP report that are relevant here are:

- The most recent generation of reanalyses (at the time of the report including CFSR/CFSv2, ERA-Interim (ERA5 was very new but some studies included it), JRA-55, MERRA-2) are substantially improved and are recommended for all studies using "full-input" reanalyses.
- In particular, the NCEP/NCAR reanalysis used in the current paper is deprecated (and has been shown to be inadequate in numerous studies dating back to at least 2002) and should not be used. The S-RIP report states "With the availability of modern reanalyses providing coverage of the pre-satellite (e.g., JRA-55 and ERA5) and the ability to obtain those reanalyses on coarser grids, there should be no reason to continue using these older reanalyses." and also – specific to this paper – that "Only the recent high-resolution reanalyses … are suitable for ExUTLS dynamical and transport studies…." Thus, if the NCEP/NCAR reanalysis is included in the final version of this paper, its well-documented deficiencies needs to be discussed, and the discussion comparing it with the two more modern reanalyses used herein (which **are** recommended for UTLS studies) should be framed that way. (I personally don't see any great value in continuing to use the NCEP/NCAR reanalysis at all given the overwhelming documentation of its failings; however, one can argue that since people continue to use it, continuing to add to that documentation is valuable; thus I don't have a strong opinion about whether or not the NCEP/NCAR results should be shown in the final paper. Some might argue that instead more of the most recent reanalyses should be included, e.g., MERRA-2, which is very widely used; while this would be desirable, I am sympathetic to the large burden in computing and storage resources adding more reanalyses entails, and, because the most modern reanalyses have been evaluated extensively, comparing even two of them as you already do is already in the range of "best practice".)
- For analyses of data during the pre-satellite era, it is important to evaluate how the particular assimilation systems and diagnostics perform without satellite inputs. Typically this is done by comparing the full input reanalysis with conventional input only reanalyses that are available for most datasets (e.g., the JRA-55C reanalysis) during the satellite era. I think doing so here would be beyond the scope of this paper, but the issue should be mentioned, and guidance about the diagnostics herein that can be found in previous work (some suggestions are given in the specific comments below) discussed.

(3) Related to (2), the paper would be much stronger if the ERA5 dataset was used for its full record (1940–present) or at the very least, for a period as long as the longest of the other reanalysis datasets used herein. This is particularly desirable since ERA5 is the most recent and highest resolution (both horizontal and vertical resolution have been shown to be key to UTLS studies) of the reanalyses used herein. (This is the point I mentioned that would require further work.)

(4) Somewhere in the discussion of the method used for distinguishing subtropical and polar jets, you should comment on any caveats for deriving trends in something that is defined using thresholds (in this case, the isentropic layers) that are themselves expected to be

changing: The potential temperature of the extra tropical tropopause is expected to be changing; thus, are the "isentropic housings" of similar jets expected to change, and how might that affect your analysis? This is a near-universal problem because so many things we want to know trends in are derived from other things that may be changing – so I am not asking for any detailed analysis, only that this issue should be mentioned and briefly discussed in the context of your results. The discussion in Weyland et al (2025, https://doi.org/10.5194/acp-25-1227-2025) may be of some general relevance to this question.

Further details related to these points are discussed in specific comments below.

Specific Comments (in order of appearance, not importance):

Abstract: Reading the abstract out of context (which is what most readers will do first, and the basis upon which many will decide whether to read the full paper), there are a couple of terms that may not be obvious to the reader: "isentropic housing" and, related, "isentropic space". I think the abstract would benefit greatly from adding a sentence or so to give a basic description of the jet identification method so the reader who is not already familiar with the method will understand what these terms mean. Also, in line 15, it is not obvious to me that the "slight increase in waviness" and "slight decrease in average speed" imply "a nearly steady circulation" (again, this may just be a matter of clarifying the wording).

Lines 29–34: This discussion of the origins of subtropical and polar jets (radiatively-driven vs eddy-driven) is oversimplified and doesn't mention the well-documented spectrum of jets that show characteristics of both. This is an important point for this paper, since it is very relevant to expectations for cases with a single jet and for the seasonal transition to a single jet. Lee & Kim (2003, https://doi.org/10.1175/1520-0469(2003)060<1490:TDRBSA>2.0.CO;2) is an important reference on this. The introduction of Manney et al (2014, already cited herein) has a discussion of this point with a number of other relevant references. Spensberger et al (2023, https://doi.org/10.1175/JCLI-D-23-0080.1) is also a very good reference for discussion of this and its relationship to the jet locations (especially theta levels), the multiplicity of jets at a given time / region, and transitions between different jet configurations.

Lines 35–61: There are numerous other references that are specifically relevant to the work in this paper and the discussion in these paragraphs. Spensberger et al (2023) is particularly relevant given their use of location in theta to distinguish subtropical and polar jets, and that and and Spensberger & Spengler (2020, https://doi.org/10.1175/JCLI-D-19-0715.1) are relevant because they use a jet characterisation scheme based on theta at the dynamical tropopause. These papers, as well as Manney et al (2014), Manney & Hegglin (2018, already cited herein), and Manney et al (2021a, https://doi.org/10.1175/JCLI-D-20-0947.1), have extensive discussion of jet locations in altitude (theta) and latitude; they also describe regional and seasonal variations in dominant jet types (e.g., subtropical, polar, both, none), jet altitude (theta) and latitude by type, and jet strength (windspeed). They discuss differences in subtropical and polar jet altitudes (and hence tropopause altitude at their locations). These papers also discuss several other works that touch on these topics, but, to my knowledge, these are the primary ones wherein the discussion of jet features and evolution in the context of their isentropic level, and of the seasonal evolution in jet types, provides additional motivation for this current work – this paper is a natural extension of the exploration of many ideas raised in those papers. Thus it would be very good to mention these here to put the current work in the context of the most relevant previous work, and to reinforce the motivations for the current work.

Lines 69–76: There are a couple of errors in this description of Manney & Hegglin's (2018) methods: The criteria for counting multiple maxima in the same "jet region" as separate jets requires not a particular lower value of the windspeed (in fact, if the windspeed between them dropped to 25 m/s, as stated in this paper, they would not have been counted as being in the same jet region to begin with), but rather a drop in windspeed from the maximum of a certain minimum amount (by default set as a drop of 30m/s or more). The sentence "The polar jet is then defined as the strongest westerly jet poleward of the subtropical jet, or poleward of 40° latitude if no subtropical jet is identified" is correct as stated, but it becomes apparent later that the authors of the current paper seem to have interpreted "strongest westerly jet" as "strongest windspeed maximum", because they state that the JETPAC method always identifies both a subtropical and a polar jet, which is wrong – because all "jets" (that is, windspeed maxima greater than 40m/s) are identified (and multiple such jets in a region with windspeed greater than 30m/s disambiguated) before applying the criteria to identify subtropical vs polar jets, if there is not a jet in either of the categories that is greater than 40m/s, no jet of that type is identified at that time / longitude (smaller windspeed maxima are cataloged, but they are not further analysed in any of the published research to date using JETPAC that I am aware of). In fact, the ability to identify whether one or the other or both is present is an important factor in the research that has been done using JETPAC results (for one explicit example, see Fig. 1 and Figs. S3–S5 in Manney et al, 2021a).

Lines 76–78: The sentence summarising Manney & Hegglin's results on this line is not very accurate, as they found very little in the way of signatures that were robust in a global and/or annual mean, but very distinct and often significant changes in individual seasons, hemispheres, and longitude regions. The most "globally" robust results that I think are relevant to this paper are an increase in altitude for both types of jets (this is associated with increasing tropopause heights, as has been widely reported), a poleward shift of the SH polar jet, an equatorward shift of the NH polar jet, and a weakening of the summertime subtropical jet. One regional / seasonal change that is very relevant to this paper (see comments below) is the increase in Asian summer monsoon area that is accompanied by a poleward shift of the summertime subtropical jet in that longitude region. (All that said, it seems important to me to note that they found – consistent with other authors who analysed both zonal mean and longitude-resolved jet changes – that most trends are not uniform or even the same sign around the globe or in the annual mean.)

Lines 93–95: This seems an overstatement given the numerous papers that look at seasonal (and regional) variations – including the Manney and Spensberger papers noted above and many others that they cite (e.g., Koch et al 2007; Peña-Ortiz et al 2013; Rothlisberger et al,

2016) as well as numerous papers that focus on aspects of summer time upper tropospheric jet streams (e.g., Schiemann et al, 2009; Banderier et al, 2025, https://doi.org/10.5194/wcd-6-715-2025; Rousi et al, 2022; Banderier et al 2026, https://doi.org/10.5194/egusphere-2026-1045 [preprint]). I would suggest focusing on what specific aspects of summertime upper tropospheric jets are understudied in this motivational statement.

Lines 107–119: As noted in the general comments, the reanalyses used should be discussed in the context of the S-RIP (and A-RIP as appropriate) findings. See that general comment for further details. As noted there, it would substantially benefit the paper to use the ERA5 reanalysis for at least a time period matching the longest of the other reanalyses used herein, or for its full, longer duration. Also, please mention whether the JRA-55 data are used on the model levels (60 levels as I recall) or on the coarser pressure level grid.

Lines 129–132: Numerous other papers use 30m/s as a cutoff for the presence of a jet (or, in the case of the Manney et al papers, the presence of a "jet region", with a higher cutoff being required for a jet core to exist). A helpful reference in summarising other jet characterisation methods might be Keel et al (2024, https://doi.org/10.5194/gmd-17-1229-2024) since they summarise many methods.

Lines 172–175 & Figure 2: Because using the NCEP/NCAR reanalysis is strongly deprecated, and it has been extensively documented to be unsuitable for UTLS studies or any other studies where the vertical structure matters – such as determining "isentropic housings" (per S-RIP/A-RIP discussion), I strongly suggest that, even if you keep the analysis of the NCEP/NCAR reanalysis in the final paper with appropriate discussion of its shortcomings and their potential impacts, all figures (especially one such as Figure 2) that don't show all three reanalyses be shown for one of the more modern reanalyses that have been shown to be suitable for such studies.

Lines 159–165: The poleward and upward shift of the warm season (in this case, Asian summer monsoon season) subtropical jet over most of the eastern hemisphere around the Asian summer monsoon is extensively documented – see, e.g., Schiemann et al (2009), Manney et al (2014), Manney et al (2021b, https://doi.org/10.1175/JCLI-D-20-0729.1), and references therein. Manney et al (2014, apparent in Figs. 2, 6, 7, 8) and Manney et al (2021a, Figs. 1, S3–4) also show the dominance of regions with only a subtropical jet in NH summer in association with this evolution (the same figures show common occurrence of a unimodal jet in summer in the SH).

Lines 185–192: See immediately previous comment. Table 1 of Manney et al (2014) also quantifies the increase in latitude, altitude, and potential temperature of the zonal mean subtropical jet in summer in both hemispheres (the summertime latitude increase is larger in the NH because of the dominance of the Asian summer monsoon circulation).

Lines 207–215: See two immediately preceding comments for previous work specifically discussing relationships of summertime jet changes to ASM circulation. The figures noted above in Manney et al (2014) do show easterly jets characteristic of the TEJ extending poleward to about 30N during the peak of the Asian summer monsoon season (JJA), consistent with this discussion. Also, is there a reason why the windspeed threshold you use doesn't also require the winds to be westerly? Wouldn't that eliminate any confounding effect of the very strong and extensive TEJ in the Asian summer monsoon region / season?

Lines 231–233: Isn't this also similar to the arguments for the jet characterisation of Spensberger & Spengler (2020)?

Line 245–246: How do the vast differences in the native vertical grids of the reanalyses used (e.g., S-RIP Final Report, Fig. 2.1) affect the robustness of the interpolation to isentropic surfaces?

Line 274: Are these distributions for the time period that is used that is common to all three reanalyses? Otherwise, this is not a fair comparison, is it? (This is one of many places where the paper would be strengthened by using at least as long a record from ERA5 as from the other reanalysis datasets.)

Lines 276–281: Given the demonstrated unsuitability of the NCEP/NCAR reanalysis for this type of studies, its outlier status should be highlighted (and is quite obvious from Fig. 9). Also, you say "...exhibit a range of **only** 1.0 PVU" (my emphasis). 1.0 PVU is around 50–90% of the values – I hardly think that qualifies as "only".

Line 284: Per above discussion, ERA5 is actually the longest of the reanalysis records used herein – you just aren't taking advantage of that. Since it is also the most modern, has the best resolution, and is increasingly being shown to out-perform earlier reanalyses in many ways (in particular, UTLS studies), you are sacrificing confidence in and robustness of your results by not taking advantage of that record.

Lines 285–286: How are the trends calculated? Is it a simple linear regression? Or?

Lines 288–289 and 298–300: As noted above, you cannot assume without further discussion and/or analysis that differences between the satellite and pre-satellite eras are physical. The most robust method of evaluating pre-satellite era analyses is by comparing "conventional input only" reanalyses to full input reanalyses for the satellite era. "Conventional input only" reanalyses comparable to JRA-55 (JRA-55C) and ERA5 do exist. I'm not asking for this analysis, I think that is, indeed, beyond the scope of this paper. But you may get some clues as to how robust the type of analysis you do here might be across that transition from previously published comparisons. As one point of reference, Manney et al (2017, https://doi.org/10.5194/acp-17-11541-2017; their Figs. 5 & 6) compared UT jets, multiple tropopauses, and "subvortex jets" in JRA-55 and JRA-55C for 1980 through 2012, which may shed some light on agreement between fields (winds, tropopause) that are relevant to your work. Hitchcock (2019, https://doi.org/10.5194/acp-19-2749-2019) discusses pre-satellite data usage extensively, but he focuses on stratospheric fields; their results may nevertheless be of some value.

Lines 309–311: The wording here is a little unclear. First, I think you are saying that the slope is nearly zero and that that slope is not significant. Are you saying this to convey that there is no evidence of a trend, or to suggest that there may be a trend because the “non-trend” is not significant? Also, the use of “alternated” in the next sentence makes it sound like it is oscillating with some semi-regular period between two extremes, which I don’t think is what you mean.

Lines 317–318: It is important to say something about the other reanalyses in this case, since, because of sparser data coverage, reanalyses (particularly in the pre-satellite era to earlier in the satellite era; and particularly for NCEP/NCAR and other obsolete reanalyses) show less consistency in the SH.

Lines 320–322: It would be nice to fill the bars for this (the “subvortex” jet, or bottom of the stratospheric vortex) in Fig. 13 with a different color, so it would stand out to the reader.

Line 326: Please give the trend values and uncertainties.

Lines 327–335: Is this consistent with previous studies that looked at summertime trends in the SH jets (I appreciate for this case there may be only a few of those, but there are some, at least for latitude and windspeed)? Also, is the increase in waviness consistent with the increase in windspeed and the slight poleward trend (I had gotten the idea that the general picture was that wavier jets would be likely to be slower and more equatorward)? Also can you say anything about why the windspeed might show less interannual variability than in the NH, while the latitude shows more interannual variability than in the NH?

Lines 336–341: This gets at my immediately previous comments / questions, but it is not very clear. Surely you can say something about whether the “the SH warm season jet exhibits, on average, a more continuous ribbon of high wind speed around the hemisphere” by either looking at the geographical distributions relevant to your jet identification method, or from previous work, or both? (E.g., Manney et al 2014, show a considerable frequency, larger than in the NH, of bimodal jets (subtropical and polar) in the SH summer. Also, that paper and several others report a subtropical jet spiraling in to meet a polar jet in some longitude regions in the SH summer as a prominent feature. These seem rather different from your results, so it might be good to comment on how they might depend on the jet characterisation methods.)

Lines 343–348: This is another place where it is important to consider reanalysis robustness and differences. The S-RIP report shows several studies that indicate aspects of the SH circulation are not well-defined in the pre-satellite era (the figures in Manney et al 2017 mentioned above for JRA-55 show a hint of this), so it is not unlikely that the pre-satellite SH results are dubious. There are also some other physical considerations, such as the impact of the Antarctic “ozone hole” and interannual variability and trends therein on the tropospheric jets, with impacts that may linger into the warm season (the “Ozone and Climate” chapters in the 2018 and/or 2022 WMO Assessments of ozone depletion should have some references on this). I’m not looking for further work here, but just a bit more discussion of possible factors.

Lines 354–355: I think “first time” is an overstatement. Spensberger & Spengler’s 2020 jet characterization method fits the description of a “tropopause-level jet identification in θ/PV space” and they do analyse the warm season. Several other papers use similar methods (e.g., Banderier et al, 2025, and references therein). If you are going to say “first time” you need to be more specific about the key details of your method that make it unique.

Lines 387–391: Per previous comments, this discussion should be written considering the context of the S-RIP Report and the work described therein.

Lines 403–404: As per previous comment on Manney et al (JETPAC) method, it not only does not guarantee the existence of a polar (or a subtropical) jet at any longitude in any season, several of the papers that have used it looked at and discussed (more or less explicitly depending on the paper) situations with dominant jets identified as one type only.

Line 406: A comparison of methods and impact of results between the JETPAC method and that used by Spensberger & Spengler (2020) in the latter paper is relevant here.

Lines 412–414: This is a place for further discussion of the Asian summer monsoon connection, as well as previous studies discussing “hybrid” subtropical and polar jets.

Line 419–418: How do you know it will expand rather than contract? What other features (including possible trends in the monsoon circulation) would argue for one over the other?

Lines 421–422: Not to mention heat domes, etc.

Figures: All figures need larger axis numbering / labels

Note: I usually add a list of typographical, grammatical, and minor wording suggestions. I would normally have some of those here, but there are not that many (primarily numerous instances of using big words when small ones will not only do but be clearer) and I feel I’ve kept you waiting long enough already.