Second review of "Variability of temperature and ozone in the upper troposphere and lower stratosphere from multi-satellite observations and reanalysis data" by Shangguan et al. [Research Article acp-2018-1237]

Review Summary:
After revisions this paper is substantially improved, especially with regards to its motivation, literature review, and context. I appreciate the care the authors took to especially improve the introduction, and to add several analyses which provide more appropriate evidence for their conclusions; thank you. I am still concerned with some grammar issues in the text, but on that note the manuscript has also improved. With some additional editing to improve readability and a more careful write-up of the ozone results section, it is my conclusion that this manuscript will be fit for publication.

Major comment:

Figure 12: This figure is an excellent addition, but still needs a bit of work. The trends in the biases (i.e. trends in the differences between the reanalyses and GPS, akin to Randel and Wu 2006, Fig. 7) over the of the period of record should be included as an additional bar in this figure, for context, and discussed in the text. You essentially have this information in previous figures, but it is needed to provide context here to these trends.

Section 3.4: I continue to be concerned with the discussion in this section. It's clear from Figure 15 that there is essentially no agreement between either the datasets or the reanalyses (in contrast to what the authors claim on pg. 13, line 20). On this basis, it is very hard to judge what are the actual trends observed over this period, and it makes evaluation of them with the model (Fig. 16) less meaningful without knowing having confidence in the observational trend. Given the disagreement, it seems unlikely the trend analyses in Fig. 16 has a clear bearing on the real-world. The one clear consistent theme is that the tropical and SM lowermost stratosphere have increasing ozone trends and temperatures, while in the NM there is a clear decreasing trend in both temperature and ozone. This is mentioned briefly in conclusions (pg. 15 line 28), but it should be highlighted before this. The concluding paragraph in section 3.4 is good, but I still need to understand how it relates to the real-world, if the model and observations are not clearly in agreement.

Line-By-Line/Figure Comments:

Pg. 1 line 6: change "by" to "with"

Pg. 2 line 5: "SST increase" --> "warming SSTs"

Pg. 2 line 9: "greatly concerned" should be something like "extensively studied"

Pg. 2 line 15: Although I can tell what you mean, this line is confusing as written. May I suggest: "It is useful, therefore, to quantify the accuracy and variability of reanalysis temperature fields."

Pg. 2 line 25: add "it not being" before "susceptible"

Pg. 2 line 28: add "a" before "new generation"

Pg. 3 line 10: Rewrite as, "Long-term trends are a key issue in UTLS studies."

Pg. 4 line 11: The phrase "fully understand the exact reason" suggests that you have conclusively determined what the driving forces are, once and for all. This isn't a defendable point, as you note (and is clear from disagreements in literature) the interactions are very complex. I would suggest instead saying something like, "... are used in this study to investigate the reason".

Pg. 5 lines 15-17: You already say the range of your study in the preceding paragraph, so this is repetitive. You should either merge this with that line, or cut this line altogether.

Pg. 6 line 1: "For" should be "To"

Fig. 1: Is the weird spiking behavior of 2006 in NM at 250hPa related to a difference in when GPS observations were integrated into ERA5? It would be helpful to mention this (maybe at the top of pg. 9?) and explain it, because it stands out when you look at the differences figure.

Pg. 9 line 29: You should note that at higher altitudes in the lowermost stratosphere (70 hPa) ERA5 actually has the largest biases (~0.5K) in SM and NM, 2002-2006. It would be useful to write a line saying that no single reanalysis is universally better than the others, but (as you note in this line) on balance ERA5 appears to be best overall.

Pg. 11 line 7: This line isn't correct, and the nuance is important. Trends in the lowermost stratosphere are only significant in each of your datasets at 70hPa (and 50hPa additionally in the case of MERRA2). Please rewrite for accuracy.

Pg. 12 line 6: "specially concerned" --> "carefully considered"

Pg. 12, signal-to-noise: Thank you for including this analysis. It's much clearer and Fig. S2 makes it obvious to the reader that 16-year tropical trends are less meaningful. I would add a statement to this effect in line 16.

Pg. 12, lines 25-29: As rewritten and explained, this is a much stronger argument and more compelling evidence for your conclusions. Thank you!

Pg. 13 line 35: "dominated for" --> "dominate"

Pg. 13 lines 15-25: You need to note somewhere in here that the reanalysis and model trends look nothing like the observations... which casts doubt on the model's abilities to capture the behavior of the observations (and the relevance of the model results for the real world).

Pg. 13 line 35: In the 30-10hPa, I actually think that the fixed SST run is a much better representation of the observational datasets. It is important to be careful when describing this, as ozone and temperature trends should be expected to be increasingly linked as you go higher into the stratosphere.