Review of **Influence of atmospheric waves and deep convection on water vapor in the equatorial lower stratosphere seen from long-duration balloon measurements** by Carbone et al.

This paper makes use of data from the STRATEOLE 2 campaign to assess the influence of atmospheric waves and deep convection of water vapor in the tropical upper troposphere and lower stratosphere. The authors' method relies on variations in the Pearson correlation coefficient to determine when atmospheric waves are drivers of variations in water vapor concentration. Overall, the dataset and method used here are interesting and novel and will be of interest to ACP readers.

I do have some concerns over the paper which I believe should be addressed before publication. Specifically, the readability of the paper can be greatly improved through reduction of redundancies and careful consideration of which details need to be included for the authors to get their message across. My detailed comments are below.

General Comments:
1. There are several cases throughout the paper where figures are described in detail in both the main text and the figure caption, for example, in Line 121. Removing some of these unnecessary details from the main text may help to improve the flow of the paper
2. There are some parts of the study which seem secondary (or tertiary) to the main goal of the paper and are perhaps better suited for supplementary material to not overwhelm a reader with details that may be less necessary for their purposes. For instance, section 3.2 and associated figures 4 and 5 do not add anything substantial to the focus of the study and therefore I suggest they be added to supplementary material instead.
3. Throughout the manuscript, the Pearson's r correlation is mentioned many times, but frequently with errors varying from "Pearson's r" to "Pearson'r" to "Pearsons' r", etc. Please review and fix these occurrences.
4. Throughout the paper, there seems to be a lack of "polish" to the writing, specifically when it comes to spelling, misplaced words, and sentence structure. I tried to point out a number of these in my technical corrections below, but I am sure there are quite a few that I missed. I believe the manuscript needs to be reviewed thoroughly by the authors to ensure their message is clear to readers.

Specific Comments:
1. Lines 15-16: A bit wordy for first sentence of abstract, especially "intended to fly over the…"
2. Line 18: Either include "(H_2O)" for consistency with other gases or get rid of the chemical formulas for the other gases
3. Line 20: Add "(r)" after Pearson correlation coefficient
4. Lines 35-38: These sentences are quite redundant; suggest combining
5. Line 39: suggest changing to "the observed increase of stratospheric…"
6. Line 40: Needs citation

7. Line 43-44: consider using the oxford comma here and throughout paper for consistency
8. Paragraph starting at Line 45: Recent literature from William Randel is very relevant and should be included in this paragraph (e.g., https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030648).
9. Line 99: Suggest adding quick general description of location of Seychelles Islands
10. Line 107: what exactly is the Zephyr gondola?
11. Line 133: Sentence starting with "This allows..." needs to be reworded, maybe something along the lines of "This allows for the dramatic reduction of the optical path length and thereby enlightens the instruments."
12. Line 139: Please define HITRAN
13. Line 148-149: This sentence is a bit unclear, suggest rewording and using a word other than "deported", perhaps "transported"?
14. Line 150-151: This sentence needs to be reworded but I also don't really understand what its trying to say. For instance, what is the compromise that is being discussed?
15. Line 157: Please include the MLS layer vertical spacing, i.e. X levels per decade of pressure
16. Paragraph starting at Line 166: I don't think the few uses of "CTH" warrant using an acronym rather than just spelling out the words
17. Figure 3a: I really love the design of this figure and think it's a very interesting way to show the comparisons of the datasets, however, the color range is a bit narrow so it is hard to fully see the differences within each MLS circle. Perhaps look at some alternative possible color tables? And ensure they are colorblind friendly
18. Tables 1-3: Why are these commas instead of periods?
19. Section 3.3: I really like this section! Figure 6 is a very pleasing result in my opinion
20. Line 203: First usage of ECMWF and ERA5, please define.
21. Figure 6 caption: Please indicate what the gray lines are
22. Section 3.4: This section was very hard for me to understand, even after reading it several times. I suggest reworking and potentially condensing as much as possible. A couple of things that I feel could benefit from clarification:
    a. What are the plus or minus ranges for all these values? A SD? 95% confidence?
    b. Be extra clear about what specific datasets are being used to calculate each pearson r. I was confused if we were comparing between observations or between one observation (which one?) and ERA 5?
    c. Speaking of ERA5, as this plays a large role in this analysis, this should be discussed/described in the data section, with a citation.
    d. Because of some of my confusion with understanding what datasets are being compared, I also don't fully understand how the Pearson r values are compared with the mean difference values and how that relates to the claimed physical mechanisms? Is it about what can be resolved in ERA5 versus what cannot? What size waves?
23. Lines 297-299: Suggest rewording to ".... Thus limiting the spectrum of atmospheric waves reproduced by the analysis to wavelengths..."

24. Line 303: The range of latitudes averaged over in ERA 5 should be added to the figure caption instead of this sentence
25. Figure 7: Think black outlines around the circles in this figure could help to make them more visible
26. Lines 312 – 336 and figure 8: This detail may be better suited for the methods section or supplementary materials, as it feels more like a methodology than results and distracts from the results you are trying to describe in this section
27. Line 360: What exactly are the "unfavorable" tape recorder dynamics?
28. Line 406: a brief discussion/description of Typhoon Rai could be helpful here for context
29. Line 463: Reword "none of them shows undoubtedly a signature of deep convection"
30. Line 523: Reword "Not seen in Fig. 11 but seen in Fig. 10b"
31. Line 526: By "vicinity of the decreasing phase", do you mean the dissipating phase of the tropical storm?
32. Line 526: Some context about tropical storm Paddie could be helpful here
33. Figure 11: I am having a very hard time reading this figure. The combination lower resolution and higher resolution cloud tops within each panels is very confusing. Why are there seemingly multiple resolutions from one product? Perhaps smoothing could help to make this figure more digestible. Additionally, I suggest making the balloon observations a bit larger so they can stand out against the background.
34. Line 545: Specify tropical cyclones, not just cyclones in general

Technical Corrections:
1. Line 31: In the light of → in light of
2. Line 40: and to the → as a result of
3. Line 45: in tropical tropopause layer → in the tropical tropopause layer
4. Line 94: accent should be removed from STRATÉOLE
5. Line 103: It is a bit hard to tell in this formatting, but there are some off occurrences of new paragraphs that should be double checked
6. Line 114: Add comma – altitude of 20.5 km thus above → altitude of 20.5 km, thus above
7. Line 127: an heritage → a heritage
8. Line 152: allows to demonstrate → demonstrates
9. Line 158: In the frame of the present study → In the present study
10. Line 271: ZEPHYR → Zephyr for consistency with the rest of the paper
11. Line 278: 8K → 8 K
12. Line 278: present are large → present large
13. Line 300: No need to mention (a,b,c,d,e) for figure 7.
14. Line 301: Strateole should be in all caps for consistency with the rest of the paper.
15. Line 395: here you stated typhoon Raï but later just say Rai, please select the correct one and use consistently throughout
16. Line 404: has been → was

17. Line 417: isentropic level → isentropic levels
18. Line 425: to a fast → with a fast
19. Line 449: week → weak
20. Line 461: "on the opposite" needs to be reworded
21. Line 503: besides → outside of
22. Line 509: to → too
23. Line 520: hours → hour
24. Line 527: Autralia → Australia
25. Line 553: Suggest changing "exposes" to a different word
26. Line 568: could → can
27.