Review of the manuscript "High-altitude atmospheric turbulence and infrasound measurements using a balloon-launched small unscrewed aircraft sustem" by A. N. Haghighi et al.

General comments:

The authors present an exciting UAV system that mainly includes a five-hole probe and an infrasonic microphone for probing turbulence in the troposphere and lower stratosphere. The technical developments are unquestionably to be welcomed. They may represent a new step towards a technology better suited to in situ turbulence and small-scale structure measurements at high altitude, particularly in the stratosphere, by providing decisive information on the coupling between the fine-scale stratification, mixing processes and gravity waves. However, the proposed article appears to have major shortcomings that should be remedied before a possible publication. If some of the issues are due to misinterpretation, the description of the methods should be clarified. There are also a number of inaccuracies or blunders which should also be corrected and some parts should be expanded to facilitate the interpretation of the results. Consequently, in view of the potential interest of this work, I propose that it be *publishable after a very thorough revision and perhaps re-evaluation of certain parameters*. The review does not comment the part relative to infrasonic measurements because I don't have enough knowledge to evaluate it.

Major comments

(1) The proposed comparisons between temperature, humidity and wind profiles from radiosondes and sUAS cannot be conclusive, as they are made with data that are separated by several hours (up to around 6-7 hours) and launch sites separated by ~ 160 km (this information only appears on line 270 when comparisons are discussed), and the fields are not stationary. Under these conditions, performance evaluation is difficult, if not impossible, and cannot "allow validation", as stated line 248, because disagreements can always be explained by the non-colocalization and non-simultaneity of the measurements. It would have been more useful to include a radiosonde under the balloon during its ascent, in order to make more appropriate wind comparisons. The PTU IMET-XF data during the ascent may provide better conditions for temperature and humidity comparisons, even if they are corrupted at small scales. On the other hand, the representation of data points with circles, rather than continuous lines, systematically used in the figures, does not allow simple comparisons between profiles and identification of the peaks referred to in the text. From a qualitative point of view, the three wind profiles measured by sUAS seem to indicate fluctuations compatible with the presence of gravity waves at *all* altitudes with a vertical wavelength of the order of 1 km or less, whereas the balloon data do not seem to reveal such fluctuations (but, once again, the graphical representation makes analysis difficult). The technical data of the radiosondes used (such as the vertical resolution) should be indicated. Current standard sondes are now able to measure profiles at 1 Hz, which does not seem to be the case here. We can wonder whether the wind fluctuations observed by the sUAS, and apparently not by the balloon, are the result of an instrumental artifact or not. A spectral analysis of the whole (sUAS and balloon) wind (and T) profiles would provide useful information to check the consistency between the data.

(2) The calculation of ε described on page 18 contains several important errors that need to be corrected. The conversion of the frequency spectra into wavenumber spectra must be made with the magnitude of the *relative air wind speed*, NOT the aircraft's ground speed. Description of the method can be found in Frehlich et al. (JAS, 60, 2487-2495, 2003) or Kantha et al. (PEPS, 4:19, 2017). As the difference can be important, especially at high wind speeds, the impact on ε should be far from negligible. As a corollary, the longest wavelength of velocity fluctuation is not given by the sole horizontal velocity of the aircraft, but by the relative velocity (lines 348-350). Because the *relative air wind speed may significantly vary with altitude* (which is why it should be quantified and shown), TKE estimated for frequencies $f < 5\ Hz$ and shown in Figure 9 is likely not correct because it indicates the total energy for *variable* wavenumber bands. In addition, k_1 as defined in (9) is NOT the component of the wavenumber vector in the direction of the flight path (line 362), but along the direction of the relative wind vector formed by the direction of flight path and horizontal wind vector. There is no difference between the two, only when the UAV flies in the direction of the wind. Therefore, TKE dissipation rates (and TKE) must be recalculated.

(3) Section 3.4

The method used to reconstruct the distributions of parameters in the $\alpha - z$ plane is not clear. The authors should very clearly explain how the interpolation method works as this is not a standard method of visualizing the data. However, it does not seem to be feasible. Because only one value is obtained for a given altitude, it is not possible to interpolate the distribution of a parameter for any value of α at this altitude. The method will systematically produce artifacts (isolated structures with vertical bands) unless the layer probed by the sUAS has a thickness at least greater than the vertical distance covered by the instrument to make 360°. It turns out that all the plots and discussions are incorrect, and section 3.4 should be deleted in its entirety, unless the authors can demonstrate the merits of their approach.

Specific comments

Line 8: see comment (1) above.

Line 34-41: Some important references on in-situ measurements of turbulence are missing, e.g., Barat and Bertin (JAS, 41, 819-827, 1984, and references therein), Bertin et al. (Radio Science, 32, 791-804, 1997), Alisse and Sidi (JFM, 402, 137-162, 2000), Gavrilov et al. (Ann. Geophys., 23, 2401–2413 2005). In addition, the manuscript ignores references from the radar literature (e.g. Sato and Woodman, JAS, 39, 2546-2552, Fukao et al., JGR, 1994, and many others). UHF or VHF clear air radars have enlightened the layering structure of the stratosphere, mentioned line 37, to be attributed to thin and horizontally extended turbulent layers or/and stable layers. Incidentally, the authors recognize that the horizontal stratification is a key feature of the stratosphere. Is this feature consistent with the "vertical structures" supposed to have been detected with the sUAS measurements in section 3.4?

Line 65-67: The reference list about turbulence measurements from sUAS must be thoroughly revised. Some of them are not about turbulence (Bäfuss et al., 2018; Rautenberg et al., 2018, Jacob et al., 2018) and many others are missing. For example: Lawrence and Balsley (JTECH, 30, 2352-2366, 2013), Balsley et al., (BLM, 147, 165–178, 2013), Balsley et al. (JTECH, 35, 619-642, 2018), Reuder et al. (Acta Geophysica, 60, 1454-1473, 2012), Shelekhov et al., (Atmos. Ocean. Phys., 57, 533-545), Kanthe et al., (PEPS, 4:19, 2017), Luce et al., (JAS, 77, 231-2326, 2020), Calmer et al., (AMT, 11, 2583-2399, 2018), among others.

Line 86-87: "However, due to the transient nature of their Lagrangian flight trajectory, balloon-based approaches are not necessarily amenable to obtaining detailed statistical descriptions of turbulence at high altitudes." The comment is unclear. What do the authors mean?

Line 91-93: The introduction of the "infrasonic microphone" has already been made in the previous paragraph. It is this redundant. In general, the various paragraphs of the manuscript should be better organized to avoid such redundancies (they occur several times). This gives the impression of a juxtaposition of paragraphs with no guiding line.

Line 121: Please convert km/h into m/s. The controllability of the sUAS is an important parameter and more information about limitations and performance should be given. We understand that the UAV can safely fly for wind conditions up to 31 m/s at least. How is the horizontal velocity of the glider controlled? It must be significantly high than 31 m/s for stability. A figure showing the ground speed of the sUAS with altitude (and the relative air speed, for the reason described in (2) Major comments) would be informative.

Line 129: please explain why the sampling is made at 10 Hz.

Line 192: The authors seem to indicate that the five-hole probe sensor had an effective time response of 0.1 s. The reason is unclear (but I do not have the background to understand). The corresponding spectra should show a gap from ~ 10 Hz. It is roughly observed in Figure 8 at z=10 km (but around 5 Hz) and there is no evidence of a transition at z=1 and 18 km. How do the authors interpret this feature?

Line 227: The reason(s) of the choice of large values in circle radius (1-5 km) is not explained.

Line 229: A figure showing the descent rates of the sUAS with altitude would be useful to figure out the conditions of sampling.

Line 243: "horizontal distances": with respect to the ground? If yes, it means the ground speed of the sUAS varied between 10 m/s and 80 m/s but this information is not provided in the manuscript. If these distances are expected to characterize the largest scales sampled during 30 sec by the instrument, it is not correct, because the relative speed should be considered (see (2) of Major comments).

Line 252-271: These two paragraphs must be re-written. They are particularly confusing and not rigorous. For example:

- the first paragraph seems to describe general properties for the 3 flights but the second paragraph focuses on flights 2 and 3. So, we *deduce* that the description made in the first paragraph is for Flight 1. In addition, almost all the statements are disputable or not well-introduced. The agreement within 10% is unclear (please add information/figure that corroborates this result) and it is not true for <U> at all altitudes for example. But the paragraph focuses on temperature profiles. So, does this quantification only apply to temperature? But then why introduce the paragraph with "with the exception of RH"…?

- The altitude of the top of the boundary layer is estimated to be “roughly lower than 3 km” in the first paragraph, but up to ~5 km in Flight 3. First, these altitudes (especially 5 km) are not realistic even for convective boundary layer. Second, the criteria used to estimate this altitude are not explained. Third, there is *no* indication consistent with these estimates in the figures.

- The figures are not correctly labeled: “Fig. 5(a,c,e,)” should be “Fig 5 (a,b,c)”. ‘Fig. 5c” should be ‘Fig 5b’. Fig.5 b, d, f should Fig 5 d.e.f. and at other places. Please check.

- What is the criterion used to define the tropopause altitude? It is found at 11 km (line 258) (presumably for Flight 1 in the first paragraph), quite consistent with figure 5a, but indicated to be at 12.5 km on line 278. It is found at 13 km for Flight 2 while the temperature inversion is actually observed at 11.5 km in figure 5b. It is stated that it is at 14 km for Flight 13, but an inversion can be found at 12 km. How are quantified the lapse rates in the troposphere and stratosphere and what is the interest to estimate such values (and tropopause altitudes) if they are not compared between the instruments? This comment also applies to humidity and velocity profiles, since the text does not describe the differences and similarities between the profiles, but rather their characteristics, which is another objective. There are too many caveats.

From lines 290: This part should be separated from the previous ones because it is not about comparisons between radiosonde- and sUAS-derived profiles anymore. It seems to me important to show N^2 (squared BV frequency) and shear profiles before describing Ri profiles, since one of the purposes of the manuscript is to assess the performance of the sUAS measurements. In addition, low Ri values can have different causes, i.e. a strong shear and/or low N^2. The knowledge of these two parameters can help the interpretation of the turbulent events.

Equation (7): In practice, the impact of the variation of g with altitude can be ignored. The error is much less than all the other uncertainties.

Line 301: "Here we assume the critical Richardson number takes on a value somewhere in the range 0.25<Ri<1". It is unclear. Ric=0.25 is a necessary condition below which air can become dynamically unstable and turbulent (if Ric>0.25, a shear instability cannot develop). Once turbulent, there is a critical Richardson number at which the flow begins to laminarize: it is generally accepted to be between 0.2 and 1 but turbulence can be found for Ri >>1 according to Galperin et al. (2007) (but the corresponding turbulent regime should strongly differ from the turbulent regime for small Ri values). In practice, these thresholds must be used with caution because the Richardson numbers estimated from in-situ data are scale-dependent, i.e. depend on the vertical resolution at which they are calculated. It is common to apply arbitrary thresholds (i.e. "Ri is minimum and small ~0.25-1")

Figure 8: the information is interesting. Why not showing the corresponding spectra for v and w? In Figure 8b and 8c, it should be F instead of Φ (y label).

Line 333: I am again skeptical about the interpretation of the increased TKE layer up to 4 km as corresponding to the CBL top

Figure 9: EDR and TKE should be presented in logarithm scale (and continuous lines), because the linear scale over-represents the maxima near the ground. Indicating that "TKE is close to 0", line 333, is symptomatic of the fact that the linear scale is unsuitable for the present purpose. Figures *showing the slopes, as calculated in Figure 8 for 3 cases, vs altitude and for the 3 wind components* should be included. It would enable us to identify the altitudes where the inertial slope is indeed observed, those where a different

regime is observed, and those where instrumental noise is dominant for all frequencies. This information is *essential* for the purpose of the manuscript. A characterization of the slopes vs other parameters (e.g. TKE, EDR, Ri, etc) could be very informative.

The dataset offers the possibility to show $\langle u'^2 \rangle$, $\langle v'^2 \rangle$ and $\langle w'^2 \rangle$ separately. Plotting, for example, $\langle u'^2 \rangle$ vs $\langle w'^2 \rangle$ would be interesting for quantifying anisotropy. A discussion of this anisotropy in light of ε, Ri, etc, would be very enlightening.

TKE and ε are related by a master length scale (e.g. Mellor and Yamada, Rev. Geophys. Space Phys., 20, 851-8751982) which of great interest for the characterization of turbulence. The dataset shown in Figure 9 offers the potential to estimate this scale.

Line 337: "…caused by *inertial* turbulence" and remove "elevated" in the same sentence because the comment is valid for all levels of TKE. The spectra with a -1 slope may either reveal another turbulent regime or be due to a white noise contamination even for f< 5 Hz when the atmospheric signal is weak. Figure 8c is apparently in favor of the first interpretation for the selected case but it is not necessarily always true, especially when the instrumental noise dominates.

Line 343: "…more active turbulence conditions during these flights". This statement should be nuanced because (1) the rejection was based on the u spectra only, (2) the non-detection of a -5/3 slope does not mean the absence of turbulence, (3) the corresponding levels of EDR of flight 1 (qualified as "weakly active") shows a significantly higher background than flight 2, indicating higher spectral levels but not consistent with an inertial subrange. As we do not know about the interpretation of the observed non-inertial subranges, turbulence activity cannot be qualified.

Line 345: Please remove this sentence. The reference of Kelvin waves is not suited here because equatorial Kelvin waves (Fujiwara et al. 2003) are waves trapped around the Equator similarly to coastally-trapped Kelvin waves.

Equation (9): The notation Φ_{11} of the spectrum may not be appropriate ($F_{11}(k_1)$ would fit better the notation used in Figure 8). Φ_{ij} generally refers to the spectral density tensor (see e.g. Doviak and Zrnic', Doppler radar and weather observations, p. 326, 1984). The sUAS is "sensitive" to the 1-D *longitudinal* spectrum (see e.g. Hocking (EPS, 1999)).

Line 367 refers to $\Phi(k_1)$ (Φ_{11}?) and F (F_{11}?). Line 368: k_1^n should be $k_1^{-5/3}$. Strictly, the power-law fitting should be applied to a limited range of k_1 since the smallest wavenumbers are not well-resolved.

Technical comments

Line 4-5: The sentence is unclear, please rephrase.

Line 20-21: please add references.

Line 24: "Despite the higher stability of the stratosphere"

Line 25, see also line 318: "…due to mechanical and thermal disturbances" -> "due to shear instabilities and gravity wave breaking". The formulation is unsuitable because the mechanical sources of turbulence refer to those produced by obstacles close to the ground. The rest of the paragraph is awkward –and not rigorous- and references of the "classical" literature should be included instead.

Line 38: "high turbulent kinetic energy dissipation rate": please be more quantitative.

Line 45: "and for identifying the inner scale of turbulence."

Line 45: "This experiment": please be more specific with references.

Line 45-47: The Richardson number is not defined and it is not explained why Ri=0.25 is an important value and why turbulence when Ri >0.25 should be noted. The authors should indicate that some LITOS results were corrupted (Soder et al., AMT, 2019) due to balloon wake and that turbulence observed when Ri >>1 was suspect. In addition, useful information on the relationship between Ri and ε can be found in earlier references mentioned in the specific comments (line 34-41).

Line 63-72. Pitot tubes are also used (e.g. Lawrence and Balsley (JTECH, 30, 2352-2366, 2013).

Line 90: “air masses” generally refer to “large bodies of air” at synoptic scales in meteorology. The term is not suitable here.

Line 90: “geostationary” usually refers to satellite orbits. Do the authors mean “relatively constant location above the ground?”

Line 92: “”traditional” -> “standard”

Paragraph 2.1: It seems more natural to present the instruments first, then the configuration of the experiment. 2.1 -> after 2.3.4 and before 2.4

Fig.1 : Please add the location of the balloon launch site (El Paso) and show the distance in km (in Fig. 4 also). The distances are crucial for the interpretation of the radiosonde and sUAS data and longitude/latitude coordinate system is of little use here.

Line 102: “Three *sUAVs* were flown” (?)

Line 106: The altitudes are given in km m.s.l but the profiles are shown from z=0 (i.e. above the ground (line 241). Please indicate the corresponding altitude AGL, even if it can be roughly deduced from Fig. 1. The third sUAS was released from 30 km m.s.l (~28.5 km AGL ?), but the profiles are shown up to 25 km. Please clarify.

Line 191: “… the actual probe *frequency* response…”

Line 194: remove “disconnected”

Lines 232-235: Please indicate Local Time instead UTC (and avoid MDT). By doing so, the reader does not need to convert by himself when interpreting the PTU profiles measured by the sUAS and the radiosondes at different times (Figure 5).

Line 246-249: The first sentence is not necessary and the second has already been written. A more detailed description of the radiosonde data is necessary (see (1) of major comments).

Figure 5: please add LT times for all the radiosonde and sUAS flights. A figure showing the trajectories of (and horizontal distance between) both instruments is necessary.

Line 258: For flight 1, “the temperature continued to decrease with altitude at a rate of 1C/km” -> the temperature continued to decrease at a mean rate of 1C/km between 11 and 19 km” (otherwise it is confusing, see specific comments also)

Line 274: please add “(not shown)”.

Line 281: please show the NOAA upper air wind maps. The absence of reference points make difficult to confirm the statements.

Line 282-285: what do the authors mean ?

Line 307: The term “potential instability” refers to “an atmospheric condition in which otherwise stable air would become unstable if forced to rise (e.g. over high ground) thereby reaching its saturation point.” See e.g. www.encyclopedia.com. Please replace “potential” by “possible shear”.

Line 309: “marginally unstable tropopause”: what do the authors mean?

Line 316: Here again, the terminology is improperly used. “An atmosphere is said to be “conditionally unstable” if the environmental lapse rate is between the moist and dry adiabatic lapse rates. This means that the buoyancy (the ability of an air parcel to rise) of an air parcel depends on whether or not it is saturated.” (see glossary of meteorology) . “suggesting the possibility of localized buoyant production”. Do the authors refer to statically unstable conditions, ie. Ri <0? If yes, it must be clearly stated and defined earlier, e.g. around line 302.

Line 318: “mechanical turbulence”: see above, comment for line 25

Line 328: Do the Hanning window preserve variance?

Line 329: Do the cut-off at 5 Hz related to the effective limited time response indicated line 192?

Line 334-335:The description is unclear because of the use of a linear scale and a dot representation (see “specific comments”)

Line 336: Please remove “although the regions of elevated k appear at different altitudes”. This comment is not useful.

Line 344: Please indicate the altitude of the tropopause in Figure 9.

Line 354: “As direct measurements of …” please explain more or add a reference

Line 355-356: The sentence indicates a condition that has no reason to exist at this stage, since inertial domains have been identified by spectral analysis.

Line 520: Please remove “flux”. The gradient Richardson number and the flux Richardson number have two distinct definitions.