Response to Referee 1

We appreciate the time taken by the referees to once again read our manuscript and provide comments and suggestions for its improvement. We have made additional revisions to the manuscript and to help clarify where changes have been made in the revised version, we have highlighted all changes made using blue text.

The only major revisions made to the current draft were: (1) the addition of an uncertainty analysis section; (2) the addition of velocity measured by the onboard GPS during balloon ascent, interpreted as a measurement of wind velocity, with associated discussion of the resulting comparison; (3) removing the moving average step during calculation of vertical gradients and presenting N and S instead of N^2 and S^2 to provide a presentation of these quantities which may be more familiar to readers; and (4) removal of the comparison of Ri and EDR as the comparison did not provide much value and the authors felt the need to reduce the length of the paper after adding an additional 9 pages of content since the original submission. Most of the remaining changes were minor textual changes to better highlight and clarify certain points brought up by the referees.

Below, we respond to the individual comments made by the referee. To do so, we have reproduced the original review, with our comments provided in blue text.

Haghighi et al. present an exciting new measurement system that allows to sample the atmosphere up to 25 km with high resolution. The manuscript has been significantly improved. The plots look more trustworthy, but some questions remain, especially with regards to the basic wind estimation.

Comments:

- I lost confidence in the calibration of the wind measurement system. When you write that yaw and pitch alignment were corrected now, I wonder how this was done.

  As noted in the manuscript the "additional estimate and correction of the misalignment of the probe axis and aircraft's body-frame axis was conducted following the approach described in" Al-Ghussain and Bailey, (2021) Atmospheric Measurement Techniques. Given that this is a somewhat lengthy procedure, we preferred to just reference the previously peer-reviewed and published work rather than reproduce the details here. Note that this technique is essentially an extension of the flight calibration approach described by van den Kroonenberg et al, but is instead applied a posteriori using a slightly different optimization approach.

  I assumed this had been done even before the first manuscript.

  Unfortunately, we did not to conduct these corrections initially following the flight campaign as the approach had only been recently published and the codes had not yet been generalized for datasets other than the ones used in the paper. Although we had intended to apply these corrections, focus of subsequent work was on data analysis and we did not revisit determination of the velocity vector until receiving reviewer feedback from the original submission.

  Have Lenschow maneuvers for in-flight calibration been done?

  Typical in-flight calibration maneuvers require flight at a constant altitude, which was not possible with this aircraft as it is a glider.

  Have not only the misalignment angles but also the airspeed factor been determined? This is essential.

  The Al-Ghussain and Bailey correction approach also accounts for influence of aircraft blockage effects on the airspeed measurement. We have added mention of this in the revised paper. Note that for other studies (using powered sUAS) we have compared the in-flight maneuver calibration approach to the Al-Ghussain and Bailey approach and found little difference between the corrected results (although this comparison was only used for internal validation and not published).

- Please show some verification that all the angles and airspeed measurements are correct now, including static pressure.

  Although we had initially discarded all data taken during ascent as the aircraft sensors (five-hole-probe, humidity, and temperature) were not configured to perform measurements on ascent, we have since realized that an estimate of the horizontal wind speed can be provided by the ground speed of the aircraft during the balloon ascent. We have included these wind speed and direction profiles for additional validation of the five-hole-probe measurements made during descent and the results will hopefully restore the referee's confidence in the five-hole-probe measurements. Note that we do not independently measure static pressure, except via the aircraft's altimeter. However, we have compared the airspeed produced by treating the five-hole-probe as a Pitot-static probe to that measured by the aircraft's Pitot-static probe and the comparison was favorable.

– An uncertainty propagation for the wind algorithm with flow probes was developed by van den Kroonenberg et al. (2008) and could be used to identify the uncertainties. It would be good to include the uncertainty in some of the plots.

  We have added a new section on uncertainty analysis to the revised manuscript, in it we implement a Monte-Carlo-based estimate for uncertainty propagation which is very similar conceptually to the approach used by van den Kroonenberg et al. (2008) but by implementing it as a Monte-Carlo analysis we do not rely on a single reference state and can also include the effects of coupling between different error sources.

– Appendix A, Fig. A1: The NOAA maps are not readable at the given resolution.

  We have updated this figure with a higher resolution version.

Response to Referee 2

We appreciate the time taken by the referees to once again read our manuscript and provide comments and suggestions for its improvement. We have made additional revisions to the manuscript and to help clarify where changes have been made in the revised version, we have highlighted all changes made using blue text.

The only major revisions made to the current draft were: (1) the addition of an uncertainty analysis section; (2) the addition of velocity measured by the onboard GPS during balloon ascent, interpreted as a measurement of wind velocity, with associated discussion of the resulting comparison; (3) removing the moving average step during calculation of vertical gradients and presenting N and S instead of N^2 and S^2 to provide a presentation of these quantities which may be more familiar to readers; and (4) removal of the comparison of Ri and EDR as the comparison did not provide much value and the authors felt the need to reduce the length of the paper after adding an additional 9 pages of content since the original submission. Most of the remaining changes were minor textual changes to better highlight and clarify certain points brought up by the referees.

Below, we respond to the individual comments made by the referee. To do so, we have reproduced the original review, with our comments provided in blue text.

The authors have thoroughly revised their manuscript but there are still many shortcomings. Some additional information is unclear and not convincing. I think that the revised version cannot be published in its current form. In this second review, I focus on the main issues only.

1. As outlined in the first review, the manuscript hesitates between the evaluation of a new technique and a scientific work on atmospheric processes, which makes the paper very unclear. The primary objective of the paper should be to provide all the information needed to give the reader a clear idea of the platform's strengths and weaknesses. It is not the case. For example, section 3.1 shows the comparisons between temperature, humidity and wind profiles obtained by Hidron and radiosondes. The differences and similarities between the two should be emphasized. Paragraph 4, page 14 (lines 354-364) is out of the topic, because it describes the temperature structure of the troposphere and stratosphere without considering the differences/consistencies between the UAV and radiosonde profiles. In addition, it is not clear why this description is useful, since it is not used afterwards.

   The intent with the original structure of the manuscript was to not only describe the measurement technique, but also illustrate some of the statistical analysis which can be conducted with the data from the measurement technique, which we believe to be a strength of the platform. We also note that in the section of text the reviewer is referring to includes "With the exception of $\langle RH \rangle$, the HiDRON H2 measurements are broadly consistent with the trends and values indicated by the radiosonde data, although the HiDRON H2 temperature shows a noticeable warm bias compared to the radiosonde above $z = 16$ km that appears most noticeable for Flight 3. Note that no density correction was applied to this sensor measurement to account for the reduced convective heat transfer at these altitudes." which we would say addresses the point of comparison between the HiDRON and radiosonde measurements and identfication of weaknesses in the platform which needs to be addressed in future flights.

   As the focus of the paper is intended to be on the turbulence statistics calculated from the HiDRON, the section being referenced had originally been written with the intent to provide a general atmospheric conditions during the measurements, including stability conditions, as reflected in the lapse rate and Richardson number and mean wind profiles, order to provide the reader with context for the turbulence statistics discussed in the later section. However, the revisions requested by the previous round of reviews has somewhat muddled this structure, as comments suggested that we add additional comparisons to the temperature and humidity profiles, move certain points of discussion, and add additional details and statistical metrics, the result being that the description of atmospheric conditions interpreted from the temperature profiles highlighted by the referee now appears out of context and the original narrative structure has become disordered. We have removed this now 'out of topic' paragraph in the revised manuscript but understand that this is only intended as an example of the reviewers disagreement with the organization of the manuscript. We have made some minor edits throughout the manuscript to try to restore, and better explain, the original narrative structure.

2. The description of Figure 7, page 17, focuses on the main trend of the wind profiles and on the presence of the jet-stream during flight 1, detected by both instruments. However, the authors do not discuss the "high-frequency" <U> fluctuations revealed by Hydron measurements, even though they are the most striking pattern of the three profiles (these fluctuations do not appear in the radiosonde data but their vertical resolution may be insufficient). Are they real or artefacts due to the helical pattern of the Hydron path, for example? Can they result from inertia gravity waves? Is a spectral analysis consistent with IGW? These questions, not addressed by the authors, are of prime importance, as their answer obviously conditions the quality of shear and Richardson number estimates.

We note several things here. First, the 'high frequency fluctuations' were initially intended to be addressed by Section 3.5, where we had pointed out that "When the measurements are examined in this manner, some of the scatter in the vertical profiles was be found to be associated with horizontal heterogeneity in the measured statistics". Again, this connection may have become lost in the revised manuscript due to additional content inserted at the request of the referees. Second, these fluctuations are not evident in the radiosonde data simply because of the lack of vertical resolution in the radiosonde wind profiles we were able to obtain. We tried to find higher resolution publicly-available data for this region and this time of day, but were unable to do so. However, although we had initially discarded all data taken during ascent as the aircraft sensors (five-hole-probe, humidity, and temperature) were not configured to perform measurements on ascent, we have since realized that an estimate of the horizontal wind speed can be provided by the GPS-measured ground speed of the aircraft during ascent. This result shows the same 'high frequency fluctuations' as measured by the five-hole probe on descent, and we have added a corresponding discussion of the oscillations observed in the profiles, which also notes that similar behavior is commonly observed in high resolution radiosonde launches. In light of these observations, we feel that adding additional gravity wave analysis would not add any value to the paper, while significantly increasing its length.

3. The newly introduced paragraph on stability conditions is unclear and not convincing. First, the text does not describe the red profiles (N2). We implicitly understand that they are obtained from the second method (described from line 407). The "noisy and spike" structure of N2 profiles and the large differences between the wind shear estimated from the two methods are suspect. Figure 9 shows Ri profiles for one method only (likely the method corresponding to the red profiles in Figure 8). Unfortunately, they are clearly not physical, as no results in the literature show such a stable atmosphere at a vertical resolution of 100 meters (or even lower). The authors should check the literature and their calculation methods.

First, we would like to point out that the profiles in Figures 8 and 9 are identified as N_t^2 (black) and N^2 (red), and S_t^2 (black) and S^2 (red) in the legend. This may not have been clear as a quirk in our graphing program results in the legend text needing to be smaller than desireable to position the legend within the whitespace of the figure. We have updated the legends with larger text where possible. We do note that the difference between N_t^2 and N^2 was provided in the text with the differences between the two being the method of with "The profiles of N^2 and S^2 calculated in this manner (using central differencing between time-adjacent statistical segments) are referred to as N_t^2 and S_t^2" where the subscripted t is intended to indicate that the calculation is conducted using differences in time. The profiles of N_z, and S_z on Figures 8 and 9 are calculated using a technique whereby we calculate the vertical gradients using the data points which were the 'nearest neighbors' in the vertical direction. Hence providing a more true estimate of the local vertical gradients of wind and temperature. In the revised manuscript, we have updated the text to use N_z^2, S_z^2 and Ri_z to avoid further confusion.

We have also done this as the distinction between calculation approaches is an important one since, as the HiDRON is a glider, for every 1 km of vertical descent there is roughly 15 km of horizontal travel. This may not have been clear from the trajectory illustrations provided in Figure 6, as the vertical axis in km was not to scale with the longitude and latitude on the horizontal axes, thus causing the orbits to appear 'tighter' than they were in reality. We have revised Figure 6 with the km axis to scale with the latitude/longitude to provide a more accurate 3D rendition of the flight profile. We have also made several small revisions in the text to highlight the same point.

We feel the distinction in these calculation approaches is relevant as horizontal heterogeneity can can play an outsized impact in the calculation of N^2, S^2 and Ri and the spiralling flight path will be much more sensitive to horizontal heterogeneity than measurement approaches designed to only measure the vertical profile. One downside of this approach is that, due to the size of the orbit, the vertical resolution is of the order of this calculation is of the order of the vertical distance between orbits, roughly 1.5 km. This is why there is an appearance of 'clearly not physical result' as the reviewer is misinterpreting the resolution as being much higher than it is. This misinterpretation is likely because the vertical gradient can be calculated for every 1500 m long (mostly horizontal) statistical window, which results the appearance of higher vertical resolution when plotted as a vertical profile. This is one of the reasons we believe that these measurements are better described using the approach provided in Section 3.5.

We apologize that these discussion points were unclear in the revised manuscript, and have made further revisions to try to make the distinction between these approaches, and the reasoning behind implementing them, and their implications, more clear. Finally, we have removed the moving average step in the calculation of S_t^2, N_t^2 and Ri_t and plot the results in terms of S and N instead of N^2 and S^2 to attempt to better bring the results of these calculations more in line with the referee's expectations.

4. The expression of ε (14), line 492 does not seem to be correct because it is equivalent to write $E_{ll}(k_l) = A$. As a consequence (?), there are inconsistencies in Figure 10. There are no substantial differences between TKE measured during flights 1, 2 and 3 (e.g. ∼0.1 m^2 s^{-2} around 20 km). However, EDR(flight 1) is about ∼5 times EDR(flight 2 and 3), i.e. ε (flight 1) is about 10^2 larger. This is not consistent with TKE. The EDR profiles show an "envelope" that should correspond to the instrumental noise floor (apparently confirmed by a spectral slope close to 0 in figure 14e). The noise level was likely much more important during flight 1. During the present jet-stream, flight 1 should be more prone to turbulence, whereas Fig 10d seems to indicate the virtual absence of inertial turbulence. It is counter-intuitive and then suspect.

There was indeed a typo in equation (14) and have corrected it in the revised manuscript. We have verified that the calculation itself was correct. As we mention in the manuscript "Note that the approach used is only an approximation of ε as the inertial subrange only rarely follows the -5/3 slope, hence it will provide a non-zero value even if no turbulence is present, and therefore some caution is required when interpreting these EDR profiles beyond being a qualitative indication of the presence of turbulence in the form of localized regions of relatively high EDR. " We generally agree with the referee's remaining assessment, which is why we had introduced the spectral slope an additional metric for the identification of regions of turbulence. We note that, as this is an in-situ approach we can expect the sensors to experience increased noise due to the impact of ionizing radiation on the sensor's electronic components.

As for the influence (or lack thereof) of the jet stream on turbulence production, we admit that we were also surprised by this result. We do note however that the time scales of the shear associated with the mean shear of the jet stream ($\sim 13000 m/25\ m\ s^{-1} \approx 500\ s$) corresponds to a square shear frequency on the order of $10^{-}6$ which is comparable to what was measured for all three flights with S_z^2 over vertical scales of 1.5 km and lower than that measured by S_t^2 at much smaller spatial scales. Hence we would hypothesize that the influence of the jet-stream was not strong enough at our measurement location to produce additional shear production, as the jet steam was centered over El Paso (EPZ), 150 km to the south (see Appendix A in the manuscript and the EPZ radiosonde profiles provided in Figure 8). Noting also that the air was relatively dry ($RH < 40\%$), and the measured lapse rate less than the dry adiabatic lapse rate, the buoyant-production would also be reduced for Flight 1 when compared to Flights 2 and 3 (as reflected in the lower N_z^2 and N_t^2 values).

5. The analysis in section 3.5 has been thoroughly revised, and considering the strong reduction of the resolution, this representation may show interesting trends. However, the description of the results are very difficult to follow. In addition, the results depend on the validity of the estimates of N^2, S^2 Ri and EDR described in the previous sections (e.g. the sudden "jump of S^2 above ∼15 km is suspect). The authors conclude that all scenarios can occur (e.g. high turbulence with high Ri, low turbulence with high shear, etc.) but I believe that the reliability of the quantitative results is not high.

We feel that this section is relevant given the glide slope of the aircraft. Particularly as we found that a not insignificant amount of the fluctuations measured in the vertical profiles can be attributed to horizontal heterogeneity, and this becomes apparent given the presentation provided in Section 3.5. The text in that section is largely intended to draw the reader's attention to certain large-scale features that are evident in the visualizations and not intended to make any concrete inferences about the nature of the turbulence. Particularly, as we note again, that much of the disconnect between Ri_z, S_z, N_z and the turbulence-related quantities is likely due to the significantly reduced vertical resolution of this approach. Although this point had been raised in the conclusions already, we have made several additions to the text to better reiterate it.