Response to Referee 3

We appreciate the time taken by the referees to once again review our manuscript and provide comments and suggestions for its improvement.

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

This paper has been improved by a tighter focus and omission of many previous vague and unsupported statements. Most of the arguments from the previous reviews have been settled by removal of the sections in question, or by clarification in this revision. Remaining issues are noted below.

Comments:

- Line 10: "nominal vertical resolution on the order of 1 m during was achieved" Please re-word.

  We have reworded this statement to 'a nominal vertical resolution'.

- Line 15: "and broadband response measured within boundary layer turbulence". Do you mean "broadband winds"?

  We have reworded this statement to "The low-frequency response of the infrasonic microphone was found to correlate to long wavelength wind velocity fluctuations measured at high altitude, and broadband frequency response of the microphone was also measured within boundary layer turbulence.'

- Line 27: "predict stratospheric and high altitude turbulence" High altitude is different than stratospheric?

  We have reworded this statement to 'predict stratospheric and upper-tropospheric turbulence'.

- Line 114: "which allowed it to measure a wide range of turbulence scales horizontally at high vertical resolution." Confusing/misleading. Implies that horizontal and vertical effects can be separated, or that vertical variations dominate horizontal ones. Statement should be qualified accordingly.

  We have reworded this statement to 'which allowed it to measure over a distance of up to 50 m horizontally for a 1 m change of altitude'.

- Line 211: "to check for any Reynolds-number dependence of C_β, C_α and C_q." Were there any?

  We have added a statement that none was found.

- Line 214: "flexible polymer tubing" could be almost anything. Usually, these tubes must be semi-rigid to properly propagate the pressure variations. Please specify.

  The manufacturer sells this product as PVC plastic tubing with a trade name of Tygon. We have updated the text to include this information.

- Line 259: "angle between the true airspeed determined from the aircraft's pitot probe, and the vertical velocity determined by the aircraft's variometer". No idea what this means. Pitot does not provide an angle, and variometer (pressure altitude rate, akin to inertial velocity) is not the same as the vertical component of relative wind.

  When the autopilot manufacturer was queried as to how the angle of attack is determined, the description above is verbatim to the response we received. We have contacted them again and have updated the description of how α was found, specifically from a combination of the true airspeed determined from the aircraft's pitot probe, the vertical speed determined by a Kalman filter fusion of the static pressure rate of change, vertical acceleration, and global positioning system velocity, along with the aircraft orientation measured by the autopilot gyroscopes which was used to transform between inertial and body-fixed coordinates.

  We have updated the text accordingly.

– Lines 293-301: Sensor and signal amplification/conditioning and quantization noises are random processes, so cannot be measured before-hand and subtracted from signals measured later. The noise removal process described is as likely to corrupt the measurement as it is to clean it up. Even if "noise" means biases, these are likely to be variable with temperature, so could not conceivably be obtained from pre-flight sampling.

This would generally be true, except that much of the high frequency noise in the pressure transducer signal is quasi-periodic noise which we believe to be introduced by a switching voltage regulator, rather than more stochastic noise sources described above. In this revision, we have added additional exposition about the noise characteristics as well as evidence that the approach we have devised is successful in reducing the impact of the quasi-periodic noise, and evidence that the noise was independent of atmospheric conditions. Specifically:

1. Excerpts from time series illustrating the noise content.
2. Frequency spectra showing that the frequency content of the transducers in an inactive environment chamber did not change between atmospheric and low pressure/low temperature conditions.
3. Frequency spectra for all altitudes, which shows that the high frequency content of the transducer signal does not change during the flight for frequencies higher than the signal from the five hole probe
4. A comparison of excerpts of the pressure time series taken during flight showing that the filter appears to successfully remove electrical noise from the pressure signal.
5. A comparison of frequency spectra before and after filtering showing that the high frequency content is reduced, and that the pressure signal more accurately reproduces a -5/3 slope in the square-root of the corresponding dynamic pressure signal (i.e. in the resulting velocity signal).

Please also note that this noise filtering approach was added in response to a previous comment that stated 'Not clear how (of if) the noise floor/noise figure is removed in the qualified data fits. The noise floor can corrupt the estimated slope, so its removal can often retrieve epsilon values that were previously rejected. This could help to explain the absence of qualified values at the higher altitudes,' As the noise floor in the wind velocity was a function of numerous inputs, we could not devise a hands-free, unbiased, approach to consistently identify or remove it from the wind estimate. Therefore, we introduced this filtering approach to at least remove some of the noise introduced by the pressure transducers. We found that its impact on the statistics presented in the remainder of the paper was negligible.

– Lines 340-341: "uncertainty in wind magnitude was found to be most dependent on the yaw angle" How? Why? Do the wind excursions in Figure 4 correlate with yaw angle?

The uncertainty dependence of wind magnitude on aircraft attitude is introduced into the wind estimate during the coordinate transformation between body-fixed and inertial coordinate systems. During flight, the yaw angle can vary from 0 to 360 degrees, whereas the sideslip, angle of attack, pitch and roll angles are near zero. The result is that the greatest contribution to the u/v wind components during transformation is the true airspeed being multiplied by the sine/cosine of yaw and cosine of the pitch. The result is that during an orbit, the horizontal velocity components will have high sensitivity to yaw error at both $0°$ degrees and $90°$. For a similar reason, the vertical component of velocity is most sensitive to error in pitch.

As illustrated by the uncertainty bounds shown in Figure 4, the wind excursions cannot be explained by error in yaw. Note also that prior to preparation of this manuscript, we conducted an informal perturbation study to determine if these excursions could be explained by measurement error and found that for them to be removed from the wind profile, the yaw error would not only have to be non-monotonically dependent on altitude (i.e. not attributable to sensor drift), but it would have to be so high that it could only be explained by failure of the system.

We have added the first statement to the manuscript, the second statement is left out due to the informal nature of the perturbation analysis that was conducted (the uncertainty analysis presented in the manuscript being the more rigorous approach).

– Lines 400-401: “and we assume that the characteristics of the atmosphere within these segments are horizontally homogeneous (i.e. they are a function of z only)”. A very loaded assumption to make with no justification!

This statement was not thought to be as loaded as as the reviewer is implying, due existing consensus that the intrinsic stability in the stratosphere and upper troposphere will promote horizontal homogeneity (as exemplified by one of the other reviewers being insistent that variability in the measured profiles must be due to measurement error or vertical stratification rather than any horizontal heterogeneity). As the aircraft approaches the surface and the radius of the orbits becomes smaller, this assumption also becomes less restrictive.

We have added our rationale for this assumption in the revised manuscript.

– Lines 440-441: “the HiDRON measurements do contain short wavelength fluctuations” How short? Please quantify.

We have added a statement that the wavelengths we are referring to are the ones on the order of 1 km or less.

– Line 444: “These low frequency waves may be bias in the wind estimate introduced by the orbital path”. Indeed, it seems like they may correspond to one turn on the helix, so may be an anomaly in the wind retrieval. Comparison to the vertical period of the helix would be important here.

We have added some quantification and discussion of these values to the text.

– Line 464: “backing with altitude”?

Backing winds are winds that change direction counterclockwise with height. We have revised the statement to remove the use of backing and simply state the wind direction changes.

– Line 470: Should note that an accurate assessment of TKE also depends on capturing the lowest wavenumber components in the inertial subrange, since they contain the largest energy per wavenumber. Did the PSDs over the statistical intervals exhibit a roll-off or flattening of the $f^{-5/3}$ slope at low wavenumbers, indicating the outer scales were captured?

A note to this effect has been added to the revised manuscript. The low frequency content more often increased, rather than rolled off, as described in the wavelet analysis, and discussion of the time series. This was also the justification for the selected segment length.

– Line 473: “over a specified frequency, f, range”. What was the range?

The frequency range determination process was described in the manuscript shortly after this statement. We have revised the text by removal of the statement above so that the question of frequency range does not come up until it is actually described.

– Line 478: “which will reach a minimum at the frequency where the noise has a greater contribution to the integration than the signal”. Why?

Compensated, or "pre-multiplied," spectra are commonly employed in turbulent boundary layer studies because they facilitate the visualization of energy spectra on semi-logarithmic axes. Specifically because they allow for a clearer representation of the frequency/wavenumber dependence of the relative contribution of each frequency/wavenumber to the overall energy content. This is because, for example,

$$d\langle u^2\rangle = F_{uu}df = fF_{uu}d(\log f). \tag{1}$$

Hence when $fFd(\log f)$ begins to increase on a semi-logarithmic plot at high frequencies, this indicates a frequency range where the energy content increases with f. Given that universal equilibrium range turbulence will decrease in energy content with f, the minimum in $fFd(\log f)$ indicates a frequency at which the noise begins to have a greater contribution to the variance than the turbulence content.

However, the use of the compensated spectrum is a procedural detail and in hindsight distracts from the overall point of this processing step, which is to determine at frequency the contribution to overall variance increases with increasing f, rather than decreases. The same result could have been achieved using the un-compensated frequency spectrum, and we only used the compensated spectrum to simplify visualizing the frequency-dependence of $\langle u^2 \rangle$, $\langle v^2 \rangle$ and $\langle w^2 \rangle$.

As including the detailed exposition above would distract from this point, while adding little value to the overall intent of its inclusion, we have simplified the discussion to simply state that the upper bound of this range was determined by identifying the frequency where the noise has a greater contribution to the integration than the signal than the velocity fluctuations.

– Line 478-479: "filter frequency was consistent with the probe's frequency response in the boundary layer and varied between 1 Hz and 20 Hz above the boundary layer". Does "filter frequency" correspond with the "frequency range" in line 473? By "probe's frequency response" hear do you mean "probe bandwidth"? How did this vary above the boundary layer?

By filter frequency, we meant the upper bound of the range described in the previous statement. We have modified 'filter frequency' to read upper bound of the frequency range.

Yes we do mean probe's bandwidth. These terms are interchangeable and, as we use frequency response at numerous points throughout the paper prior to using it in this sentence, would prefer not to change it to 'probe bandwidth'. We have clarified that we are referring to the maximum frequency response, as frequency response in general can also refer to the Bode plot of the probe's response to excitation.

As noted in the text, the frequency at which the noise exceeded the signal varied between 1 Hz and 20 Hz. We do not describe a trend above the boundary layer since no trends were evident, being dependent on the presence of low-frequency energy content. We have altered the text to now read 'with the higher upper frequency bounds corresponding to instances where there was increased low frequency content in F_{uu}, F_{vv} and F_{ww}.'

– Line 514-516: "due to the statistical segment length used for averaging, the value of $\langle k \rangle$ will be biased to wavelengths smaller than the statistical segment length" Why? Do you mean smaller than the outer scale (as noted for Line 470 above)? Is this what "and therefore may not completely describe the actual energy content of the turbulence" is alluding to? This could be said much more clearly.

That is exactly what we meant. Due to the segment size of ~ 3 km we cannot resolve wavenumber content larger than the length of the statistical segment (or frequency content below $O(0.01)$ Hz). The point of the above statement is to note that this is insufficient to capture any sort of outer scale/low frequency energy contribution below of wavelengths longer than the segment length. We have changed this statement to note that, in addition to the implicit assumptions made when calculating $\langle \varepsilon \rangle$, the method used to calculate $\langle k \rangle$ reflects only the energy content corresponding wavelengths smaller than the statistical segment length (or frequencies higher than the the inverse of the time taken to traverse that segment length).

– Figure 10: It would be helpful to draw a $\langle k \rangle^{3/2}$ line on the plot for reference.

This line has been added.

– Line 524: "Above the boundary layer turbulence $\langle k \rangle$ and $\langle EDR \rangle$ are largely in agreement" This is hard so see, since $\langle k \rangle$ should be proportional to $\langle EDR \rangle^2$ but Figure 11 compares $\langle k \rangle$ to $\langle EDR \rangle$.

We have changed this figure and corresponding references in the text to $\langle EDR \rangle^2$

– Line 535: "Nyquist frequency of the minimum probe response". No idea what this means. Nyquist relates to the sampling frequency, not the frequency response.

We have changed this sentence to refer half the maximum frequency response of the probe.

– Lines 538-540: "Noticeable in Figs. 11b, d, and f is the significant long wavelength content for $\kappa\ell < 0.003$ (wavelengths larger than 2 km) when $z > 10$ km." I don't see this. It would help to show a color bar. Looks to me like there is significant long wavelength content below .0003 rad/m over all altitudes. And why does the plot have a curved boundary on the left and a straight boundary on the right? Seems that the "time" variable discussed in the wavelet transform is really altitude z here. Correct? This whole discussion is very terse for readers unfamiliar with wavelets.

The long wavelength content for $\kappa\ell < 0.003$ when $z > 10$ km is more evident when plotting the wavelet coefficient on non-logarithmic contours (which does a poor job of visualizing the short wavelength distribution of the coefficient), or when using a non-colorblind friendly colormap.

We have modified this figure by changing the colormap, added a colorbar, and adjusted the horizontal axis to better constrain the wavenumbers to those less than an orbit. However, the difference below 10 km and above 10 km is still subtle, therefore we have clarified within the text that the contributions for altitudes greater than 10 km are most noticeable for Flights 2 and 3

The wavelet transform is calculated in the time/frequency domain, but since altitude is a function of time (which we had tried to indicate by referring to $z(t)$), it is then plotted as a function of altitude. We have rephrased this sentence to be more clear.

The curved boundary on the left of the figure is due to the time-frequency nature of the wavelet transform. At the start and end of the time series, there is insufficient information to resolve the low frequency content. However, towards the central part of the time series, the maximum low frequency content can be resolved, resulting in the curved boundary on the left (low frequency) end of the figure. The right of the figure, representing the high frequency content, is not subject to such resolution issues and therefore has a straighter boundary (although in this presentation, since the frequency has been transformed to wavenumber using Taylor's hypothesis, the highest wavenumber resolved is a function of the airspeed, which decreases with altitude, resulting in the slanted boundary on the right of the figure.)

We have updated the text to provide more description of the wavelet transform and its features.

- Lines 541-547: What are the implications of these observations from the wavelet transform? Why is the frequency content behind turbulent parameterization important? Usually this is constrained by the inertial cascade.

Our primary rationalization for examining the low frequency content is that it helps us to explain the differences between k and EDR, highlighting that EDR does not capture the low wavenumber content. This analysis also helps us to justify and understand the use of the turbulent kinetic energy estimate using lower frequency energy content than used for the initial $\langle k \rangle$ estimate.

We have modified this paragraph to better highlight some of the above points.

- Lines 557-558: "and therefore is attributed to increased atmospheric absorption due to the increase in molecular mean free path with altitude". Seems this could also be due to the decrease in coupling coefficient to the microphone diaphragm due to lower density.

It is not clear to us how the density can affect the efficiency of conversion of mechanical energy to electrical energy. Assuming the intention was to describe the decrease in mechanical forcing which could be expected due to lower density, we then would expect the variance of pressure measured by the microphone to scale with P^2, rather than P. However, we found that it does not scale with P^2. In addition to various other normalizations, we also tried scaling the variance of the microphone signal with ρc (corresponding to the expected change in sound intensity with altitude), which was also was unsuccessful. The only scaling we found that provided some success was the normalization of the the variance of microphone signal with P, as noted in the paper, and this result was consistent with the expected attenuation due to increase in mean free path, following the discussion presented in the cited reference (Bass 2007).

- Line 562: "The altitude attenuation will be dependent on the local temperature as well". Why?

In the paragraph before the referenced statement we had attributed the altitude dependence to the mean free path, which is a function of pressure and temperature. We have added this statement explicitly to the above sentence.

- Lines 565-566: "the resulting infrasonic amplitude profile can be observed to strongly correlate with [k]". Depends what you mean by "correlate". The variations in $[k]$ are not correlated with the normalized infrasound variance, only the large scale means seem to correlate. The infrasound signal looks like a LP filtered version of the TKE. Seems the infrasound signal (even filtered at 20Hz) should be able to follow the variations in $[k]$ that occur over km of vertical intervals. Why does it not?

The use of the word correlate to indicate correspondence was a poor choice and has been changed. The $[k]$ measurement is an in-situ measurement, whereas the infrasonic microphone is a remote measurement. Therefore there is no reason to expect exact correlation between the two measurement approaches. Indeed, at least ideally, the infrasonic microphone will detect the turbulence before the aircraft enters it, acting as a filter to the 'spikiness' of the $[k]$ measured by the in-situ sensor. Not only that, but the microphone will also detect any sound generated by nearby turbulent patches that the aircraft does not fly through. The net result can thus be expected to be a low-pass version of the $[k]$ profile, as the sound propagation will be be omnidirectional and emitted from numerous locations, whereas the $[k]$ can be expected to be constrained to stratified vertical layers.

We have added these discussion points to the revised manuscript.

- Line 569: Do you mean $[\sigma]_{LF}$ instead of $[\sigma_f]$? Also, I don't understand why these two would increase at the same rate if high frequency energy is primarily increasing, as supposed as being "most likely".

  Yes, we meant $[\sigma]_{LF}$ and have fixed this typo.

  Rate is probably the wrong word for what we are trying to describe. We have reworded this statement to try to more clearly describe our rationale for why the ratio $[\sigma^2]_{LF}/[\sigma^2]$ will remain constant in the boundary layer.

- Line 575: "This is due to the helical flight path". It is really due to the small flight path angle. A steep helix would not be as susceptible to horizontal gradients.

  We were referring to the specific helical flight path flown during these flights, not helical flight paths in general. We have reworded the statement to 'This is due to the shallow glide slope of the particular helical flight path flown in these experiments'.

- Line 585: "These values were then re-interpolated to each statistical segment." Don't know what this means, exactly. Averaged for each segment? Interpolated how? And why were the 200m smoothed data averaged again (binned) at 100m intervals?

  In this context, the smoothing is applied as a low-pass filter, and therefore the time-series post-smoothing has the same number of data points as pre-smoothed. The bin averaging over 100 m intervals was done to facilitate the differencing across the 100 m interval used for gradient calculation (i.e. equivalent to downsampling the signal). In the past we calculated a difference across $\Delta z = 100$ m for each point in the time series, but a different reviewer took exception to this approach in a previous revision of this manuscript, stating it obscures the vertical resolution of the differencing. Finally, since the $\Delta z = 100$ m downsampled data points at which gradients were calculated do not coincide with the z locations of the statistical segments over which $\langle \theta_v \rangle$ was calculated, we had to interpolate the gradients to locations of the statistical segments. It is a messy process, but seems to be necessary to achieve Ri profiles similar to what is observed from radiosonde measurements.

- Line 601-602: "effectively reproduces the N2 profiles calculated along the flight path". Only if by "effectively" you mean a highly smoothed version of N2.

  Yes, we meant that (particularly when compared to the S^2 profiles) this approach produced a highly smoothed version of the N^2 profile determined with other approach. We have reworded this statement to 'effectively reproduces the trend of the N^2 vertical profiles calculated along the flight path.'

- Line 617: "Wind profiles were in good agreement with the available National Weather Service radiosonde profiles". Please quantify "good". Likewise, quantify "best comparison" in the next sentence.

  We have added a discussion quantifying the difference in the body of the paper while discussing the wind profiles and summarized this in the last section.

– Lines 619-620: “over a large horizontal wavelength range with high vertical resolution”. Confusing/misleading. Implies that horizontal and vertical effects can be separated, or that vertical variations dominate horizontal ones. Statement should be qualified accordingly (as was done in the body).

We have reworded this statement to ‘over a large horizontal distance relative to the vertical distance traveled.’

– Line 637: “fror example”. Typo.

This typo has been fixed.

– Line 643: “Additional flight patterns can also be designed with tighter helical descent can be designed”. Awkward.

Reworded to ‘Additional flight patterns can also be designed with tighter helical descent’

– Appendix A: “all transducers” is vague. Wind retrieval is a combination of many transducers, including relative wind, attitude, and inertial velocity. It would be clearer to use the more specific nomenclature from the body of the paper.

We have removed the line legend and reference the line colors in the figure caption. To use the nomenclature from the body of the paper would have required shrinking the font size below acceptable limits (at least for the graphing software that we used when generating the bulk of the figures used in the paper).

Response to Referee 4

We appreciate the time taken by the referee to review our manuscript and provide comments and suggestions for its improvement.

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

This article presents an exciting new gliding UAS platform for conducting potentially low-cost meteorological observations up to 30 [km] MSL. A calibrated five-hole pressure probe was employed to quantify turbulence characteristics and assess the effectiveness of an infrasonic microphone to qualitatively observe atmospheric turbulence. While the discussions presented in the article adequately support the effectiveness of the platform's capabilities to observe stratospheric environments reliably, the limited discussion detailing the processing of data needs further refinement before publication.

- 98: "Within the atmospheric boundary layer, the infrasound energy from ground-based arrays has been found to correspond to the turbulent kinetic energy in the atmosphere, particularly when buoyantly-produced 100 convective turbulence is present (Cuxart et al., 2015). The infrasound energy levels were also found to increase in the presence of elevated jets or turbulence above the measurement height, which was thought to be caused by the sound generated at higher altitudes reaching the microphones."

  "found to correspond to turbulence" how? A working hypothesis relating infrasound measurements to turbulence is warranted if any quantitative assessments of turbulence characteristics are to be derived from the infrasonic microphone measurements.

  The authors of the cited study specifically show that the infrasound energy measured by integration of the energy spectrum of the recorded microphone signal in the range 0.01 Hz to 15 Hz increases with the turbulent kinetic energy measured by a co-located sonic anemometer. The authors refer to the microphone signal amplitude as being a surrogate for turbulent kinetic energy but the relationship was non-linear, and the authors of the cited study related the turbulent kinetic energy measured by the sonic anemometer to the voltage content in the microphone signal, not the sound pressure level or pressure itself (which will vary with make and manufacture). Hence we described this relationship as a qualitative correspondence rather than a correlation or quantitative relationship.

- The objective of using an infrasound instrument for turbulence characterization is unclear here. Do the authors intend to simply use infrasound measurements for qualitative turbulence detection? Or quantify turbulence characteristics? A decisive discussion helps clarify the objectives of using the infrasound instrument.

  The motivation for the inclusion of the microphone in this study was indeed for qualitative turbulence detection, motivated by the potential usage of infrasonic microphones to detect clear air turbulence. The in-situ sensors on the UAS were intended to provide the quantification of the local turbulence characteristics allowing a comparison of the microphone measurements to the quantified turbulence. This is because the use of infrasonic microphones for clear air turbulence detection is still in its infancy, with little known about how this particular microphone might respond at altitude, or whether a signal indicating the presence of turbulence can even be detected. Given that the difference in remote vs in-situ sensing modalities are very different, we are still at the point of qualitatively connecting measured infrasonic energy to turbulent kinetic energy.

  We have re-written this section to try to be more clear.

- 220-224: "During flight, the autopilot maintained flight speeds sufficient to produce pressure differences well within the range of the low-sensitivity transducers (i.e. the dynamic pressure was maintained between 100 Pa and 200 Pa) which exceeded the range of the high sensitivity transducer connected to ΔP_1. Hence, only the readings from the low-sensitivity sensors were used for data analysis. However, the high sensitivity transducers provided a means to estimate the uncertainty of the pressure measurement, as will be described later."

If I understand this correctly, the wording suggests that the high-sensitivity pressure transducer was used for uncertainty estimation only and not for scientific analysis. This raises questions about the increasing instrument noise floor with altitude. Have the authors modeled/empirically identified the five-hole probe instrument's noise characteristics as a function of altitude? Was the high-sensitivity transducer saturated frequently in flight and mostly the data unusable?

Specifically, the high-sensitivity ΔP_1 transducer (which is sensitive to dynamic pressure) was saturated during the entire forward flight portion of the experiment, making recovery of any relative wind velocity vector impossible, The high-sensitivity ΔP_{32} and ΔP_{54} transducers, which are sensitive to α and β, were not saturated, since the α and β angles are close to zero throughout the entire flight. As described in the uncertainty analysis section, this allowed us to compare the high- and low-sensitivity ΔP_{32} and ΔP_{54} transducer pressure readings during the entire flight and use the difference between them for uncertainty estimation. We have updated the paragraph referenced above to clarify.

As will be discussed later in this response, the noise was found to be independent of altitude.

– 243: "The time-dependent horizontal wind velocity magnitude and direction could then found from" modified to "The time-dependent horizontal wind velocity magnitude and direction could then be found from"

This has been corrected. Thank you.

– Figure A1. "Figures showing comparison of (a) horizontal wind velocity magnitude, (b) horizontal wind direction, and (c) vertical component of wind velocity calculated using all transducers to find Q, α, and β and using only two transducers to calculate Q and β with α determined from the aircraft angle of attack measurement. Comparison of resulting (a) ⟨u′2⟩, (b) ⟨v′2⟩, and (c) ⟨w′2⟩ Reynolds stress tensor components." I believe that the figure caption has mislabelled tiles "(a) ⟨u′2⟩, (b) ⟨v′2⟩, and (c) ⟨w′2⟩". Shouldn't it be "(d) ⟨u′2⟩, (e) ⟨v′2⟩, and (cf) ⟨w′2⟩"?

This has been corrected. Thank you.

– 293 - 298: "To minimize the impact of electrical noise introduced into the pressure signals by the sensors and data acquisition system, during post-processing a background noise subtraction procedure was conducted on the digitized voltage signals prior to scaling them to Pascals. This process involved identifying a 5 minute long segment of the signal measured prior to balloon launch when a cover was present over the five-hole-probe (Fig. 1a) and the infrasonic signal was quiescent. This portion of the time series was assumed to be representative of the background electrical noise and therefore subtracted from the full time series in the Fourier domain in 5 minute long segments."

The authors assume, without adequate justification, that the instrument noise characteristics are independent of flight dynamics. This choice, without a discussion or proper justification, is speculative and questionable, and the representativeness of the instrument noise measured pre-flight in quiescent conditions warrants reevaluation.

Further, it is claimed that electrical noise is "minimized" without a presentation/discussion identifying/stating (with references to studies in literature if any) the implications of noise on the measurements or derived data products.

In this revision, we have added additional justification that the approach we have devised is successful in reducing the impact of high frequency periodic noise content that was present. Specifically, the revised manuscript now includes:

1. Excerpts from time series illustrating the noise content.
2. Frequency spectra showing that the frequency content of the transducers in an inactive environment chamber did not change between atmospheric and low pressure/low temperature conditions.
3. Frequency spectra for all altitudes, which shows that the high frequency content of the transducer signal does not change during the flight for frequencies higher than the signal from the five hole probe
4. A comparison of excerpts of the pressure time series taken during flight showing that the filter appears to successfully remove electrical noise from the pressure signal.
5. A comparison of frequency spectra before and after filtering showing that the high frequency content is reduced, and that the pressure signal more accurately reproduces a -5/3 slope in the square-root of the corresponding dynamic pressure signal (i.e. in the resulting velocity signal).

Note also that despite this evidence indicating that the electrical noise removal was successful, we found that its impact on the statistics presented in the remainder of the paper was negligible. However, the noise filtering approach was added in response to a prior reviewer who felt that not having such an approach would bias the velocity spectra in the inertial subrange.

– The discussion of pressure measurement errors due to noise is out-of-place in section 2.2.1. It is recommended that the authors discuss the instrument noise/characteristics in section 2.2.1 instead.

  We had initially included this discussion in 2.2.4 as it cannot be determined if the noise source is the transducers, data acquisition system, or embedded computer and the embedded computer is the final link in this chain. However, given that the approach had its greatest impact on the five-hole-probe pressure measurement, we have moved this discussion to 2.2.1 as suggested.

– 359: "Comprise" modified to "compromise".

  This has been corrected. Thank you.

– 382: "This latter constraint is introduced since ΔP_1 exceeded the transducer's range shortly after the aircraft was released and started its flight towards the helical orbit location." Is the transducer under consideration here the high- or low-sensitivity transducer? Please clarify.

  We have clarified by rewriting this statement to read 'This latter constraint is introduced since even the low-sensitivity ΔP_1 exceeded the transducer's range shortly after the aircraft was released and started its flight towards the helical orbit location. By the time the aircraft reached the orbit location, ΔP_1 had returned to a range measurable by the low-sensitivity transducer, although it never reduced to a value measurable by the high-sensitivity ΔP_1 transducer.'

– 401: "In order to decrease the vertical spacing between statistical values, each segment is overlapped with its neighbor by 50%, thereby decreasing the spacing of statistical quantities to 75 m vertically)." It is unclear as to where the bracketed text begins. Please clarify.

  The bracket has been removed.

– Figure 9: The horizontal wind data presented here exhibits periodic motions (on visual inspection) on 1 km vertical scales (Ex: 5 - 10 km in Flight 2). It is not uncommon to expect periodic artifacts in horizontal wind estimates derived from flights following periodic/helical tracks. It is recommended that the authors present a brief discussion to clarify the quality of the estimated horizontal winds with emphasis on the impact (if any) of periodic artifacts related to the flight trajectory on the horizontal wind components. Any evidence showing that the periodic artifacts (if any are present) are uncorrelated to flight dynamics will aid in resolving questions on the quality of horizontal wind estimates.

  We are familiar with this issue with winds measured by orbiting UAVs which, from our experience, is typically introduced through error propagation and amplification in the conversion from body-fixed, to inertial coordinate systems, with error in the magnetometer reading the greatest culprit. This concern has also been raised by several other referees in early versions of this manuscript. To address this issue, prior to preparation of the initial submission and over the course of several revisions we have:

  1. compared periodicity in the wind estimate to that of the helical flight path's pitch and observed that they do not exactly correspond to the orbital pitch, which means that if there is a bias it is altitude dependent However, the similarity is close enough that we do not want to highlight the difference in the manuscript as it is weak justification. Note also that the agreement between periodicity and orbital pitch is really only evident during Flight 1, as the aircraft passes through the edge of the jet stream.
  2. noted that very similar periodicity is evident in the ABQ and TUS radiosonde measurements. The introduction of multiple radiosondes into the comparison also shows that the HiDRON measurement trends are within the variability in the radiosonde wind measurements. However, these radiosonde measurements are not coincident/co-located, nor are they high resolution measurements, indeed it appears that the wind measurements in the radiosonde

profiles were downsampled, which may artificially introduce the periodicity observed. Therefore we do not feel confident that this is sufficient evidence that this periodicity is natural and not measurement bias.

3. introduced the uncertainty analysis which shows that the observed periodicity in the winds far exceeds the expected uncertainty, but this analysis only accounts for known errors and cannot incorporate unknown biases (i.e. magnetometer drift). However, prior to preparation of this manuscript, we conducted an informal perturbation study to determine if the periodicity in the wind estimate could be explained by measurement error but found that for them to be removed from the wind profile, the yaw error would not only have to be non-monotonically dependent on altitude (i.e. if it is due to sensor drift, it would have to be drifting significantly away from and then back toward the correct value), and the magnitude of the error would have to be so high that it would be noticeable by the ground operators. Given the informality of the perturbation analysis and that we cannot know beyond a reasonable doubt that there wasn't drift/failure of the autopilot yaw measurement to this degree of error, we elected only to include the formal uncertainty analysis.

4. added in the wind estimated by the aircraft's drift measured by GPS on ascent. These are only roughly co-located measurements (see Figure 6), but there are certain parts of each of the profiles that quite strongly agree between the two wind estimates. Most notably at high altitudes during Flight 3, where the ascent/descent phases of the flight were most co-located and coincident. Introduction of this comparison appears to have done the most to assuage the previous reviewers' concerns.

Most of these arguments have already been noted in the text, although relatively weak when taken individually, as a whole they provide us with some confidence in the wind estimates. The net result is that the discussion we have provided in the manuscript acknowledges that there may be bias introduced by the flight path. However, we intentionally keep the discussion vague since we cannot be 100% certain that the wind estimate is/isn't affected by this bias.

We also hypothesize that it is possible that there there may be significant horizontal heterogeneity in the winds at these altitudes given the amount of mean kinetic energy, large-scale shear, and static stability present, which could lead to large-scale quasi-two-dimensional inactive instabilities at the edge of the jet stream. The glider will be much more sensitive to such horizontal heterogeniety than the balloon-based measurements, which will drift with the air parcel and therefore not experience as much horizontal shear. However, at least one previous reviewer was resistant to these suggestions, particularly as we have no evidence that such horizontal heterogeneity is present. We therefore have limited this point to a suggestion of its possibility in the text.

In the revised manuscript we have added some additional exposition to the text in the discussion of uncertainty about the source of periodic artifacts and in the discussion of the wind estimates about the possibility that periodicity in the wind profiles can be attributed to this source of uncertainty.

– 438: "In general, the wind magnitude and direction measured by the HiDRON H2 are within the bounds provided by the radiosonde soundings, with the wind direction measured during descent producing good agreement with that reported by the radiosondes and by the GPS on ascent." This statement is vague. It is not clear what the constitutes "good" agreement here especially when the nearest radiosonde sources for comparison are spatially separated on scales as large as 100s of km.

  We have added a discussion quantifying the difference in the body of the paper while discussing the wind profiles and summarized this in the last section.

– Section 3.3: Infrasonic Detection of Turbulence. The discussion presented in this section suggests that the authors use the infrasonic microphone for purely qualitative analyses. It is vaguely stated that the measured infrasonic amplitude correlates to TKE and no further insights were offered to reinforce the effectiveness of utilizing the instrument for turbulence observations. It is recommended that the authors provide comments on the effectiveness of the infrasonic microphone data for any quantitative turbulence characterization.

  As we do not have a comprehensive spatio-temporal map of the sound-generating turbulence during these flights, we cannot yet provide a quantitative empirical connection between the turbulence in the atmosphere and the infrasonic signal measured by the microphone. Although there are ongoing efforts seeking to produce this relationship, we cannot include them in this manuscript. We therefore have to relegate our discussion and conclusions to a qualitative evaluation of the microphone response, and have tried to rework the relevant sections of the manuscript accordingly.