Response to Referee 3

We once again appreciate the time taken by the referee to further 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. As the new referee comments are integrated into the original review, the original review by referee is provided in black text, the original response by authors is provided in blue text, the latest review by Referee 3 is provided in green text, and the response from the authors to these latest comments is provided in red text. Note that below we have removed portions of the original response which Referee 3 has found acceptable

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.

Reviewer comments for this latest revision are colored in green. Items without additional green comments are acceptable in the revision.

This version provides more information on the filtering scheme for multi-hole probe pressure measurements, but the noise-subtraction approach remains highly suspect. Within this, however, is a basis for determining the noise floor of the measurements, which could then be used to select spectral data above the floor by a suitable factor for turbulence characterization. The estimation process for vertical wind remains unclear, particularly for flights 2 and 3 where one of the pressure sensors was disconnected. As this is not really needed for the results of the paper, perhaps this could be omitted if it cannot be clarified.

Comments:

- 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.

  Text still incorrectly implies that vertical aircraft speed is the same as vertical wind, e.g. lines 325-326: “This orientation data allowed transformation of the vertical speed from inertial to body-fixed coordinates”. How do you get angle of attack from vertical (body) velocity and pitch angle and pitot airspeed? Pitch angle is not the same as angle of attack, and flight path angle is not either, unless you assume vertical air motion is zero. The description of “a combination of factors” is still very vague, making it hard to gauge the integrity of the data. Later, lines 419-421, implies (incorrectly) that the vertical wind is found from the sine of the pitch angle multiplied by the airspeed. It would be much clearer if the equations used to make these estimates and error bounds were provided.

  Although we believe we understand the calculation, we do not have access to the exact form of these equations being used (being contained within the manufacturer’s software), therefore we cannot reproduce them within this manuscript. However, we confirmed that the calculation does indeed assume the vertical velocity of the air relative to the ground is negligible relative to the vertical velocity of the aircraft relative to the ground. We had thought that this assumption was introduced in the prior version of the manuscript but now realize this was not the case. Thank you for identifying this omission.

  We have revised this manuscript to include this statement. Note that, with this assumption, α can calculated following the standard form used in flight dynamics.

– 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.

The noise shown in Figure 4(a) does indeed appear to have a set of spikes, with a repetition frequency of about 1 Hz. The fact that there is no prominent, persistent spectral component near 1 Hz (Fig 4(b), but the noise spectrum is very broad, suggests that these spikes are not regular. Read: unpredictable in amplitude and phase, even on a short time scale. Hence subtraction of a pre-recorded signal, in the time domain or frequency domain, cannot be expected to remove noise in measured signals. To the extent that this subtraction reduces the spectral component amplitudes, it can be expected to reduce signal variation in the time domain, and appear "filtered", but both signal and noise are affected equally. At frequencies where the signal component amplitudes are large, the subtraction of a small noise changes the signal very little. But at frequencies where the component amplitudes are not large relative to the noise, then both signal and noise are altered greatly, sometimes reducing and sometimes increasing the component magnitudes, depending on the relative phase. Because of this random effect, it would be better to only accept spectral components for analysis that are significantly larger than the noise spectrum, e.g. by 1 order of magnitude. All others are corrupted by noise, and doubly corrupted by the filtering technique proffered here. A test for this corruption would be to compare time domain signals pointwise before and after filtering at a scale where the frequencies of interest (e.g. up to 20 Hz) can be seen. The spectra of Fig 4(f) show another way to assess this, but it would be helpful to show the noise spectrum, too.

We agree with this commentary, and were actually surprised that the process worked as well as it did. We also note that while vetting this procedure, we did indeed closely compare time domain signals before and after filtering at a scale where the frequency content on the order of 20 Hz is evident seen and did not observe evidence of corruption. However, we obviously could not provide such a detailed comparison in the revised manuscript as it would be impossible to do so over the length of the entire flight. Also, selectively presenting a 1 s portion of the signal showing no corruption provides little value as evidence that such corruption was not present elsewhere.

This is why we provided the spectra of Fig. 4(c) which shows *all* one-sided power spectra calculated from the flight data, and compares them to the spectrum which is used to represent background noise, as the referee is suggesting we do in Fig. 4(f).

The difference between Fig. 4(c) and Fig. 4(f) is that for Fig. 4(f) the pressure signal has been transformed to an approximated velocity signal prior to calculation of the spectrum, specifically to illustrate how the agreement with the -5/3 inertial subrange slope is slightly improved with inclusion of the filtering process. For this figure, a single spectrum has been isolated to maintain readabilty of the figure. However, given the request of the referee, we have added the corresponding background noise spectrum to the revised manuscript so that a reader can compare the magnitude of the noise to useful signal if so desired.

If the noise spectrum is well below the signal spectrum up to where the inertial cascade is fit, then this indicates that that spectrum should not be unduly influenced by noise, or the filtering method. But how representative is this one case? Otherwise, it is very difficult to believe that the resulting modified signals are reasonably free of artifacts, since the basic filtering approach does not make good sense.

Referring to Fig. 4(c) and, as should be evident throughout the manuscript, there are portions of the flight where there is little-to-no high frequency fluctuations in the air to measure (e.g. where there is no turbulence to measure). In these instances, there is no reason to expect the signal content to increase above the background noise spectrum (i.e. all the signal is in the DC portion of the signal, which is subtracted out when calculating the energy spectrum). In these cases, the time-domain signal can fully be expected to be corrupted by the noise subtraction process, but since there was no information contained within the high frequency content of the signal (only noise), this corruption would have little impact on the measurement results. Where there *is* useful information in the signal then that information is typically over an order of magnitude above the background noise spectrum.

Another approach would be to remove the spikes (if they are the main problem) by an outlier removal method in the time domain.

This would be a challenging approach to implement in practice. As illustrated in Fig. 4d, the noise is smaller than the signal fluctuations and most outlier detection schemes use the standard deviation of the signal to detect outliers. Hence, this process would require a significant amount of effort to implement correctly.

Lastly, the hypothesis that this noise is due to the (switching) power supply seems unlikely, given that these switch at hundreds of kilohertz, and, due to the filtering inherent in the technique, produce supply line noise primarily at the fundamental switching frequency. So, even in the presence of aliasing, this would not be expected to produce such spiky noise. Are there high-current loads in the avionics that have narrow pulses at about 1 Hz, e.g. telemetry transmissions?

Our thought was that the high-frequency switching noise might somehow be aliasing into the sampled lower frequencies. However, we are not confident in this assumption, as it would require the noise to bypass the anti-aliasing filters, which is why in the manuscript itself, we only refer to the noise being present in the power supply line since we measure the noise in this signal. We cannot attribute the noise source to the avionics, since it appears even when the system is outside the aircraft (e.g. as evident by comparing the noise spectrum from the environment chamber Fig. 4b to that measured during flight Fig. 4c).

We reiterate that the only reason we implement this background noise subtraction approach is due to the referee's original comment that background subtraction should be applied to the energy spectrum prior to fitting the -5/3 inertial subrange slope used for the dissipation rate calculation. However, such background subtraction is only truly possible in the pressure transducer spectra (e.g. Fig. 4c) since that is the only point in the process where the noise is easily identified. Once the pressure signals from the different transducers are convoluted with the directional calibration and aircraft kinematics, it is much harder to confidently discriminate between useful signal and noise.

Regardless its source, the actual influence of the background noise on the measurement results is quite small. Prior to its implementation we compared statistics with and without the filtering applied and found that it has very little impact on the derived statistics, with the most noticeable impact being reduced scatter in some derived quantities, such as the turbulent kinetic energy dissipation rate and Reynolds stresses. This can be expected, as such statistics (e.g. mean, variance, power spectra) will average out the phase distortion that will be introduced when the signal and noise have similar amplitude, while still reflecting the subtraction of the magnitude of the noise contribution to the signal.

However, we agree that this process is not ideal, and there will likely be phase distortions introduced into the time-dependent signal which may not be reflected in the derived statistics. We therefore have added additional text to the manuscript which acknowledges the likelihood of phase distortion being introduced through this process. If this is deemed an insufficient response to the referee's concerns, we can easily revert back to the statistics calculated from the unfiltered signal without any impact on the results and discussion within the body of the paper. However, in such a case we will not be able to address the referee's initial comment regarding background noise subtraction.

– 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.

This argument oversimplifies the wind estimation problem. The vector conversion of relative wind from body to inertial coordinates is equally sensitive to yaw errors at all yaw angles. It is the vector combination of relative wind and inertial velocity in the wind triangle that produces sensitivity variation on the orbit, since when the plane has the highest inertial velocity (flying downwind), its rate of change of attitude is greatest on the circle, and this can expose errors due to timing mismatches in the various sensors. The paper uses a complicated timing recovery scheme, since the data is not time stamped with a common reference, so this may be where to look for periodic wind estimate excursions. Do they occur around this point on the circle? Or do they occur where the orbit nears the jet? Or are they at 0 and 90 deg? (I can't tell from the paper).

The statement referenced above directly addresses the question as to why uncertainty in yaw can specifically lead to periodicity in the wind estimate and why it correlates to the yaw angle of the aircraft. It is also consistent with (van den Kroonenberg, A., Martin, T., Buschmann, M., Bange, J., and Vörsmann, P.: Measuring the Wind Vector Using the Autonomous Mini Aerial Vehicle M2AV, J. Atmos. Oceanic Technol., 25, 1969–1982, 2008) as noted in the manuscript. As noted in the statement, the sensitivity the referee is describing above is manifested as long-wavelengths in the wind estimate through the coordinate transformation, most notably the contribution from yaw, as we note in the text.

As to the influence of systematic errors, we are aware of these issues, and in the current study these were addressed by applying the referenced correction procedure of (Al-Ghussain, L. and Bailey, S. C. C.: An approach to minimize aircraft motion bias in multi-hole probe wind measurements made by small unmanned aerial systems, Atmospheric Measurement Techniques, 14, 173–184, https://doi.org/10.5194/amt-14-173-2021, 2021.) this optimization procedure is specifically designed to minimize the type of errors being referenced by identifying and corrects systematic bias in yaw, pitch, roll, as well as any residual timing mismatches and undetected airframe distortion. Use of this approach is already referenced in the manuscript. As for specific timing mismatches, recall that we have redundant measurement of the relative airspeed through the pitot probe (logged by the autopilot) and the five hole probe (logged by the on-board data acquisition system). These redundant measurements allow for precise intercomparison of the time lags between the two systems, which was examined carefully during our initial vetting of the results prior to preparation of the original submission. The importance of timing mismatches was also brought up in review several revisions ago and, as noted in the response to that comment, we detected a very slight mismatch in the system clock rates and re-processed the data accordingly. Note that correcting this mismatch did not have a noticable impact on the measured wind profiles, indicating that it was not a significant source of error.

Of course, it is impossible to remove all systematic bias error, and therefore we also note that our error estimation includes the effects of bias error in Equation 11 (and as discussed in the corresponding discussion) through the term E_B. This term is intended to characterize the impact of unaccounted for bias error.

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.

I think you mean Figure 5.

Yes, you are correct, the response was written before adding the figure addressing the electrical noise filter. It's inclusion shifted the figure numbers resulting in the aforementioned typo.

Your uncertainty analysis only models the random forms of error, not the systematic ones such as that mentioned in the response above. So it is premature to say that these are covered by the doubling mentioned in the text.

As mentioned above, the reviewer is incorrect and the uncertainty analysis does indeed include the systematic (i.e. bias) error. Of course, the magnitude of the undetected bias errors are truly unknowable, however we believe that we have exercised all due diligence to eliminate the known biases, and estimate the uncertainty introduced by the unknown biases on the wind estimates.

The real test would be to plot the winds as a function of azimuth on the circle, as well as noting the mean wind direction on each circle and the location of the jet, to see if aircraft motions are correlated with the measurements.

Indeed, and during our vetting of the data prior to preparing the original submission (and at many points between then and now) we have indeed examined the results in such a manner to ensure that we were not misrepresenting the results. We include one such figure below as figure 1, which illustrates that the wind excursions/long wavelengths which the referee is referencing are not dependent on the yaw angle and vary in spatial location with altitude. We have not included this figure in the revised manuscript as it implies that these wind excursions are due to spatial heterogeneity. Discussion of any observations of spatial heterogeneity in previous versions of this manuscript was not well received by another reviewer. Given that our only evidence that the spatial heterogeneity was present are the results from the aircraft measurements we cannot provide any evidence that such spatial heterogeneity was present during our measurements.

We do note that the Flight 1 results in figure 1 does show lower winds at $\pm 90°$, which may be a signuature of systematic bias having an influence on the estimate due to the high winds present during that flight. We acknowledge in the manuscript that the wind excursions observed in the profiles, particularly for this flight, could be due to these types of errors.

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).

More rigorous, yes, given the assumptions, but perhaps leaving out the largest effects!

As noted above, we believe that we have exercised all due diligence to identify and eliminate the effect of biases, and to estimate the uncertainty of the unknown biases, on the wind estimates. We also acknowledge in the text that the wind excursions observed in the profiles could be due to these types of errors. Given this, it is not clear what additional revisions we can make to the manuscript to address the reviewers concerns.

However, should the editor feel that a version of figure 1 is important enough to include in the manuscript, we are happy to include it, but have not added it to the revised manuscript for the reasons noted above.

– 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!

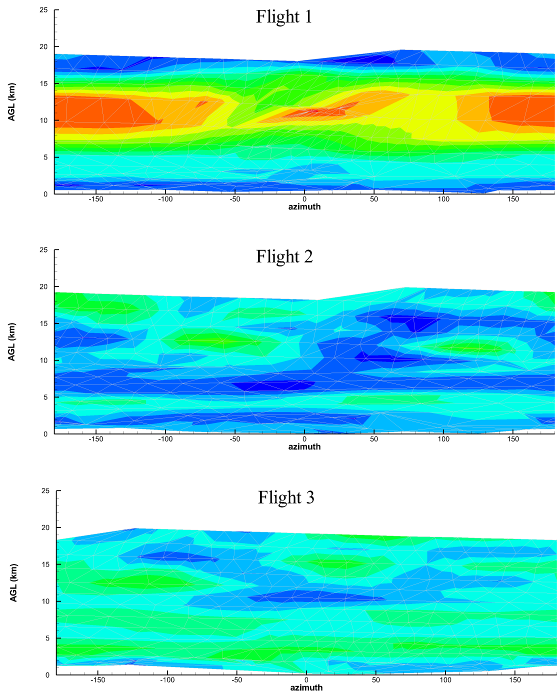

Figure 1. Isocontours of wind magnitude for each flight shown as a function of azimuthal angle of aircraft position relative to the center of the aircraft's flight path. Also shown is the triangular mesh used for determination of the contours. The isocontours illustrate that the majority of the long wavelength periodicity shown in the profiles is not correlated to aircraft yaw angle.

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.

"Promote horizontal homogeneity" does not mean that the layer structures are perfectly horizontal and homogeneous over arbitrary distances. Does this extend over hundreds of m, km, tens of km? Gravity wave activity, for one, can upset this ideal situation, as intimated later from the data in lines 537-539, possibly contradicting this assumption. Also, nearer the surface, does the orbit radius decrease faster than the scale of horizontal variations? Basically, there is still no quantitative support for this assumption, and it should not be treated as common or obvious.

We agree with the referee that there are many potential sources of horizontal heterogeneity, and as evident in Figure 1 above, this assumption is not strictly supported by our measurements. It's inclusion was intended to satisfy concerns of another referee regarding the Richardson number estimation.

Upon reflection, we now realize that the assumption referenced above is neither justified, nor necessary, and therefore the simplest way to address the referee's concerns is to remove this statement from the revised manuscript.

– 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.

The only "quantification" I see is the statement "that the periodicity is shorter than the pitch of the helical flight path". How much shorter? Does this hold for all orbits?

As the period/wavelength of these waves is altitude and flight dependent, their quantification can only be provided in an altitude-dependent analysis (e.g. as done with the wavelet analysis presented later). the statement we introduced at this point in the manuscript was intended to specifically address the referee's concern that the periodicity was correlated to the pitch of the helix.

In the revised manuscript, we have added a statement at this point in the manuscript referencing the wavelet analysis that appears later. The wavelet analysis allows the reader to see the energy content of the wind as a function of altitude and wavelength.

- 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.

I thought the explanation above was fine. But why make the detail vague by leaving it out?

As we mention in the original response, including the detailed exposition above would distract from this point, while adding little value to the overall intent of its inclusion. There are approaches which could have been used that achieve the same outcome and hence the exposition provided above is unnecessary for repeatability or understanding of the paper. We thought this addressed the referee's 'why?' question above.

As we recognize the importance of precision in scientific writing, and that the referee has identified this as a point where such precision is necessary, we have added the above description to the revised manuscript.

- 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.

A frequency response (amplitude curve as a function of frequency) does not have a "maximum frequency of response". This is a smoothly varying function with no lower bound, so some standard point on this curve is picked to indicate "bandwidth" of response. This is typically the -3dB roll-off frequency. This bandwidth (particularly if you say 3dB bandwidth) would be widely understood, whereas your "maximum frequency response" could be confused with other things, such as the frequency where the response is maximum.

We had incorrectly made the assumption that our use of the phrase 'maximum frequency response' implied that we were referring to the -3dB roll-off frequency. We have added a clarification in the revised manuscript at the location where we first use this term to indicate what it is referring to.

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}.'

The frequency where the noise exceeds the signal is entirely dependent on the turbulent energy in the flow, so no trend (say with altitude) would be expected.

The phrase being referenced above was specifically introduced to answer the referee's original question 'How did this vary above the boundary layer?' However, on re-reading the original comment, we think that it's possible that the referee was referring to how the maximum frequency response varied above the boundary layer, instead of how the filter frequency varied above the boundary layer? If the former then, as already described in the manuscript, we only conducted frequency response measurement tests at ambient conditions and in an environment chamber at conditions close to those experienced in the stratosphere. Given that the maximum frequency response change between these two conditions was very similar (20 Hz and 10 Hz respectively) we can only infer that the maximum frequency response variation was small enough that it can be assumed constant above the boundary layer.

- 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.

Lines 713-715: garbled revised sentence.

We have revised the sentence spanning Lines 713-715.

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 comment made by the referee on this version of the manuscript.

I suggest that the authors amend the title of the article, for better clarity, to "High-altitude atmospheric turbulence measurements and qualitative infrasound observations using a balloon-launched small uncrewed aircraft system"

We have updated the title following this suggestion. However, we note that while the comparison between infrasonic energy and turbulent kinetic energy was qualitative, the infrasound measurements themselves were not qualitative. We therefore have modified the suggestion to 'High-altitude balloon-launched uncrewed aircraft system measurements of atmospheric turbulence and qualitative comparison with infrasound microphone response' which we believe better captures the intent of the suggested title change.