Reply to Reviewer #1 (Author's response is in red font)

General comments

The manuscript uses aircraft measurements from a recent field campaign (PERCUSION) in the region of the Atlantic ITCZ to characterise the turbulence observed at flight altitudes of mostly 13–15 km. The key hypothesis tested is that the doldrums — regions of low surface wind speed in the deep tropics — are associated with decreased upper-tropospheric turbulence. The authors show some indication that this is the case, but also note that the relationship appears necessary rather than sufficient: turbulence is reduced when surface winds are weak, but reduced turbulence (or column-integrated divergence within the ITCZ) does not imply weak surface winds. In addition, the paper characterises the turbulence during the campaign more generally and assesses the suitability of 10 Hz vs. 100 Hz measurements for turbulence analysis, drawing on two other aircraft campaigns (EUREC4A and SOUTHTRAC).

This is an interesting and, as far as I am aware, novel research question, and the data set (183 constant-altitude legs, 89 circles) is a valuable resource. My main concerns — detailed in the major comments below — relate to the scope and structure of the manuscript, to the choice of data sets used to characterise the atmospheric conditions, and to the statistical support for and interpretation of the key result. Overall, I think the paper has the makings of a solid contribution once the scope is sharpened and the analysis choices are better justified and more consistently applied. Major comments are listed below, followed by minor comments organised by section.

We would like to thank the reviewer for the careful review of the manuscript and the encouraging, positive assessment. We must also admit that the original structure of the manuscript was not easy to digest. For this reason, we have restructured the entire paper and also included the dropsonde data in the analysis.

Major comments

1. **Two papers in one.** The title, abstract, and first half of the introduction motivate the doldrums/turbulence question, while a second research strand — the characterisation of turbulence during PERCUSION together with the comparison against SOUTHTRAC and EUREC4A and the assessment of 10 Hz vs. 100 Hz measurements (based mostly on these two other campaigns) — is only loosely connected to it. The connection between the two strands is never made explicit, which makes the paper difficult to follow and obscures what the central result is. Please clarify the relationship between the two aims, or separate them: decide which story is the main one, restructure the manuscript around it, and either shorten or move to the supplement the material supporting the secondary story. If the cross-campaign comparison remains, it should be reflected in the title and abstract and carried through consistently (see the comment on Section 3 below, where PERCUSION itself is not actually compared against the other two campaigns).

OK, accepted. We agree and have sharpened the manuscript's focus on the doldrums hypothesis as a central, novel contribution, while retaining the characterization of turbulence as necessary supporting material rather than as a second, independent objective. We have decided not to remove the characterization entirely, as it illustrates both the reliability of the 10-Hz BAHAMAS data and the turbulence intensities derived from it, on which the doldrums analysis is based. However, we have now reduced this part to what is strictly needed, removed some specific paragraphs of the turbulence characterization, and moved the 10 Hz/100 Hz comparison for the EUREC^{4}A data to the Supplement.

Specifically, the following changes are made in the manuscript:

The literature review in the Introduction (**Section 1**) was shortened from six to three paragraphs, with each paragraph now explicitly linked to a subsequent part of the analysis.

Section 2.2 was streamlined: The regression analysis of the TKE–ε slope, as well as the supplementary material on length scales (Kolmogorov microscale, Ozmidov scale, buoyancy scale, Equations 4–7, and the corresponding figures of the Sections 3 and 4.2), were completely removed from the main text, as they are not directly used in the doldrums test in Section 5.

Section 2.3 was revised to more clearly highlight the difference between the ERA5 analyses and the high-resolution ECMWF IFS data set, composed of hourly operational analyses and short-range forecasts, for the doldrums analysis, see Major Comment 2.

Section 2.5 is new and presents the methodology of the dropsonde analysis from the BEACH dataset, see Major Comment 2.

Sections 3 (EUREC^{4}A validation) and **4.2** (PERCUSION turbulence characterization) were each reduced to a brief summary, retaining only the facts needed later in the manuscript (reliability of the 10-Hz data; the east-west asymmetry of turbulence mentioned in the conclusion). For EUREC^{4}A, only the overview of the flight segments, the power spectra, the spatial EDR/TKE distributions, and the PDFs of ε_w and are now in the Supplement (Table S1 and Figs. S1–S3).

Section 5 is now divided into four subsections where the new subsection 5.2 explicitly refers to the results of the dropsonde analysis.

Section 6 (Summary and Conclusion) was streamlined as well.

2. **Choice of data set for divergence/convergence.** The doldrums-related analysis is based on the circle legs, yet the divergence/convergence used to characterise the doldrums (VIMAD) is computed from ERA5 rather than from the observed circle products derived from the dropsonde measurements (e.g. via the regression method; see the BEACH data set). Since the aircraft observations are the whole basis of the paper, using the measured profiles of convergence and divergence seems the more natural — and more convincing — choice. Related to this, ERA5 vertical velocity and divergence are not generally considered reliable in this region; please either justify their use or discuss this limitation explicitly.

This is a valid point and we fully agree that this was the most important gap in the manuscript's analysis. The BEACH Level-4 dataset (Gloeckner et al., 2026) was not yet published at the time of the original submission; it has since become available, and we used it as an independent, purely observational cross-check of the operational IFS-based divergence signal. The first author acknowledges the help of Claude (Anthropic) in coding the data retrieval and data analysis in Python.

Changes in manuscript:

New Section 2.5 describes the BEACH dropsonde methodology for the 89 circles flown during

PERCUSION and for two independent doldrums criteria named as circle-mean and sonde-count.

New Section 5.3 presents circle-averaged vertical velocity profiles (Fig. 10) and a dropsonde-derived w_{TOP}-vs-$|u_{10m}|$ analogue to VIMAD (Fig. 11), both of which reproduce the IFS-based finding of a weaker, predominantly divergent overturning circulation above the doldrums. These are new and independent findings which are obtained entirely without the ECMWF's operational analyses. This cross-validation is now also summarized in the Abstract and Conclusion.

In addition, **Section 2.3** was clarified to explicitly distinguish the two different ECMWF IFS products used in the manuscript: the ERA5 reanalysis (~30 km horizontal resolution, 137 model levels) supplies N, S, Ri, and the cold-point tropopause height along the flight tracks, whereas $|u_{10m}|$, TCW, and VIMAD — the fields underlying the doldrums classification itself — are taken from the higher-resolution (~9 km) operational IFS analyses and short-range forecasts. This distinction was not made explicit in the original submission, where the text could be read as implying that all 2D fields, including VIMAD, came from the coarser resolved ERA5 product. We believe this contributed to the reviewer's original concern about ERA5's reliability for divergence in this region. We additionally updated the captions of Figs. 7–9 to explicitly read "Hourly *operational* IFS analyses interpolated...", so that the distinction is visible even to a reader who only scans the figures.

3. **Underuse of the straight-leg data.** Related to the above, it is unclear why only the circle legs are used to address the central question rather than all 183 constant-altitude legs. If the circle products are not used — though that would be the natural choice — the analysis could be extended to all legs, since mean surface conditions in the vicinity of the flight track can equally be inferred from ERA5 (e.g. by identifying a surrounding region and evaluating whether it satisfies the doldrums criterion). This would increase the sample size behind the paper's key result.

Another valid remark: We did not extend the doldrums analysis to straight flight segments. The cross-check based on dropsondes (Major Comment 2) is available only at the 89 circular flight points, since dropsondes were deployed exclusively during circular flight patterns. Extending the ECMWF's IFS-only analysis to straight legs was considered but would not benefit from the same independent observational cross-check that we consider the stronger response to this concern. We acknowledge this as a valid remaining limitation. In addition, the check of ECMWF's interpolated $|u_{10m}|$ on the 23 flight tracks reveals only very limited number periods (mostly very short) where $|u_{10m}| \leq 3$ m s^{-1}, the longest actually occurring on the transfer flight from Cape Verde to Barbados.

4. **Dissipation-rate/spectral analysis (Sects. 2.2, 3, 4).** A large fraction of the paper is devoted to the energy dissipation rate, the TKE–ε relationship, and the 10 Hz vs. 100 Hz comparison, but it is not clear how — or whether — this material is used to address the central hypothesis about doldrums and turbulence. If it is not needed for the main result, I suggest moving most of it to the supplement, retaining only what is strictly necessary for interpreting the doldrums analysis.

Agreed. We went beyond simply streamlining this material, we removed it.

The theoretical TKE–ε relationship, the corresponding regression analysis of the slope for both EUREC^{4}A and PERCUSION, and the supplementary material on length scales (Kolmogorov microscale, Ozmidov scale, buoyancy scale) no longer appear in either the manuscript or the Supplement. This quantitative characterization of turbulence is more detailed than is necessary for the present study, which focuses on the doldrums, and we now plan to present it in a separate manuscript dedicated to this topic.

Section 2.2 has been shortened (material on length scales and the TKE–ε regression has been removed) and it now contains only the definitions of TKE and ε_w (Equations 1–2), which are used directly in the doldrums test in **Section 5.4**.

Sections 3 and **4.2** have been condensed, see Major Comment 1.

5. **Statistical support for the key result.** The claim that TKE_{max} is reduced above the doldrums (Fig. 15) rests on a purely visual/qualitative comparison. It would benefit from a more formal statistical treatment — e.g. a test of whether the low TKE_{max} values associated with doldrums conditions are unlikely to be drawn from the same distribution as those associated with $|u_{10m}| > 3$ m s^{-1}.

We agree and have added a one-sided Mann–Whitney U test, which does not assume equal group sizes or normality. The results of this test was incorporated in the manuscript (**Section 5.4**): "...a one-sided Mann–Whitney U test was applied, as turbulence-related quantities in stratified media are generally not normally distributed but strongly right-skewed (cf. the ε_w and EDR_w distributions in Fig. 5); this test confirms that TKE_{MAX} is significantly lower over the doldrums (median = 0.16 m^2 s^{-2}, n = 13) than over the remaining circles (median = 0.78 m^2 s^{-2}, n = 76; Mann–Whitney U test, U = 205, $p < 0.001$)."

This result is now also stated in the Abstract and Conclusion.

6. **Interpretation of the main finding.** One of the more interesting outcomes is that the original hypothesis — reduced turbulence due to a thermally more stable upper troposphere — is not supported (N is in fact lowest over the doldrums), and that the reduction instead appears related to weaker upper-tropospheric shear. This is a nice result, but it is under-discussed. The conclusions speculate that the reduced shear might itself be linked to reduced convective activity; could this be tested directly, for example with an analysis analogous to that of Fig. 11?

We agree that this is an important, somewhat counterintuitive result, and we now explicitly highlight it rather than merely mentioning it in passing. We did not additionally investigate whether the reduced shear itself correlates with the distance from convection. This task remains an open question, probably, better answered by using high-resolution numerical simulations instead of data from operational NWP models. Nevertheless, the argument regarding the east-west asymmetry of convection (which is already included in the manuscript) suggests a qualitative, not quantitative, link.

Abstract and Conclusion now state explicitly: "...indicating that reduced shear, rather than enhanced static stability, is the more likely proximate driver of the reduced turbulence above the doldrums."

Minor comments

Title/Abstract

– The title poses a single, specific question, but the abstract (and later the paper) covers noticeably more ground (see Major comment 1).

- It would help to state the underlying hypothesis briefly in the abstract — i.e. why subsidence over the doldrums might be expected to reduce turbulence; the explanation currently appears only in the main text.
- It is not clear from the abstract why the doldrums analysis is restricted to the circle legs rather than making use of all 183 constant-altitude legs (see Major comment 3).

The abstract has been revised (see Major Comments 1, 2, and 5); the underlying hypothesis regarding subsidence above doldrums was already stated in the first sentence of the abstract referring to doldrums. The analysis of the 89 circular flight legs to analyze doldrums is mentioned explicitly now.

Section 1 (Introduction)

- The introduction falls into two parts that are not well connected: the first on the doldrums, the second largely a literature survey of turbulence measurements from various campaigns and altitudes.
- The second part is long and describes a number of individual studies in detail without always making clear why each is relevant to the present question. I recommend condensing it considerably and focusing on what is needed to motivate the analysis that follows.
- EUREC4A is mentioned only once in the introduction, yet later becomes a central point of comparison (Sect. 3).

We have shortened the literature review from six paragraphs to three, and the content is now explicitly linked to the subsequent analysis (Major Comment 1). Although EUREC^{4}A is still introduced once in Section 1, its role is now clearer, since Section 3 is now explicitly defined as a data quality review using the EUREC^{4}A flights.

Section 2.1

- Lines 140–142 already state, citing Appendix B of Dörnbrack et al. (2022), that the 10 Hz and 100 Hz spectra are very similar within the 0.4–4 Hz analysis interval used to determine ε. Given this, the added value of the full 10 Hz vs. 100 Hz comparison in Sect. 3 is unclear; this supports moving that material to the supplement (see Major comment 4 and the comments on Sect. 3 below).

As stated above, we support the decision to condense Section 3 and have moved the key findings to the Supplement, see Major Comments No. 1 and 4.

Section 2.2

- It is not clear why the energy dissipation rate is needed for the paper's central (doldrums) question. Several length scales are introduced in connection with the theoretical TKE–ε relationship (integral, Kolmogorov, Ozmidov, buoyancy), but it is not clear how they are subsequently used or compared.
- The distinction drawn between TKE and dissipation rate as having "slightly different meanings" (l. 189 ff.) could be clarified further — do the authors mean that one integrates over the full spectrum while the other reflects only part of it?

This issue was resolved more quickly than originally planned. Instead of clarifying the details regarding the length scales, we removed them from the main text (Equations 4–7 and the corresponding results), since they are not used in the doldrums test. Section 2.2 now covers only what is necessary for Section 5.4.

Regarding the second point, we clarified the text which is now in line 147: "The two quantities TKE and ε_w, determined from the aircraft measurements, have slightly different meanings. While ε_w represents the rate of energy transported downscale in the energy cascade, TKE is a measure of the total intensity of turbulent eddies in the segments under consideration. TKE is the half sum of all the velocity variances both in physical and spectral space. In contrast, ε_w is proportional to S_w only in the selected and limited frequency (wavenumber) range and its estimate assumes a $k^{-5/3}$ spectral slope, Eq. 2."

Section 2.3

- Please clarify why only one of the data sets collected during PERCUSION is used, given that the circle products derived from the dropsondes would also provide the relevant information (see Major comment 2).
- Column water vapour (CWV), rather than total column water (TCW), is the more standard quantity for defining the ITCZ. Please clarify the choice and its consequences.
- If ERA5 is used for these diagnostics, how is the associated uncertainty accounted for?

In the revised version, we added the BEACH data set to cross-check the ECMWF IFS analysis, see Major Comment 2. The use of TCW remained unchanged and is consistent with earlier ITCZ-threshold convention in the manuscript. In an early preparation phase of the ECMWF analysis, CWV fields were plotted and compared with TCW and no significant differences in identifying the ITCZ were identified for the PERCUSION flights. The uncertainty of the IFS analyses is now addressed indirectly via the independent confirmation by the BEACH dataset.

Section 3 (EUREC4A)

- Since this section concerns a different campaign and motivates a data-quality point (validity of the 10 Hz data) rather than a result of the present study, I suggest moving it to the supplement (up to around l. 270).
- Two points if it is retained: (a) how might the different meteorological conditions sampled by PERCUSION (inside the ITCZ) vs. EUREC4A (outside) affect this assessment? (b) Can the agreement between ε at 10 Hz and 100 Hz be quantified (e.g. with a correlation coefficient) rather than left to visual comparison? I would have expected the flattening of the spectral slope to appear as a systematic bias — is that the case?
- From around l. 270 onwards, PERCUSION itself is not compared against the other two campaigns, even though characterising the PERCUSION turbulence is one of the paper's stated objectives. The purpose of the EUREC4A–SOUTHTRAC comparison at this point is unclear; please clarify or reconsider whether it is needed in the main text.

The essential, key content of the original Section 3 referring to EUREC^{4}A was moved to the

Supplement (see Major Comment 1). A quantitative correlation coefficient between 10 and 100 Hz EDR/TKE, and an explicit systematic-bias check of the spectral flattening, were still not added in this revision round as it requires more analysis of the raw BAHAMAS data which is currently not feasible. The ε_w distribution comparison across the campaigns SOUTHTRAC and EUREC^{4}A and PERCUSION is retained (now Fig. 5 in the main text) because it summarizes statistics which is used in Section 6 (Summary and Conclusions).

Section 4 (PERCUSION)

– The opening paragraph largely repeats the campaign description given earlier; consider condensing.

The opening paragraph was removed. Section 4 was condensed further, see Major Comments 1 and 4.

– The discussion of the ERA5 cold-point tropopause (l. 314 ff.) reads as a data-set description and would fit better in the data section than in the results.

The general statements about the ERA5 CPT were moved to Section 2 (Methodology).

– Line 341: what is the physical reason for the flattening of the spectra at higher frequencies?

Good point. We don't know and have to investigate this in more detail, see our answer to Section 3 above.

– Figure 8: the many overlapping lines make the individual contributions hard to distinguish.

This is Figure 4 in the revised manuscript. We have decided not to change the color scale. The figure is intended to provide a general overview of the geographic distribution of TKE and EDR_w. Essentially, it focuses on the overall distribution described in the text rather than on the behavior of individual flights.

Section 5 (Doldrums and upper-level turbulence)

– The number of doldrums circles is stated inconsistently: Fig. 12 and footnote 3 refer to 15 circles satisfying the strict criterion, yet only 13 circles have wind speeds less than 3m/s in Fig. 13.

– Please clarify what is meant by the "northern edge of the ITCZ".

– Figure 12 is missing a colour bar. The figure is interesting, but its connection to the paper's central question about upper-level turbulence could be made more explicit.

– The comparison with Klocke et al. (2017) is not fully convincing to me: since a vanishing meridional wind is itself a defining criterion for the ITCZ centre in their method, v_{10m} being close to zero there follows almost by construction — and it is referred to as the ITCZ centre there, rather than its northern edge. Please reconcile this with the framing used here.

- Line 408: could this be shown explicitly by restricting the Fig. 12 analysis to the western part of the domain?

We have identified the source of the discrepancy. The revised manuscript in fact reports four related but distinct circle counts, drawn from two different datasets and averaging conventions, which were not clearly distinguished from one another in the original submission but are made explicit in the revised text:

- **13 – ECMWF/IFS:** $|u_{10m}| \leq 3$ m s^{-1}, averaged over the square area enclosing each circle (used in Sect. 5.2, Fig. 8/9, the Abstract, and the Sect. 5.4 statistical test).
- **15 – ECMWF/IFS:** $|u_{10m}| \leq 3$ m s^{-1}, averaged along the circular flight track itself (Sect. 5.1, Fig. 7, footnote 3).
- **12 – BEACH dropsondes:** circle-mean criterion (Sect. 2.5) — closely matches the IFS area-averaged count of 13.
- **18 – BEACH dropsondes:** sonde-count criterion (Sect. 2.5): at least 5 individual dropsondes per circle recording ≤ 3 m s^{-1}, irrespective of the circle mean.

In **Section 5.1** (first mention of "15 circular flight segments") we added: "Note that this along-track average yields a slightly different circle count than the area-averaged criterion used in Sect. 5.2 (13 circles, Fig. 8), since the two quantities are computed over different spatial domains (flight track vs. enclosing square)."

In **Section 5.2** we replaced "This number is smaller than the 18 circles resulting from the average along the individual circular segments." with "This number is somewhat smaller than the 15 circles obtained when $|u_{10m}|$ is instead averaged along the circular flight track itself (Sect. 5.1, Fig. 7)."

In **Section 5.3** (first mention of the sonde-count criterion's 18 circles) we explicitly refer to the ECMWF IFS values by " … (a dropsonde-based count, distinct from the ECMWF-derived 13/15 discussed in Sect. 5.1–5.2)".

Section 6 (Summary and conclusions)

- The conclusions reflect the paper's dual focus noted above (Major comment 1) — the text moves between the doldrums question and the broader turbulence-characterisation/campaign-comparison material without always making the connection clear. I suggest revising this section once the overall scope of the paper has been settled.

The "Summary and Conclusion" section has been completely revised (see Major Comments 1, 2, 5, and 6). After a brief survey of the data set and the applied methods, it now begins with the findings on the doldrums, incorporates the statistical test and the BEACH cross-validation, and highlights the findings regarding the relationship between shear and stability.