**[ Responses to the Comment by the Anonymous Referee #4 ]**

>> We deeply appreciate the referee #4 for providing constructive comments. The manuscript is revised following the comments below.

This article investigates the applicability of Quick Access Recorder (QAR) in-situ aircraft data obtained at a 1 Hz sampling rate as a cost-effective alternative to derive energy dissipation rates (EDR). The authors compare 5 methods to estimate EDR using QAR data collected during 2012 from Korean-based air carriers (B737 and B777 aircraft). The study is overall well organized and written. Plus, the topic itself it is certainly of great value to the aviation industry and the turbulence research community at large. However, there are a number of major aspects that need to be addressed upon publication at the journal of Atmospheric Measurement Techniques. Given the magnitude of these issues, it may be more appropriate for the authors to consider resubmission in the future. Here below I provide both general and specific comments with the hope that these can help the authors get to a publication stage.

**General comments:**

1) Most of manuscript is devoted to the description of the different methods used to derive EDR. While this is of value, none of these approaches are firstly introduced herein, so there needs to be significantly more weighting in the analysis and quantification of the usability of QAR (see other comments below).

➔ Thank you very much for the good comment. We reflected the reviewer's general comments in the revised manuscript. Also, the abstract is also reshaped to emphasize the analyses on characteristics of the EDR estimates. [Page 1, Lines 7-27]

2) A main drawback of the current study is the lack of higher-rate data to use as reference in the estimation of EDR. Even if the 5 methods produce similar EDRs, all estimates can be biased. It is impossible to use such comparison in the lack of a higher-accuracy reference (e.g., 10 Hz data) to determine the adequacy of 1 Hz QAR-derived EDR estimates. Otherwise, one can only comment on the differences between the 5 methods, which has already been reported for most of them in the literature, and that is not the aim of the current manuscript [Page 3, lines 21-22: "The main purpose of this study is to examine the feasibility of various objective EDR estimations using the 1-Hz (coarser frequency) flight data for possible sources of atmospheric turbulence in cruising altitudes"]. If simultaneous retrievals of 10 Hz and QAR are not available,

1

➔ Thank you very much for the good comment. Unfortunately, we do not have any higher-frequency (e.g., 10-Hz) QAR data. Therefore, we applied our method to the 20-Hz wind data obtained from the Boseong Meteorological Observatory (BMO), South Korea for 10 days (between 0000 LST 1 October and 0000 LST 10 October 2019). For a direct comparison, raw 20-Hz wind data are subsampled to the 1-Hz wind data using Reynolds averaging. It is noted that Kim et al. (2021) derived the energy dissipation rate (EDR) using the BMO high-frequency sonic data. In general, 1-Hz data shows similar distribution with the 20-Hz data although some peaks revealed in the 20-Hz wind data are smoothed in the 1-Hz wind data (Fig. A1). In the energy spectrum of streamwise component of 20-Hz and 1-Hz BMO wind (U-wind) data at 300 m above ground level at a certain time (Fig. A2), both spectra show a similar distribution, following a theoretical slope (-5/3) especially in the overlapped range of frequency (0.1-0.5 s$^{-1}$). Although two spectra have minor discrepancy in the overlapped frequency range of 0.1-0.5 s$^{-1}$, we confirmed that 1-Hz data have quite similar energy spectra with higher-rate data.
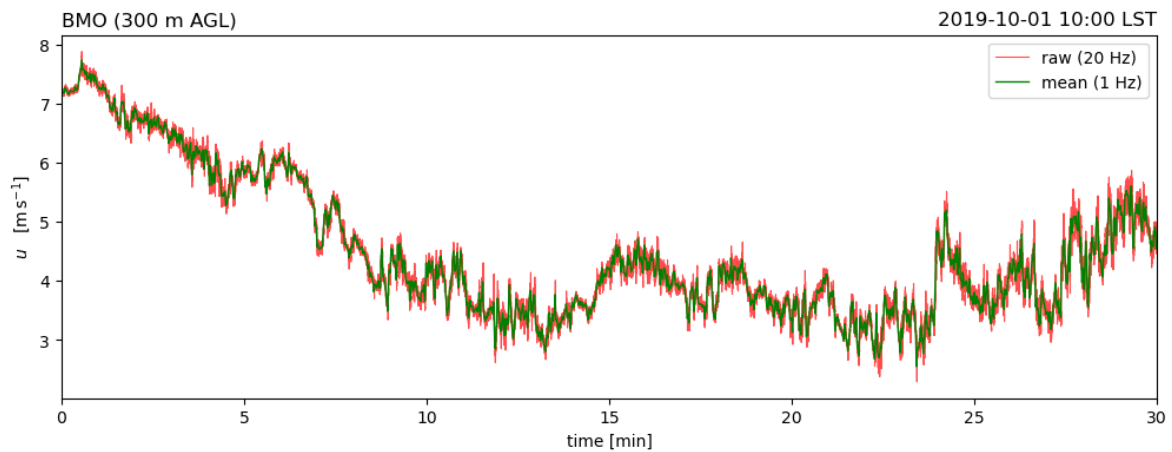


**Figure A1.** Timeseries of streamwise component of wind data (U) obtained from the BMO between 1000 LST and 1030 LST 1 October 2019 at 300 m AGL. Red and green lines indicate raw 20-Hz wind data and mean 1-Hz wind data, respectively.
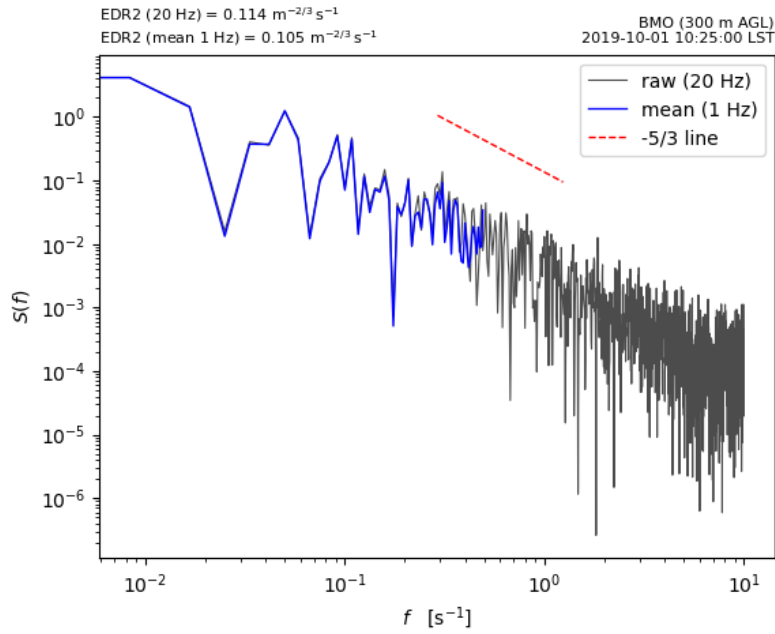
**Figure A2.** The energy spectrum of U-wind obtained from raw 20-Hz (black) and subsampled 1-Hz (blue) BMO data at 1025 LST 1 October 2019 at 300 m AGL. The dashed line represents the theoretical slope ($f^{5/3}$) in the frequency domain.

We also calculated the EDR based on three inertial dissipation method (EDR1, EDR2, and EDR3) using the zonal wind of 20-Hz and 1-Hz BMO data for 10 days (between 0000 LST 1 October and 0000 LST 10 October 2019). When the timeseries of EDRs calculated using two datasets are examined, variations in the EDRs are quite similar in 20-Hz and 1-Hz BMO data although there are some underestimation or overestimation in the magnitude of the EDR (not shown). In scatter density plots of 20-Hz EDR and 1-Hz EDRs from the BMO datasets (Fig. A3), it is found that 1-Hz EDR data is highly correlated with 20-Hz EDR data (r > 0.92), implying that the 1-Hz data can provide a reliable information (timing and location) of atmospheric turbulence. Based on this result, we can confirm that the use of 1-Hz flight data to estimate the EDR is reliable, although the magnitude of the EDR is somewhat underestimated in some cases. Intercomparison between 1-Hz and higher frequency aircraft data will also be conducted in the future. The abovementioned statements are included and Figs. A2 and A3 are included as a new figure (Figure 5) in the revised manuscript. [Page 8, Lines 14-28]

During the revision process, the statements on a comparison of EDR climatology between the higher-rate (e.g., 10-Hz or 20-Hz) data and 1-Hz QAR data are deleted in the revised manuscript. In the revised manuscript, we include the statements that 1-Hz QAR data can be an additional source for measuring EDR when the high-frequency in-situ aircraft measurement is not available and this can contribute to expand more EDR information at the cruising

altitudes in the world. [Page 3, Lines 20-24; Page 15, Lines 17-19]
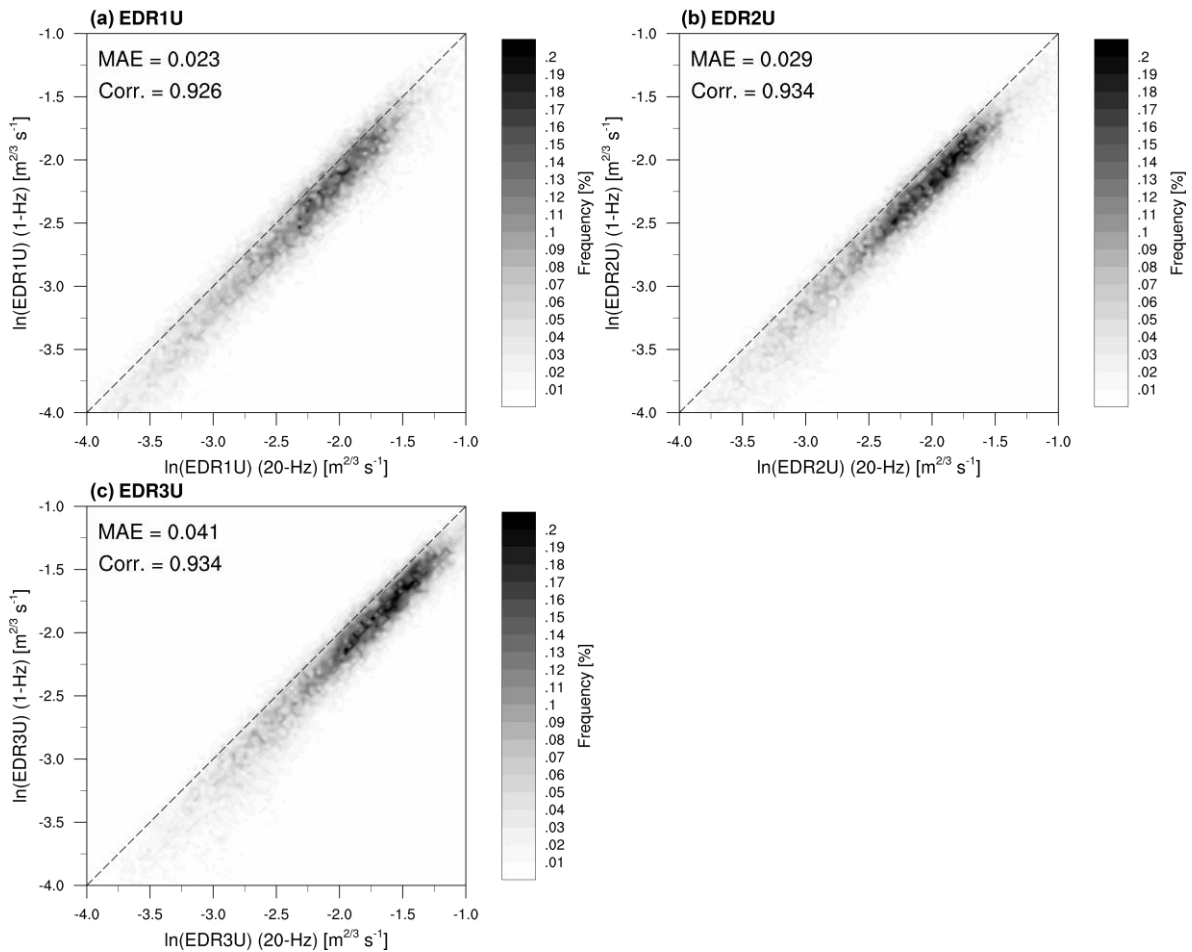


**Figure A3.** Scatter density plots of the EDRs calculated using the 20-Hz wind data (x-axis) and the 1-Hz wind data (y-axis) obtained from the 10-day BMO data. Pearson correlation and MAE are given in the top-left corner of each panel.

3) The use of a climatological log-normal distribution as reference may not be adequate in this case neither is supported by the presented results. There are reasons why the PDFs of EDR do not behave as a log-normal distribution. These can be related to the limited data availability (seasonal effects) and/or geographical considerations. To me, a higher probability of null-light turbulence that decreases as EDR increases makes good sense, as turbulence encounters are a rare event in a climatological sense. Therefore, the assumption that the QAR-derived PDF of EDR has to be log-normal is not necessarily a good way to assess the performance of this lower-frequency data.

➔ Thank you very much for pointing out such an important issue, which we did not properly mention in the original manuscript. We did not intend that the lognormality is used for

4

demonstrating the reliability of data in the original manuscript. Instead, we intended to report characteristics revealed in the QAR-derived probability density functions (PDFs) in the original manuscript. To avoid any confusion, the statement is modified in the revised manuscript. [Page 10, Lines 32-33; Page 11, Lines 1-2; Page 15, Lines 12-13]


3-1) Page 10, Lines 5-6: This is questionable. In the majority of cases there is not a reasonable fit for the EDR values at the peak and below, so I do not think these types of distributions are log-normal. Also, the higher probability bins should have a larger impact on the fit, so the fact that some of these needed to be excluded appears to be the result of a clear departure from a log-normal distribution. Maybe there is some issue with the particular optimization method selected (Levenberg-Marquardt) or its implementation by the authors. Perhaps trying another least-square minimization algorithm will help understanding where the actual problem comes from. Some level of discrepancy among the PDF and the best is to be expected, specially toward the tails of the distribution, but not for the area of larger probability if the selected reference is good descriptor of the observed quantity.

➔ Thank you for the good comment. During the revision process, we conducted lognormal fit obtained using other optimization methods [Powell's method (Press et al. 1992) and Nelder-Mead method (Gao and Han 2012)] to the EDR (Figures A4b and A4c, respectively). Compared to the Levenberg-Marquardt method used in the original manuscript (Fig. A4a), mean and standard deviation of the natural logarithm of the EDR (EDR1V) have similar value depending on each optimization method ($C_1$ and $C_2$ in Fig. A4, respectively) and this can also be found in all five EDR estimation methods (not shown). In this regard, it is considered that the choice of minimization algorithm will not have a significant impact on the resultant lognormal fit. These statements are included in the revised manuscript. [Page 10, Lines 30-32]
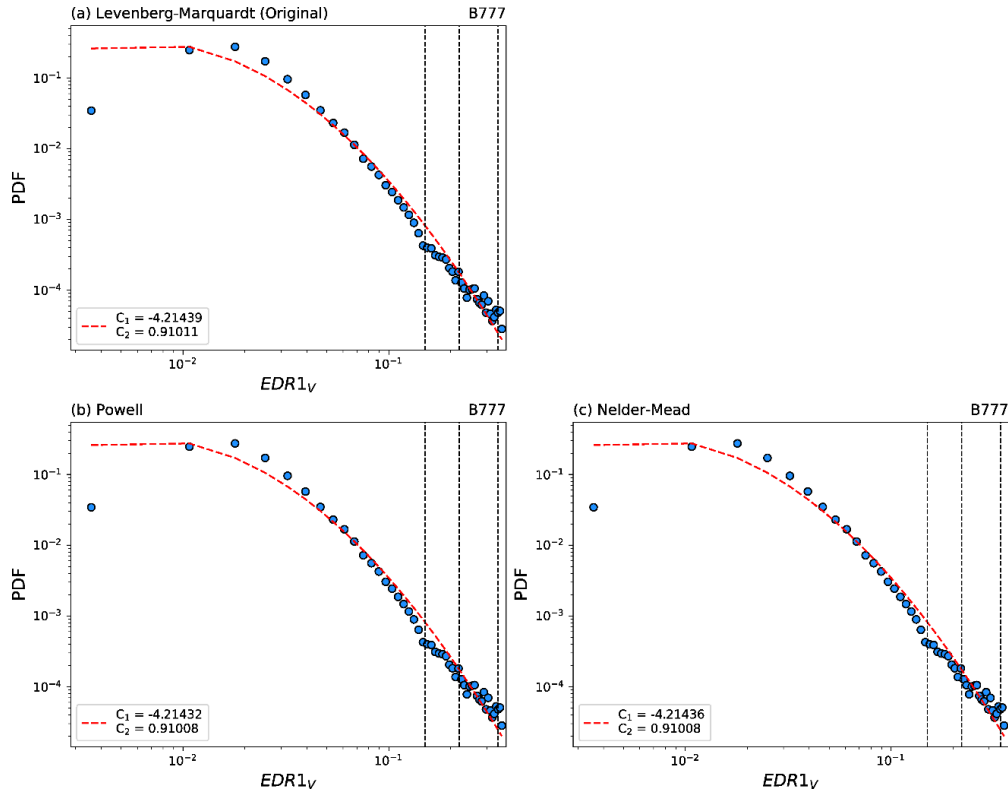
**Figure A4.** The PDFs (circles) of the EDR calculated using meridional wind (EDR1V) and lognormal fit (dashed line) over the EDRs obtained from the B777 data. The filled circles indicate data that were used in the fit, and the open circles indicate data that are excluded from the fit. The lognormal fitting is conducted via (a) Levenberg-Marquardt algorithm (used in the original manuscript), (b) Powell's method, and (c) Nelder-Mead method. The mean and standard deviation of the natural logarithm are written in each figure ($C_1$ and $C_2$, respectively).

3-2) Page 10, Lines 11-14: This claim is not supported by the data. I agree that the interest is more on larger EDR values, but the presented results do not justify the claim of a 'reasonable' log-normal distribution. This seems to be a key point of the manuscript to prove the validity of the QAR-derived EDR estimates, which makes me wonder about the value of these study lacking a reference (ideally 10 Hz data) to compare to.

➔ Thank you for pointing out this. As responded above, we included the validity of the 1-Hz EDR estimates using the BMO data in the revised manuscript [Page 8, Lines 14-28]. The statement related to the lognormal distribution is modified in the revised manuscript. [Page 10, Lines 32-33; Page 11, Lines 1-2]

3-3) Page 10, Line 18: Having similar statistics (mean and std) does not guarantee any skill in any of them (see general comment #1).

➔ Thank you for the comment. As responded in your comment 3), we did not intend that the

lognormality is used for demonstrating the reliability of data in the original manuscript. The statements related to the PDFs are modified in the revised manuscript [Page 10, Lines 32-33; Page 11, Lines 1-2]. Also, as responded above, we applied our method to the 20-Hz BMO data and subsampled 1-Hz BMO data and it is shown that 1-Hz EDR data is highly correlated with 20-Hz EDR data, although the magnitude of the EDR can be underestimated in some cases. These statements are also included in the revised manuscript. [Page 8, Lines 14-28].

4) The 'Results: Case analysis' section is highly subjective and does not allow to quantitatively assess the QAR-based EDR estimates. These are interesting cases, including the discussion of the associated potential large-scale mechanisms leading to turbulence. However, without a reference EDR to compare to, they do not add much to the verification and establishment of QAR data as a reliable source of data to derive EDR, neither to the actual strengths/downsides of each of the 5 exercised approaches to estimate EDR.

➔ Thank you for the good comment. As responded above, we indirectly tested the validity of the 1-Hz EDR by applying the same methods to both higher frequency (20-Hz) raw data and subsampled 1-Hz data from the BMO data. Statements related to this are included in the revised manuscript [Page 8, Lines 14-28]. Based on this result, it can be confirmed that the use of 1-Hz flight data to estimate the EDR is comparable. Therefore, we kept 'Results: Case analysis' section in the revised manuscript.

5) There are some details regarding the application of some of the techniques used to derive EDR that need attention, including the use of B737 data at 0.25 Hz. See the specific comments down below.

➔ Thank you very much for the good comments and suggestions. We addressed the specific comments below. Regarding to the original statement of the use of Boeing (B) 737 data, both the B737 and B777 data are recorded every second. To avoid any confusion, we modified the original sentence in the revised manuscript. [Page 4, Lines 4-5]

**Specific comments:**
1) Page 1, Lines 18-20: This statement is not supported by the results (See general comment #3).

➔ During the revision process, this statement is deleted.

2) Page 1, Lines 20-22: Or maybe inappropriate range of the spectrum being captured. This is just a speculation that needs further support.

➔ During the revision process, this statement is deleted.

3) Page 2, Line 24: What does the term 'off-line' refers to here? Post-flight? Please explain.

➔ The term 'off-line' in the original manuscript is changed into 'post-flight' in the revised manuscript. [Page 2, Line 26]

4) Page 3, Lines 13-14: What are the advantages/disadvantages of using QAR versus EHS and ADS-B? It would be good to put QAR in the context of data quality, availability and access compared to these other 1Hz sources, as well as elaborate on the fundamental principles and assumptions used to derive EDR in each case. Maybe this discussion is more appropriate in the methods section.

➔ We are sorry if the statement of the original manuscript brought a confusion. The main purpose of this study is not comparing the QAR data with Mode-S Enhanced Surveillance (Mode-S EHS) and Automatic Dependent Surveillance-Broadcast (ADS-B) data, but trying to find out an additional source of EDR estimations by applying all possible EDR methods to the currently available sources of low-frequency data (1-Hz QAR data in this study). To avoid any confusion, the sentence is modified in the revised manuscript. [Page 3, Lines 24-25]

5) Page 3, Lines 26-27: I would suggest removing this sentence from here since it is mentioned in the methods section.

➔ The sentence is deleted as suggested.

6) Page 3, Line 30: "MOG-level turbulent cases". Based on what? It should be an independent method… PIREPs?

➔ The moderate-or-greater (MOG)-level turbulence cases were determined based on derived equivalent vertical gust (DEVG $\geq 4.5$ m s$^{-1}$, Gill 2014). The statement is included in the revised manuscript. [Page 3, Line 31]

7) Page 6, Line 22: This may be a bit too large of a scale for the aircraft to feel since wind span is less than 100 m (~50 m).

➔ Thank you for the comment. Yes, we partly agree with your opinion. Given the low-

frequency (1-Hz) data, we chose predefined inertial range that has the minimum size of horizontal wavelengths. In addition to that, we found the inertial range that minimizes errors between theoretical slope and observed slope. This statement is included in the revised manuscript. [Page 6, Lines 23-26]

8) Page 6, Line 28: Should mention that the range is 4-5s for the B737.

➔ We are sorry for if the statement of the original manuscript brought a confusion. Given that the recording frequency of the B737 data is also 1-Hz, the same inertial range is applied. The related statement is modified in the revised manuscript. [Page 4, Lines 4-5]

9) Page 7, Line 1: How about the 0.25 Hz?

➔ We are sorry again for if the statement of the original manuscript brought a confusion. The recording frequency of the B737 data is also 1-Hz. The related statement is modified in the revised manuscript. [Page 4, Lines 4-5]

10) Page 7, Line 10: "no window". This seems contradictory. The next sentence indicates that a window has been applied. Do the authors mean 'no windowing'? If so, a 2-min FFT can be rather noisy (as clearly evident Fig. 4), tending to over-estimate EDR (e.g., Munoz-Esparza et al. MWR2018 in the references). The same would apply to method 3 (von Karman spectrum). Please comment on this.

➔ Thank you for pointing out this. We conducted a fast Fourier transform using a Welch window with no overlap. The original term "no window" is changed into "no overlap" in the revised manuscript. [Page 7, Line 12]

11) Page 7, Lines 16-17: The von Karman model is meant to capture larger scales, so it seems like the authors would need to consider a wider frequency range than 0.2-0.5 Hz.

➔ Thank you for the good comment. As responded to the other reviewer's major comment, we additionally conducted the sensitivity test on the inertial range. When the EDRs with the fixed inertial range (EDR1, EDR2, and EDR3) and those with dynamically selected range (EDR1-opt, EDR2-opt, and EDR3-opt) that minimizes the discrepancy between the theoretical and observed power laws within a given range are computed, it is found that there exist high correlations ($r > 0.97$ for the B777) and low MAEs (between 0.001-0.003 $m^{2/3}$ $s^{-1}$ for B777) between EDRs using the fixed range and dynamically selected range (Figure A5). In the present

9

study, the fixed inertial range is considered, regardless of an underestimation in the magnitude of some EDRs (e.g., EDR1U and EDR1V), as it can be more computationally efficient in calculating the EDR. Further investigation, however, may be required using more and longer data in the future. This statement is included and Fig. A5 is included as a new figure (Figure 4) in the revised manuscript. [Page 8, Lines 4-13]
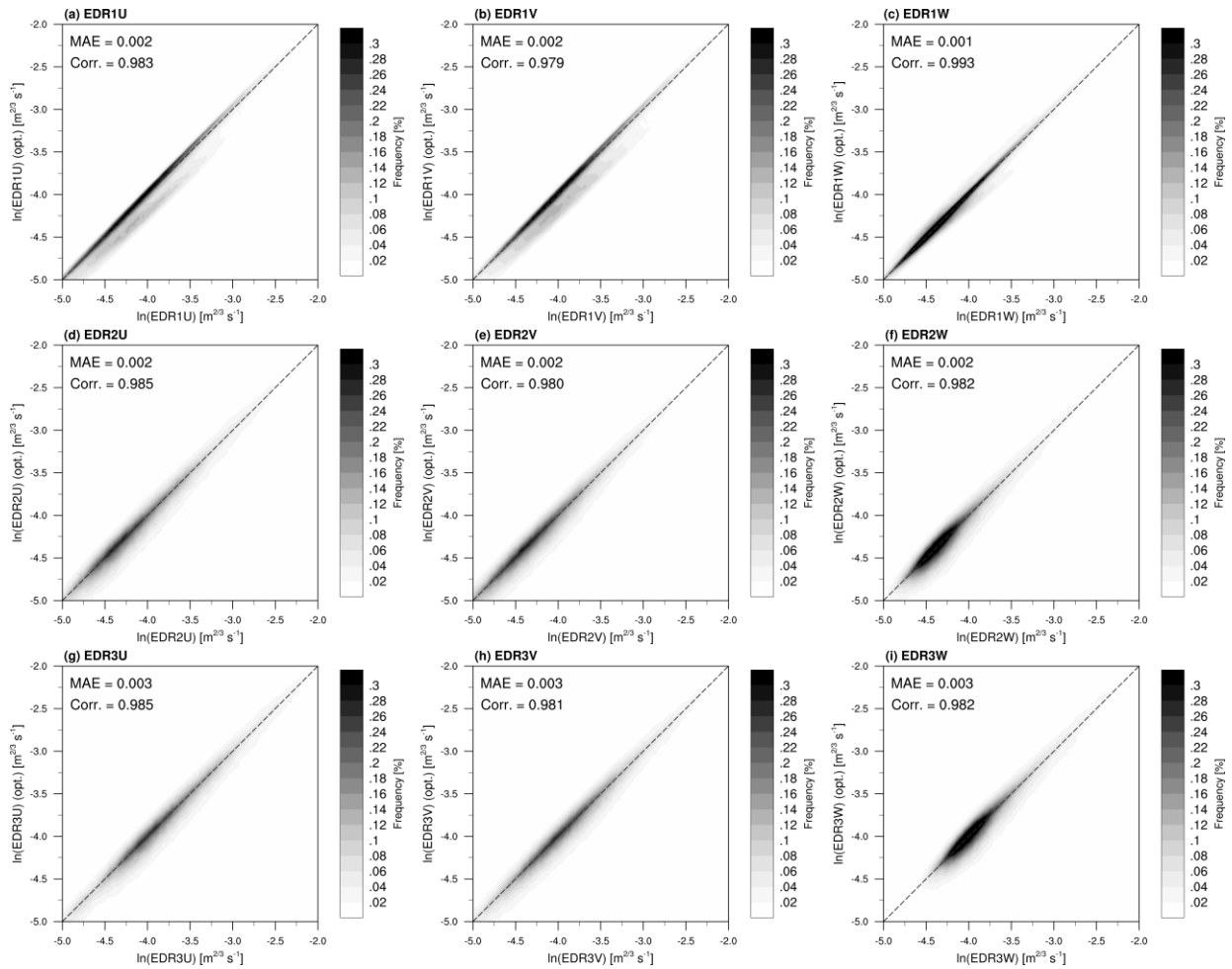


**Figure A5.** Scatter density plots of EDRs using the fixed and dynamical inertial ranges for the B777 data. Pearson correlation and MAE are given in the top-left corner of each panel.

12) Figure 4: The labels on Fig. 4 are confusing. The two models can be -5/3 since the lines are not parallel in the plot. Also, it would be good to add what the actual EDR values are in the figure for the readers benefit. Also, the B737 will have a Nyquist frequency of 0.125 Hz, which is out of the time interval for the -5/3 fit that the authors propose (2 – 5 s).

→ Figure 4 is modified to Figure 6 of the revised manuscript, as suggested.

13) Page 8, Lines 4-5: Not sure this is reasonable to state without further proof.

10

➔ The sentence is deleted during the revision process.

14) Page 9, Lines 9-12: These equations are redundant, as they are the exact same as in Eq. 11. Please remove.

➔ The equations are deleted as suggested.

15) Figure 5: I recommend the plot the x-axis (DEVG) in log scale, since that will help evaluate the nature of the fits and PDFs better. Also, it seems only half of the distribution is plotted. The point of the log-normal distribution is that there are instances to the left of the mean of the distribution (lower values), which are symmetric to the larger values (see Sharman and Pearson 2017, their Fig). However, this shows the peak occurs at the lowest DEVG values. It seems like there may be better distributions to use that the log-normal for DEVG, which may potentially lead to a better fit.

➔ The x-axis of Fig. 7 is changed into the log scale in the revised manuscript as suggested. Also, a whole distribution was plotted in Figure 5 of the original manuscript, which is a similar plot to Figure 9 of Kim et al. (2020).

16) Page 9, Line 20: "wide coverage". I assume you mean over time, but it would be good to specify.

➔ Thank you for the comment. The statement is modified in the revised manuscript. [Page 10, Lines 11-12]

17) Figures 6 and 7: Here the x-axis needs to be in log-scale. Many of the distributions do not look log-normal, but maybe the x-log will help. Also, please use the same range of log(EDR) for all panels. Otherwise, it is very difficult to compare them.

➔ The x-axis of Figures 8 and 9 is changed into the log scale in the revised manuscript as suggested. Also, although we respect the reviewer's suggestion, we decided to keep the different range of the EDR for better representation. Instead, we include the notification of different range of the x-axis in the figure caption.

18) Page 15, Lines 12-31: This appears to be out of the scope and context of the current manuscript, since forecasting is not the topic of this work.

➔ Thank you for your comment. As one of other reviewers (Reviewer 2) recommended that it

is worth to discuss, we decide to keep this context in the revised manuscript with modifications to clearly point out the messages in summary. As shown in the case analyses in this study, the characteristics of the EDR observations from different wind components (U, V, and W) are highly depending on the possible sources [Convectively induced turbulence (CIT), clear air turbulence (CAT), and mountain wave turbulence (MWT)] of turbulence in the upper troposphere and lower stratosphere. This can be eventually useful for the situational awareness of cruising aircraft and tactical avoidance for turbulence, and for producing a better climatology of aviation turbulence. These statements are included in the revised manuscript. [Page 16, Lines 5-24]

**References**

Gao, F., and Han, L.: Implementing the Nelder-Mead simplex algorithm with adaptive parameters. Comput. Optim. Appl., 51, 259-277, 2012.

Gill, P. G.: Objective verification of World Area Forecast Centre clear air turbulence forecasts, Meteor. Appl., 21, 3-11, 2014.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: Numerical Recipes: The Art of Scientific Computing. 2nd ed. Cambridge University Press, 963 pp., 1992.

Kim, J., Kim, J.-H., and Sharman, R. D.: Characteristics of energy dissipation rate observed from the high-frequency sonic anemometer at Boseong, South Korea, Atmosphere, 12, 837, 2021.

Kim, S.-H., Chun, H.-Y., Kim, J.-H., Sharman, R. D., and Strahan, M.: Retrieval of eddy dissipation rate from derived equivalent vertical gust included in Aircraft Meteorological Data Relay (AMDAR). Atmos. Meas. Tech., 13, 1383-1385, 2020.