Response to Reviewer 1's Comments

Han-Chang Ko, Hye-Yeong Chun

February 19, 2026

< Overview >

This paper has a potential of significant and timely contribution that effectively leverages high vertical-resolution radiosonde data (HVRRD) to create a unique global turbulence dataset. The proposed Ri_min method correctly addresses the limitations of the Thorpe method by including turbulence detection in statically stable, strong shear layers (0 < Ri < 0.25). The paper is well-written and the core methodology is sound, satisfying the main scientific requirements for publication in Earth System Science Data (ESSD). I recommend a Minor Revision to address documentation clarity, data usability, and minor scientific detail, ensuring the dataset meets the high standards of long-term data archiving.

We sincerely appreciate your thoughtful review and valuable comments on this manuscript. We seriously considered your comments and posted responses to your comments and suggestions. Below, we indicate the original comment of the respective reviewer in blue and our answer in black. In addition, we provide a tracked-changes version of the manuscript.

< Minor Comments >

1. Justification of Two Practical Considerations
   While the methodology for deriving EDR from Ri_min is detailed, two practical considerations of using L and VWS instead of Def need to be justified more clearly.
   - The authors mentioned that the Thorpe method has a limitation to explicitly consider convectively unstable condition (Ri < 0), which could have a relatively large mixing length (convective overturning of wave or turbulent eddy). But, shear driven KHI is very intermittent, which would have a small mixing length (small-scale eddy). So, it might be necessary to justify more specifically why the authors adapted the length scale from Thorpe method here?

   We appreciate the reviewer's comments. We would like to clarify that the present study does not adopt the Thorpe length scale L_T. In our formulation, the length scale L is defined as the vertical thickness of a continuous layer in which Ri < 0.25 (Lines 142–143 in the original manuscript). This definition differs fundamentally from the Thorpe length scale L_T, which is determined from reordered potential temperature. We acknowledge that the sentence in the original manuscript, Line 144: "*Note that L is the same as the thickness of the turbulence layer that is estimated from the Thorpe method (Ko et al., 2019)*" may have led to confusion. To avoid any confusion, this sentence has been removed in the revised manuscript.

   - The authors substitute the def in Eq. (8) by VWS based on simple synoptic-scale analysis, which would make sense in some part. But, it would not be applicable in a case where strong VWS in anticyclonic shear and curvature jet stream (e.g., Knox 1997). Given the simple assumption that the cyclonic and anticyclonic jet streams are equally happening in mid-latitude, anticyclonic curvature jet stream is theoretically stronger than cyclonic jet based on the gradient wind balance. I think the authors need to more carefully justify (or provide some limitations of current method) using VWS instead of DEF in this study.

We thank the reviewer for raising an important dynamical concern regarding the approximation of total deformation (Def) by vertical wind shear (VWS), particularly in environments such as strong anticyclonic shear and curvature jet streams. In the original manuscript, the approximation was justified using a scale analysis based on large-scale synoptic conditions (horizontal scale 10^6 m). Under this assumption, the dominant contribution to Def arises from the $\partial u/\partial z$ ~ 10^{-3} s^{-1}, whereas $\partial u/\partial x$ ~ 10^{-5} s^{-1}. This implies that neglecting horizontal shear introduces an error on the order of ~1%.

To address the reviewer's concern, we further considered a scale representative of strong anticyclonic shear and curvature jets, adopting a reduced horizontal scale of 10^5 m following Holton and Hakim (2013). In this case, the horizontal shear term increases to order 10^{-4} s^{-1}, while the vertical shear term remains dominant ($\partial u/\partial z$ ~ 10^{-3} s^{-1}). The relative contribution of horizontal shear therefore increases, leading to a potential uncertainty of approximately 10%. Although $\partial u/\partial z$ remains the leading-order term, the approximation becomes less accurate under such strong and curved jet structures.

We have added a discussion in the revised manuscript (Lines 153–166) explicitly acknowledging this limitation.

2. Illustrative Case Study for 0 < Ri < 0.25 Turbulence vs in situ EDR near jet stream
The core scientific advancement of this work is the ability of the Ri_min method to detect turbulence in the statically stable, high-shear region (0 < Ri < 0.25) where the Thorpe method fails. This claim, while theoretically sound, needs compelling direct comparison of Ri_min method with in situ EDR obs. I know it is difficult to find a case that has collocated pairs of HVSSD and in situ EDR. Considering that jet stream is synoptic scale, it will be great to find a CAT outbreak day (last 1-2 days) near jet stream to be captured by both HVRRD and in situ EDR. Then, it will be more convincing that this new method really show a good performance for KHI near upper-level jet system.

We thank the reviewer for this valuable and constructive suggestion! We agree that an illustrative comparison between the Ri_min-based turbulence and in-situ flight-EDR observations in a statically stable, high-shear environment near a jet stream would provide a more compelling demonstration of the method's capability to detect shear-driven turbulence (0 < Ri < 0.25), which is not captured by the Thorpe method. Following this suggestion, we conducted a case analysis and the result is shown below in Figure R1 (Figure 3 of the revised manuscript). The selected case occurred at 00 UTC on 13 January 2024, when a pronounced trough developed over the United States at 200 hPa, accompanied by a strong and curved jet stream on the southeastern flank of the trough (Figure R1a). This synoptic configuration is dynamically favorable for enhanced vertical wind shear and Kelvin–Helmholtz instability in the statically stable upper troposphere.

A zoomed-in view of the jet region (Figure R1b) presents collocated HVRRD-derived Ri_min-based turbulence (circles) and in-situ flight-EDR reports (triangles) within the 10–12 km altitude range. In the region highlighted by the red box, both datasets consistently indicate severe-intensity turbulence in close spatial proximity, demonstrating that the Ri_min method successfully captures strong turbulence embedded within the strong-shear jet environment. In contrast, in the two regions marked by yellow boxes, both datasets indicate null turbulence intensity, further supporting the consistency between the Ri_min-based estimates and aircraft observations.

The new figure (now included as Figure 3 in the revised manuscript) and the corresponding discussion (Lines 271–292) have been added to strengthen the observational support for the proposed Ri_min method.

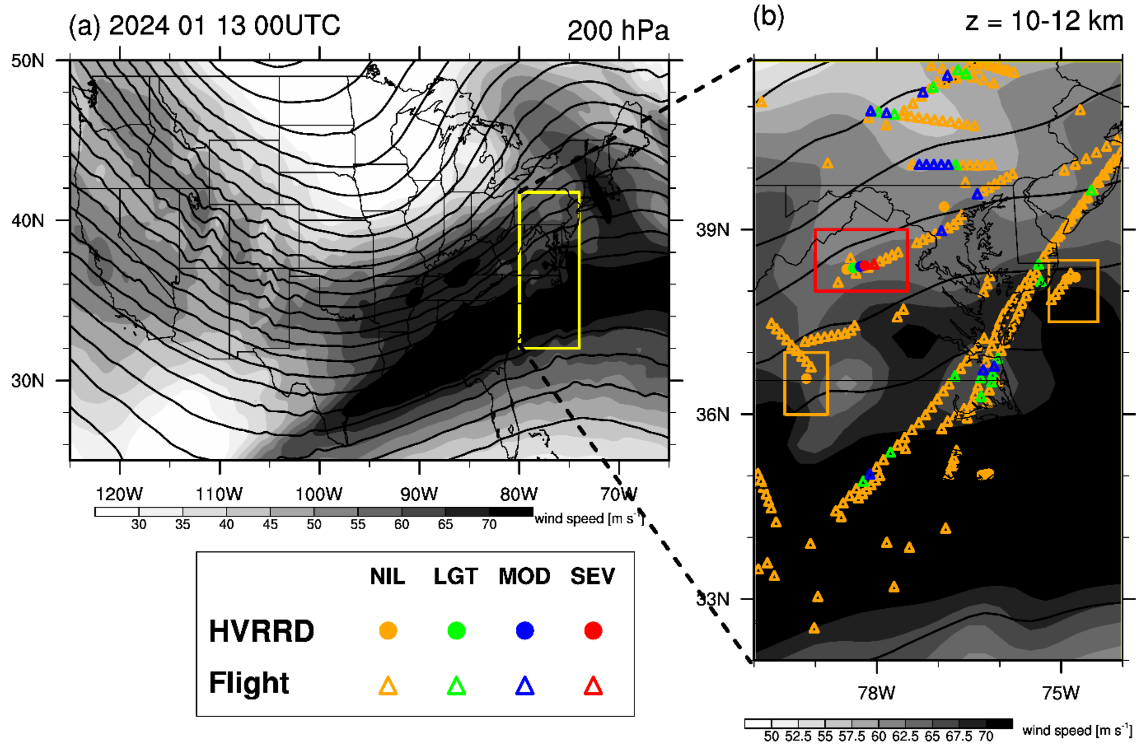

Figure R1. (a) Horizontal distribution of 200-hPa wind speed (shaded) and geopotential height (contours) at 00 UTC on 13 January 2024. (b) Zoomed-in view of wind speed at 200 hPa with overlaid turbulence observations at z = 10–12 km. Filled circles and open triangles denote the turbulence estimated from HVRRD-Ri_{min} and that observed from in-situ flight-EDR, respectively. Colors represent turbulence intensity categories (null; NIL, light; LGT, moderate; MOD, and severe; SEV). The red box highlights a region where both datasets consistently indicate SEV turbulence in close proximity, while the yellow boxes mark regions where both datasets indicate NIL turbulence.

3. Explicit Declaration of Data Output Format and Internal Structure
For publication in ESSD, data usability is paramount. The manuscript must explicitly declare the chosen long-term archiving data format and its internal structure for user accessibility. Please state the definitive file format (e.g., NetCDF, HDF5, or another community standard) that will be used for the final dataset. More importantly, provide a clear, dedicated table listing all variable names (e.g., turbulence_edr), their units (e.g., m^2/3 s^-1), and the corresponding CF-compliant metadata for each variable provided in the data files.

We thank the reviewer for this important and constructive suggestion. We fully agree that clear documentation of the data format and internal structure is essential to ensure usability and long-term accessibility, particularly for publication in ESSD. Therefore, we have explicitly specified the definitive file format (NetCDF) of the final archived dataset in Section 6 (Data availability) of the revised manuscript (Lines 482–483).

Furthermore, we have added a dedicated table summarizing all variables included in the dataset, including their variable names, physical units, and CF-compliant metadata attributes in the new Table S1 in the Supplementary Material. A concise description of this table is now included in Section 6 of the revised manuscript (Lines 482–483).

4. Clearer Statement on Spatio-Temporal Data Heterogeneity
The dataset is unique because it is global, but its temporal and horizontal resolution is fundamentally constrained by the heterogeneous global radiosonde network (typically

00/12 UTC, geographically sparse). The current description emphasizes the high vertical resolution but downplays the horizontal and temporal sparseness of the overall product. A single sentence or footnote in the Abstract or Data Description section must clearly state that the dataset's spatial and temporal coverage reflects the limitations and heterogeneity of the operational global radiosonde observing network.

We thank the reviewer for this comment. We agree that, although the dataset provides high vertical resolution globally, its horizontal and temporal sampling is inherently constrained by the heterogeneous structure of the operational radiosonde network. In response, we have added a statement in Section 3.1 (High vertical-resolution radiosonde data; Lines 184–188 in the revised manuscript) explicitly noting that the dataset's horizontal and temporal coverage reflects the limitations and inhomogeneity of the global radiosonde observing system. In particular, we clarify that observations are typically available only at 00 and 12 UTC and that station distribution is geographically uneven, with denser coverage over mid-latitude continental regions and sparser sampling over oceans and parts of the tropics and polar regions.

5. Final Archival Repository and Versioning Commitment
ESSD requires a strong commitment to long-term data preservation and traceability. Please specify the exact permanent digital repository (e.g., a recognized data center like NOAA NCEI, EUMETSAT, or a major institutional repository with DOIs) where the final dataset will be lodged. Furthermore, provide a clear statement regarding the versioning scheme (e.g., "The dataset presented here will be designated Version 1.0 (v1.0). All future updates will be released as incremental versions, v1.1, v2.0, etc., with associated Digital Object Identifiers (DOIs).").

We thank the reviewer for this important suggestion regarding long-term preservation and version control. The final dataset is archived in a permanent, DOI-assigning repository (Zenodo), and the corresponding DOIs are explicitly provided in both the Abstract and the Data availability section of the original manuscript. Zenodo ensures long-term preservation, public accessibility, and persistent DOIs for the deposited dataset.

In addition, we have now clearly specified the versioning scheme in the Data availability section (Line 479 in the revised manuscript). The dataset presented in this paper is designated as Version 1.0 (v1.0). Any future updates, corrections, or extensions will be released as incremental versions (e.g., v1.1, v2.0), each with an associated DOI to ensure transparency and reproducibility. Version history and change logs will be documented within the repository metadata to maintain full traceability of dataset evolution.