Review of *“The interaction of warm conveyor belt outflows with the upper-level waveguide: a four-type climatological classification”* by Vishnupriya et al.

Summary

Vishnupriya et al. present a diagnostic study that classifies warm conveyor belt outflow interactions with the upper-level jet stream (Rossby waveguide). Using ERA5 and Lagrangian tracking, the authors systematically combine multi-decade seasonal climatology of WCB–waveguide interaction types, extending prior case-specific research by introducing an objective classification. The large sample lends confidence that the reported frequencies and patterns are robust climate statistics rather than anecdotal findings. The methodology, from using the well-known LAGRANTO tool for WCB tracking, to identification of waveguide disturbances, is rigorous and consistent with previous literature. The case studies (Fig. 4) and schematic drawing (Fig. 11) are helpful in argumentative demonstrations. Despite the complexity of the subject, the manuscript is generally well-organized and written. Overall, the paper’s structure (methods, case examples, climatology, composites, summary) makes it easy to follow the logical progression from methodology to key conclusions. Therefore it has the potential to be published in WCD, but some clarifications and revisions are still needed.

Major Comments

- **Threshold Sensitivity.** The classification relies on specific PV threshold criteria that, while grounded in prior studies, are somewhat subjective. For example, ridges are defined by a PV anomaly < –1 PVU and cutoffs by PV < 2 PVU. Likewise, the blocking definition requires a –1.3 PVU anomaly persisting 5 days. It is not fully explored how sensitive the results (especially the relative frequencies of interaction types) are to these threshold values. The authors note that certain methodological elements are subjective but claim the main results are not sensitive to those choices; however, the paper would benefit from evidence of this (e.g. a brief sensitivity test varying the PV anomaly cutoff by some amount). As it stands, it is hard to know if slightly different thresholds might change an event from “weak interaction” to “ridge” or alter the 58.7% ridge frequency.

- **What about the stratosphere?** In the conclusion part (line 585-586), the authors find that in boreal winter, “tropospheric WCB outflows most frequently result in ridge interactions (58.7%), followed by no-interaction (27.7%), block (9.7%), and cutoff (3.9%)

interactions." However, such numbers come from a normalization stated in Section 3 (line 250-252): "ridge interacting type (54.0%), followed by no interaction (25.5%), while block (8.9%) and cutoff (3.6%)". So there are about 100% - (54 + 25.5 + 8.9 + 3.6)% = 8% of stratospheric interactions resulting from WCB outflows that are excluded from this study. This 8% is comparable to the 8.9% blocking type and way larger than the 3.6% cutoff type. While it's understandable to focus on tropospheric impacts, these excluded cases (gray dots in their Fig. 2b) are interesting for completeness, especially given that some literatures argue stratosphere-troposphere coupling may be important and modifying blocking frequency (e.g., Davini et al., 2014), which potentially indicates the left-out stratospheric outflows to be a non-negligible category. Hence, it'd be best for the authors to answer if stratospheric WCB outflows are truly negligible in number and in analyses? Also it'd be best for the authors to include the "tropospheric" constraint in line 15 for WCB outflows.

- **A block could be double counted as a ridge!** Computationally, the authors still enforce mutual exclusiveness by applying a strict hierarchy when they tag each trajectory, so every WCB parcel ends up in one and only one class. Yet, a single trajectory's starting point can coincide with multiple feature types, say a block would typically also satisfy a ridge criteria. This hierarchy, while physically reasoned (a cutoff indicates a more intense wave breaking than a ridge), could lead to ambiguous cases being classified as the highest-ranking feature present. It would be useful to clarify whether the hierarchy ever overrides what a meteorologist might consider the primary interaction. The authors should ensure this automated decision-making doesn't misclassify borderline situations or transitions of types, such as an initial ridge may become a block ***after*** temporal persistence threshold being met. What's the proportion of such borderline transitions in the study? Wouldn't this also lead to a developed block being slightly older than spawned ridges, in contrast to Fig. 7 panel c, where block is argued to be of the youngest WCB outflow age at the point of interaction? For air parcels in a block, how many of them are of "young age" captured in Fig. 7c, how many of them have stayed in the upper level for more than 5 days (not captured in Fig. 7c)?

- **Timing of Interaction Assessment.** The classification is determined at the backward trajectory starting point, which is termed the "point of interaction," but this choice may not capture delayed or downstream impacts. If a WCB outflow does not immediately coincide with a ridge/block at the initial time but contributes to one a day later, such an effect would be missed by the algorithm. In other words, some WCBs might be labeled "no interaction" at the start point but go on to amplify a wave downstream. The paper would benefit from a discussion of this limitation – essentially, the method diagnoses interactions at a fixed time (when trajectories hit the 2-PVU surface) and may not track the subsequent evolution. This is partly addressed by analyzing the eastward advection of no-interaction vs. stagnation of block-type air parcels, but it remains possible that a WCB initially categorized as non-interacting could induce a ridge slightly later. Clarifying how the results might change if the interaction were evaluated over a time window (rather than an instantaneous point) would be useful.

- **Causality vs Colocation:** The study assumes that if a WCB outflow is co-located with a ridge, block, or cutoff, then the WCB ***interacts*** with the waveguide to produce that feature. This is a reasonable interpretation, but it is essentially inferred rather than directly proven. The authors offer a posteriori justification by noting the much weaker PV anomalies for trajectories classified as non-interacting, implying those indeed had minimal effect. Still, the methodology identifies ***associations*** between WCB outflows and PV disturbances – it does not demonstrate that the WCB caused the disturbance. It would strengthen the paper if the authors could argue more explicitly that the identified ridges/blocks were actually enhanced by the WCB. As an example, one could ask: might a ridge have existed anyway, with or without the WCB, and the WCB simply happened under it? The implicit assumption of causality could be more critically discussed. This point is important for interpreting the climatology: the paper shows where WCBs *co-occur* with certain waveguide disruptions, but future work (perhaps using modeling experiments) would be needed to confirm the extent of the WCB's causal influence.

- **Statistical / significance testing.** While the dataset is large, the manuscript does not report formal significance testing for differences between categories. Phrases like "differ significantly" (line 616) in the composite analysis appear to be used qualitatively. It would improve the rigor if the authors could demonstrate that key distinctions (e.g. the differences in outflow latitude or PV between categories) are statistically significant. Similarly, for the composite maps (Fig. 9 and 10), showing stippling or some indication of significance for anomalies would help support statements that, say, blocks are preceded by significantly higher eddy kinetic energy than no-interaction cases. If significance testing was not conducted, the authors should temper the language or clarify that "significantly" is meant in a qualitative sense. Adding some basic statistical analysis would bolster the conclusions that the observed differences are robust and not artifacts of variability.

- **What happens from Mid Ascent to End of Ascent?** One thing stands out in Fig. 6 is that the four types of interactions could hardly be distinguished from each other in panel 6c Mid Ascent, but there are significant spatial differences demonstrated in panel 6b End of Ascent. What in this ascending process is causing the difference? Could you connect Fig. 9 and Fig. 10 arguments with Fig. 6 panel b and c differences? If possible, could you expand in details about the governing mechanism?

- **Composite Selection Bias:** The method for constructing composites introduces additional subjective criteria: the authors only composite time steps when a given interaction type is sufficiently dominant (>40% of WCB trajectories in the region). This ensures “pure” cases but might bias the composites towards extreme examples. For instance, a time step with 39% no-interaction, 46% ridge, 10% block, 5% cutoff might be excluded entirely, whereas a time step with 41% no-interaction triggers inclusion as a “no-interaction case”. Such hard thresholds (40% for one type, and a secondary 10% cutoff criterion for the cutoff type) could skew the sample of events used for composites. The authors should justify the choice of 40% – presumably to get a decent sample size while maintaining category signal – and perhaps test that varying this threshold (30% vs 50%) does not qualitatively change the composite patterns. If a more inclusive compositing approach yields similar patterns, that would alleviate concern that the composite results are dependent on this filtering.

Minor Comments

- **Fig. 4.** The abscissa and ordinate in Fig. 4 are wrong. You cannot have two 80°N in one map, nor 0-60°W being perpendicular to 0-80°E.

- **Fig. 8.** Panel a does not explain the letter M in the total number of trajectories.

- **Fig. 9-10.** The caption of labelling (a,e) as point of interaction is not consistent with the figure labelling (a,e) as start of ascent. Same inconsistency happens for all rows in the plots.

- **Composites of cutoff-interaction time steps should be included in the manuscript, not in supplementary materials**. The difference for block interactions to ridge of having “intense negative PV anomaly and strong ridge” (line 549) are important, but the cutoff interactions also share this difference and signify a more important sign change of EKE anomaly throughout the ascending time range. I would suggest swap Fig. S15 with Fig. 10.

Recommendation: *Major Revision*.