We thank the referee for the insightful comments. Below, we repeat the reviewer's remarks in red italics, and add our respective responses in normal text.

Reviewer 2

General Comments:

This paper utilizes an unsupervised, hierarchical clustering technique to define geometric clusters representing the behavior of the stratospheric polar vortex as defined by boundary objects and features from a FFT of 10-hPa GPH fields. Thereafter, this work examines the stratospheric polar vortex response to gravity wave forcings in climate model scenarios, with a focus on three hotspot regions: East Asia, the Himalayas, and North America. Their results conclusively show that the hotspot forcings reduce amplitudes of planetary wave size 1 features, with the largest of those changes occurring in the Himalayan hotspot. Additionally, the observed gravity wave forcings hold implications for geometric responses in the stratospheric polar vortex. The results from this paper are compelling and provide interesting contributions to the discussion of stratospheric polar vortex morphology. Outcomes from this paper, specifically regarding the clustering of the vortex, hold applicability for examinations of subseasonal stratospheric-tropospheric teleconnections.

The paper itself is technically dense and difficult to follow at times. It reads like a chapter of a dissertation. The paper would improve greatly with some changes to the wording and extrapolation of specific methodologies to help the content stand as an individual publication. Additionally, the author should explain the underlying motivation for using Fourier descriptors to define the vortex boundary in place of pre-existing, similar methodologies.

We thank the reviewer for their thorough and constructive feedback, as well as for recognizing the contributions of our study. In response to these comments, we have revised the manuscript extensively. These revisions have substantially improved the overall clarity and quality of the manuscript.

I would recommend publication following the completion of these major and minor revisions.

Major Comments:

1. *There is a lot of "hand waving" done with respect to explaining the specifics of this experiment and its methodology to the point that a lot of backtracking was required to make sense of what was being discussed in Section 3. This is particularly an issue for anything surrounding the model simulations since the author assumes the reader has read Mehrdad et al. 2025a. It is not necessary to include all clarifying information about the specifics of the original experiment in Mehrdad et al. 2025a, but enough detail is needed so that an individual can read this paper on their own and know what is happening. I would suggest that you include more specifics about the methodology*

surrounding model simulations and data in Section 2. Clarified methodological information will improve this section of the paper substantially.

We appreciate the reviewer's comment and agree that providing clearer context for the model simulations improves readability. The present paper builds directly on Mehrdad et al. (2025a), which provides a full technical description of the experimental design and model configuration. The essential information required to understand the current study—model type, boundary and initial conditions, ensemble size, grid resolution, forcing strength, and the spatial definition of the hotspot regions—is included in Section 2.1 (lines 81-101).

To further clarify the setup while minimizing repetition from the earlier paper, we now provide a concise summary of the distribution of the SSO-induced zonal-wind tendencies by hotspot in Appendix A (lines 486–506 of the revised manuscript). We also added a one-sentence pointer in Section 2.1 directing readers to Appendix A for this summary of the imposed SSO forcing.

Figure A1 in the revised manuscript presents the column-integrated (200–1 hPa) maps of the SSO-induced zonal drag during NDJFM, highlighting where the forcing is applied. Figure A2 shows the corresponding vertical structure of the SSO-induced zonal-mean zonal-wind tendencies for the control and for anomalies in the sensitivity experiments (HI–C, NA–C, EA–C). We believe these additions provide sufficient methodological clarity while avoiding unnecessary redundancy.

2. *This paper would benefit from an explanation and motivation for the choice to use Fourier descriptors for diagnosing the vortex boundary as opposed to using pre-existing vortex diagnostic methods, like those from Seviour et al. 2013, or k-means clustering. The FFT method presented here is very compelling, but what is the underlying motivation for using it? Seviour, W. J. M., D. M. Mitchell, and L. J. Gray (2013), A practical method to identify displaced and split stratospheric polar vortex events, Geophys. Res. Lett., 40, 5268-5273 doi:10.1002/grl.50927.*

We appreciate this comment and have expanded the motivation in Section 2.3. Our choice of Fourier descriptors is driven by the physics we aim to diagnose: vortex morphology governs where and how planetary and gravity waves interact with the SPV edge. Fourier descriptors offer a compact, geometry-native representation of the boundary in which each coefficient has geometric meaning (roundness, elongation, orientation; finer deformations at higher orders). By retaining only the leading modes, we filter small-scale, transient irregularities and emphasize the large-scale morphology that is dynamically relevant for wave–mean flow coupling, while keeping dimensionality low for robust clustering.

In contrast, threshold diagnostics (e.g., centroid, aspect ratio) provide only a few scalars and cannot distinguish higher-order shape attributes; and gridpoint-wise clustering (e.g., k-means on dynamical fields) measures similarity at fixed locations, which makes it sensitive to modest displacements/rotations and can separate geometrically similar vortices. Our approach therefore retains physical interpretability, increases geometric fidelity, and reduces sensitivity to incidental shifts, which is particularly useful for isolating geometry-conditioned responses to regional GW forcing. (See Section 2.3, lines 112-133).

3. *While there is a lot of useful information in the discussion of the sensitivity experiments, sections 3.3-3.5 of this paper suffer from an oversaturation of material. I found it difficult at times to maintain focus on the key takeaways from the figures. This quality of writing is ideal for a dissertation but not necessary for a publication. I would suggest thinning out the text from this area, focusing on the primary results. I may also suggest modifying the figures here to show only the most relevant clusters. For example, the panels in Figures 10-18 for C1, C7, and C10 are generally empty.*

We thank the reviewer for this helpful suggestion. In response, we have substantially revised Sections 3.3–3.5 to improve focus and readability. The revised text now emphasizes the primary and consistent signals while foregrounding the key messages from each figure (see revised Sections 3.3–3.5, lines 294-365).

To reduce the density of material, we have also thinned the text and moved the former Figure 8 to Appendix D (now Figure D1; lines 519–532), together with its accompanying description. Furthermore, to enhance readability, we have divided the discussion section into shorter subsections, each focused on a specific aspect while maintaining clear links among them.

Regarding the figures, we fully appreciate the motivation to simplify them by removing panels that convey negligible signals. After careful consideration, we chose to retain the panels for all clusters in each experiment, as they convey meaningful information and are needed for the completeness of the analysis. Showing all clusters allows readers to verify that certain clusters indeed make minimal or inconsistent contributions, which is an important part of the scientific conclusion. Additionally, readers who may be interested in accessing the full set of class contributions—not all of which are mentioned in the text, as we focused the discussion only on the most relevant and consistent clusters—can do so directly from the figures. This approach reduces cognitive load in the narrative while preserving the full evidentiary record.

Overall, these revisions improve clarity and narrative focus while preserving the full evidentiary information conveyed by the figures.

Minor Comments:

Line 1-2 – I would personally include a reference to what months specifically define the "winter pole" in the Northern Hemisphere.

We thank the reviewer for this suggestion. To address it, we have modified the introduction (line 18 in the revised manuscript).

Line 30 – You may also want to cite this paper as well: Butchart, N. 2022: The Stratosphere: A Review of the Dynamics and Variability. J. of Weather and Climate Dynamics. 3(4), 1237-1272. https://doi.org/10.5194/wcd-3-1237-2022

We thank the reviewer for this helpful suggestion. The reference to Butchart (2022) has been added in the revised manuscript (line 21).

Line 47 – This could be a technical comment, but I would be inclined to put the initial reference of "compensation mechanisms" in quotes, since it is a reference to terminology.

We thank the reviewer for this suggestion. We have placed "compensation mechanisms" in quotation marks in the revised manuscript (line 47).

Line 48 – "These" referring to what? The compensation mechanisms?

We thank the reviewer for this comment. We have clarified the reference by revising the sentence to read "Among these compensation mechanisms is …" in the revised manuscript (line 48).

Line 53 – What is meant by "inter-model comparisons of stratospheric dynamics"? Would it be possible to provide an example?

We thank the reviewer for this helpful comment. By "inter-model comparisons of stratospheric dynamics," we refer to the large spread that models exhibit in (i) SPV morphological diagnostics (aspect ratio, excess kurtosis, centroid latitude) and (ii) the frequency and ratio of vortex-split versus vortex-displacement SSWs (Sigmond et al., 2023; Kuchar et al., 2024). We have revised the beginning of the paragraph to make these explicit and to improve the logical flow of the paragraph (lines 53–58 in the revised manuscript).

Line 59 – "They" referring to what?

We thank the reviewer for pointing this out. We have revised the sentence by replacing "They" with "These impacts" to clarify the reference and ensure it clearly points to the impacts discussed in the preceding sentence (line 60).

Line 60-62 - "hotspot-induced changes" … to GWs? There is a lack of specific clarifying information in this sentence.

We thank the reviewer for this helpful comment. We have revised the sentence for clarity, changing it from "Previous studies have documented hotspot-induced changes in the zonal-mean context" to "Previous studies have documented the atmospheric response to enhanced GW drag in hotspot regions within a zonal-mean framework" (lines 62-63 in the revised manuscript).

Line 65-72 – As stated in Major Comment 1, some readers will find this paper on its own and may not read Mehrdad et al. 2025a. Silly as it may sound, it is necessary to provide a small amount of context to the reader in the Introduction, one or two sentences, by giving a general overview of Mehrdad et al. 2025a's general goals and key takeaways. The author does a good job of explaining how this specific paper has different motivations, continuing in this paragraph, but does not fully address the previous paper.

We thank the reviewer for this helpful suggestion. We acknowledge that adding this summary improves the paragraph and the overall flow of the introduction. In response, we have inserted a concise overview of Mehrdad et al. (2025a)'s key findings at the beginning of the last paragraph of the Introduction in the revised manuscript (lines 67-71). Additionally, as noted in our response to your earlier comment, Appendix A provides further details on the GW forcing.

Line 65 – Since this is the first mention of UA-ICON, I would reference it in full: "high-top UA-ICON global circulation model experiments."

We thank the reviewer for this helpful suggestion. We have revised the text to read "high-top UA-ICON (ICOsahedral Nonhydrostatic model with Upper-Atmosphere extension)" in the revised manuscript (lines 67-68). We have modify it to high-top UA-ICON (ICOsahedral Nonhydrostatic model with Upper-Atmosphere extension)

Line 67-69 – Something about point (ii), starting with a verb, while points (i) and (iii) start with "we" disrupts the flow of this paragraph and makes it difficult to digest.

We thank the reviewer for this comment. We have revised point (ii) to begin consistently with "we," aligning its structure with points (i) and (iii) (lines 73-75 in the revised manuscript).

Line 71 – I might replace "Arctic" with "Northern Hemisphere" or similar. Or reflect on future nomenclature used in the paper and try to maintain consistency.

We thank the reviewer for this comment. The term "Arctic" has been replaced with "Northern Hemisphere" in the revised manuscript (line 78)

Line 93 – What was the period of record for this winter season data? 30-year simulations are indicated, but what is the start date? Again, this information may be included in the previous paper from Mehrdad et al. 2025a, but it is still important to include clarifying details about the specifics of this experiment as it is its own paper.

We thank the reviewer for this helpful comment. In response, we have added further details about the boundary-condition period, the initialization of the simulations (lines 88-90), and the generation of the ensemble members (lines 97-98). Each experiment consists of six ensemble members, each representing a 30-year simulation (180 model years in total; 175 years for NA see Mehrdad et al. 2025a for more details). After excluding the first simulation year as spin-up, this results in 174 years of analyzed data (169 years for NA). Because the simulations were initialized from and driven by annually repeating present-day climatological conditions, the absolute start date of the integrations is arbitrary and does not affect the results.

Line 110 – Explain why you retain data northward of 25°N. Similarly, why 18%?

We thank the reviewer for this comment. We have clarified that data were restricted to latitudes north of 25° N to isolate the Northern Hemisphere extratropical region where the SPV is located (lines 136-137). The choice of the 18 % threshold was determined empirically.

Line 135 – Perhaps be a little more direct/descriptive beyond saying "this feature space" to provide further clarity to the reader.

We thank the reviewer for this suggestion, which enhances clarity. In response, we have revised the sentence on line 163 to more explicitly reference the 24 features.

Line 168-170 – C1 is distinguished for unstable-vortex clusters, in this case meaning more than two or no SPV boundaries. C7-C10 are split vortex events. Would C2-C6 contain a cluster specified for displaced vortex events? What about strong vortex events? You may mention this in detail later, but it is good to give some preliminary indication of what these clusters might contain.

We thank the reviewer for this insightful comment. As the introduced clustering method is hierarchical, and here we only present the top level in the hierarchy, we have retained the content as is. However, to enhance readability and provide a forward reference without preempting the detailed discussion in the results section, we have added a pointer to the relevant part (lines 198-199). This points the reader to the relevant section, where visual content (beyond the dendrogram branching) and in-depth description are presented, maintaining the manuscript's logical flow.

Figure 4 – The cyan lines are difficult to read in this figure. I may suggest using a different colored contour. Additionally, I may separate Panels d and h into their own figures. Included here, the panels are a little small, and their inclusion makes this figure a bit too busy.

Figure 6 – Similar comments to Figure 4. The coloring is difficult to see, and the subplots on this figure are small and difficult to read. Breaking up or structuring this figure differently may make it easier to read.

We thank the reviewers for their constructive feedback. In response, we have revised Figure 4 by removing excess free space, tightening the layout, and enlarging the subplots to improve readability. This adjustment enhances overall clarity without necessitating the separation of panels d and h in Figure 4, as the larger subplots now make the details more visible. We also make the former Figure 6 (now Figure 5 in the revised manuscript) slightly larger in scale in the revised manuscript. Regarding the cyan contours, we evaluated alternative colours but retained the current ones, as they provide the best contrast across the full range of positive and negative anomalies. While they are less distinct over near-zero (white) regions, darker alternatives become indistinguishable against strong anomalies. These changes improve the figures while preserving their comprehensive structure.

Line 234 – This sentence may not be necessary.

Thank you for the comment, the sentence is removed in the revised manuscript (line 260).

Line 262 – What exactly is an "occurrence frequency"/how is it calculated?

We thank the reviewer for pointing out the need for clarity on the term "occurrence frequency." To address this, we have revised the paragraph (lines 278-279 in the revised manuscript) by adding an explanation of how it is calculated: specifically, as the percentage of days each cluster occurs relative to the total days in the extended winter period.

Line 375, 377 – "gravity wave" may be replaced with GW.

We thank the reviewer for this comment. The use of the abbreviation "GW" has been reviewed and applied consistently throughout the revised manuscript

Line 391 – Which configuration? Be a little more specific.

We thank the reviewer for this suggestion, which improves the specificity of the description. To address it, we have revised the sentence (lines 392-393 in the revised manuscript) by clarifying "this configuration" to refer explicitly to the equatorward position of the positive anomaly band (north of the forcing region) in the HI experiment.

Line 393 – At this point within the discussion section, I may remind the reader what each of the clusters (C5, C6, etc.) physically represents.

We thank the reviewer for this helpful suggestion, which reinforces the connection between the discussion and the earlier description of the clusters. To address this, we have added a brief reference at the beginning of the paragraph (revised lines 395–396) directing the reader to the section where the cluster characteristics are described in detail. Moreover, we have divided the discussion section into shorter subsections, each focusing on a specific aspect while maintaining clear links among them. These changes improve readability and help the reader more easily follow the key messages and connections across sections.

Technical Comments:

In general, I noticed many tense inconsistencies (e.g., present vs. past tense within the same sentence, etc.) Choose one tense to use throughout the paper.

We thank the reviewer for this very helpful comment. We have carefully reviewed the manuscript to ensure consistency in verb tense throughout.

All of the figures would benefit from being increased in resolution. Figure formatting within the text should also be changed to keep the discussion collocated with the image as close as possible.

We thank the reviewer for this comment. We have reviewed all figures in the manuscript and improved their quality where possible by increasing resolution, enlarging scales, and tightening the layout (e.g., Figures 3, 4, B1, and C1). Regarding figure placement, the text discussing each figure is already located as close as possible to the corresponding image in the LaTeX source. We anticipate that any remaining minor spacing issues will be resolved during the journal's typesetting and proofreading process.

Line 45 – "exert" should be "exerts"

We thank the reviewer for this observation. Upon review, we confirm that the original verb "exert" is correct, as the compound subject "the location and distribution" is plural and therefore takes a plural verb. Because "location" and "distribution" refer to distinct elements, no change was made to the sentence on line 45.

Line 133 – "features 4 to 23" may be replaced with "features four to twenty-three" for consistency

We thank the reviewer for this suggestion. In the manuscript, we consistently use numerical notation when referring to feature indices. To maintain this consistency, we have chosen to retain the numerical format here (line 160).

Figure 5 – Increase figure size.

The figure (now Figure C1 in the revised manuscript) has been enlarged in the revised manuscript.

Section titles 3.3, 3.4, 3.5 – Use the unabbreviated versions of the sensitivity experiments for the titles.

We thank the reviewer for this helpful suggestion. The subsection titles have been revised to use the unabbreviated names of the sensitivity experiments. They now read: Himalayas (HI), Northwest America (NA), and East Asia (EA).

Line 318-320 – "consistent" is used three times in some form here. I may suggest rewording.

We thank the reviewer for this observation. This section has been completely rewritten in the revised manuscript in response to the reviewer's major comment.

Line 431, 448 – "zonal mean" should be "zonal-mean" here

We thank the reviewer for noting this. The term has been corrected to "zonal-mean" throughout the manuscript for consistency.

Line 461 – "help" should be "helps"

We thank the reviewer for this observation. Upon review, we confirm that the original verb "help" is correct, as the compound subject "the shape-based clustering and class contribution framework" is plural and therefore takes a plural verb. Because "clustering" and "framework" refer to distinct elements, no change was made to the sentence (line 474).

References

Butchart, N.: The stratosphere: A review of the dynamics and variability, Weather and Climate Dynamics, 3, 1237–1272, https://doi.org/10.5194/wcd-3-1237-2022, 2022.

Mehrdad, S., Marjani, S., Handorf, D., and Jacobi, C.: Non-zonal gravity wave forcing of the Northern Hemisphere winter circulation and effects on middle atmosphere dynamics, Weather Clim. Dynam., 6, 1491–1514, https://doi.org/10.5194/wcd-6-1491-2025, 2025a.

Seviour, W. J. M., D. M. Mitchell, and L. J. Gray (2013), A practical method to identify displaced and split stratospheric polar vortex events, *Geophys. Res. Lett.*, 40, 5268-5273 doi:10.1002/grl.50927.