

Detailed point by point response to all referee comments.

Here listed are the referee comments, our replies and changes made.

1. Comments from Referees
2. Author response
3. Changes in Manuscript

Reviewer 1

This manuscript explores regional trends in mountain wave turbulence (MWT), linked to climate change, in high-resolution global climate models. They use a suite of indices and models to explore these trends and also link them to changes in low-level winds. This is an important topic, and the manuscript adds useful information. The results demonstrate significant dependence on the model, region, and index used, highlighting significant uncertainty in any conclusions about future trends. Only a few locations consistently demonstrate statistically significant trends (some negative, some positive), which itself is an important result. My main concerns with the manuscript are that it is missing important detail on the methods use, and the discussion could be expanded on and improved in some areas, with limitations further explored.

Response: We thank the reviewer for these comments and for their time spent reading and reviewing the paper. We have edited and updated key areas in the methods and expanded on the discussion. We agree that there is uncertainty within our results due to the high dependence on model, index and region. We have restructured our results section to explore these uncertainties first in section 3.2 , and then summaries overall trends in Section 3.3.

Major comments:

1 – What thresholds are used to diagnose MWT for each index to create Figure 1 (and are then used for the remainder of the manuscript)? This is not included anywhere and expanded detail on the methods are necessary for this work to be reproducible and complete. What do the frequency distributions of these indices look like? I suggest that the methods section needs to be expanded to include this detail.

Response: Yes, a good point as we mention using the 98th percentile but not the actual threshold values. Due to the number of thresholds, as its model and resolution dependent it is a large list. A note has been added in the appendix to

contact the author for these thresholds. Instead of having it in the supplementary information, we think it's best to have contact the author, so the author can highlight directly to researchers not to use if they are applying a different model.

Change: Section 2.2 (Methodology; Diagnosing turbulence) lines 106 to 110 and appendix B: Clear air turbulence indices, lines 439 - 441

2 – There should be some additional discussion about the appropriateness of using these indices in climate models as the majority of cited work (except Kim et al. 2023) apply the indices to higher resolution NWP models. Does the very smooth topography of climate models pose a significant impediment to doing this properly (even for the highest resolution grids)? I assume that for the indices you are using the native model topography to define the terrain gradient (though this is not stated anywhere and should be). What would the results look like if you used the model fields to define the CAT index, but a higher resolution topography dataset to define the terrain slope (i.e., the same terrain slope for every model)? Would that change the results? My concern is that smooth topography doesn't identify the complexity of the terrain, reducing the overall area of identified MWT locations (Greenland is a good example, where MWT is only identified in near-coastal regions, whereas previous studies, e.g., Doyle et al. 2005 J. Atmos. Sci., have observed MWT closer to the centre of the continent, i.e., not only along the steepest slopes).

Response: This is a very good point; we have now stated that our climate models have limitations due to the native model topography (methodology). We agree that it would be beneficial to use a high-resolution topography dataset, but we like the approach of using the model topography resolution, as it uses the same grids as our CAT indices and helps highlight the uncertainty in future projections of MWT. We think that a secondary paper could be written to compare using the model vs a high-resolution topography dataset to further explain the model representation of MWT. We have written a few lines within the methodology to explain these limitations and mention that our results are related to the terrain slope and MWT is likely to occur downstream over our MWT regions such as the case in Greenland.

Change: Section 2.1 (Methodology; CMIP6 HighResMIP global climate models) lines 72 to 83, a new paragraph to comment on this limitation of the study.

Q3 – What statistical technique is being used to determine statistical significance of the trend lines? As far as I can tell, significance is determined if the range of slopes (determined by error bars) are entirely positive or negative. But what is used to determine the error bars? It is also worth commenting more clearly about the fact that most of the linear fits appear as significant (i.e., strong correlation, which would probably pass statistical tests), and the ‘significance’ testing referred to is with regards to the gradient of these fitted lines being significantly different from zero.

Response: Yes, we use linear regression to determine the trend in time. We now have our figures showing p-value to assess the statistical significance of trends. This is a clear oversight and has been updated on figures and within text. Our methodology has been updated with a data processing section to highlight the statistical tests used. Some of our results have now changed by including the p-values, so a critical point we thank the reviewer for noticing.

Change: Section 2.2 (Methodology; Diagnosing turbulence) lines 110- 113.

Q4 – I had difficulty ascertaining exactly where there were actual significant trends (either positive or negative) and suggest that Figure 11 could be modified to make this clearer. Perhaps a 4th column for the overall results (combining from the three models) which identify the mean trend (whether it be positive, negative, or close to zero) and the reliability or significance of this trend. Once this is done, I suspect there will be few locations with a significant trend – if this is true, this ‘null’ result still has a lot of value.

Response: Yes, we agree our end figures were confusing. From this advice, and from the other reviewer, we have decided to update all figures in the sub-continent analysis. Figure 12 (no longer 11) shows the average slopes, over all GCMs, for each region (as you suggest). These new figures are clearer and easier to gauge the overall trend in time.

Change: all of section 3.3 (Results; Multi-model MWT regional analysis) has been updated with new summary figures and updated text, lines 244 to 336.

Other minor comments.

Minor 1: Line 35. Suggest you cite Clark et al. (2000, J. Atmos. Sci., 1105-1131) here.

Response: This has been added.

Change: line 32 in Section 1: Introduction

Minor 2: Line 37. Suggest you cite Foudad et al. (2024, J. Geophys. Res., <https://doi.org/10.1029/2023JD040261>) here, noting the uncertainty in future trends due to model variations etc.

Response: Yes, this paper should have been included, this has been added.

Change: line 35 and 49 in Section 1: Introduction

Minor 3: Line 65-68. Unclear of purpose of these sentences. Could be removed or re-written.

Response: This has been updated.

Change: Lines 68- 69 in Section 2.1: Methodology

Minor 4: Line 110-115. Are these differences purely related to topography resolution or are they impacted by impact of model resolution on flow features

Response: You are right this is not clear. We assume by flow features, you mean the upper-level CAT projections or the low-level wind speed used in creating Ds. In Smith et al. (2023), which uses the same GCMs, model resolution did not have a large difference across CAT projected trends. The low-level wind speed may slightly differ but the real key attributing limiting factor is the terrain gradient threshold, as the indices are so dependent on it. We have updated this, and made it clear it is due to our methodology and may differ in a study that used NWP terrain gradients.

Change: Line 130 in Section 3.1

Minor 4: Line 142 and numerous occurrences elsewhere. The word 'propagate' seems like the wrong word to use here to talk about trends.

Response: Yes, agreed and fixed.

Change: Word propagate removed everywhere but introduction where it makes sense.

Minor 5: Line 177-182. Could be rewritten to be clearer. The two conditions could be described earlier in the methods section.

Response: Rewritten with a new paragraph in methods explaining this model agreement method.

Change: No longer in text

Minor 6: Line 274. Suggest cite Doyle et al. 2005 here

Response: This has been added.

Change: Line 258 in Section 3.3 Multi-model regional analysis

Minor 7: Line 290. 'previously touched' – unclear, suggest rewrite

Response: Yes, agreed and fixed.

Change: Removed from text

Minor 8: Line 335. 'if applicable' – unclear, suggest rewrite

Response: This has been rewritten.

Change: Removed from text

Minor 9: Line 347. 'tug of war...' - unclear, suggest rewrite

Response: This has been rewritten.

Change: Removed from this section as unclear, but use of tug of war used 306 but written in a better way to highlight the overall trend uncertainty.

Minor 10: Line 510. It would be useful to describe a more detailed comparison between your results and those published by Kim et al. How are trends different etc.

Response: Yes, this has been added at the end of our results section.

Change: Line 325-331 in Section 3.3 Multi-model regional analysis

Minor 11: Figure quality. Many of the figures were difficult to interpret as the axes ranges were often too large. For example, if the y-range of Fig. 8 was changed to -100 to -50, the data and trends would be easier to interpret.

Response: Yes, thank you for noticing this. We have left some figures with a wide y-axis to see the variation and spread of MWT projections for each indices, but we notice a few of our figures had an unnecessary wide range, so we have updated them.

Change: Across results section

Reviewer 2

Overall comments:

This study investigates global changes in the occurrence of strong-intensity mountain wave turbulence (MWT) using high-resolution GCM data. The use of relatively fine-resolution datasets to project future changes in MWT is interesting, and the focus on resolution dependence as well as regional and seasonal variations offers potential scientific value. However, several methodological aspects are insufficiently explained, and the figures showing the main results are difficult to interpret. Moreover, the discussion section does not clearly state the key findings and their broader implications for this topic. In its current form, the paper suffers from methodological uncertainty and a somewhat disorganized logical structure and thus requires substantial revision before it can be considered for publication. Detailed comments are provided below.

Response: We thank the reviewer for these comments and for their time spent reading and reviewing the paper. We have updated the structure, revised our methodology and figures.

Major comments:

1) Clarity of data: The description of the datasets and method is not sufficiently clear. It would be helpful to specify the vertical resolution of each dataset, particularly the number of vertical levels used in calculating the turbulence indices. The readability of the model data are also quite limited. Additionally, please provide proper citations for each model dataset used. In addition, the period of the data used should be described in the “Data and Methods” section rather than in the “Results” section.

Response: Thank you for your comment, we have updated the methodology to further explain the data used and our methods for calculating MWT. Due to the limited vertical resolution offered in global climate models, our focus for the comparison across our models is the horizontal resolution, with a patch of turbulent air very thin but wide. However, you make a good point that some of our indices use a vertical component, so in our appendix we have added the equation of each index and the height levels used in our analysis.

Change: Section 2.1; CMIP6 HighResMIP global climate models created due to this comment.

2) Methods: Before comparing the MWT results among the models, the authors should clearly explain how the MWT-prone regions were defined in a comparable manner. Regarding the criterion of “ds” (specifically, whether ds is nonzero or zero), the choice of threshold should account for model resolution to ensure that known mountainous regions can be adequately represented. Therefore, it may be necessary to adjust the ds threshold depending on the spatial resolution. For example, Kim et al. (2023) mentioned that the ds threshold was tuned to adequately capture typical MWT-prone regions such as the Rocky Mountains, Andes, and Himalayas. In addition, the Methods section in Kim et al. (2023) actually indicates the use of low-level wind speed rather than vertical velocity. The corresponding sentence in the manuscript should be revised to avoid any potential misunderstanding.

Response: Thank you for your valuable comment, we can see how our methodology is confusing. Ds is dependent on the individual model and index, and so is a

nonzero quantity, we have rewritten this section to make it clearer and to avoid any misunderstanding.

Change: Section 2.2 Diagnosing turbulence lines 93 to 95

3) Following the previous comment, Figure 1 shows that, for example, only parts of the western US mountain ranges are identified, while most of the expected mountainous regions appear to have no MWT computed. Can the authors provide evidence, such as case studies or other previous observational studies, that supports the reasonableness of the predefined MWT-prone regions used in this manuscript? Without such justification, it is difficult to assess whether the spatial filtering applied here appropriately captures areas that are susceptible to MWT.

Response: Yes, it is a limiting factor and arises due to our modelled terrain gradients, we have written a section about this in our methods to highlight why some prone areas are missing.

Change: Section 2.1 Methodology; CMIP6 HighResMIP global climate models lines 81-86 and Section 3.1- Two-decade comparison lines 130-135

4) Figure quality: The overall figure quality is too low, and the font sizes are too small. It is quite difficult to match the figures with the corresponding descriptions in the main text. Figures should allow readers to grasp the key results even without reading the full text, but in their current form, this is severely limited. Please improve the figure quality, increase font readability, and ensure that figure annotations are clearly written.

Response: Thank you, this has all been updated to be clearer.

Change: Throughout paper figures have been updated.

5) Figure presentation: The figures present the results from each model or region in a purely enumerative manner. Instead of displaying all results, it may be more effective to summarize the results in a more integrated and comparative format. The current figures are highly repetitive, and the small size makes it difficult to discern the key patterns. Please consider reorganizing the figures to highlight model similarities and/or differences more clearly.

Response: We agree that the results section needed reformatting and needed a summary to effectively display the results. We did not remove section 3.2 plots, as we like to highlight the variability in index, region and model. We have made them clearer, through slope trend colouring and reducing the wide y-axis spread for some continents. We have also included a summary plot (figure 10) at the end of section 3.2. For section 3.3, we have edited all figures, so they show the average over indices and GCMs, so they are clear overall multi-model trends. Our paper first discusses the model and index uncertainty for each continent, and seasonal differences in Section 3.2. Then next in Section 3.3, it says here is the multi-model average but please do take into consideration the uncertainty highlighted in Section 3.2. We think this is a better format for the paper.

Change: all result section has been updated.

Minor comments:

1) There is inconsistency in the spacing between numbers and units throughout the manuscript. Please ensure consistent formatting.

Response: This has been corrected. A formatting issue on latex.

Change: Throughout paper

2) Line 4: A space is missing between "EC-Earth-3P-HR" and "(36 km)". Please correct this.

Response: This has been corrected

Change: Line 4 of Abstract

3) Line 8: Instead of simply writing “101 years”, please specify the exact period. This will help readers clearly understand the time span of the current analysis.

Response: no longer in abstract

Change: not applicable

4) Line 10: The expression “regions in Asia” is vague. Please clarify which specific regions are included.

Response: This has been updated to list several counties.

Change: line 8

5) Line 18: The phrase “vertical and horizontal atmospheric waves” is awkward. It seems that the authors intended to refer to vertically and horizontally propagating atmospheric (internal) gravity waves. Please revise the wording accordingly for scientific accuracy and clarity.

Response: This has been corrected

Change: line 15

6) Line 20: Please clarify what “this boundary” refers to. The current wording is ambiguous.

Response: This has been corrected

Change: line 17

7) Line 20: The phrase “typically lower at the tropopause” may not always be true, as it depends on the background wind conditions. If critical-level filtering by background wind does not occur, GWs can propagate to higher altitudes. Please revise this sentence.

Response: This has been revised

Change: no longer in text

8) Line 22: "lane et.al" should be corrected to "Lane et al. (2009)". In addition, there are multiple typographical errors in in-text citations throughout the manuscript. Please carefully check and correct them.

Response: Thanks so much, this has been corrected

Change: Line 22

9) Line 22: Instead of simply saying "this height", it would be more informative to include an example altitude range.

Response: This has been corrected

Change: line 19

10) Line 28: Please clarify what is meant by "substantial shift in airspeed". Does "airspeed" refer to true airspeed? It is unclear how airspeed changes as a result of turbulence encounters.

Response: Yes, it was supposed to be related to the evidence that a quick height adjustment can lead to an aircraft stalling, but it's a bit of unnecessary information so has been removed.

Change: removed from text

11) Line 30: Convectively induced turbulence (CIT) includes not only turbulence within convective clouds but also near-cloud or out-of-cloud CIT occurring in cloud-free air.

Response: Yes, there's different definitions with it being under the umbrella of CIT and CAT, as occurring above from active convection but yes we agree that near-cloud turbulence (NCT) should be mentioned, and so it is now included within the introduction.

Change: lines 27 to 29

12) Line 31: Out-of-cloud CIT can also be undetectable by onboard radar. It can occur in cloud-free air due to convectively induced background enhancement or vertically and horizontally propagating convective GWs. Please revise the original statement to avoid any confusion.

Response: Removed from text in updates

Change: removed from text

13) Line 52: Please provide the full name of “PIREPs” when first introduced (e.g., pilot’s verbal reports (PIREPs)).

Response: Yes this has been added.

Change: lines 51-52

14) Line 54: Please add citations to support the statement “Many previous studies”.

Response: Yes, section has been updated.

Change: lines 49-50

15) Line 53: I could not find “Gill (2018)” in the reference list. Also, Sharman et al. (2014) compared PIREPs with automated aircraft data and discussed timing and location discrepancies of PIREPs. I’d suggest including this reference.

Response: This has been included.

Change: lines 54-55

16) Line 65: Please clarify whether the model data were analyzed on native vertical levels or isobaric levels.

Response: Yes, this has been included.

Change: line 110

17) Lines 82-83: According to Kim et al. (2023), low-level wind within a specific vertical range was used for the δs calculation. Please verify and clarify this point.

Response: Yes, this has been clarified.

Change: line 94

18) Lines 83-86: Please clarify how the consideration of terrain gradient is connected to linear theory. As linear wave theory is not limited to mountain waves, so the current statement may require modifications.

Response: Yes we were relating it to the fact terrain-gradient is a condition for applying linear theory, but we agree it's not very related and has been removed

Change: removed from text

19) Lines 83-84: The phrase "the wave of a gravity wave amplitude" is unclear. Please rewrite it for clarity. Moreover, Wolff and Sharman (2008) may not be the most appropriate reference for this statement.

Response: This section has been removed.

Change: removed from text

20) Lines 92-93: In Sharman and Pearson (2017), the values of 0.989 and 0.991 correspond to PODN, not AUC. The verification table does not show AUC values this high. Please correct this misinterpretation.

Response: Yes, that is a clear mistake thank you for noticing, this has been corrected.

Change: lines 104-108

21) Lines 97-98: Please clarify what the "median" and "maximum" values are calculated relative to (e.g., temporal, model, spatial distributions).

Response: This has been corrected

Change: lines 114 -116

22) Line 110: Please add a reference to support the statement about previous PIREP analyses.

Response: We have added Sharman et. al 2014 as it is using PIREPs over the US, and our point related to the fact no MWT is found over north America in our coarser models.

Change: lines 128-129

23) Section 3: The figures are very difficult to read. Please increase the font size and consider summarizing key results in a few representative figures, moving the rest to an Appendix.

Response: Yes, we agree that the figures needed to be revised. We have updated the figures in Section 3.2 to have clearer slope trends, and created a summary plot (figure 10). In section 3.3 we have updated all figures so they are clearer and summaries the multi-model MWT trends.

Change: All figures edited

24) Before analyzing MWT results by model resolution and model difference, please ensure that the predetermined mountainous regions defined for each model are comparable. The analysis should ideally be conducted over similar regions for a fair comparison.

Response: We assume this point is related to if ds is model dependent. Yes, this has been added by confirming in methodology. We think it is appropriate to compare as our ds is model dependent.

Change: lines 93 -95

25) Line 124: Please clarify the phrase "In order of most turbulent".

Response: This section has been rewritten so phrase no longer included.

Change: removed from text

26) Lines 130-131: The statement “time has no linear effect on MWT” is unclear. Please explain what is meant by “no linear effect”.

Response: This section has been rewritten so phrase no longer included.

Change: removed from text

27) Line 136: It seems this should refer to “Figures 2 and 3” rather than “Figures 2 to 5”, since Figures 4 and 5 correspond to South America and Europe. Please clarify.

Response: Yes, that has been updated

Change: checks made to make sure all figs numbers are correct

28) Line 145: Instead of stating “meteorological NH’s winter”, it would be clearer to specify the exact period (e.g., DJF).

Response: Yes, section has been updated.

Change: Line 152

29) Line 155: Please discuss the physical meaning of why MWT_HTG shows a significant decline while other indices exhibit increases. What mechanism could explain this discrepancy?

Response: A weakening of the HTG compared to other mechanism is likely associated with polar amplification and the weakening of the meridional temperature gradients. This has been added in appendix.

Change: lines 400-401