Review for

Evidence of gravity wave contribution to vertical shear and mixing in the lower stratosphere: a WISE case study

by

Umbarkar et al.

Summary:

In a case study based on a WISE aircraft measurement, Umbarkar et al consider the occurrence and relevance of gravity waves on the mixing in the lower stratosphere. The WISE measurements are used to identify gravity waves (GW) along the flight track, and to study how the tracers (CO and NO2) are mixed across the tropopause. Three NWP models (ERA5, IFS, ICON), with differing horizontal/vertical resolutions but also differing parameterizations of GW and turbulence, are used to study the meteorological conditions leading to the turbulence and mixing in the lower stratosphere.

The research topic, GW and how they influence the upper troposphere / lower stratosphere (UTLS) is certainly of great interest and asks for further investigations. This study contributes to this knowledge gap, however, the presentations lacks at many places clarity and I was partly confused (or simply could not follow) the arguments. I don't think that new analysis is needed, but the presentation must be substantially improved to make the study publishable

In the following, I will address in greater detail my concerns, and thus explain why I think that major revisions are needed.

Major concerns:

1. Clarity: There were several places in the manuscript where I could not follow the argument, or where the overall relevance of the statements remains unclear to me. I list a few examples, but encourage the authors in general to carefully improve the storyline and clarity of their arguments.

- L139-142: "we restrict ourself here to the presentation of gravity wave vertical fluxes of horizontal momentum in terms of distribution of probability density functions (PDFs). The vertical transport due to small–scale processes in the ExUTLS is investigated via absolute momentum flux (AMF)" -> this is a preamble to the following paragraphs, but ' in terms of distribution of probability density functions (PDFs)' remains rather unclear.

- L280-261: "All of these findings are potential indications that these air masses have been subject to a process which breaks the relation between CO, N2O, and Θ; one potential process which can lead these correlations is mixing" -> Referring to the relation between the three fields is unclear to me; this relationship should be much more clearly introduced in the preceding text.

- L269-271: “Particularly, the small transition between slopes observed in N2O spectra around 10−2 Hz (green and blue lines), where the slopes of w indicates turbulent energy source, and here, the N2O hints the turbulent behavior of small scales, e.g., those related to GWs, might be substantial to explain the dynamics in the lower stratosphere.” -> What small transitions? How does the N2O hints to the turbulent behavior of the small scales? Why do the slopes of w indicate a turbulent energy source? This is all difficult to understand. Please rephrase in easier way, and add some explanations.

- L289-290: “Due to distinct sampling frequencies and noise characteristics, separate filtering methods were applied to flight and model datasets to optimally isolate GW related perturbations” -> This remains rather vague? Please be more specific.

- L336-337: “Moreover, the large amplitude wave signatures pointing towards the strong upward motions across the tropopause, that could influence the vertical shear, and in turn turbulence generation in this region” -> Please rephrase in clearer way. Do you, essentially, want to say that GW of large amplitude are discernible at and across the tropopause and that they locally enlarge vertical wind shear, which in turn increases the probability of turbulence?

- L370-372: “ In continuation to turbulence analysis, this section addresses the occurrence of dynamic instability associated with GWs, particularly in the LMS, as well as the relation between the occurrences of GWs and shear perturbations in terms of 2D histograms.” -> Mentioning ‘in terms of 2D histograms’ confuses more than it helps in this preamble to the section.

- L396-400: Here, suddenly inertial instability and their association to GW is introduced. So far, inertial instability was not a key aspect of the study, and I was astonished them to appear here ‘out of the blue’. Possibly, this link between GW and inertial instability is important, but then I would expect it to be introduced and discussed already in the introduction.

- L433-434: “which further induce strong potential instability and ultimately result in the development of higher risk CAT” -> What do you exactly mean with ‘potential instability’? And what is ‘higher risk CAT’? Please rephrase in clearer way.

- L444-445: “One plausible explanation is that local flow deformation in ICON exhibit enhanced values during turbulent events in the LMS, partly compensating for weaker resolved shear, thereby causing higher abundance of TI1 in the vicinity of GWs.” -> difficult to understand!

- L448-451: “Overall, following the positive relation between small–scale shear and AMF in the LMS above the North Atlantic as discussed in Sect. 4.3, we can go even one step further. The relationship observed between the turbulence indices, TI1 and TI2, and GWs, leads us to conclude that GW play a role for the occurrence of CAT in our analysis region, as evidenced by the positive correlation between turbulence diagnostics and the AMF” -> The statement starts with ‘positive relation between small-scale shear and AMF’, and it ends with a ‘positive correlation between turbulence diagnostics and AMF’. This is confusing, at least to me, and I struggle to get the key message from the statement.

- L468-479: Switching back and forth between TI1 and TI2 is quite confusing, and at the end hindered me to get the key message of the paragraph. For example, first is mentioned that TI1 and TI2 ‘exhibit identical patterns’, but a few lines below it is ‘with wider spread out in TI2’…

- L510-511: “We note that this is key for the potential occurrence of clear air turbulence, as highly transient yet frequent mixing processes in the extratropical lowermost stratosphere” -> Is this sentence complete?

2. ERA5/IFS/ICON: Three models are included in the analysis, with different spatial and temporal resolutions, but also differing in the way how turbulence is handled. I am not completely convinced that keeping all three models in the study helps the storyline, or whether it is more confusing. Some points to consider: (i) The difference in the models is at several places in the manuscript highlighted, but it should more systematically be discussed in the methodology section; (ii) Why is only ERA5 discussed in section 3.2, and the other, higher-resolved IFS and ICON are not? (iii) At some places it is unclear whether the study is about a model intercomparison, or whether the physical processes are the focus.

Some of the conclusions remains somewhat vague. As an example, it is written in L387-389: “ Note here that the ERA5 and IFS shows quite similar relation when looked at the occurrences at the range of 10−3 to 10−2 (see the green colors), while the difference in ICON likely arise mainly from the occurrence of low Ri values (see Sect. 4.2). Nonetheless, ERA5 shows fewer occurrences of maximum shear than IFS. This could be due to its coarser horizontal resolution, which might lead to the limited representation of (partially resolved) GW spectrum in the UTLS.” -> This statement seems plausible, but remains also speculative.

I think it is worthwhile to study the case in different models, however, I would appreciate if the key process analysis is based on one model, and that all the sensitivity (model dependence) is handles in a separate section that is only dedicated to this model comparison. In this way, process study could more clearly be separated from intercomparison, and so the storyline would become easier to follow for the reader.

Finally, note also that some of the differences in GW/turbulence between ERA5, IFS and ICON might originate from differences in the large-scale evolution of the models. In particular, the evolution of ICON (in Figure 1 but also in Figure 7 & 9) seems to be somewhat different compared to ERA5 and IFS. This aspect should be discussed when comparing the GW and the turbulence indicators in the three models.

3. GW source and evolution: GW is the key topic of the study. Figure 6 nicely shows that a GW indeed is present and propagates from the mid-tropospheric levels to the tropopause, and from there further propagates with a different angle into the stratosphere. Given that this GW is key for the study, it would be good to know what the origin of the waves actually is. Since the WISE flight is over Greenland, I assume that the GW source are topographic flows?! In short, I would suggest to more clearly investigate the source of the wave. Furthermore, it would also be interesting to see, based on the models, how spatially and temporally persistent the wave is. What is its lifetime? Does it change amplitude, propagation angle over time?

Additionally, the link between GW and vertical wind shear is highlighted several times in the manuscript. I fully agree that GW locally modify the background flow, and so can lead to enhanced vertical wind shear, with possible triggering of Kelvin-Helmholtz instability. However, vertical wind shear is only one part in the Richardson number, the other one being vertical stratification, which has a stabilizing effect on the flow. Hence, I wonder whether the modification of the stratification by GW activity must also be considered. So, why not more systematically consider Richardson number (with wind shear and stratification included), instead of only wind shear.

Next, I do not clearly understand how the absolute momentum flux (AMF) is used in the line of arguments. I see that it is a metric to quantify the momentum transport due to gravity waves. What its specific role in the discussion remains somewhat unclear. I see that there is a gravity wave associated to the WISE flight, that it can be characterized with respect to the propagation direction (as in Figure 5), and that the gravity waves leads to regions of enhanced vertical wind shear, which in turn lead to turbulence due to KHI. But why is the AMF really needed in the overall storyline? It is stated that in L390-391: "Overall, these results pinpoints the crucial role of GW-induced shear to the potential turbulence occurrence, and subsequently representing the important role of GWs to transporting the momentum across the tropopause". But why this is important for the storyline, remains unclear to me. In this sense, I also do not clearly understand what we can learn from the AMF-S2 plots in figure 8 and, later, in figure 11 from the AMF-TI1/2 plots. Especially, in figure 11 the distributions of the two turbulence indicators (TI1 and TI2) look very similar to me.

4. TI1 vs TI2 & introduction: The turbulence indicators TI1 and TI2 are compared at several places. In L191-194 it is written: "The inclusion of CVG in TI2 offers an advantage over TI1 by capturing the small–scale flow features such as those associated with GWs or upper level frontal structures. These region often exhibit enhanced shear and convergence making TI2 more sensitive to turbulence driven by mixing processes". Is there a reference that supports this statement about TI1 and TI2? I am not completely convinced that the difference between TI1 and TI2 should 'easily' be used to separate between more GW-related turbulence and less GW-related. At least, I would like to see this aspects to be discussed in the introduction.

This may be lead to a deficiency of the introduction. Turbulence indicators (ZI1 and TI2), vertical wind shear, absolute momentum fluxes (AMF) could also be discussed in the introduction. So, for example, in L176 it is written: "While Ri captures dynamic instability, it may not fully represent turbulence under complex flow conditions. To account this limitation, we also employ the Ellord-Knapp turbulence indices, TI1 and TI2, which are empirical diagnostics specifically developed to identify CAT.". Hence, since all these indicators are extensively used in the manuscript, a careful introduction to them from the beginning would be appropriate.

5. Relevance of synoptic situation / target region: The study first uses the WISE flight data to identify turbulence segments along the flight track and then to see whether GW are associated (or leading) to this turbulence. The discussion of the WISE measurements is mostly confined to sections 3.2 and 3.3, whereas the later sections use the NWP data to get a more comprehensive picture of the GW activity along the WISE flight. In this sense, the vertical cross-sections in Figure 6 are very helpful, and also the figures 7 and 9 are important to get a more comprehensive figure. I wonder, however, whether all figures should focus (zoom in) more to the region explicitly around the WISE flight. At them moment, part of the discussion deals with features (e.g., tropopause fold) that are farther away from the WISE flight path. Possibly, the study gets some more focus if it restricted more strongly at and around the flight.

To me, it is also not so clear why the turbulence/GW analysis is performed ‘within a baroclinic system. How relevant is it really for the analysis that the flight and GW encounter happened within a ridge (baroclinic system). Maybe, the motivation is that these regions are often sources of GW? This is hinted to in the introduction: (L30-33): “In the extratropics, baroclinic waves represent an important, albeit less well understood source of GWs. Surface fronts and upper level jet streams associated with baroclinic wave development generate GWs, primarilythrough spontaneous imbalance, i.e., deviations from balanced flow trigger wave emission (Plougonven and Zhang, 2014; Zhang et al., 2015b). Regions of baroclinic instability, particularly along jet streaks and frontal zones, are hotspots for non-orographic GW activity. If so, I would expect that the origin of the GWs in this study are more clearly related to these baroclinic processes? For example, the origin of the GW remains rather unclear. It ssems to originate from lower-to-mid tropospheric levels, i.e., possibly of orographic origin and not directly related to a baroclinic wave.

Specific comments:

- L11: “Further analysis of turbulence diagnostics suggests” -> Please be more specific!

-L85-83: This sentence tries to summarize in one statement the goal of the study, which leads to a complicated sentence structure. Please rephrase in easier way.

- L132: A warm-conveyor belt is mentioned here.Okay, but the WCB is mentioned gor the first time and comes here somewhat ‘out of the blue’.

- L143: “Note that,.... scale GW” -> Incomplete sentence?!

- In Figure 4a, the w spectra flatten out towards low frequencies, while this is not the case for potential temperature and the two tracers. Why does w behave differently compared to potential temperature?

- L342: Minor detail, but you are referring to upper and lower panels *before* referring to the figure

- L358: Why ‘in principle’?