Authors response to egusphere-2025-5142
"Evidence of gravity wave contribution to vertical shear and mixing in the lower stratosphere: a WISE case study"

Umbarkar et al.

We thank the referees for their comprehensive comments on the manuscript. We have carefully addressed all points and revised the manuscript accordingly. As the manuscript has undergone major revision, we first list the central changes in the revised version of the manuscript to give an overview.

Reviewer comments are given in *italic*, replies in standard and corresponding revised text passages in *blue*.

Central changes in the revised manuscript

– We have modified the title from "Evidence of gravity wave contribution to vertical shear and mixing in the lower stratosphere: a WISE case study" to "**Evidence of gravity wave contribution to vertical shear and mixing in the lower stratosphere**".

– We want to note that the main goal of this study is to determine whether previous findings based on idealized simulations (Umbarkar and Kunkel, 2025) can be confirmed in forecast and reanalysis data. A dedicated focus is on the relation between shear and turbulence occurrences in regions of gravity wave activity in the region of a baroclinic wave. We decided to use a case study of a baroclinic wave over the North Atlantic. This case study has several benefits: (i) observations exists which show the existence of a gravity wave related to the baroclinic wave and mixing occurring in the vicinity of the gravity wave; (ii) a comprehensive model forecast dedicated for this case with ICON. In addition to ERA5 reanalysis data and IFS forecast data, the latter allows to study this case also in a comprehensive ICON model forecast which has been dedicated for this case study. It also allows us to uniquely study the relation between gravity waves, shear and turbulence in various model resolutions. We thus have a suite of comprehensive models to test whether our predictions from the idealized world hold true in the "real" world (real in terms of comprehensive models and during a baroclinic life cycle over the North Atlantic).

– Following the suggestions by two referees to enhance the clarity of the manuscript and to separate the analysis between ECMWF and ICON products, we changed the flow of result sections. We start with the observations and present the synoptic situation based on ERA5. We then introduced the gravity wave, turbulence and shear diagnostics first for ERA5. After that we put these into perspective, once compared to IFS data, and then for the ICON simulations. We thus gain insights explicitly how ERA5 sees the situation, then a comparison for an ECMWF product with finer horizontal grid resolution. Finally, the ICON simulations, despite their different lead time, provide a unique insight into the situation, since we have the same situation in three model domains (with and without convective parameterization) and also in relation to our previous results from idealized ICON model simulations. Comparing ICON and ECMWF also shows whether there are general differences between a model with a non-hydrostatic and a model with a hydrostatic dynamical core and with varying parameterizations.

– **Add-ons**:
(i) Discussion based on source mechanism of observed gravity waves is now included in the section 4.
(ii) A brief introduction of turbulence/CAT indices is provided in part of introduction section.
(iii) We also discuss the applicability and limitations of hybrid approach used to compute GW momentum fluxes and turbulence diagnostics.

Response to Reviewer 1

Major comments:

1. Introduction / representation of turbulence:

Comment: *It should be stated more clearly that Kelvin-Helmholtz instability and turbulence are not explicitly resolved in ICON, IFS or ERA5 even if the gravity waves themselves are; instead, these processes are represented through subgrid-scale parameterizations. It would therefore be useful to analyse the contributions from these subgrid-scale processes explicitly. This distinction should be discussed, and where possible the relevant tendencies should be diagnosed. For example, in ERA5 these can be obtained from MARS using short forecasts on model levels, with fields such as the time-mean eastward and northward wind tendencies due to parameterizations.*

» Thank you for pointing this out. The discussion on representation of turbulence is now added to section 2.5. In this study, our analysis focuses on large-scale diagnostics of turbulence (*Ri*, shear) to assess the dynamical conditions conducive to turbulence and its relation to gravity wave activity. We state now more clearly that Kelvin-Helmholtz instability and turbulence are subgrid scale processes in the models used in this study and are only indirectly diagnosed via the large-scale diagnostics. We appreciate the suggestion to analyse additional parameters (tendencies) from the ERA5 forecast. We will keep this suggestion in mind for a future analysis. However, for this study we decided to discuss the turbulence diagnostics based on the resolved wind field to be able to determine the differences between the models, since the tendencies are neither available for the IFS, nor for ICON. We include in the manuscript:

It should be noted that Kelvin-Helmholtz instability and turbulence are not explicitly resolved in the used reanalyses as well as forecast even if the gravity waves themselves are. Instead, these processes are represented through subgrid-scale parameterizations. Our analysis is based on diagnostics of the resolved fields in ERA5, IFS and ICON. These diagnostics highlight *regions which have conditions favourable for the occurrence of turbulence. We therefore regard these regions as regions of potential turbulence occurrence based on the resolved fields of the respective model.*

2. Origin of the gravity waves:

Comment: *The source of the turbulence-inducing GWs is unclear from the analysis and discussion. Do the authors attribute these waves primarily to baroclinic instability, or to flow over orography? From the figures, flow over orography appears at least as plausible as baroclinic instability, particularly given that the GW signal occurs near the transition between ocean and Icelandic topography. The presence of a front does not necessarily imply GW generation via baroclinic instability; instead, the front may simply provide favourable background wind conditions allowing surface-generated GWs to propagate into the UTLS. It would therefore be useful to extend Fig. 6df down to the surface and to explicitly mark the underlying topography.*

» We agree. We extended Fig. 6 in the original manuscript to show the entire vertical range down to the surface. From this it seems that the flow associated with the baroclinic wave over the topography of Iceland is the source of the GW. This information is added to the manuscript, but we note that the source of the GW is not the primary focus of our analysis. The modified figure is shown below.

The gravity wave field is primary noticeable over the Icelandic highlands. The flow associated with the baroclinic system over the topography of Iceland leads to the emission of the gravity wave. Its signal is already evident in the troposphere and changes its appearance in term of wavelength over the tropopause. This change is expected due to changes in stability across the tropopause. The flow over orography appears to be source mechanism for the generation of GWs, particularly given that the GW signal occurs near the transition between ocean and Icelandic topography.
Moreover, there is presence of upward propagating wave signatures with short horizontal wavelength that are excited in the upper troposphere above the location of troposphere jet. This indicates that the GWs that are originated from surface orography further emitted away from their source region and propagate upward toward the lowermost stratosphere.

3. Comparison of ICON with IFS and ERA5:

Comment: *The ICON forecasts analysed here are at lead times of 3-4 days, whereas the IFS forecasts are at lead times of 0-1 day and therefore much closer to the analysis/ERA5. This difference should be explicitly acknowledged when comparing the two models. A major reason for the discrepancies between ICON and IFS is likely the different background flows, as ICON will have drifted further from the analysis. This is already evident in the large-scale fields shown in Figs. 1, 6, 7, 9, 12, and 13, and the discussion of these figures should reflect this. For a more direct one-to-one comparison, the two models should ideally be compared at the same forecast lead time.*

» We thank the reviewer for noticing this important point. We have now explicitly clarified the difference in forecast lead times between ERA5/IFS and ICON and revised the manuscript accordingly, separating the analysis of ECMWF products and ICON output. These changes are detailed in central changes.

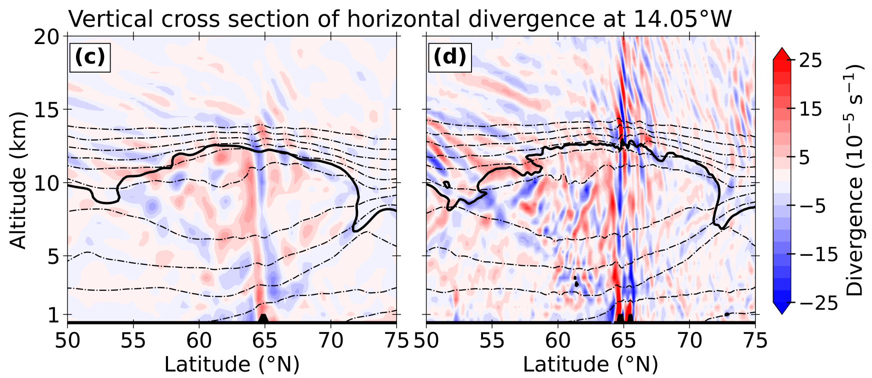

Figure 1. Vertical cross sections of horizontal divergence at 14.05° W is for c) ERA5 and c) IFS datasets. The solid black line represent the 3.5 PVU as a dynamical tropopause whereas dashed lines in lower panel represent the potential temperature starting from 280 K (bottom) to 380 K (top) with 10 K increments. (Figure 6 (lower panel) in original manuscript).

Comment: *In Sect. 2.4, please also discuss the vertical resolution of ICON in the analysed region, as this is important for comparing vertical wavelengths with IFS (L331-332).*

» We now added vertical resolution of ICON in the analysed region in the manuscript section 2.4.

In the vertical the ICON simulations have 90 levels from the surface up to 23 km altitude with a vertical grid spacing of roughly 250-300 m in the UTLS.

Comment: *Regarding Fig. 6, it is notable that ICON and IFS show very similar GW amplitudes, even though ICONs horizontal resolution in this region is about three times finer than that of IFS (3.3 km versus 9 km). One might expect ICON to produce larger-amplitude resolved GWs, consistent with the statement in L323 that nominal resolution controls resolved GW amplitude. This does not appear to be the case in Fig. 6. This could again be related to differences in the background flow. However, later in Figs. 8 and 11, ICON does show larger GW momentum fluxes than IFS, as expected given its higher resolution. Please comment on this apparent inconsistency. I am also somewhat confused by the ICON figures in the supplement and the discussion on L409-410: It was my understanding that you are analysing the nested simulations for ICON throughout the manuscript with the region of interest having 3km horizontal resolution. Is this not the case? If not you need to clarify section 2.4.*

» The reviewer raises an important point regarding the comparison between ICON and IFS GW amplitudes. We note that the ICON results analysed (in the original manuscript) correspond to the ∼13 km horizontal resolution simulation, rather than the ∼3 km horizontal resolution simulation. Because the analysed ∼13 km ICON simulation has a horizontal resolution closer to that of IFS, the similar GW amplitudes shown in Fig. 6 were therefore consistent. This may have caused confusion in the original manuscript. **We have now included analysis of all three domains of ICON simulation in the revised manuscript.** For more details, see response in central changes.

4. Scale separation in the UTLS (L159-161):

Comment: *The authors state that their scale-separation approach follows commonly used GW separation methods and cite several previous studies. While this is indeed true in the stratosphere, where there is a clear scale separation between planetary waves and gravity waves (and where studies such as Gupta et al. and Stephan et al. apply these methods), the situation is less clear in the UTLS. In this region, the separation between GWs and other mesoscale structures is more ambiguous. Please comment on the applicability and limitations of this approach in the UTLS.*

» We assume the reviewer is referring to the wavenumber-based scale separation, which is consistent with Gupta et al. (2024); Stephan et al. (2019). We agree that the separation between GWs and other mesoscale structures was noted to be ambiguous in the UTLS. However, Wei et al. (2022) suggested that the using divergent wind as a proxy provides more reliable estimates of GW momentum flux, particularly in the upper troposphere. Also, recent work by Zhang et al. (2025) has evaluated similar scale-separation approaches at 200 hPa, i.e., at comparable altitudes to those considered here. Moreover, our analysis primarily focuses on the lowermost stratosphere where scale separation is comparatively expected to be more robust. We have clarified these applicability and limitations in the revised manuscript.

Recent study by Zhang et al. (2025) suggest that such hybrid scale-separation approaches can be applied at lowermost stratosphere levels (e.g., around 200 hPa). Moreover, Wei et al. (2022); Zhang et al. (2025) further suggest that the use of divergent wind as a proxy for separating GWs provides more reliable estimates of GW momentum flux, particularly in the upper troposphere. In this study, the focus is primarily on the lowermost stratosphere, where the scale separation is comparatively more robust, and the limitations of the method are therefore reduced.

5. Figure 1 (PV field):

Comment: *Are regions of negative potential vorticity present, as suggested by the colour-bar range? If so, could these be highlighted using a distinct colour rather than shades of blue? Negative PV is dynamically significant and would indicate regions of instability.*

» Yes, regions of negative potential vorticity are now highlighted in magenta color, e.g. in figure 2. Also negative PV values are often related to cloud processes.

6. Wavelength estimates (L302-303):

Comment: *Please state the estimated vertical and horizontal wavelengths explicitly here, so that they can be clearly compared with model-resolved scales later in the paper.*

» In this study, hodographs are drawn from the flight data using profiles at multiple times at one flight altitude, 12.5 km. The limited altitude region here means that only one or two vertical wavelengths are sampled. Under the assumption of quasi-stationary wave propagation, this approach can provide indirect information on wave characteristics in horizontal, yet the estimation of vertical wavelength adds some uncertainty to the analysis.

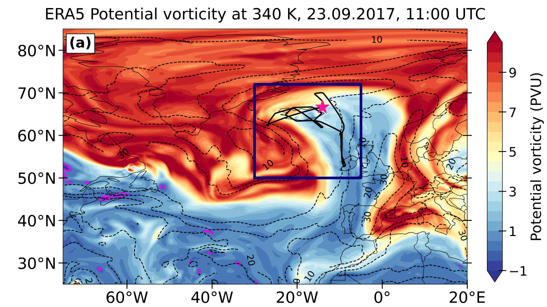

Figure 2. Potential vorticity distribution (in PVU) at the 340 K isentropic surface on 23.09.2017 at 11:00 UTC for ERA5 reanalysis. The dashed lines show the horizontal wind (in m s^{-1}) on 340 K. The thick black line shows the flight path of RF05 and the star marks the aircraft position at 11 UTC. Magenta color shows the negative PV values.

The text "Moreover, as we will show later in the analysis of the model data, these profiles give an estimates of horizontal and vertical wavelength that are similar to those determined from the models" is modified to

Moreover, as we will show later in the analysis of the model data, these profiles give an estimates of at least horizontal wavelength that are similar to those determined from the models. The range of horizontal wavelength estimated is roughly 200-250 km.

7. Interpretation of Fig. 6d-f and GW breaking:

Comment: *If irreversible mixing is attributed to GW-induced KH instability, why does the GW appear to continue propagating beyond the UTLS in Fig. 6d-f? This would suggest that the GW has not undergone saturation or breaking, which would normally be required for GW-induced CAT as argued in Sect. 4.4. Alternatively, this may not be the intended interpretation. The discussion in L361-368 is somewhat unclear and appears to suggest that only reversible GW propagation contributes to shear generation. Please clarify this point.*

» The superposition of GW-induced shear with the background flow can locally enhance the total vertical shear, thereby reducing the Richardson number and favoring the onset of Kelvin-Helmholtz instability. In such cases, wave breaking and turbulence occur intermittently and locally, while other parts of the wave packet or additional wave components can continue to propagate vertically and transport momentum (e.g., Fritts and Alexander, 2003; Tsuda, 2014; Kunkel et al., 2014; Achatz et al., 2024). Note that the KHI is based on the ERA5 field along the flight path. The low values of Ri suggest that KHI might occur and serve as the process which leads to the irreversible mixing. The irreversible mixing is diagnosed from the observed quantities and chances are high that the gravity wave is related to this mixing. Locally, not only the tracer show this sign for mixing but also state variables highlight that turbulence occurs (see Fig. 4). As noted by the reviewer, this does not lead to the total breakdown of the GW in the models but to a further propagation into the stratosphere. This is an interesting fact and the question remains whether this is the lack of the models to fully represent this process or whether as stated before that parts of the wave packet or additional wave components continue their propagation.

8. Figure 7 and GW contribution to shear:

Comment: *Since the authors argue that the GW contribution to shear is important for mixing, it would be very useful to show an analogue of Fig. 7 for S', i.e. the component of the shear associated with gravity waves. This would help to quantify how much GWs actually contribute to the shear and would also allow one to verify that this contribution is confined to regions where GWs can be clearly identified.*

» Done! We included the distribution of $S^{2'}$ to appendix.

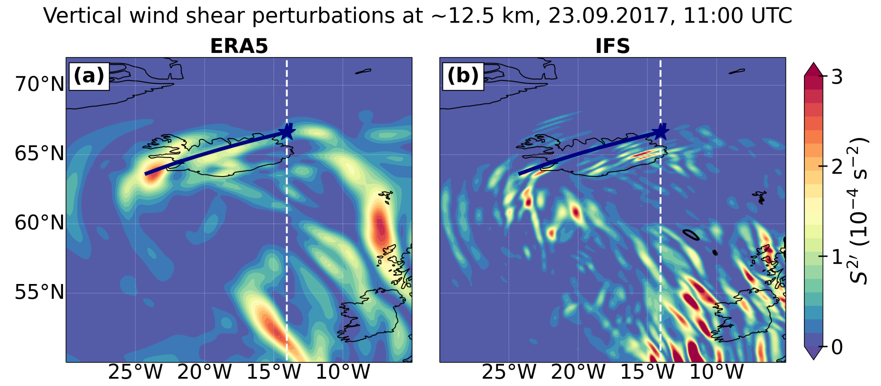

Figure 3. Distribution of vertical wind shear perturbations $S^{2'}$ at 12.5 km altitude for (a) ERA5 and (b) IFS.

9. Section 4.5 and momentum-flux deposition:

Comment: *To complement Figs. 12 - 13, it would be useful to show the vertical gradient of the resolved GW momentum flux without any additional modulation. One would expect momentum-flux deposition in the region where GW-induced KH instability, CAT, and turbulence occur. Is this observed?*

» To further assess momentum deposition, we now examined the vertical structure of the GW absolute momentum flux. In Figure 4, the largest magnitudes of AMF appear to be concentrated in the upper troposphere. ERA5 exhibit a pattern similar to IFS, whereas, ICON exhibits more localized but intense AMF that extend upward into the lower stratosphere (Fig. 4c). The AMF exhibits high magnitudes and sharp gradients in the vertical in regions of strong shear $S^2 \geq S_t^2$ and potential turbulence, particularly in the lowermost stratosphere. This figure has been added to the appendix.

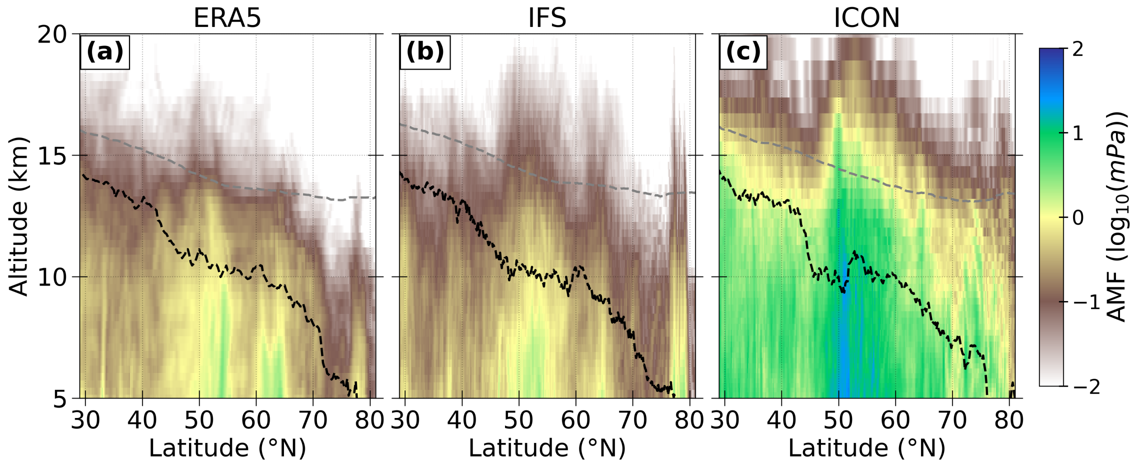

Figure 4. Zonal mean resolved GW absolute momentum flux (AMF, $log_{10}(mPa)$) for (a) ERA5, (b) IFS and (c) ICON.

Minor comments from the marked-up PDF:

Comment: *L120: Could you please clarify what you mean by 'derived'? Do you mean the forecasts have been initialized from the operational analysis? Or does this refer to the 137 levels, in which case these are not really derived from operational analysis data, but are the same vertical resolution used in analysis and in forecasts.*

» We intended to refer to the standard model configuration, where the 137 vertical levels are used consistently in both the operational analysis and forecast. The sentence modified to

... and incorporates 137 vertical levels which are the same vertical resolution used in analysis and in forecasts.

Comment: *"Particularly, the small transition between slopes observed in* N_2O *spectra around* 10^{-2} *Hz (green and blue lines), where the slopes of w indicates turbulent energy source, and here, the* N_2O *hints the turbulent behavior of small scales, e.g., those related to GWs, might be substantial to explain the dynamics in the lower stratosphere. This sentence is not at all clear. What do the authors want to say here?*

» Modified as:

Moreover, N_2O *and* CO *also have a slope close to -5/3 for intermediate frequencies i.e., for smaller wavelengths. From* 10^{-2} *Hz onward to larger scales, the* N_2O *spectra follow a -5/3 slope which suggests that small-scale processes, potentially those related to GWs, contribute to turbulent mixing in the lower stratosphere.*

Comment: *What is the correlation in these? Could you provide some numbers please.*

» The shown Fig. 8 (in the original manuscript) is not targetted for the correlation, instead, it is the colocation or co-occurrences of GWs and shear perturbations under the potential turbulent episodes.

Comment: *Do you mean shorter vertical or horizontal wavelength GWs? Usually very short vertical wavelength gravity waves carry less momentum than long vertical wave length gravity waves because momentum flux goes like k/m (see eqn 7 in Ern et al: https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2004JD004752)*

» Right. We meant GWs with relatively short horizontal but large vertical wavelengths. Horizontally propagating GWs with large horizontal wavelengths tend to dominate the horizontal derivatives of momentum fluxes, whereas, upward-propagating large-amplitude GWs (i.e., with large vertical wavelength) carry large momentum, which then contribute to the vertical derivatives of the wind components.

Comment: *What is the relative contribution of this GW-induced shear to the total shear? This is not really discussed anywhere and would be an important to assess.*

» The central focus of the study is to understand the role of small-scale processes in shear generation. While a direct quantification of the relative contribution of GW-induced shear to the total shear is not explicitly performed, its influence is indirectly assessed through the shear perturbation field ($S^{2'}$, see also Fig. 3) and analysis of momentum flux filtered for regions with high shear occurrences (Fig. 13 in the original manuscript). A more detailed quantitative decomposition is beyond the scope of the present study but represents an important direction for future work.

Comment: *L563 "flow imbalance" This has not been made fully clear in this paper. Are the authors very sure the waves are not of topographic origin?*

» Please see our response to comment 2. The sentence now read as

...inertia GWs associated with the flow over orography...

Response to Reviewer 2

Major comments:

1. Clarity:

Comment: *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.*

» we meant "in terms of occurrences." The text has been modified to

In order to understand momentum transport due to small scales and occurrence of GWs, we present the gravity wave vertical fluxes of horizontal momentum in form of probability density functions (PDFs).

Comment: *L260-261: All of these findings are potential indications that these air masses have been subject to a process which breaks the relation between CO, N_2O, 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.*

» This relation between CO, N_2O, and Θ has been stated on line 249-251 of the original manuscript. For clarity, we have now included the following statement.

Note that irreversible mixing can be followed from the interrelation of CO, N_2O and potential temperature. Due to the chemical life times of the tracers we can infer that N_2O is almost constant in the troposphere with a sink in the upper stratosphere and thus a minor negative gradient with altitude. The same is true for CO, with the difference that the mixing ratios of CO are less constant in the troposphere and the gradient in the stratosphere is stronger due to stronger chemical sink. However, their tracer-tracer relation is positive. While both tracers decrease with altitude, potential temperature increases, which in turn gives specific background relation of these three species. A deviation from the stated background condition is then only possible in the presence of non-conservative processes which also leads to mixing. This is the case here, since we see ...

Comment: *L269-271: Particularly, the small transition between slopes observed in N_2O spectra around 10–2 Hz (green and blue lines), where the slopes of w indicates turbulent energy source, and here, the N_2O 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 N_2O 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.*

» We have clarified this in more details and added the following text.

The w spectra shows a slope of -3 for the smaller wavelengths (i.e., for frequency <0.8 Hz), in agreement with Lachnitt et al. (2023). For higher frequencies (shorter wavelengths), the increase of power spectra to 0.5 Hz could be indication of turbulent processes below the synoptic scale. Overall, the w and Θ *spectra follows the slope close to* $k = -5/3$*, can be related to three dimensional isotropic turbulence. This is an indication of dynamic processes on or below the mesoscale affecting the flow. Moreover,* N_2O *and* CO *also have a slope close to -5/3 for intermediate frequencies i.e., for smaller wavelengths. From* 10^{-2} *Hz onward to larger scales, the* N_2O *spectra follow a -5/3 slope which suggests that small-scale processes, potentially those related to GWs, contribute to turbulent mixing in the lower stratosphere.*

Comment: *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.*

» Modified to

Note that, due to distinct sampling frequency and noise characteristics in the measurements, the flight data is filtered using a 4^{th} *order Butterworth bandpass filter (Butterworth, 1930), while the model data is processed using the Helmholtz decomposition approach described in Sect. 2.5, in order to isolate GW related perturbations.*

Comment: *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?*

» We meant

Horizontally propagating GWs with large horizontal wavelengths tend to dominate the horizontal derivatives of momentum fluxes, whereas, upward-propagating large-amplitude GWs (i.e., with large vertical wavelength) carry large momentum, which then contribute to the vertical derivatives of the wind components. This, in turn, leads to higher shear values and lower Richardson numbers, that increases the probability of turbulence.

Comment: *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.*

» Removed in the revised text.

Comment: *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.*

» Right. Thanks for pointing it out. Here, we refer more generally to dynamical instability. A detailed investigation of inertial instability is rather beyond the scope of this study. We replaced it with "dynamical instability" in the revised text.

Comment: *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.*

» the word "potential" throughout the manuscript is referred to " potential for turbulence occurrence or the potential for dynamic instability to occur" as $Ri_c \leq 1$ is considered. From these we can only infer that the background conditions are conducive for turbulence. this is revised in the manuscript (L199). We erased "higher risk". This was placed incorrectly in this sentence.

Comment: *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!*

» This part is fully modified in the revised manuscript. See central changes.

A possible explanation is that, in ICON, enhanced local flow deformation during turbulent events in the LMS partly compensates for weaker resolved shear, but leading to a higher occurrence of TI1 in regions influenced by GWs.

Comment: *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.*

» Modified to

Overall, following the positive relation between small-scale shear and AMF in the LMS above the North Atlantic (Sect. 4.3), we can go even one step further. The co-occurrence of enhanced AMF with both increased shear and elevated turbulence indices, TI1 and TI2, leads us to the conclusion that GW play a role for the occurrence of CAT in our analysis region.

Comment: *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...*

» The entire section 4.5 has been modified in the revised manuscript. See central changes for more details.

Comment: *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?*

» Modified to

We note that this is key for the potential occurrence of clear air turbulence, which manifests as highly transient yet frequent mixing processes in the extratropical lowermost stratosphere.

2. ERA5/IFS/ICON:

Comment: *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 103 to 102 (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.

» Thank you for the suggestions. This is also pointed out by another referee. We have revised the manuscript accordingly by presenting the analysis of the key processes based on ERA5, and all the sensitivity (model dependence) handled in a separate section. In this context, we separated the discussion of ECMWF products and ICON simulations throughout the manuscript. We hope that this modifications addresses the issues raised. Please see central changes (and revised manuscript) for more details.

3. GW source and evolution:

Comment: *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.

» We thank the reviewer for this suggestion. To address the first part of comment, we initially show the evolution of GW fields and the key processes using ERA5 data in the revised manuscript. Regarding the origin of GWs, this issue was also raised by other referee. Upon detailed investigation, we attribute the source of GWs to be flow over orography. These GWs further altered in shape at the tropopause due to the background conditions. We extended Fig. 6 to the surface and included following text in the manuscript.

The GW field is primary noticeable over the Icelandic highlands. The flow associated with the baroclinic system over the topography of Iceland leads to the emission of the GWs. Its signal is already evident in the troposphere and changes its appearance in term of wavelength over the tropopause. This change is expected due to changes in stability across the tropopause. The flow over orography appears to be source mechanism for the generation of GWs, particularly given that the GW signal occurs near the transition between ocean and Icelandic topography.
Moreover, there is presence of upward propagating wave signatures with short horizontal wavelength that are excited in the up-

per troposphere above the location of troposphere jet. This indicates that the GWs that are originated from surface orography further emitted away from their source region and propagate upward toward the lowermost stratosphere.

» Regarding the second part of comment, we agree that the modification of vertical stratification by GWs is possible. The role of GWs in modification of stratification has been discussed in previous studies (e.g., Kunkel et al., 2014; Kaluza et al., 2019, 2021). Here, our intention is to clearly address the role of GWs to generation of enhanced shear (which further result in the formation of tropopause shear layer). We also note that analysis of Richardson number with both wind shear and stratification is already included in section 4.2 and 4.3 (Fig. 7 and Fig. 8). The analysis of potential turbulence provided in order to understand how the GWs and GW-induced shear results in turbulence occurrence. Following the reviewer's suggestion, we now included a brief discussion on this point in relevant section.

» Concerning the third part of your comment, the use of absolute momentum flux is a pragmatic choice for a simple reason that *one would expect momentum-flux deposition by GWs in the region where GW-induced instability and turbulence occur (Fritts and Alexander, 2003).* We now included this line in the relevant section.
Figure 8 (AMF-$S^{2'}$ pair plot) and, figure 11 (AMF-TI1/2 pair plots) are implied to represent the relation between GWs and shear perturbations during the potential turbulent event. The examination of AMF-S2 pairs simply show that large AMF values, and hence GWs, correlate with large shear perturbation values. This tells that GW induced shear is an important driver that lead to turbulence generation in the subsequent region. Next, AMF-TI1/2 pairs show the relation between GWs and CAT indices, where there is role of GWs to the development of CAT. It turns out that, GWs on smaller scale that induces shear can play as a key driver of the onset of CAT locally. We believe that this is an important supportive implication by not only showing the dynamical and physical evidence of co-location of GWs, shear and turbulence, but also provide statistical evidence for the formation and **co-occurrences** of the shear and potential turbulence that arises in association with GWs.

4. TI1 vs TI2 & introduction:

Comment: *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 (TI1 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.*

» This is an important point, and we thank you for pushing for greater clarity on it. In our analysis, the inclusion of the convergence (CVG) term in TI2 provides additional sensitivity to local flow deformation and convergence associated with small-scale perturbations. Since turbulence diagnostics are computed from **small-scale perturbation** fields, we expect TI2 to be more responsive to features such as GW-induced convergence, localized deformation and shear. However, we acknowledge that this does not uniquely isolate GW-related turbulence, and other processes (e.g., frontal dynamics) may also contribute. We have therefore revised the text to clarify this interpretation and included a brief discussion in the introduction. This now reads:

In addition to vertical shear-based diagnostics (e.g. gradient Richardson number), empirical turbulence indices such as TI1 *and* TI2 *are widely used in CAT identification (Ellrod and Knapp, 1992; Sharman et al., 2006; Jaeger and Sprenger, 2007).* TI1 *is a semi-empirical diagnostic; although it neglects the influence of the stratification, it has been shown to perform well within the hierarchy of turbulence predictors (Sharman and Pearson, 2017). Both Ellord indices combine vertical wind shear and horizontal deformation, while* TI2 *additionally includes a convergence term, enhancing its sensitivity to localized flow features associated with small-scale perturbations.*

5. Relevance of synoptic situation / target region:

Comment: *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, primarily through 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 seems to originate from lower-to-mid tropospheric levels, i.e., possibly of orographic origin and not directly related to a baroclinic wave.

» As mentioned in central changes, the primary aim of this study is to determine whether the findings from the idealized cases (Umbarkar and Kunkel, 2025) are also evident under more realistic conditions. Note again that the WISE flight took place in a baroclinic system which is our system of reference. In this context, the flight data is used to confirm the presence of GWs and to identify GW-induced mixing over the Iceland. Meanwhile the broader model analysis allows us to investigate their spatial extent, evolution and their influence on shear and potential turbulence occurrence. Therefore, we intentionally retain a larger spatial extent in the figures to capture the full dynamical environment (e.g., tropopause structures) in which these processes occur.

» Regarding second part of the comment, please see our response to comment 3. This information is added to the manuscript. Note that we are not only interested in the GW itself but also in the larger scale structure around the GWs.

Specific comments:

Comment: *L11: Further analysis of turbulence diagnostics suggests -> Please be more specific!*

» Modified to

*Further analysis of turbulence diagnostics including Richardson number on large-scale and turbulence indices (*TI1, TI2*) on small-scale suggests that GWs produce shear which leads to the occurrence of clear air turbulence.*

Comment: *L85-87: 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.*

» Modified!

Our primary goal of this study is based on recent revelations from idealized simulations of baroclinic life cycles (Umbarkar and Kunkel, 2025). In these simulations, a clear relation between GW occurrence, enhanced shear and turbulence occurrence in the lower stratosphere was found. The question is now whether this relation is also evident within an extratropical life cycle over the North Atlantic.

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

» Warm-conveyor belt is not important here, hence removed.

The simulation was initialized with the operational ICON global analysis at 00:00 UTC on 20 September 2017, shortly before the cyclone developed, and ran for 96 hours, ending at 00:00 UTC on 24 September 2017, when the life cycle of cyclone was in its final stage (Schwenk and Miltenberger, 2024).

Comment: *L143: Note that,.... scale GW -> Incomplete sentence?*

» Rephrased as:

Note that the momentum flux at sub-synoptic scales is used here as a first proxy for the momentum flux due to small–scale GWs.

Comment: *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?*

» Please see our response to comment 1 part 3.

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

» Corrected!

Comment: *L358: Why in principle?*

» Removed!

Response to Reviewer 3

Major concerns

Comment: *1. The link between gravity waves and the occurrence of turbulence is expected from dynamical reasons, with the GW fluctuations enhancing the shear and modifying stability. However, the choice to use the Absolute Momentum Flux (AMF) to quantify GW when relating turbulence to gravity waves (l140-141). There is not necessarily a better choice for quantitatively relating the two. Nonetheless, the choice to use AMF should be better explained. What is the theoretical justification to use absolute momentum fluxes (AMF) in order to investigate "vertical transport due to small-scale processes"? For conservative waves (i.e. in the absence of dissipation), it is possible to have large fluxes associated to displacements which are reversible, and hence do not lead to net transport. AMF is a quadratic quantity, and it depends on the frequency of the waves, with the low-frequency waves carrying less momentum than high-frequency waves. The impacts of waves on stratification and on shear will also involve both their amplitudes, and their frequency.*

» One of the key properties of GWs is their ability to transport horizontal momentum vertically, then depositing it into the mean flow upon breaking. The use of absolute momentum flux in this study is a pragmatic choice for a simple reason that *one would expect momentum-flux deposition by GWs in the region where GW-induced instability and turbulence occur (Fritts and Alexander, 2003; Rapp et al., 2004).* This has been included in the revised manuscript.

Comment: *2. About the discussion and conclusions: the authors benefit from a case which has been documented by airborne observations. Regarding the results that concern turbulence and mixing, how much of the conclusions are supported by evidence from the measurements? It is not sufficiently clear, in the concluding section, which statements are based on inferences based on modelled diagnostics (Richardson number, Turbulence Indices...) and which statements (regarding mixing and turbulence) are based on the measurements.*

» In the manuscript, the airborne observations are primarily used to confirm the presence of GWs and turbulence, as well as to identify GW-induced mixing event over the Iceland. In contrast, the model-based analysis is performed to investigate their spatial extent, evolution and their influence on shear and potential turbulence occurrence. We clarify that the first point in the conclusions is directly supported by the observational evidence (including ERA5 data interpolated along the flight track), whereas the subsequent conclusions are based on model-derived diagnostics such as GW momentum flux, shear, Richardson number and turbulence indices. This is clarified in the revised manuscript and results from the observations are more directly discernible from the model results.

Minor concerns

Comment: *l16-18: 'Large amounts of momentum and energy can be propagated and transferred..' The sentence should be rephrased. There are two concerns: one should not overstate the role of gravity waves in the circulation, and the formulation*

is clumsy.
Gravity waves matter for the general circulation. Their vertical propagation induces momentum fluxes that contribute to driving the mean circulation in the middle atmosphere (stratosphere and mesosphere). In the mid-latitudes, this contribution can be important but is secondary relative to that of Rossby waves, in particular planetary waves. In the mesosphere, the contribution of gravity waves is essential. The reference cited rather emphasizes the contribution of large-scale, balanced waves.

» Rephrased as:

GWs redistribute momentum and energy through vertical propagation, which affects the general circulation...

Comment: *l32-34: moist processes turn out to play an important role in the generation of GW from mid-latitude and high-latitude weather systems (e.g. Wei and Zhang 2014, Plougonven et al 2015, Holt et al 2017)*

» Thanks for pointing this. It is now included in the text.

Moreover, tropospheric moist processes turn out to play an important role in the generation of GW from mid-latitude and high-latitude weather systems (e.g. Wei and Zhang, 2014; Plougonven et al., 2015; Holt et al., 2017).

Comment: *36-38: The sentence on baroclinic life cycles is ambiguous: is this a statement on the contribution of weather systems (i.e. synoptic activity)? While weather systems are essential to the tropospheric circulation and zonally-averaged zonal flow, their influence does not necessarily penetrate very far into the stratosphere; the contribution from larger-scale Rossby waves is more important for the Brewer-Dobson Circulation (line 37). However, it seems that the topic of these two sentences should rather be 'Gravity waves excited from tropospheric weather systems (or synoptic eddies)...' given that the references given on line 38 are reviews on gravity waves. With gravity waves in mind, one should more clearly separate the different impacts:*
1. the vertical transfer of momentum; regarding this, what is in fact essential is the ultimate dissipation of the momentum fluxes associated to gravity waves, leading to forcing of the background flow at upper levels. Models to not take into account a 'depletion' of momentum at the launch level of GW parameterization. In this sense, what matters really is the forcing of the gravity waves in the middle atmosphere, not a 'transfer of momentum', since the source region is not affected. This forcing of the middle atmosphere's circulation is the main motivation and main purpose of parameterizations.
2. there are local impacts due to gravity waves, such as enhancement of local shear, and onset of turbulence where the waves break. This has implications for vertical mixing. These effects are known to exist, but are not as high on the list as the forcing of the circulation in the stratosphere and mesosphere. This should be reflected in the way these implications are put forward in the introduction.

» L36-38 "Baroclinic life cycles, as persistent large–scale midlatitude wave patterns, play a pivotal role in modulating large–scale circulations. They influence the Brewer-Dobson circulation by vertically transporting momentum and energy, influencing jet dynamics, and affecting the potential vorticity distribution in the UTLS (Alexander et al., 2010; Achatz et al., 2024, and references therein)." modified to

Gravity waves excited from tropospheric weather systems (or synoptic eddies) can propagate into the UTLS and middle atmosphere. Upon dissipation, these waves deposit momentum and energy, leading to forcing of the background flow at upper levels and thereby contribute to large-scale dynamics such as the BrewerDobson circulation (Alexander et al., 2010; Achatz et al., 2024, and references therein).

Comment: *l48-49: the effect of GWs on the thermodynamic structure near the tropopause should be made more precise. GWs that do not dissipate induce fluctuations, which are reversible. Hence, although they modify locally the shear and stratification, this is averages out in a time average, a priori. The situation is different when dissipation is present: mixing due to breaking induces irreversible transport. Dissipation of momentum or heat fluxes induce a forcing of the background flow. What the authors have in mind here should be better explained.*

» We have modified the text for clarity. This reads

GWs likewise play a key role in shaping the thermodynamic structure of the UTLS by either enhancing vertical wind shear or reducing static stability (Kunkel et al., 2016; Kaluza et al., 2021; Dörnbrack, 2024) through reversible fluctuations; however, when they dissipate or break, they can induce irreversible mixing and flux convergence, thereby modifying the background flow on longer time scales.

Comment: *l75-78: There is a problem in the discussion of the motivations of this study. The different processes and different scales through which gravity wave may have impacts are stated in a way that includes repetition and which could be better organized, here and in other parts of the introduction. For example, in lines 75-78, the first sentence of the paragraph highlights the impact of gravity wave breaking for "vertical mixing of key tracers". The third sentence begins by "Beyond its relevance to aviation", although aviation has not been highlighted in the previous sentences, and "cross-tropopause transport" is highlighted - although this is very related to "vertical mixing of key tracers"...*

» We have modified the entire paragraph for better readability.

Comment: *l86-87: the status of the idealized baroclinic lifecycles is ambiguous from the phrasing: it may be understood that the impacts of gravity waves will be investigated in idealized baroclinic lifecycles. The next sentence clearly indicates that observations and a real case study will be considered, in apparent contradiction.*

» Corrected accordingly.

Our primary goal of this study is based on recent revelations from idealized simulations of baroclinic life cycles (Umbarkar and Kunkel, 2025). In these simulations, a clear relation between GW occurrence and shear in the lower stratosphere was found. The question is now whether this relation is also evident within an extratropical life cycle over the North Atlantic.

Comment: *l91: "whether ERA5 is good enough resolved to study GWs, shear and mixing in the UTLS": to be rephrased, e.g. "which aspects of the gravity wave field can be reliably investigated in ERA5 fields, given their resolution."*

» Rephrased as:

The last point is related to the question which aspects of the GW field can be reliably investigated in ERA5 fields, given their resolution, and whether resolved spectrum of GWs is enough to study their impact on shear and mixing in the UTLS in case studies, and thus potentially may be suited for climatological analyses on this topic.

Comment: *l139: "ourself" -> "ourselves"?*

» Done!

Comment: *l139-141: not sure to understand the logic of restraining to the consideration of PDFs; could the authors explain?*

» It was a writing issue. We wanted to mention that we analyse GW momentum flux using PDFs, as the goal is to analyse the **occurrences of GWs** and not the characteristics. Rephrased for clarity.

Comment: *l155-161: while the methodology appears reasonable, and indeed consistent with what has been done in other studies, it does mix different methods (spectral approach to obtain the perturbation quantities, then lowpass Gaussian filter to average over several wavelengths and obtain a smooth estimate of the fluxes.*

» Right. The point is to mention that we consider **hybrid approach** following Wei et al. (2022) but not the same methods. However, it is mentioned in the steps to smoothen the averaged fields either by applying Gaussian lowpass filters or or areal smoothing (Kruse and Smith, 2015). This has been clarified in the manuscript.

Comment: *l176: "to account [FOR] this limitation"?*

» Corrected.

Comment: *l178-179: the structure of the sentence is not clear, to be rewritten*

» Rewritten as:

The Ellord-Knapp Turbulent Index (TI1) has previously been shown to be capable of detecting 70%-84% of CAT occurrences (e.g., Ellrod and Knapp, 1992; Sharman and Pearson, 2017; Kim et al., 2018; Gultepe et al., 2019; Thompson and Schultz, 2021).

Comment: *l209: word missing? "between two equatorward [?] reaching..." or placed later in the sentence ("streamers")?*

» Nope. It reads as "two equatorward reaching stratospheric streamers"

Comment: *l229: "ascend" -> "ascent"?*

» Corrected.

Comment: *l276-277: the paragraph discusses the measurements. Suddenly,*

» The sentence related to temporal resolution in models is only to explain the limitations against observations and why we attempted to look for gravity wave signatures in the measurement.

Comment: *l292-293: "which rotates anticyclonically (cyclonically) for an upward (downward) propagating wave" -> "which rotates anticyclonically for an upward propagating wave, and cyclonically for a downward propagating wave."?*

» Done!

Comment: *l334-335: the conclusion that the IFS manages to capture the appearance of inertia-gravity waves excited by jet-front systems, despite a limited resolution, is consistent with findings of previous studies that have investigated gravity waves in ECMWF products and reanalyses in particular (see Jewtooukoff et al, 2O15, and references therein).*

» Thanks for the suggestion. We have now cited Jewtoukoff et al. (2015).

Comment: *l343: the authors emphasize the shear in Fig. 7, "in particular in ICON". This is somewhat at odds with the figure, in which the shear in ICON seems weaker.*

» Right. These differences might be due either to the representation within ICON or to the level we are focusing on. Another possibility is that this reflects a systematic difference between ICON and IFS. As also noted by another reviewer, this discrepancy might be due to the fact that ICON forecasts analysed here are at lead times of 3-4 days, whereas the IFS forecasts are at lead times of 0-1 day and therefore much closer to the analysis/ERA5. Following this, we have separated discussion of ECMWF products and ICON simulation in the revised manuscript. These modifications are detailed in central changes.

Comment: *l371-391: as far as I understand, the relation between AMF (absolute Momentum Flux) and the shear is somewhat opportunistic, rather than a fundamental relation. Momentum fluxes are used as a proxy for the presence of waves, and these contribute to shear. Figure 8, perhaps because of the logarithmic axis and logarithmic colorbar, does not seem so compelling regarding the correlation between shear and AMF. The text and comments around this figure should clearly recognize that this relation is tentative.*

» While we agree that AMF serves as a proxy for GW activity, the relationship between AMF and shear **perturbations** in our analysis is not purely opportunistic. In addition to the evidence of single event presented here, a consistent relationship between GW activity and enhanced small-scale shear is also found in a analysis of annual cycle (see also, Umbarkar et al., 2026), which supports the robustness of this result. Nevertheless, we acknowledge that the relationship may depend on the underlying flow conditions. The text has been revised to reflect this interpretation.

Comment: *l535-537: it is worth noting that there are several different uses that can be made of ERA5 for studies of GWs. It is very challenging and demanding to expect the representation of small-scale shear to be accurate, let alone the representation of the resulting turbulence. Providing insights about regions favorable to the occurrence of gravity waves is more reasonable. In the assessment of the value of ERA5 for GW studies, it should be mentioned that there are different levels of information that may be targeted.*

» We have included the following text:

Nonetheless, ERA5 does capture a large portion of GW spectrum and most of the enhanced shear encounters. As such the reanalysis data provides valuable information on the spectral characteristics of GWs and is therefore suitable for long term analysis of GW-induced shear, at least over the North Atlantic. In this context, different levels of information can be targeted: ERA5 is well suited to characterize the large-scale environment and regions favourable for GW activity, and to some extent their resolved spectral characteristics, whereas its limited effective resolution constrains its ability to accurately represent small-scale shear and turbulence.

Comment: *l540: "small-scale gravity waves": it is worthwhile being more precise here. It is possible to have low-frequency waves which are small-scale in the vertical (vertical wavelength of a few hundred meters) yet retain horizontal wavelength on the order of several tens to a hundred of kilometers. It is possible to have higher frequency waves that have horizontal wavelengths only of a few kilometers (and similar vertical wavelengths). The frequency (which can be estimated from the tilt of the phase lines) matters very much in the description of the waves.*

» Right. Text modified to

...to multiscale interactions dominated by upward-propagating large-amplitude GWs with relatively short horizontal but large vertical wavelengths.

Comment: *l541-544: the relationship highlighted here is not clear: do the authors highlight GWs producing pockets of turbulence, through local enhancement of shear? Or do they have evidence for secondary generation of smaller-scale gravity waves where a primary wave-packet breaks? The second interpretation does not appear very likely, but the structure of the sentence is ambiguous (∼ "Turbulence (...) correlates with appearance of (...) small-scale gravity waves"..*

» Text modified to

GW–induced enhanced shear, which are conducive to turbulence, highly correlates with appearance of envelopes (or packets) of small–scale GWs of ∼100 km of wavelength in the LMS with the same characteristics.

Comment: *l560-561: on the difficulty of convergence with resolution, the study by Kruse et al. (2022) is also relevant.*

» Thank you for the suggestion. We have now cited Kruse et al. (2022).

References

Achatz, U., Alexander, M. J., Becker, E., Chun, H. Y., Dörnbrack, A., Holt, L., Plougonven, R., Polichtchouk, I., Sato, K., Sheshadri, A., Stephan, C. C., Van Niekerk, A., and Wright, C. J.: Atmospheric Gravity Waves: Processes and Parameterization, Journal of the Atmospheric Sciences, 81, 237–262, https://doi.org/10.1175/JAS-D-23-0210.1, 2024.

Alexander, M. J., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., Sato, K., Eckermann, S., Ern, M., Hertzog, A., Kawatani, Y., Pulido, M., Shaw, T. A., Sigmond, M., Vincent, R., and Watanabe, S.: Recent developments in gravity-wave effects in climatemodels and the global distribution of gravity-wavemomentum flux from observations and models, Quarterly Journal of the Royal Meteorological Society, 136, 1103–1124, https://doi.org/10.1002/qj.637, 2010.

Butterworth, S.: On the theory of filter amplifiers, Exp. Wireless Wireless Eng., 7, 536–541, 1930.

Dörnbrack, A.: Transient Tropopause Waves, Journal of the Atmospheric Sciences, 81, 16471668, https://doi.org/10.1175/jas-d-24-0037.1, 2024.

Ellrod, G. P. and Knapp, D. I.: An Objective Clear-Air Turbulence Forecasting Technique: Verification and Operational Use, Weather and Forecasting, 7, 150165, https://doi.org/10.1175/1520-0434(1992)007<0150:aocatf>2.0.co;2, 1992.

Fritts, D. C. and Alexander, M. J.: Gravity wave dynamics and effects in the middle atmosphere, Reviews of Geophysics, 41, 1–64, https://doi.org/10.1029/2001RG000106, 2003.

Gultepe, I., Sharman, R., Williams, P. D., Zhou, B., Ellrod, G., Minnis, P., Trier, S., Griffin, S., Yum, S. S., Gharabaghi, B., Feltz, W., Temimi, M., Pu, Z., Storer, L. N., Kneringer, P., Weston, M. J., Chuang, H.-y., Thobois, L., Dimri, A. P., Dietz, S. J., França, G. B., Almeida, M. V., and Neto, F. L. A.: A Review of High Impact Weather for Aviation Meteorology, Pure and Applied Geophysics, 176, 18691921, https://doi.org/10.1007/s00024-019-02168-6, 2019.

Gupta, A., Sheshadri, A., Alexander, M. J., and Birner, T.: Insights on Lateral Gravity Wave Propagation in the Extratropical Stratosphere From 44 Years of ERA5 Data, Geophysical Research Letters, 51, https://doi.org/10.1029/2024gl108541, 2024.

Holt, L. A., Alexander, M. J., Coy, L., Liu, C., Molod, A., Putman, W., and Pawson, S.: An evaluation of gravity waves and gravity wave sources in the Southern Hemisphere in a 7 km global climate simulation, Quarterly Journal of the Royal Meteorological Society, 143, 24812495, https://doi.org/10.1002/qj.3101, 2017.

Jaeger, E. B. and Sprenger, M.: A Northern Hemispheric climatology of indices for clear air turbulence in the tropopause region derived from ERA40 reanalysis data, Journal of Geophysical Research Atmospheres, 112, 1–13, https://doi.org/10.1029/2006JD008189, 2007.

Jewtoukoff, V., Hertzog, A., Plougonven, R., Cámara, A. d. l., and Lott, F.: Comparison of Gravity Waves in the Southern Hemisphere Derived from Balloon Observations and the ECMWF Analyses, Journal of the Atmospheric Sciences, 72, 34493468, https://doi.org/10.1175/jas-d-14-0324.1, 2015.

Kaluza, T., Kunkel, D., and Hoor, P.: Composite analysis of the tropopause inversion layer in extratropical baroclinic waves, Atmospheric Chemistry and Physics, 19, 6621–6636, https://doi.org/10.5194/acp-19-6621-2019, 2019.

Kaluza, T., Kunkel, D., and Hoor, P.: On the occurrence of strong vertical wind shear in the tropopause region: a 10-year ERA5 northern hemispheric study, Weather and Climate Dynamics, 2, 631–651, https://doi.org/10.5194/wcd-2-631-2021, 2021.

Kim, J.-H., Sharman, R., Strahan, M., Scheck, J. W., Bartholomew, C., Cheung, J. C. H., Buchanan, P., and Gait, N.: Improvements in Nonconvective Aviation Turbulence Prediction for the World Area Forecast System, Bulletin of the American Meteorological Society, 99, 22952311, https://doi.org/10.1175/bams-d-17-0117.1, 2018.

Kruse, C. G. and Smith, R. B.: Gravity Wave Diagnostics and Characteristics in Mesoscale Fields, Journal of the Atmospheric Sciences, 72, 43724392, https://doi.org/10.1175/jas-d-15-0079.1, 2015.

Kruse, C. G., Alexander, M. J., Hoffmann, L., van Niekerk, A., Polichtchouk, I., Bacmeister, J. T., Holt, L., Plougonven, R., ácha, P., Wright, C., Sato, K., Shibuya, R., Gisinger, S., Ern, M., Meyer, C. I., and Stein, O.: Observed and Modeled Mountain Waves from the Surface to the Mesosphere near the Drake Passage, Journal of the Atmospheric Sciences, 79, 909932, https://doi.org/10.1175/jas-d-21-0252.1, 2022.

Kunkel, D., Hoor, P., and Wirth, V.: Can inertia-gravity waves persistently alter the tropopause inversion layer?, Geophysical Research Letters, 41, 7822–7829, https://doi.org/10.1002/2014GL061970, 2014.

Kunkel, D., Hoor, P., and Wirth, V.: The tropopause inversion layer in baroclinic life-cycle experiments: The role of diabatic processes, Atmospheric Chemistry and Physics, 16, 541–560, https://doi.org/10.5194/acp-16-541-2016, 2016.

Lachnitt, H. C., Hoor, P., Kunkel, D., Bramberger, M., Dörnbrack, A., Müller, S., Reutter, P., Giez, A., Kaluza, T., and Rapp, M.: Gravity-wave-induced cross-isentropic mixing: A DEEPWAVE case study, Atmospheric Chemistry and Physics, 23, 355–373, https://doi.org/10.5194/acp-23-355-2023, 2023.

Plougonven, R., Hertzog, A., and Alexander, M. J.: Case studies of nonorographic gravity waves over the Southern Ocean emphasize the role of moisture, Journal of Geophysical Research: Atmospheres, 120, 12781299, https://doi.org/10.1002/2014jd022332, 2015.

Rapp, M., Strelnikov, B., Müllemann, A., Lübken, F., and Fritts, D. C.: Turbulence measurements and implications for gravity wave dissipation during the MaCWAVE/MIDAS rocket program, Geophysical Research Letters, 31, https://doi.org/10.1029/2003gl019325, 2004.

Schwenk, C. and Miltenberger, A.: The role of ascent timescales for warm conveyor belt (WCB) moisture transport into the upper troposphere and lower stratosphere (UTLS), Atmospheric Chemistry and Physics, 24, 14 07314 099, https://doi.org/10.5194/acp-24-14073-2024, 2024.

Sharman, R., Tebaldi, C., Wiener, G., and Wolff, J.: An Integrated Approach to Mid- and Upper-Level Turbulence Forecasting, Weather and Forecasting, 21, 268287, https://doi.org/10.1175/waf924.1, 2006.

Sharman, R. D. and Pearson, J. M.: Prediction of Energy Dissipation Rates for Aviation Turbulence. Part I: Forecasting Nonconvective Turbulence, Journal of Applied Meteorology and Climatology, 56, 317337, https://doi.org/10.1175/jamc-d-16-0205.1, 2017.

Stephan, C. C., Strube, C., Klocke, D., Ern, M., Hoffmann, L., Preusse, P., and Schmidt, H.: Intercomparison of Gravity Waves in Global Convection-Permitting Models, Journal of the Atmospheric Sciences, 76, 27392759, https://doi.org/10.1175/jas-d-19-0040.1, 2019.

Thompson, C. F. and Schultz, D. M.: The Release of Inertial Instability Near an Idealized Zonal Jet, Geophysical Research Letters, 48, https://doi.org/10.1029/2021gl092649, 2021.

Tsuda, T.: Characteristics of atmospheric gravity waves observed using the MU (Middle and Upper atmosphere) radar and GPS (Global Positioning System) radio occultation, Proceedings of the Japan Academy, Series B, 90, 1227, https://doi.org/10.2183/pjab.90.12, 2014.

Umbarkar, M. and Kunkel, D.: Contribution of gravity waves to shear in the extratropical lowermost stratosphere: insights from idealized baroclinic life cycle experiments, Atmospheric Chemistry and Physics, 25, 10 159–10 182, https://doi.org/10.5194/acp-25-10159-2025, 2025.

Umbarkar, M., Kunkel, D., and Achatz, U.: Co-occurrence of gravity waves, vertical wind shear and turbulence in the lowermost stratosphere over the North Atlantic, Weather and Climate Dynamics, https://doi.org/10.5194/egusphere-2026-413, preprint, 2026.

Wei, J. and Zhang, F.: Mesoscale Gravity Waves in Moist Baroclinic JetFront Systems, Journal of the Atmospheric Sciences, 71, 929952, https://doi.org/10.1175/jas-d-13-0171.1, 2014.

Wei, J., Zhang, F., Richter, J. H., Alexander, M. J., and Sun, Y. Q.: Global Distributions of Tropospheric and Stratospheric Gravity Wave Momentum Fluxes Resolved by the 9-km ECMWF Experiments, Journal of the Atmospheric Sciences, 79, 26212644, https://doi.org/10.1175/jas-d-21-0173.1, 2022.

Zhang, W., Peng, J., Zhang, W., Wang, S., Li, Z., and Hanyan, W.: Sensitivity of Gravity Wave Momentum Flux Estimates on Separation Methods, Monthly Weather Review, 153, 17211742, https://doi.org/10.1175/mwr-d-24-0272.1, 2025.