Response to Referee #2

Dear referee,

thank you again for this detailed review of our manuscript and your comments. We would like to emphasise that it wasn't our intention to use ERA-Interim to validate our simulations and corrected all statements in the manuscript with respect to this. It should rather show the large-scale meteorological conditions during the mountain wave event.

Major changes include:

- replacing the temperature in Fig. 4 by the near-surface wind speed
- moving the AIRS comparison between Fig. 4 and 5 in Sect. 5.1

Yours sincerely and on behalf of all co-authors,
Michael Weimer

1 Major comments

1.1 - Figure 5: You have mentioned that reanalysis resolves the gravity wave / temperature amplitude (e.g. your Figure 4). If you wanted to properly test this (and I would have thought show conclusively that it only partially resolves this, at least when compared to the 40 km model), then you could include a column in Figure 5 that shows results based on reanalysis. Or perhaps, this could be included as a supplementary figure. I have to say that the paper does give a rather confusing message as to what are the limitations of reanalysis / coarse resolution (of order ∼100 km), so this would go a long way to clarifying this. However, if you would rather not do this, then please include clear evidence in the manuscript in the form of citations that details the ability of modern reanalysis to capture small-scale orographic gravity waves. Four other comments related to this issue are also on lines #195, #230, #243, and #375 as well as Section 5.3. Even better, as I explain below, it would surely be much better to just stick to comparing the model against AIRS, which shows great promise. This is honestly all that you need, and any other argument just gets confusing and convoluted.

As stated above, it wasn't our intention to validate our simulations with ERA-Interim. We want to show the meteorological conditions that led to the mountain wave event. Therefore, we replaced the 25 km temperature by the 10 m wind speed which is now colour-coded in Fig. 4. Additionally, we removed every statement comparing ICON-ART with ERA-Interim from the manuscript.

1.2 - Line #195: There seems to some confusion here. You wisely have included AIRS to validate the model representation of the gravity wave event (as done elsewhere), particularly I assume the temperature perturbation induced by the wave. Therefore, why is there considerable mention of the model being compared to the reanalysis representation of the wave event? Surely this is flawed and/or redundant?

We hope that our explanation to Comment 1.1 and the changes made in the manuscript removed this confusion.

1.3 - Line #230: Here you are explaining that (relatively coarse resolution, resolution ∼80 km) ERAI is able to resolve the temperature minimum induced by gravity waves to the lee of the Peninsula. How can a grid spacing of 80 km resolve a mountain wave of around 300 km wave length? Especially as your Figure 5 shows a marked difference between 80 and 40 km results. Please modify/soften the language so that this remark is less glaring. Maybe comment that there is evidence of a wave in ERAI, but given the resolution it is likely poorly resolved, amplitude underestimated, etc.... See Alexander and Teitelbaum (2007), which you refer to. See also comment above.

As mentioned above, we want to show the meteorological conditions and not validate our simulations with ERA-Interim.

1.4 - Line #243: Why would you expect temperatures to be in agreement with ERA-Interim? As mentioned above, your 40 v 80 km results are not in agreement, so why should ERAI v 40 km results be in agreement? Perhaps the agreement is because the location is over the base of the Peninsula, so conducive to more broader horizonal scale waves that reanalysis products can resolve, compared to waves forced by the much narrower Peninsula region further north (width ∼100km). My understanding is that it is well established that reanalysis products fail to capture the detailed temperature structure associated with gravity waves, which is why you have included AIRS data in your analysis. Maybe Alexander and Teitelbaum (2007) is appropriate, as it states that the fine details apparent in AIRS is not evident in ECMWF data. See also comment above.

We removed this statement.

1.5 - Section 5.3: My initial reaction to section 5.3 is that is a repeat of 5.1, and why are two separate sub-sections required? Surely a better place for the AIRS comparison would be 5.1, so that the dynamics / simulation of the mountain wave is dealt with in one place. This would also be a lot cleaner and enable you to easily remove all comparisons/mention of agreement with reanalysis that are currently in 5.1. In any case, how can Sect. 5.1 improve on 5.3? Surely a state of the art comparison with AIRS in 5.3 makes the results of 5.1 largely redundant, and in doing so lessens the impact of 5.3?

We followed your suggestion and moved the AIRS comparison to Sect. 5.1. With this, the evaluation with measurements always precedes the model domain comparisons. In addition, the section with the CALIOP comparison can build on the AIRS results which could be a better indicator that a higher resulution could improve the results.

1.6 - Line #375: Here you categorically state that the model matches AIRS. Excellent result, which negates the need for any earlier and confusing mentioning of the good agreement with reanalysis.

We agree with this statement and moved the AIRS comparison to Sect. 5.1.

2 Minor comments/changes

2.1 - Line #2: Spelling. Surfaces.

Corrected to "surface".

2.2 - Line #4: Please mention the possible role of unresolved non-orographic waves somewhere in the manuscript, and how the ICON-ART model would possibly result in an improvement in this aspect. Please see Tritscher et al. (2021) for a (short) review/discussion of the possible role of non-orographic waves and PSCs.

We added a sentence at the very end of the conclusions that they are not considered in our simulations.

2.3 - Line #29: This statement is not quite true. The requirement is that a sizeable component of the wind is perpendicular to the barrier, not all of the large scale flow. This is why south- easterly winds, which are common over the Peninsula, are closely associated with forcing orographic gravity waves.

We added this to the statement.

2.4 - Line #60: Please revise sentence beginning "Thus, a low resolution ....". Its currently unclear.

We revised the sentence to "Thus, a global low-resolution simulation provides boundary conditions for a region with refined grid, similar to mesoscale models."

2.5 - Line #75: Please revise use of “large” here. Horizontal wavelengths of 300 km are not large, but mesoscale.

We replaced “large” by “mesoscale”.

2.6 - Line #183: The list of PSC types given here could do with a little more detail/explanation, rather than expecting the reader to work this out for themselves via Pitts et al. 2018. For example, what are enhanced NAT mixtures, Wave-ice, etc? How reliable are CALIOP measurements of the different PSC types?

We included some statements explaining the categories and their reliability based on Pitts et al. (2018).

2.7 - Figure 4 caption: What does “dark shadow” refer to? It made me think of the Lord of the Rings!! I think this is a translation issue from German to English. Please correct.

We obviously don’t want to be Sauron or one of his followers and corrected the figure caption.

2.8 - Line #239: I mentioned this in my first review. Please be careful how you word this. The interaction between the flow and the detailed orography is better captured at higher resolution, but you are saying something different (flow over the mountain in the Antarctic Peninsula nest is improved). Also, even if the model orography was to converge at higher resolution (ie differences in orography between the different resolutions becomes small), you might still expect differences in the representation of the key features of the orographic gravity wave due to differences in the grid spacing / resolving finer details of the dynamics. Such features are reviewed by Smith et al. 1989. Smith, R. B. (1989). Hydrostatic airflow over mountains. Advances in Geophysics, 31, 1-41, https://doi.org/10.1016/S0065-2687(08)60052-7.

We corrected the statement.

2.9 - Line #248: This argument is rather confused here. Is it the 40 km horizonal resolution or the 500 m vertical resolution that are important, or both. Please clarify.

We clarified the statement. It's both horizontal and vertical resolution that is important.

2.10 - Line #257: Include a reference here after "consistent with theory". Smith (1989), mentioned above, is often acknowledged as being the classic paper for mountain waves.

We already mentioned Queney (1947) as reference some lines above, but added Smith (1989), too, now at both lines, so that it is clear what is our reference.

2.11 - Line #284: Have these parameters (particularly the subscripts) been previously defined? What does the subscript "NAT|ice" refer to? Please make sure that all parameters are clearly defined.

Yes, they are defined in Sect. 4.1 with the CALIOP description. We added a reference to this section when mentioning these thresholds again.

2.12 - Figure 6: I think it is worth commenting that in both the 40 km model and CALIOP that the fraction of Wave-ice PSCs is very small. Is this to be expected? How reliable are CALIOP measurements of Wave-ice PSCs?

We moved the discussion of the Wave-ice category, which was at the end of the section, up to the discussion of Fig. 6 (now Fig. 8) and added a sentence about the reliability of the Wave-ice category, based on the statements above.

2.13 - Line #321: I don't think that you can say with any certainty that a higher resolution is required to get the number of ice PSCs correct. Also, its not clear what would be the impact of higher resolution? Are you suggesting that the temperature perturbation amplitude would be larger at higher resolution, and hence more realistic? If so, then please make this clear. You also made a similar comment at line #339 and #348, so please amend this also.

We amended all three statements and included the reference to Orr et al. (2020) as suggested some comments below.

2.14 - Line #330: You have identified a deficiency in the simulation of STS. But no explanation is given as to why this is occurring, or indication that this will be looked at later. Please amend this.

We added a sentence that it should be analysed in future simulations.

2.15 - Line #364: Again, you can not state that a higher resolution would resolve this issue. You don't know this for certain and its beyond the scope of your work to examine this. You need to soften this language, so say that it is perhaps probable that this is a resolution issue (or consistent with a resolution dependence issue), which would perhaps be corrected by going to higher grid spacings, as shown by ...

We corrected the statement accordingly.

2.16 - Line #367: Please refer to Figures 5, 6 and 7 of Orr et al. 2020, which examine the impact of a higher resolution simulation (via a parameterisation) on the simulated temperatures relative to the ice temperature threshold, as well as the formation potential of ice PSCs.

We included the reference.

2.17 - Line #370: I would suggest prefacing the reference to Figure 9 with something like "Following Orr et al. (2015) ...", as this is the same model v AIRS comparison as that study undertook.

We added a sentence at the beginning of this section (originally line 354) saying that the same methodology was also applied in Orr et al. (2015).

2.18 - Line #372: You could perhaps mention that an analogous study by Orr et al. (2015) using a 4 km model seems to suggest a better match between observed v simulated wave amplitudes over the Antarctic Peninsula, ie more realistic and larger at the higher resolution. This is perhaps the evidence that you need to suggest that a further extension of your study could be to perhaps go to higher grid spacings.

We added this to the paragraph.

2.19 - Line #404: Meteorological models do not have "ice schemes". Not sure what you are meaning here, so please revise.

We replaced it by "hydrometeor microphysics".

2.20 - Figure 14: Perhaps worth comparing this to the results of Figure 10 in Orr et al. (2020), which also looked at the effects on chlorine activation. Similarly, Figure 15 could also be compared to Fig 11 in Orr et al., which also looked at the impacts on ozone.

Since the figure with NAT differences in Orr et al. (2020) shows slightly different things (temporal variability and surface concentration in contrast to our analysis) we added a qualitative statement comparing the altitude regions for late July. We included a sentence comparing the ozone result with Orr et al. (2020).

2.21 - Conclusions/outlook: Maybe a mention of the importance of unresolved non-orographic gravity waves could be included, e.g. from SH storm tracks.

We added this at the end of the conclusions.