

Interactive comment on "Mesospheric gravity wave activity estimated via airglow imagery, multistatic meteor radar, and SABER data taken during the SIMONE-2018 campaign" by Fabio Vargas et al.

Anonymous Referee #1

Received and published: 5 October 2020

This manuscript presents a fairly detailed investigation of gravity waves observed in airglow image data obtained during an observation campaign in November 2018, from Northern Germany. It is well-written and provides a good assessment of the effects of GW on the mesosphere, even if it is during a limited period of time. However, the authors should address the following would comments:

- We thank you for raising these important comments and suggestions. We address your comments the better we can.

- It is a little bit surprising that there are no waves with intrinsic period < 20 min. Most of the waves should be propagating against the wind, therefore their intrinsic phase speed should be larger than the observed one, and their intrinsic period should be smaller than their observed one.

- In figure 5f, the reviewer can see that there are waves traveling against the wind, waves traveling nearly perpendicular to the wind, and in few cases, waves propagating to the same wind direction. Although there are a few waves traveling to the same wind direction because their larger phase speed, these waves do not transport significant momentum flux, as presented in figure 5a showing large momentum flux waves moving against the wind.
- We did not detect waves with periods < 20 min, although they must be present certainly. Due to the filter wheel cycle time, a given airglow emission is sampled every 10 min, which allows only the detection of waves having periods > 20 due to the long sampling period of 10 min each image is acquired.

- lines 255-271: It is complicated to compare the results with Li et al., 2011, because of the time resolution (~ 10 min) which does not allow for detecting most short-wavelength, short-period waves. As also noticed in the manuscript, this campaign covered only a few days, while Li et al. results represent the whole year. Same problems with Li, 2011.

- We did the comparison because the paper utilizes the same autodetection technique. we focus on the results of Li et al which are comparable to the wave scales we have detected. A stronger correspondence between the results is found in the longer horizontal wavelength range (100 kilometers or longer), which are associated to convective sources far away from the observation site. In this way, we can verify that our findings are in agreement with those of Li et al. (2011), even in the case we do not detect the shorter period waves, shorter horizontal scale waves.

- lines 284-285: These mean values are very small, especially considering the error. What does -0.36 ± 1.51 mean? $-0.36 + 1.51 = 1.15$, so the MF is carried in the other direction? Or you are sure about the wave direction of propagation, and so the minimum possible value is 0? I am not sure how it is possible to measure such small MF values giving the uncertainties in the measurements.

- The summary box in Fig. 5 shows the MF mean and MF standard deviation of the sample, not the error in the MF determination. Notes that the error in momentum flux must take into consideration the errors in all the wave parameters determined by our auto detection method for wave event (see Vargas, 2019). As we showed and discussed in the manuscript, most of the waves carry small momentum flux, and this fact reflects on the mean, which is very small. The standard deviation shows more or less the spread in the data, that is, the distribution of waves going to one side against those going to the opposite side.

- line 286: The total value is very misleading and cannot be used for any future comparisons. With the analysis method used in this paper, GWs are measured for each image, therefore the more images, the larger the total MF. If the imager cycle was 1 min instead of 10 min, the total MF would be 10x larger because 10x more waves parameters would have been measured! An average MF value for each wave and its duration could be more useful to assess its impact. The statistical results (11% of the waves responsible for carrying 50% of the MF...) are still relevant but the authors have to be very careful when giving the "total" MF. -

- I agree the total MF value means nothing and will cause confusion. To give a meaningful estimation of the total momentum flux, we should have found the number of independent waves (see our reply to your next comment). We have removed the total MF from the paper.

The number of waves detected (362) is also controversial. There were NOT 362 separate GWs propagating over the observation site during these 4 nights! There were 362 wave measurements, which is different. Some of the waves probably lasted for hours and their parameters were measured several times. Of course, these parameters can evolve depending on the forcing and the background atmosphere conditions, but there were still the same GWs coming from the same sources

- The waves are not necessarily independent. We have defined as a gravity wave event the output of the cross-spectrum of two time-difference images when the special peak is larger than 10% of the total energy of the spectrum. two time-difference images are generated from three light frames of the airglow. Thus, if a wave is detected in a given set of three light images, it is considered an independent wave detection. If, in the next set, another wave detection is made, the only way to tell the two events correspond to the same gravity wave is by comparing their wave parameters, which vary among image sets. Now, the momentum flux of a wave varies as it goes through the field of view as well. Thus, considering the mean MF of all the waves detected during the campaign would be the same if we cluster the corresponding wave events into a distinct wave, average the momentum flux for that specific wave for the duration of the event over the airglow images where it shows up, and then average over all the distinct wave events found. This less efficient and more laborious though. We have a clustering algorithm that does that work for us, but still, we choose not to do it in the present manuscript.
- As we rely on a set of three images at the time from where we obtain two time-difference images, and finally the cross spectra from the 2DFFT of the time difference images. So, the time span of each set is about 30 min, and most of waves would have completely disappeared the imager FOV within that time span because the duration of a gravity wave is short based on our experience. This way, it would be more likely that each detection corresponds to an independent wave detection.

The GWs were observed in 4 different airglow layers, at 4 different altitudes. Could you give more information about that? Which waves were observed simultaneously? It would also be interesting to see the difference in MF between the airglow layers.

- The momentum flux of waves in different layers can be given and we have estimated it separately, but because the detections in an individual layer were not too many, we showed the results together. We have provided separated plots analogous to figure 5 as supplementary material where the reviewer can see the wave detections over time in individual layers.

The MF divergence would be very interesting, as well. - You should mention ducted waves. Even if the waves presented in this paper are unlikely to be ducted because of the minimum intrinsic periods, these waves exist but do not carry any vertical MF, so they would bias your results. Again, a comparison between the different airglow layers would help

- Estimating the momentum flux divergence between two layers is possible and we have done that in the past (Vargas et al, 2015, <http://dx.doi.org/10.1016/j.asr.2015.07.040>). The simultaneous identification of a wave in two layers is more demanding, and we do that for specific, individual cases. Instead of that, we have estimated the flux divergence from our model (Vargas et al, 2007, JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 112, D14102, doi:10.1029/2006JD007642, 2007) based on the distribution of periods and wavelengths of the observed waves. In this paper, we have done that as if the wave amplitude was measured in the OH and O(1S) layers simultaneously, with no change in amplitude detected (saturated waves).

- It is probably beyond the scope of this paper, but given the small number of nights, looking for sources (tropospheric weather, convection, fronts...) would be interesting.

- We did not address the sources of the waves in this paper but will do it in a follow up publication exploring the sources of primary and secondary waves based on the possibility of secondary wave generation in the range of 30-40 km as suggested by lidar temperature data.

Minor comments:

line 12: can you add the number of nights/hours of observation? It would give a better idea of the importance of the MF values.

- The number of hours per night can be seen an estimated from figure 1 keograms. We have estimated that and added to the text in line 10 (four nights, 45 hours nighttime observations).

line 26: remove "successfully"

- We have removed that.

line 93: long-period variations Section 2.2: what are the vertical and temporal resolutions of the radar system?

- The MSMR data was processed with vertical resolution of 1 km and temporal resolution of 30 min for the total wind field (Fig2a). The fluctuations (Fig2b) were calculated by subtracting the total wind field from a heavily averaged wind with 4-km vertical resolution and 4-hours temporal resolution over 400 km FOV of the MSMR system.

line 137: FWHM

- We have corrected that.

line 175: "duration" instead of "length"

- We have corrected that.

line 177: The MF equation is proportional to L_z/L_x , so it is not that surprising that large horizontal wavelength waves carry less MF. It is more surprising that large vertical wavelength waves carry less MF, though.

- From Vargas et al. (2007), we see that the momentum flux (Eq. 15) is given by $F_M/\rho_u = -\frac{1}{2} \frac{\omega^2 g^2 m}{N^4 k} \left(\frac{l'/l_0}{CF_l}\right)^2$, and other variables like the cancelation factor (vertical wavelength dependent) and the wave amplitude also are taken into account. Observe that FM has a square dependency on these variables as well as the wave frequency. All in all, because MF is multidimensional in the wave parameter variables and CF, there is no easy way to visualize it in a plot. Maybe would be optimum to show layers where two FM parameters vary while the others are kept constant.

line 191: You must have done your filtering wrong if the cutoff was 5 hours and you still see that many peaks, especially the ones around 6, 8, and 12 hours! The peak at 8.9 hr is pretty close to the peak at 8 hr. You have to give the resolution of the spectra, otherwise this is not very convincing.

- We have not explained correctly the way the wind fluctuations were calculated (it is corrected now in the manuscript). To obtain the wave fluctuations (Fig2b), we first calculate 4-hours, 4-km temporal and spatial windows to average the wind over 400km field of view. then we subtract the result from the total wind field. The total wind field was calculated using much smaller temporal and spatial windows, that is, 30 min and 1 km, respectively. Because that, the resulting fluctuations still keep some of the energy of larger periods waves as can be seen in the spectrum.
- Although we are reducing the effects of tides and waves with periods larger than 4 hours using: (a) temporal filter of 4 hours, (b) altitude filter of 4 kms, and (c) the inherent horizontal filter of about 400 km or so, we can see that signature of 8 hours with large vertical wavelengths and very large horizontal wavelengths are still visible.
- Looking to the spectrum alone we cannot tell apart tides to gravity waves. However, we can identify larger periods gravity waves in the airglow because their horizontal wavelength is ~ 1000 km whereas the tidal components have much larger horizontal scales. This way, the ~ 1000 km, 8 hours waves can be seen in the airglow keograms as the tilted brightness structures as those in Fig. 8b.

line 204: The keograms could be improved by flat fielding the images first. Not sure if this had been done, but it would improve the signal at the zenith compared to the edges. Figure 8a bottom is misleading. It looks like the

wave tilt (which is related to the direction of propagation and phase speed) corresponds to the decrease in meridional wind intensity. Not sure how to avoid that. You should add the directions of propagation in Table 3.

- We perform all the pre steps, flat fielding included, before calculating the 2DFFT of the keograms. By simulating a 3D gravity wave perturbing the airglow and using the gravity wave polarization relations from the linear theory, we would be able to confirm that the airglow intensity increases with the decreasing of the meridional wind.

line 221: What is the error on the horizontal wavelength measurements?

- The uncertainty in the horizontal direction can be determined from the horizontal wavenumber spectrum where the spread of the wavenumber energy that leaks to adjacent wavenumbers. From our calculations in Vargas 2019, we have estimated more or less 10% uncertainty (included in the text now) in the horizontal wavelength, that is, ~130 km for the wave seen on Nov. 3-4.

line 235: 8 hours, could be a tidal component.

- As pointed out above, the large-scale wave of Nov 6-7 does not correspond to a tidal component since the horizontal wave structure can be fully seen in the keogram of Fig.8b.

line 293-294: The last sentence is enigmatic. If these larger amplitude waves are only seen in the O2 emission but not below (OH, Na) or above (OI), it is quite puzzling.

- We also cannot explain this properly without relying on heavy speculation of what has happened.

Lines 363-365: Can you rephrase this sentence? It is not very clear.

- We have removed the referred sentence because it was too speculative.

Line 378: The links are missing! Figure 4: (d) Calculated volume emission rates... Figure 7: Can you add a detection threshold?

- We have fixed that. We have added a horizontal line as a reference for the horizontal threshold, but we have also calculated the spectrum from the total wind field now, not only for the fluctuations and adjusted the text accordingly.

The comparisons with previous publications are lacking some relevant studies: eg, Bossert et al., 2015, for MF of long period waves, Cao and Liu, 2016, for the impact of large vs small MF GWs, and even on the same topic Hertzog et al., 2012, Plougoen et al., 2013, or Wright et al., 2013. The last 3 describing stratospheric measurements, though. The comparisons with Li, 2011, and Li et al., 2011, are not very relevant except because they used a similar analysis method.

- We explained earlier and offered other reasons of why we did compare our results with those of Li et al. (2011) and Li, 2011.
- We have looked into the suggested references and have decided to use that of Cao and Liu, 2015 and Bossert et al. (2015) to compare with our results. We appreciate the suggestions.

Interactive comment on "Mesospheric gravity wave activity estimated via airglow imagery, multistatic meteor radar, and SABER data taken during the SIMONE–2018 campaign" by Fabio Vargas et al.

Anonymous Referee #2

Received and published: 15 October 2020

This manuscript deals with gravity waves detected using an airglow imager and discusses their effect on the background wind, combining collocating meteor radar network data and satellite data as background information. Their effort to make full use of the obtained data during the limited campaign period is much appreciated: expanding their analyses to waves with horizontal scales larger than the field of view of the imager. However, the overall analysis methods often lack detailed explanation and hard to follow, and some part of them even seem inadequately done. The authors are advised to revise the manuscript substantially by addressing the following comments.

- We cannot thank enough the reviewer for taking his/her time to read our manuscript and bring up important comments and suggestions to improve it.
- We recognize we need to expand the description of the data analysis to make it clear for the readers. Also, we believe the keogram spectral analysis would be understandable as we wrote, but it is obvious we need to specify it in more detail in the manuscript.

Major comments:

The sequence interval of the imager observation is 10 min, and the exposure time is mostly 2 min. This means that possible aliasing should be taken into account in the analyses although it is not described in the manuscript. Because of the 2 min exposure, any structure with a period shorter than 4 min will be smoothed out and not be resolved, but those with a period between 4–20 min can easily alias into slower period motions, and significantly affect the results.

- We believe that, because every image of a given airglow layer is taken at 10 minutes pace (the imager filter wheel cycle), we are only able to resolve waves with 20 minutes or longer. However, we believe the aliasing is minimal in this case because the small-scale wave structures (horizontal wavelength less than 725 km) seen in the airglow images are sharp (no smudging).

In detecting wave structures from wind data, a high-pass filter with a cut-off of 5 hr is used. However, the detected wave periods are around 4 and 8 hr, which are close to or longer than the cut-off. The wave amplitudes and also phase structures will be largely affected. Even a spectral analysis is done using the filtered wind, and fluctuation components with periods longer than 4 hr are discussed (Fig 7), leading to the subsequent analyses of those components based on the filtered wind as shown in Figures 8 and 10. Although the authors mention these peaks as 'leakage' (line 190), I am afraid that such data treatment is not acceptable at all. What the authors can remove before the spectral analysis in the present case would be only the background wind (mean wind) and trend. Since the imager data is unfiltered, wind values should be treated in a similar manner in comparison.

The details of 2D spectral analysis method shown in Fig 9 are not clear. The keograms shown in Figure 8 are thought to be used, of course, but how is the calculation done? The borders of the keograms can significantly affect the results if miss handled. Isn't there a possibility that the evening twilight seen around 5 pm affects the results? In other words, what kind of spectral window is applied in the analysis?

- We have addressed the issue with the filtering of the horizontal wind for the wind fluctuations. The filtered wind is used to compare with the large scale waves in the airglow keograms. For correcting the wave period Doppler Shift, we use the total wind at the moment of wave detection, not the filtered wind nor the fluctuations.

Minor comments: Abstract

Line 1 'large and small'

- We have changed that.

Not quantitative. More specific description is wanted. The same comments apply to the other 'small' and 'large' in the manuscript.

- We have provided a quantitative estimation of what large and small horizontal scales mean in this paper (line 130-135).
- The small-scale events refer to waves with horizontal wavelength < 725 km, while large-scale events are considered here as having horizontal wavelength > 725 km. The 725 km range corresponds to the diagonal line across the field of view of an image mapped into a 512×512 km² grid. This way, a 725 km wave would still be fully seen within the image frame (on crest, one trough).

Line 6 'flux divergence estimations'

Momentum flux is estimated, but the divergence does not seem estimated, instead MF is just assumed to be deposited around the mesopause.

- We have calculated the momentum flux of each wave detected in airglow during the four nights of the campaign because the auto-detection method outputs all the wave parameters necessary for the MF calculation. Thus, the momentum flux was calculated for each wave individually. However, the flux divergence was estimated from our simulations carried out in Vargas et al. 2007. The divergence estimations assume the waves are saturated, meaning that the wave amplitude doesn't change between two airglow layers. The assumption that the majority of waves are saturated is based on lidar measurements performed over Maui and SOR Observatory in New Mexico. In our reply to reviewer#1, we have explained how we would calculate the flux divergence directly from the airglow images. We would have to follow an individual wave to verify if it's visible in two layers simultaneously, then we would measure the MF of this wave in each layer, and finally subtract these values to have an estimation of the divergence for that specific wave. However, we have not done that in this paper.

Line 11 Mean MF, total MF and number of detected events

585.96/362 is not 0.88.

- The reviewer is correct. We have updated the estimation to 1.62 ± 2.70 m²/s².
- We have realized that showing the total momentum flux of all waves detected brings more confusion (see reply to reviewer#1) and decided to remove that from the manuscript.

How is the mean estimated? Is the detected number 362 the number of independent wave events, or multi-counted number? The same comment to the lines 285-286.

- The waves are not necessarily independent, but it would be more likely that each detection corresponds to an independent wave detection.
- As we rely on a set of three images at the time from where we obtain two time-difference images, and finally the cross spectra from the 2DFFT of the time difference images. So, the time span of each set is about 30 min, and most of waves would have completely disappeared the imager FOV within that time span because the duration of a gravity wave is short based on our experience.

Lines 17-18

The lidar data in the upper stratosphere is not actually analyzed in the manuscript, but only qualitatively mentioned. This statement does not seem appropriate in the abstract.

- We have removed that.

The main body of the manuscript

Line 52 'modes'

Meaning not clear.

- We have reworded the sentence for clarity.

Line 58 'large amplitude filtered wind fluctuations' Meaning not clear.

- We have removed the word filtered to avoid confusion.

Line 84 'long period oscillations'

The wave structure is not very clear. Even one cycle is hard to distinguish, especially for the Nov. 06-07 case.

- The Nov 6-7 case is less pronounced, but can be seen in the zonal keogram around 1200 UTC. We have overlaid ellipses to make it easier for the reader to identify the airglow structures. It is clear that we have only considered these structures because they have also caused pronounced fluctuations on the wind fields and are associated with those large-scale waves, otherwise it would be hard to prove and explain that they would be gravity waves indeed.

Line 91 'time difference'

Is this done using pairs of the adjacent 10 min separated images? The detail should be written.

- We have put details of the image processing in the appendix, one for the auto-detection method, and other for the keogram analysis. We hope this makes this part clear for the reader.

Line 116 'smaller' > shorter

- We have rewritten the sentence.

Line 116 'simultaneously the back.....'

'with' after 'simultaneously'

- We have rewritten the sentence.

Line 118 'Fig. 1c'

Meant to be Fig. 1d?

- It should be meridional keogram in Fig. 1c. The figure 1d refers to the small-scale waves ($l_h < 725$ km).

Section 3

The definition of 'small' and 'large' should be described at the beginning of Section 3. Sub-section 3.2

- We have done that as asked. The small-scale events refer to waves with horizontal wavelength < 725 km, while large-scale events are considered here as having horizontal wavelength > 725 km. The 725 km range corresponds to the diagonal line across the field of view of an image mapped into a 512×512 km² grid. This way, a 725 km wave would still be fully seen within the image frame (on crest, one trough).

The meaning of 'fluctuation' at line 182 is not clear although it is inferred later at lines 190-191. It is confusing. The filter characteristics also should be written in the caption of Figure 6.

- We have removed the word 'fluctuation' for simplicity since the result of the weighted wind calculation can be seen in the figure. The details of how the wind fluctuations were obtained are also added to the caption of Fig. 6.

Lines 210, 211, 219, 220, 224, 234, 235 316, 317, 320 and 321

Why are the uncertainty of all these values '+-1'?

- The uncertainty corresponds to the smallest division in the keogram temporal axis.

Line 244 'amplitude'

Since the background wind is not thought to be a wave, rewording will be appropriate; something as magnitude.

- We agree with the reviewer and have changed that.

Line 245

Absorption occurs at a critical level, but reflection occurs at a turning level. Different wind structures are responsible, perhaps as you know.

- There is confusion in the literature. I think as critical level the altitude where the wave could be absorbed or reflected, as the word 'critical' suits both. If we would specify further, then we should use turning level to the altitude where the wave is reflected, and absorption level the altitude where the wave is absorbed. It is much clear and gives no room for confusion. We have removed "critical" and use only absorption or turning level to be precise of the meaning we want to convey.

Lines 283-284 'toward the west of -0.36 ± 1.51 m²/s²' Very confusing description.

- We have detailed further the explanation.

Lines 286-288 The finding of a small number of waves carrying a large portion of MF is a good point. I suggest that the authors include discussion on GW intermittency and its importance in model atmosphere studies.

- We have added the discussion about wave intermittency.

Lines 289-294

This can largely affect the obtained results. Hopefully more discussion wanted. Isn't there any possibility that the layer thickness is related, or is it already taken into account in the analysis? It seems from Figure 4 that O2 layer is the thinnest.

- Wave detections in separate layers do not rely on the information about layer thickness or layer altitudes but only in the amplitude of the wave that must be large enough to produce a prominent peak in the cross spectrogram of amplitude. Thus, those features (layer thickness and altitude) hardly would influence the results. However, the fact that large amplitude waves are seen in the O2 layer is surprising given the layer overlapping structures. This must be investigated separately. At this point we don't have a good explanation and providing one right now would be pure speculation. However, the O2 has the narrower estimated FWHM for the campaign, and that would allow shorter vertical scale waves to be seen in the O2 images, and consequently larger momentum flux waves would be measured there (see lines 311 -313).

Line 296

Yes, if the acceleration or deceleration occurs all along the same latitude circle, to my understanding.

- We could not fully understand the reviewer point of view here. We could provide a better answer if the reviewer is willing to elaborate further this comment.

Lines 339-340

I think this is a nice point. In the radar studies, a long-lasting topic has been which period range of GW is most responsible for MF deposition. Short period GWs were thought to be responsible before, but recent studies suggest that longer period waves are responsible. One example I know is Sato et al, 2017 (<https://doi.org/10.1002/2016JD025834>). There can be more. Including such preceding studies will be beneficial for even more fruitful discussion.

- We have added a discussion as suggested and also added the Sato et al. (2017) paper to the reference list.

Figure 2 '> 5 hours' should be '< 5 hours'?

- The reviewer is correct, although as explained before, there is not a hard cut off in the filtered periods given the way the fluctuations were calculated, that is, Wind fluctuations were obtained by first applying a 4-hours, 4-km temporal and vertical windows to get winds representing scales larger than these values. Then we subtract these estimates from the total wind field.

Figure 5

Are these dots independent wave events with each other?

- The dots are not necessarily independent wave events but are treated here as independent waves here. That does not affect the average momentum flux because, by clustering the wave events detections into independent wave events, we would have to calculate the mean momentum flux of each individual independent wave and then average the momentum flux of these independent waves in order to know the mean momentum flux during campaign or each night. We have indeed a clustering algorithm that does the job for us, but we decided not to do so in this manuscript.
- As we rely on a set of three images at the time from where we obtain two time-difference images, and finally the cross spectra from the 2DFFT of the time difference images. So, the time span of each set is about 30 min, and most of waves would have completely disappeared the imager FOV within that time span because the duration of a gravity wave is short based on our experience. This way, it would be more likely that each detection corresponds to an independent wave detection.

Figure 6

Meaning of 'fluctuation' should be described clearly. High pass filtered?

- As per request of the reviewer, we have added to the figure caption a short explanation of how the fluctuations have been obtained from the total wind field.

Figure 7

I am afraid that a spectral analysis after applying a filtering is not acceptable.

- We believe the spectral analysis of the wind fluctuations is not faulty. We did not explain correctly how the wind fluctuations were calculated, but this issue is fixed now. This present figure shows indeed the relevant wave amplitudes during the campaign, even though we could not suppress completely the tidal modes. A significance threshold line (99% confidence) has been added to the plot in figure 7 to show the relevant wave amplitudes and periods

Figure 8

Image data and wind data should be treated in a similar manner; detrend, background removal or whatever.

- The images were detrended before spectral analysis, and also the keograms. The Wing fluctuations already have mean close to zero and no trend, as it is seen in figure 6b.

Figure 9

The way of 2D spectral analysis should be described in the manuscript.

- We Have added the 2D spectral analysis of keograms in the appendix, where we describe completely how the analysis is carried out. we have also added another appendix session to explain further how the auto detection method works. We hope that this will make the paper more self-explanatory, avoiding to relying on separated publications.

Figure 10

If all the figures are processed with the high pass filter with the cutoff of 5 hr, reprocessing is suggested, at least, for the November 6-7 case.

- That is not really a digital filtering operation. We have included the following in the text to make clear how we carried out the wind processing: "To obtain the wave fluctuations (Fig2b), we first use 4-hours, 4-km temporal and vertical windows to get winds representing scales larger than these values. Then we subtract these estimates from the total wind field. The total wind field was calculated using smaller temporal and vertical windows, that is, 30 min and 1 km, respectively. By doing that we reduce the contributions of large-scale winds, nonetheless the resulting fluctuations still keep some of the energy of larger periods waves as can be seen in the spectrum."
- As we have explained before, the Nov 6-7 case is less pronounced, but can be seen in the zonal keogram around 1200 UTC. It is clear that we have only considered these structures because they have also caused pronounced fluctuations on the wind fields which are associated with those large-scale waves, otherwise it would be hard to prove and explain that they would be indeed gravity waves

Interactive comment on "Mesospheric gravity wave activity estimated via airglow imagery, multistatic meteor radar, and SABER data taken during the SIMONE–2018 campaign" by Fabio Vargas et al.

Anonymous Referee #3

Received and published: 23 October 2020

The comment was uploaded in the form of a supplement: <https://acp.copernicus.org/preprints/acp-2020-896/acp-2020-896-RC3-supplement.pdf>

General comment:

The authors describe different measurements they carried out during the SIMONE-campaign in November 2018 at or around Kühlungsborn, Germany. Supplementing satellite-based data are used. They combine the measurements in order to fully characterize gravity waves and discriminate between large and small-scale waves. However, they do not explain what they understand under small and large scale.

From my point of view, there are two main results. The authors found out that approximately 11% or all detected waves carry ca. 50% of the total momentum flux and they had a specific look on two large wave events. They motivate that these structures might be due to secondary gravity waves. Nevertheless, this remains speculative. The manuscript is well-structured and sections 1 to 4 are easy to read in the first approach. However, on closer inspection important information is missing. In section 5, the formulas disturbed my flow of reading. Following the classical approach, they should appear earlier in the manuscript which would help here. The data and the methods used to measure and identify the gravity are appropriate, however especially the data section could have some more details. My main point concerning the section results refers to error bars and/or significance levels: I didn't find a single one.

I propose to accept the manuscript after some major revision.

- We thank the reviewer for raising important comments about our paper and for taking the time to do so. We appreciate your help very much.

Major points

The manuscript is written in a rather abbreviated style.

The most important point in this context is that the manuscript suffers from a lack of information about the quality of data and analyses.

- We understand the reviewer's concern about having written the manuscript in an abbreviate style. Maybe because we wanted to make it concise, we have failed while describing important parts of the analysis and processing, which must be done precisely to permit the reproduction of the results.

I found neither a single error bar nor a significance test in the section results (in the discussion part there are very few error bars). For clarity issues, it is probably not helpful to provide this information directly in all figures (for example I have no idea how to incorporate error bars in keograms), but it can be done in some cases (FFT) and in the others the information can be given in the text. Parts of the main results depend on statistical analyses.

- The incorporation of error bars in the spectral analysis could be found from our simulations of the error propagation into gravity wave parameters in our recent paper (Uncertainties in gravity wave parameters, momentum fluxes, and flux divergences estimated from multi-layer measurements of mesospheric nightglow layers, <https://doi.org/10.1016/j.asr.2018.09.039>.)

A second point in the context here is that the authors do not explain how they define wave event and which consequences this definition has on their results. If I operate an imaging system which takes a picture every ten minutes, is each result of the spectral analysis of each image defined as a wave event or is a wave event an oscillation in space which shows nearly the same horizontal wavelengths over some time (so in some images). Fast waves would be underrepresented in the first case. Does this have any influence on the results?

- For the small-scale analysis using the autodetection method, we have defined as a gravity wave event the output of the cross-spectrum of two time-difference images when the special peak is larger than 10% of the total energy of the spectrum. Two time-difference images are generated from three light frames of the

airglow (please take a look at fig 01 for Vargas et al 2009
(<https://angeo.copernicus.org/articles/27/2361/2009/>) showed here below:

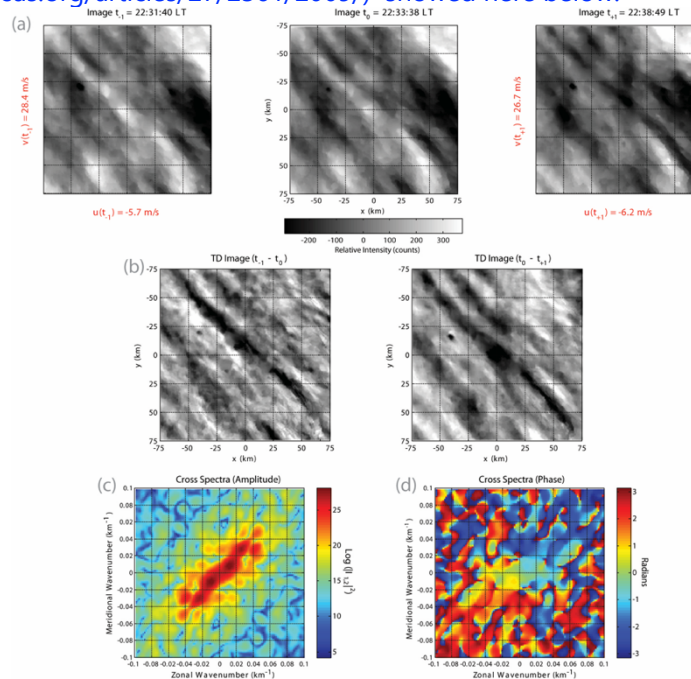


Figure 1 Processing of a set of three airglow images in order to obtain the cross spectra. In (a) it is showed a set of sequential OH images presenting GW structures. (b) Time difference images obtained from the set of images. (c) Amplitude cross-spectra of the TD images, from where we estimate the wave amplitude, propagation direction and horizontal wavelength. (d) Phase cross-spectra of the TD images showed in (a).

- Thus, if a wave is detected in a given set, it is considered an independent wave detection. Now, if in the next set another wave detection is made, the only way to tell the two events to correspond to the same gravity wave is by comparing their wave parameters. Now, the momentum flux of a wave varies as it goes through the field of view. thus, considering the momentum flux average of all the waves detected during the campaign would be the same as clustering the corresponding wave events into a distinct wave, and then averaging the momentum flux for that specific one for the duration of the event over the airglow images where it shows up, and then averaging over all the distinct wave events, just would be more laborious. We have a clustering algorithm that does work for us, but still, we choose not to do it here.
- Again, we rely on a set of three images at the time from where we obtain two time-difference images, and finally the cross spectra from the 2DFFT of the time difference images. So, the time span of each set is about 30 min, and most of waves would have completely disappeared the imager FOV within that time span because the duration of a gravity wave is short based on our experience. This way, it would be more likely that each detection corresponds to an independent wave detection.

Another point is: It is really strange to see a paper where TIMED-SABER data are used and Martin Mlynzack and James Russell (or at least the SABER team in general) are neither co-author nor mentioned in the acknowledgements. There isn't even a single citation of them (e.g. Russell et al., 1999, <https://doi.org/10.1117/12.366382> or Mlynzack, 1997, [https://doi.org/10.1016/S0273-1177\(97\)00769-2](https://doi.org/10.1016/S0273-1177(97)00769-2)). On the other hand, there exist co-authors here who also "only" delivered data. This is not consistent. Maybe the authors offered co-authorship to the SABER people and they declined, I don't know that, but is there really nothing to mention in the literature list or in the acknowledgements?

- Regarding the SABER data, we did not acknowledge the SABER team in the proper section, but we should have. We clearly thank the SABER team now in the acknowledgements statement. We have not offered coauthorship to the SABER team indeed and are not aware coauthorship is required to use SABER data. We

think an acknowledgment note will suffice. We have also included the suggested references to the list and made citations at proper locations.

Minor points

Section 1

General comment: The introduction is quite general. In my opinion, the technical focus of this manuscript is on the combination of different measurements to describe gravity waves as comprehensively as possible. Airglow imagers are used in order to derive horizontal wavelengths and periods. Wind information come from radar data. Temperature information are based on SABER data. Other groups have already done such studies or similar ones - just enter 'airglow radar SABER gravity waves' in google scholar. However, there is no reference in your manuscript. One of the main topics of the manuscript is the derivation of the momentum flux. There are already other publications on this topic which are not mentioned here or in the discussion (e.g., Ern et al., 2011, <https://doi.org/10.1029/2011JD015821> derived the momentum flux globally based on satellite data). Please include some citations and put your work into the context.

- We have included citation of other references including momentum flux derivation in the introduction section, including that on of Ern et al, 2011

p. 2 l. 48 Can you please provide the exact campaigns period?

- We have done so.

p.3 l. 55 & 57 What is small and large scale for you? This comments also refers to the headings of section 3.1 and 3.2.

- The small-scale events refer to waves with horizontal wavelength <725 km, while large-scale events are considered here as having horizontal wavelength >725 km. The 725 km range corresponds to the diagonal across the field of view of an image mapped into a 512x512 km² geographical grid. This way, a 725 km wave would still be fully seen within the image frame (on crest, one trough). We have also included these specifications in the text.

Section 2

From my point of view, the level of detail of the different subsection varies (2.2 has more details than the two other subsections). Additionally, I do not get all "classical" information for all instruments such as spatial and temporal resolution or some information about the data quality (not necessarily bias and precision but for example comparisons with other instruments, for SABER such studies are available).

I also miss further literature for ASI and SABER where the reader can look up (technical and data processing) details about the instruments (e.g. concerning ASI the detector size etc., concerning SABER the retrieval etc.).

For SABER, different data versions exist. I conclude from the legend of figure 4 that you use the newest one (2.07) but not every reader might be familiar with the "SABER notation", so please provide this information also in section 2.3.

- We have done our best to attend your request.

p. 4 l. 87 In the text you use UTC, in figure 1 you use p.m. If it was consistent it would be easier to read.

- We understand your concern, but, respectfully, we have decided to keep it as is because this minor issue will not affect the interpretation of the plot and the results of the paper.

p. 4 l. 94 You write that the result of a time-difference operation is an image where the contrast of small-period, small-scale oscillations is enhanced. I think with small-scale you mean the spatial dimension. In this case, I do not fully agree with you. Whether you enhance a signature or not should depends on the speed of the signature: if it is near zero you won't see it in your time-difference image because you just subtract it. The phase speed is directly proportional the wavelength (in this case the horizontal one) and indirectly proportional to the wave period. The phase speed significantly differs from zero, the larger the wavelength or / and the smaller the period is. So, you enhance small period but large-scale oscillations.

- The reviewer has presented a good point. Our definition of small horizontal wavelength (lh) wave considers $lh < 725$ km. Once the detections correspond to waves of small periods (< 60 min) and $lh < 150$ km, our statement of small period small horizontal scale waves is consistent. The plot below (figure 2) shows a calculation of how the TD operation would affect the wave amplitude as the wave period vary. It shows the ratio between the actual wave amplitude (A) in the original image to that on the TD image (A_T). We considered the sampling period of 2 minutes. In the plot, we see the wave amplitude decays to $< \frac{1}{2}$ of that the original wave amplitude for periods larger than ~ 25 minutes. So, the larger periods waves must present strong amplitudes to be detected by the time-difference method.

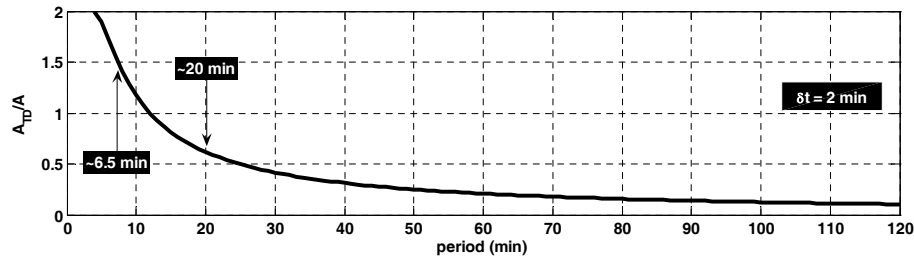


Figure 2 Ratio between wave amplitudes observed in a TD image and in an original image as a function of the intrinsic period of the wave. The sampling period used to compose this graph was 2 minutes.

p. 4 l. 115 Which high-pass filter exactly? If you say you use a five-hour high pass filter, it means for my that all periods longer than five hours can pass. Later in the document, it becomes clear that the opposite is true and you probably referred with the term “high-pass” to the frequency domain. Can you please clarify here which periods are still present in the filtered data. Additionally, in the spectra in figure 7 you still see a rather dominant p. 4 l. 118 Are you sure that you really see exactly this oscillation?

- We should not have referred to that operation as filtering because reader will think as a digital filter operation. Here is how we did it: The total wind field was obtained by processing the raw meteor radar data using a 30-min, 1 km window. To calculate the wind fluctuations, the raw meteor radar data was processed using 4-hours, 4-km averaging windows. Then result was than subtracted from the total wind field resulting in the wind fluctuations associated with gravity waves. We have corrected this in the paper to make it clear for the reader what has been actually done to obtain the wind fluctuations.
- From this, it is clear that oscillations of periods longer than 4 hours will be suppressed, but not completely eliminated. Also, waves having vertical scales < 4 km will be suppressed, but not completely eliminated. The spectra in figure 7 reflects the fact that not all longer period oscillations were eliminated, but it doesn't mean the spectra is incorrect. We also can tell what peaks are associated with gravity waves once waves of similar periods are also seen in the keograms.

p. 5 l. 120 The abbreviation SABER is not explained in the manuscript.

- We have explained the abbreviation earlier in the manuscript when we first mention the acronym SABER in line 59-60.

p.5 l. 127 Please concretize the lapse rate you mean.

- The reviewers has asked to explicitly give the value of the lapse rate, is that correct?

p. 5 l. 127 Please be a little bit more precise here. It is not important that the orbit of TIMED is exactly over the observatory, the field of view of the instrument has to measure over or near the observatory.

- We have elaborated more. The specific fragment now reads “Even though the satellite orbits registered during SIMONe-2018 were not exactly over the observatory, the instrument measurements are performed in limb vertical plan that is near or within the field of view of the imager (Fig. \ref{fig03}), where the colored dots indicate where the measurements are being made, not the satellite position”

p. 5 l. 128 From figure 3 it looks like you use also profiles which were not within the field of view of the imager.

- The reviewer is correct. We have used all the measurements we could find to diagnose the background atmosphere stat in the nighttime period during the campaign.

p. 5 l. 133 It is great that you are able to calculate the VER profiles for the different species but why don't you use the ones provided by SABER?

- The reviewer is correct. That is because we did not want to mix in the analysis the use of measured profiles with the calculated profiles. It would not cause a big change in the results; thus, we carried out the analysis using only the calculated profiles.

p.5 l 137 Please correct the letter shift (FWMH→FWHM) and explain the abbreviation.

- We have done so.

Section 3.1

The way to present such campaign results is always a tightrope walk. If the reader is flooded with details, he loses the overview, if it is too short, questions remain open. For me questions remain open.

I miss two formulas (dispersion relation for high-frequency waves and vertical flux of horizontal momentum, or at least citations of literature where the reader can look them up) and in the case of the dispersion relation I would like to read some information about the Brunt-Väsälä frequency (probably calculated based on SABER, VER weighted?).

- We have addressed the lack of explanation by providing two appendix sections where we detail the auto-detection method and the keogram spectral analysis for the readers.

Please also provide the information that the cross-spectrum is the Fourier transform of the cross-covariance function, since you treat pictures you probably use the 2D version. Another useful information would be the sensitivity of the analysis. You calculate the cross-covariance of two TD images. The calculation of TD images already means filtering the data for fast waves. The calculation of the cross-covariance provides you information about the waves which change from one TD image to the other. So, there are two stages of filtering and it would be interesting to provide information which waves pass this filter, and which don't.

- The reviewer is correct, we have indeed used do 2D Fourier transform in each time difference image and then combine the spectrum of these two obtaining the cross spectrum the figure 1 above shows the process clearly.
- One thing that is not clear in the figure is that we have correct the images (shifting pixels) to account for the wind velocity across the field of view, that is, we have performed the Doppler shift correction in the wave period already. This process is depicted in the figure 3 bellow. Notice the red arrows indicating the shift in a bright spot of the image.

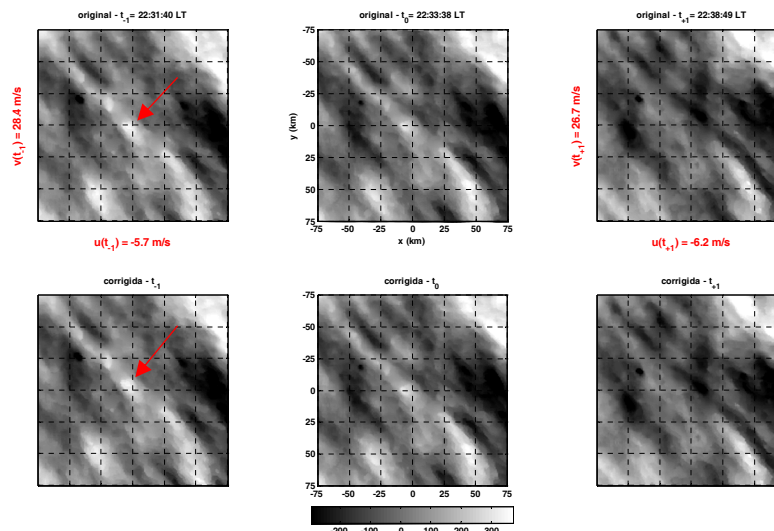


Figure 3 Doppler shift correction by moving pixels in the processed image set

p. 6 l. 165: I think you mean 5j instead of 5e and the wavelengths range between 10 km and 40 km.

- The reviewer is correct. We have changed that accordingly.

p. 6 l. 169: I see 100 m/s instead of 80 m/s.

- We have changed that accordingly.

p. 6 l. 176: "larger momentum waves on Nov. 2-3 and Nov. 3-4." Ok, it's a log plot but the maxima of these two nights are not so much higher than the maximum of Nov. 1-2, for example, which brings me to the question of error bars. If you plot them, figure 5l might become a little bit busy, so this might not be the optimal solution, but in any case you should give information about the error bars somewhere (also concerning the other parameters shown in figure 5) and state whether the sentence about the date of appearance of the larger momentum waves is still valid in this case.

- We have calculated the error bars of each parameter of each wave event measured. These errors are calculated using the methodology developed in Vargas 2019 (<https://doi.org/10.1016/j.asr.2018.09.039>). We will provide the average error of these parameter in each plot for completeness. As the reviewer says, the plots will become busy if we insert error bars in each data point.

Section 3.2

p. 7 l. 184 "once the layer peaks within +/-2 km from each other"? The wind fluctuations weighted with the VER of O(1S) and OH(8-3), so figure 6 d and f, look really similar, even though the respective peak heights are more than 4 km away from each other.

- WE have doubled check our calculations and verify they are correct. The weighted wind fluctuations are similar indeed, but not identical. The similarity is related with the long vertical wavelength fluctuations that can be seen in Figure 2c and Figure 2d. we have added more explanation to the text (line 191).

p. 7 l. 185 The citation you provide refer to the equatorial region, you should mention at least one citation referring to mid latitudes (e.g., Wüst et al., 2017, <https://doi.org/10.5194/amt-10-4895-2017>).

- We have inserted the suggested citation.

In any case, 15 km is indeed large. When I look at figure 4 c (VER measured by SABER), I find values of 9 km to 11 km which is much more in the range of other authors. Your simulated profiles in figure 4d (thick lines) show higher values but none of them reaches 15 km. So, first I wonder where the 15 km FWHM comes from and then I would be interested in why is there a difference between simulated and measured values? Which one can I believe or are both ok within certain limitations? And now we are again in the range of error bars, quality of the data etc.

- We have explained in the text that the FWHM was calculated from a gaussian fit of each VER. The layer thickness is in a table 2 for all the simulated and observed VER. It is clear that we if we proceed in the same way as the reviewer, by measuring the FWHM directly from the graphs, we will have much shorter or narrower FWHM. We wanted to be consistent while composing Table 2 by using the FWHM from the gaussian fit of the VER curves.
- Also, the measured profiles could be affected by long period waves which would push the layer up and down and also boot cause it to be narrower or thicker depending on the phase of the perturbation affecting the layer. See figure 4 below of our simulation of how a wave causes the depicted effects in the redline layer from Vargas 2019 (<https://doi.org/10.1029/2019JA027356>):

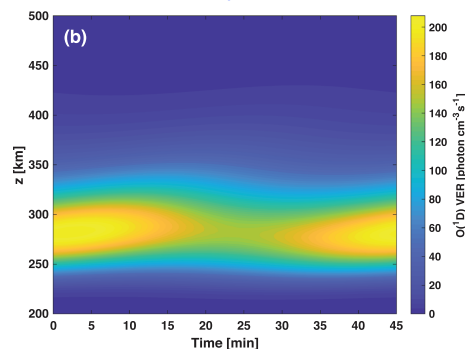


Figure 4 Time versus altitude cross section of the (b) perturbed O(1D) volume emission rate,

p. 7 l. 187 The values after the +/- are what?

- The value refers to the spread in the rms value of the wind.

p. 7 l. 189 I miss a significance level in figure 7.

- We have replotted figure 7 using the total wind field showed in figures 6a, 6b, and 6c. The significance level has also been added and the text has been modified accordingly.

p. 7 l. 193 Please omit "salient" – this peak is as salient as the 12 h peak for example, which is only due to leakage as the authors say. By the way, I find it strange that peaks which are due to leakage are as prominent as others. Here, a significance level and more information about the filter as already mentioned before would help.

- We have removed the word salient and have added further details of how the fluctuations are obtained (line 196).

p. 7 l. 211 I find it difficult to see a period of $8 \text{ h} \pm 1 \text{ h}$.

- Evidently the wave is propagating zonally eastwards. The tilt in the brightness patch indicates how fast the wave is traveling in that direction, and because the tilt is small, we can only see that it is a small speed wave. To obtain the wave periodicity we had to rely on the keogram spectral analysis, but the keogram helps to verify the brightness patch is due to an 8-hours large-scale gravity wave, not an 8-hour tidal structure.

p. 7 l. 215 Which spectral analysis? Please concretize.

- We have added more explanation in the appendix section to show how the spectral analysis of keograms was carried out. We also have put that on the text.

p. 7 l. 217 The keogram analysis does not only provide horizontal wave numbers, you also mention the wave frequency in figure 9. Can you please concretize this in the text? Furthermore, you mention the observed frequency $\hat{\omega}$ in the text but the frequency ω in figure 9. You probably mean the observed frequency in figure 9, don't you?

- We have done as asked. Yes, figure 9 refers to the observed frequency. We have corrected the symbol in the figure and in the figure caption. We perform doppler correction in the wave period later on using the background wind in order to estimate the large-scale waves momentum flux.

p. 8 l. 118 &v226: Is there a difference between ~ 0 and 0.0×10^{-3} within the unknown accuracy of your analysis and the unmentioned significance levels?

- Based on our auto-detection method, the significance level for a spectral peak to be considered as a wave event is that the sum of the spectral energy around the spectral peak is larger than the total energy of the spectrum. Fig 9b and 9c show that for different waves, the spread of energy around the peak is different, so the uncertainty of each one depends on the spread, where the threshold wavenumber is at where the energy sum around the peak is $< 10\%$. For more information, please refer to Vargas 2019 (Advances in Space Research 63 (2019) 967–985)

p. 8 l. 228 & 230: You mention that the phase is propagating downward but the wave is propagating upward? Don't you mean the energy is propagating upward?

- We know that gravity waves are dispersive type of waves, and because that, if the wave packet is propagating upwards, that is also the direction of energy propagation. In our view, we could use one way or another interchangeably.

Section 4

p. 8 first paragraph: you argue that the atmosphere is convectively stable using the mean temperature profiles over some days. When discussing dynamic instability, you look at individual measurements and argue that they can reach 50 m/s and more. This is not consistent having the mean on one side and the individual profiles on the other. Please go a little more into detail here. Furthermore, you write that since the wind field is rather strong it could cause critical levels – wind fields can always cause critical levels. It just depends on the phase velocity of the wave. You probably mean that large wind speeds are more likely to cause critical levels.

- The reviewer is completely right regarding the discussion about the wind velocity. Actually, the wind shear above 40 meters per second per kilometer would be the cause of dynamic instabilities. We have rephrased the sentence to read "Thus, gravity wave dissipation due to convective instabilities would not affect the vertical

evolution of the gravity wave field during the campaign. We have also verified that dynamic instabilities did not occur because the wind shear was <30 m/s/km most of the time during the campaign. On the other hand, because the horizontal winds occasionally achieved relatively large magnitudes >50 m/s (Fig.02a and Fig.02b), the wind field could have caused absorption of waves having phase speed <50 m/s traveling to the same direction of the wind vector."

- We have removed "critical" and use only absorption or turning level to be precise of the meaning we want to convey, once we believe "critical" causes confusion for readers not aware of the details of gravity wave propagation in vertically structured atmosphere.

p. 8 second paragraph: You write that the wind is controlling the propagation of southeastward waves via dynamical filtering and draw this conclusion from the comparison of two figures, one addressing the momentum flux direction (figure 5a) and the other showing the wind direction (figure 5k). Why don't you conclude this from the comparison of figure 5f and k? Figure 5a tells you only that there is not much momentum transported in southeastward direction, but this can be due to too "weak" waves or too few waves or a mixture of both.

- WE could have done so from figure 5f and k, but we wanted to show that looking at the momentum flux azimuth chart is not enough to tell that the wind is controlling the horizontal propagation direction, once the wave propagation direction azimuthal histogram still shows waves moving into the wind.

p.9. l. 253 You mention a wind speed of ca. 39.3 m/s. Is this the mean over the background wind or over the total wind (including the gravity wave signatures)?

- 39.3 ± 18.9 m/s refers to the wind speed mean and standard deviation in the southeast quadrant sector (270 to 360) in Fig. 5 k. (we have now stated that in the text in lines 264-265).
- The mean wind speed on the 270-360 degrees quadrant corresponds to the background wind.

p. 9 l. 253 & 254: the number behind the plus-minus sign is the standard deviation?

- It is the standard deviation of the sample.

p.9. l. 255-281: you compare your results to two other publications, one of them is a PhD thesis and not peer-reviewed I assume. It is not clear to me why only these two publications are suitable for comparison. There exists a number of other studies since 2011 just type 'airglow gravity waves horizontal wavelength' in google scholar, even if you restrict your search to 2016 to 2020, you still get many results - not all of them might deliver all wave parameters you derived but even in this case there exist some. So, could you please extend your comparison here accordingly?

- We have done so as per your suggestion and the other reviewers as well.

p.9. l. 256: You write that your results are directly comparable to Li et al. (2011) since they used a similar auto-detection method. However, you say nothing about the sensitivity of the instruments, which is used by those authors, to the different horizontal wavelengths. The sensitivity of an airglow imager should be determined by the field of view and the number of pixels your imager has, for example. As you mention in line 265 also other parameters such as filter wheel cycles or integration time determine the sensitivity. Can you please adapt the formulation in the manuscript accordingly?

- As far as we know, the Li imaging system was not calibrated for sensitivity, Notice the Li system is the same used in ALO, Chile, since 2009, and we have total understanding of the system details. The IAP imager system is also not calibrate for sensitivity.

p.9 l. 267: Please replace min by minutes

- We have replaced "min" to "minutes" in the whole manuscript.

p.9. l. 269: You write that the filter wheel cycle which is used does not only influence the derived periods, it seems to influence also intrinsic phase speeds. It should! Phase speed is related to the ratio of wavelength and period. The filter wheel cycle influences the sensitivity of your measurement system in the temporal domain, it does not affect the spatial domain. If you can detect waves which are characterized by smaller periods but by horizontal wavelengths in the same range than before, you will automatically see faster waves.

- We have included that in the text to make the readers aware of that as well in lines 284-286.

p. 9 l. 275-277 Can you please provide citations here or is this information taken out of Li (2011)?

- The statement is based on our personal experience observing gravity waves with ALO airglow imager at ALO, Chile, over 10 years, although Li (2011) provides the only study to date discussing the small-scale waves. Another study from our collaboration with a Chilean university has been submitted but is still under review, and maybe can be include in the citation list if gets published sooner.

p.9. l. 280 Convective sources during winter? Ok, we could think about strong low-pressure system but since you do not bring further arguments here, can you please mark this sentence as speculation?

- We have in fact found satellite images for some of the days during the campaign presenting large convective structures to the north (over Oslo) and west (over Netherlands) of the observatory. We did not used those figures in this paper because we want to write another paper that questions the origin of the observed waves, including the possibility of secondary wave mechanism (based in lidar measurements) generating the observed large-scale ones.

p.10 l. 293 & 294 Your vertical wavelength is in the range of the FWHM. According to table 2, the O2 layer is characterized by the smallest FWHM compared to the other layers you observed. So, I would assume the strongest signal in this layer.

- We have added the following to that sentence: "It is not clear why the enhanced waves are seen most in the O2 emission once the layer's peaks nearly overlap, but this could be related to the fact that the O2 VER has the smallest FWHM (see Table 2). These shorter λ_z waves would be seen in images of the O2 emission primarily, and their momentum flux would be larger for it increases as λ_z decreases (see Fig. 5i)".

p. 10 l. 208 You left out $1/4H^2$ where H is the scale height. Please mention this. Did this omitted factor contribute to the error bar (given in l. 316) or is it not necessary?

- Using the full dispersion relation, the error in the vertical wavelength (l_z) has been estimated in our work Vargas et al. 2019 (Advances in Space Research 63 (2019) 967–985). We have found there that the error in l_z can is large because it depends on other variables that are also estimated and have errors. For instance, for a 10 % error in the horizontal wavelength and 10% error in period, the error in l_z would be about ~20%. Also, we don't know exactly where the airglow layers peak, and the altitude of the layer also impacts on the l_z estimation and in the l_z error. We have omitted the term $1/4h^2$ because it causes only ~5% difference compared to the reported value, although we have revised the error to be 20% at this time.

p. 10 l. 316 I wonder where the error bar comes from since you didn't provide any error bars in your analysis. Please comment.

- WE did a visual calculation of the vertical wavelength and have assigned an error of 1 km to that measurement that is ~2x the size of the vertical resolution (0.5 km) in Fig.10.

p. 11 l. 328 Please substitute m/s by m/s² when writing about g.

- We have corrected that.

Section 5

p. 12 l. 362 Where do these numbers come from? I find 22-41 m/s/day in the manuscript and this is for the assumption that the wave breaking continues for 24.

- The numbers come from Figure 9e of Vargas et al. (2007) for waves having $l_z = 20-25$ km and $l_h \sim 100$ km that are the 11% of the observed small-scale waves carrying more than 50% of the total observed flux in our study. We have explicitly state that in the paper now.
- Yes, that is in the case the wave breaks for 24 hours, which is an estimation. But it is likely the waves will die out earlier, so the impact of a specific wave does not last longer.

Section Acknowledgements

p.12 l. 378 Which ftp?

- We have included the link for downloading the data.

Figures

Figure 2

It would be nice to provide the hint that the y-axis of a and b are different to the ones of c and d. Non- radar people might further appreciate a note explaining why there is a daily cycle in the length of the vertical profiles.

- The daily cycle is associated to the number of meteors been measured at time throughout the day. The meteor density is larger at earlier morning ours and smaller at afternoon hours. Because the wind calculation relies on the number of meteors to obtain a quality wind data, when there are not enough meteors the wind cannot be calculated. This is reflected in the plot as the given daily cycle at certain altitudes ($z < 80$ km and > 100 km). We have added explanation of that in the paper.

Figure 3

This figure is a challenge. I hope I can explain my confusion and I think you can solve most of it by just showing the respective fields of view and not the values measured within the fields of view. First of all: what do the colours in the rectangle and in the ellipse mean, you do not have a colour bar or maybe two?

- The colors in the rectangle is the pixel intensity of the airglow image (we can provide a color bar for this one). The ellipse shows the radar FOV and its colors reflects the meteor density (maybe not a good idea to show the color bar for this one, which would make the plot a little busy).

The measurements shown are taken at which day?

- The measurement corresponds to different SABER orbits during the week of the campaign, specifically, DoY 307, 308, and 312.

I think the rectangle shows the imager but isn't the field of view of an all-sky imager a circle or a kind of ellipse when projecting it on a map with an equidistant lat-lon grid? You probably have just cut the area.

- The FOV of raw airglow images are indeed like a circle. During pre-process, the image is mapped into a 512×512 km², rectangular grid of pixels with resolution of 1 km/pixel. The lat-lon grid in the figure was then scaled to match the mapped airglow image as 1 degree in latitude and 1 degree in longitude do not correspond to the same distance.

Then: on the right hand-side of the plot where only the ellipse covers an area (not the rectangle), there is a kind of border. This effect also appears the bottom of the plot. It does not appear where the ellipse ends and only the rectangle covers an area. Why? Maybe you plotted the rectangle over the ellipse.

- We want to preserve the shape of the rectangle to show that the imager FOV. The ellipse is the MR FOV and the center is not the same of that the imager once the MR system operates at other location away from the imager site. We again scaled the MR FOV onto the image rectangle.

Last: Please insert a y-axis labelling. Puh!

- We have done that.

Figure 4

a) Can you please provide the different SABER profiles in different colors or line styles or ...? I wanted to find out more about the static stability at individual days, but this is not possible if all profiles have the same color and line style.

- We have added colors and line styles to help telling apart the different profiles.

Please provide explicitly day of the year and UTC. I can extract it from the legend, but this might not be clear for everybody.

- We have done that.

c) Can you please scale the x-axis in a way that all profiles also the individual ones are completely within the plot?

- We have done that as well.

Figure 5

a) What do you mean with momentum flux versus propagation direction? If I plot y versus x, I have two axis. You also have two "axis" (one is the radius and one the direction) but the radius seems to show the number of waves or wave events.

- The radial coordinate shows the magnitude of the momentum flux in Fig. 5a. The momentum flux scale is written between 60 and 90 degrees. Notice the larger momentum flux events are moving towards 120 degrees, having magnitudes of ~ 25 m²/s².

c) The value after the +/- is the standard deviation? About how many waves do we talk?

- The reviewer is correct. We have measured 362 events during the campaign, although some wave events are not necessarily independent. Some events could have been clustered as a single, individual event since the same wave could be seen in multiple image sets (1 image set = 3 sequential images). We have not clustered the events for this study.

d) j) and m) What is a wave event, so is it the result of one FFT or do you group the results of your FFT, e.g. if a wave is found in one FFT and a wave with rather similar parameters is found in the next FFT, is it the same wave or do you count it twice? Would it make sense to count it twice or would you include a bias in your statistics when counting it twice since fast waves will be counted less often than slow waves?

- Again, some wave events are not necessarily independent. As the reviewer notice, a bias could have happened if we were to analyze and FFT single images at the time. However, we rely on a set of three images at the time from where we obtain two time-difference images, and finally the cross spectra from the 2DFFT of the time difference images. So, the time span of each set is about 30 min, and most of waves would have completely disappeared the imager FOV within that time span, meaning that the same wave would likely not be detected in the next set of three images. This way, it would be more likely that each detection corresponds to a different, independent wave detection (line 315-320).

m) There is a strange white triangle in the lower right corner.

- It must have been a glitch of the software while coloring the bar of the histogram.

Figure 6

It would be better to extend the y-axis range especially in part (d) to (f) since especially at Nov. 7th the values are not shown in the plot anymore.

- WE have fixed that.

Tables

Table 2

The FWHM I read from figure 4c is roughly 9 km and 11 km for SABER mean VER profiles (thick green and thick black line) – why is there a difference to the values given here?

Here, I come back to one of my earlier comments. You simulate the layer but you also show values measured by SABER. How do I have to judge the differences? For example, you simulate another OH transition than you measure. The different OH transition dominate at slightly different heights (see e.g. von Savigny et al., 2012, <https://doi.org/10.5194/acp-12-8813-2012>).

- We have simulate the VER profile for the OH(8,3) using the SABER mean temperature and atomic oxygen for the campaign. I believe the difference between the simulated and measured lies on the average we have used in the simulation. I mean, to calculate the OH(8,3) VER, we rely on the mean temperature and atomic oxygen profiles. I believe the SABER OH and simulated OH VER would be much closer in structure if we have used single T and O profiles for the simulation, although we have not checked that here. We should include this explanation in the paper as well.