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.

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. 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. 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?

Another point is: It is really strange to see a paper where TIMED-SABER data are used and Martin Mlynczack 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 Mlynczak, 1997, 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?

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.

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

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.

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.

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.

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.

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?

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

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

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.

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.

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?

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

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?).

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.

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

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

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.

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.

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). 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.

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

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

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.

p. 7 l. 211 I find it difficult to see a period of 8 h±1 h.

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

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 $\widehat{\omega}$ in the text but the frequency ω in figure 9. You probably mean the observed frequency in figure 9, don’t you?

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?

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?

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.

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.

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)?

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

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?

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?

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

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.

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

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?

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.

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?

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.

p. 11 l. 328 Please substitute m/s by m/s^2 when writing about g.

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

Section Acknowledgements

p.12 l. 378 Which ftp?

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.

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 measurements shown are taken at which day? 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. 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. Last: Please insert a y-axis labelling. Puh!

Figure 4

a) Can you please provide the different SABER profiles in different colours 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 colour and line style.
   Please provide explicitly day of the year and UTC. I can extract it from the legend but this might not be clear for everybody.

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

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.

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

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?

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

Figure 6

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

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).