Review of: ”Medium-scale gravity waves observational methodology for antarctic airglow observations” by Gabriel Augusto Giongo et al.

Overview

This manuscript introduces an improved method for detecting medium-scale gravity waves from ground-based airglow imagers. The method is based on keogram analysis, it is designed to automatically detect wave events and determine the horizontal wavelength, azimuth, period and ground-based phase velocity of the waves. It was tested using simulated data as well as actual airglow images. Authors report reliable wave identification and low errors of the retrieved wave parameters.

Unfortunately, the new methods, although they seem to contain good ideas for dealing with many aspects of the problem, are not explained clearly enough, to the point that I had difficulties interpreting some of the main results (see major comments below). I have also outlined some possible issues with the method itself and made a few suggestions, I hope they can be of some use. Since introduction of the improved methodology is the main goal of this paper, I cannot recommend it for publication in its current state.

General/major comments

1. **Severe lack of clarity in mathematical methodology and result presentation.** This manuscript cites quite a few previous works in relation to various methods used in wave analysis and error estimation. While it is the accepted norm for *derivations* of those results and their finer detail, the authors of this manuscript rely solely on the references and very general methodological concepts (e.g. “error propagation”) to describe the key quantitative parameters and mathematical methods used in their work. Because of this, I had difficulty understanding which physical or statistical quantities are described by the provided numerical values. Some of the most important results are presented in figures with undefined color scales, or as a percentage of an unstated quantity or as a standard deviation, without a clear a description of the data set for which the standard deviation was calculated (see major comment 2 and specific comments for concrete examples). I must apologize if I overlooked something or if some of the subsequent comments are simply a result of me interpreting some physical quantity not in the way that the authors intended, but, in my opinion, reader should not need to guess quite so often.

2. **Simulated keogram analysis results.** The authors mention “error propagation” in several places in the manuscript. For example, they state that “It is important to emphasize that this study is not trying to validate the wavelet transform or other well-known mathematical procedures, like linear fitting. Still, it tracks errors throughout the combination of processes employed. Errors due to any transform or mathematical procedure exist and can be studied or found in other publications (Liu et al., 2007; Xu et al., 2024).” This leads me to believe that quantities described as “propagated errors” are simply a result of statistical errors of well-know data analysis procedures (e.g. linear fitting, wavelet analysis) combined using the standard methods of calculating the error of a derived quantity from the errors of the parameters the quantity was derived from. On the other hand, running the analysis on simulated data provides an *independent* method of estimating the error: one can simply compare the results of analysis to the known parameters of the simulated data. Therefore:

   (a) I do not understand what authors mean in the beginning of Section 3, when they state that “Artificial images made as arrays containing known wave patterns were used to verify the quality of the analysis, track the error propagation through the procedure, and estimate future errors generated throughout the process”. Does this mean that the error estimates for real data run (i.e. "propagated error") was somehow based on the results of the simulated data run? If yes, then how? Was the simulated data used only for setting the threshold for acceptable linear phase fits, or for something else?

(b) Authors only provide one example of a simulated keogram, while "more than 800" were used in total. More information on this keogram set should be provided. Did every simulated keogram have two waves in it? What were the wavelengths and periods? What were the relative amplitudes?

(c) **(Important!)** Table 2 is extremely confusing. Firstly, I have no idea what the "standard deviation" means here. I did not find an explanation anywhere in the main text. Variance of the linear fit for phase is discussed in detail in this section, but it is not clear how it would be used to obtain the standard deviations for all the measured parameters. Are these just the standard deviations of the parameters used for all the different simulated waves? Secondly, what kind of error estimates are given in the table? Since this is a test on simulated data, it is vital to provide a clear overview of how wave analysis results compare to the actual parameters of simulated waves. Yet the table suggests that this is "propagated error". If this is just the propagated error in the sense I defined it at the beginning of this comment, what was the point of the simulated data run and why are the authors not showing how the analysis results compare to the actual wave parameters? If Table 2 does indeed show this comparison, then how do these errors compare to the errors provided for the real keograms (which are the "propagated errors", right?), where the wave parameters were not known in advance?

3. **Collocation of zonal and meridional wavelength estimates.**

(a) As one can see from figures 4 and 5, the wavelengths can end up being retrieved from a part of keogram that does not include the center (zenith) pixel. This can lead to a situation when zonal and meridional wavelengths are retrieved from completely disjoint parts of the image, neither of which contain the zenith pixel (which is used to identify temporal peaks in wave activity). In my opinion, this a major flaw of the method. Many structures seen, for example, in Figure 4, do not span the whole spatial extent of the keogram: there are major differences in phase structure and phase tilt between North and South, as well as East and West, parts of the image. Therefore, I would assume that method would often estimate the meridional and zonal wavelengths of completely different wave packets. The authors also seem to be aware of this: they state in Section 5 that waves not spanning entire keogram are a problem and that "The resulting period can be closer to one wave while the resulting wavelength is closer to another". It therefore hard for me to understand why so little care is taken to ensure the collocation of both wavelength estimations and the wave activity peak identification. Point (b) outlines one possible way to resolve that, just as an example.

(b) As far as I understand, the method described in the manuscript only makes use of one row and one column from each image. Furthermore, the peaks of wave activity are only identified from a time series obtained from a single pixel. It seems to me, that a simple way to improve the method would be to pick a few (let us denote their number by n) equally spaced rows and the same number equally spaced columns in the image. Then one would apply the wavelet power spectrum analysis not just for the zenith pixel, but for the n^2 intersections of rows and columns. The strongest peaks in these power spectra could then be used to identify wave events, and wavelengths could be estimated from the meridional and zonal keograms that intersect at the point with a strong power peak. This way, wave events that happen off-zenith could be identified. Even more importantly, one could apply stricter colocation criteria for phase fitting (e.g. only use the parts of the two keograms that contain their intersection point, where the power spectrum peak was identified) without loosing too many wave events. Finally, if this method ends up detecting waves in close spatial and temporal proximity to one another, one could use this as a consistency check. As far as I understand, none of the methods described in this manuscript are very computationally expensive, so it should not be a problem to run this for $n = 5$, for example. There could be, of course, many more (and probably better) ways to resolve the colocation problems and make better use of data from the whole image.

4. **The need for manual intervention.** In the beginning of Section 2.2 the authors state "Although the procedure is based on Fourier Transform as in previous works, the new feature of this methodology is that it automatically selects the waves along the keogram and verifies the quality of the oscillatory signal, resulting in less user bias on its usage". Therefore, the automatic nature of the new analysis method is the key novel aspect in this work. However, Section 5 implies that some of the most important flaws of the method could be adressed by manual inspection of results. In particular, phase lines of valid wave fits "can be bent but not tortuous, which is a consequence of a close superposition of the waves.". In my understanding, distinguishing a "tortuous" phase line from a bent one should not be too difficult to accomplish automatically: one could consider, for example, checking whether some norm (e.g. the Euclidean norm) of the second temporal derivative of the phase line exceeds a threshold. That would, of course, introduce another ad-hoc parameter to the analysis (besides the currently used threshold for the variance of linear fit), but this would still be more objective than manual inspection of results, that the authors have set out to avoid. Also, was any manual selection of wave fits used to obtain any of the results shown in the manuscript?

5. **Validation against whole images.** Table 2 shows the parameters of some retrieved waves. Those waves have wavelengths that are generally similar, or smaller, than the spatial extent of the airglow image (512×512 km). Would it not be possible to verify some of the image parameters obtained from keograms against the corresponding images? I understand that some of the images would be affected by the Milky Way and various other light sources, but would it not be possible to obtain at least some images, where the medium-scale waves discussed would be visible (perhaps after applying low-pass filter to suppress small scale structure)? I agree that keogram analysis has many advantages over this approach, but could individual images be useful at least for demonstrating that wave azimuth and wavelength were retrieved correctly? If this is indeed very difficult, maybe authors could have validated their new method against previous, more manual keogram-based wave analysis techniques?

Minor/specific comments

1. L29: I do not understand this sentence. The geographic location and altitude of a particular satellite observation (which is typically meant by the term "geolocation") can typically be determined pretty accurately. In particular, nadir-viewing satellite instruments (e.g. AIRS, AWE missions), offer similar geolocation accuracy as ground-based airglow imagers. Maybe the authors are talking about the spatial data coverage, and not geolocation here? Also, "local features and properties" is an extremely vague description given that methodology is the main topic of this paper! Authors should explain this statement in a lot more detail or remove it.

2. L106: "The peak identification procedure needs an equally dimensioned array". Why is that? Is that a fundamental requirement of the method, or is this just needed for compatibility with some particular implementation, software library, etc.?

3. Figure 2: The color scale in this figure is not defined or explained anywhere. This must be fixed.

4. Figure 3: Percentage of what? Power spectrum is normalized with respect to what exactly?

5. Figure 4: Units for deviation are not given. Also, the choice of color scale is very unfortunate, as there is very little contrast for deviations between 0 and 15, hence only the negative phases of most waves are visible.

6. Equation (5): while it is, of course, OK to refer the reader to the study by Xu for *derivation* of this equation, authors should explain clearly what C_f is and how it was used in this work.

7. L216: "Waves for a broad spectrum of velocity propagate in all directions without apparent anisotropy. However, more waves are seen with zonal components preferentially in the eastward direction compared to the westward direction". In my understanding, the second sentence directly contradicts the first: more waves with zonal components in the eastward direction is anisotropy.

Minor typos and suggestions

This is a list of typos that I noticed and minor, mostly language related, suggestions. Point-by-point replies to these are not necessary.

1. L11: I would suggest "The observed parameters of the waves [...]"

2. L15: I would suggest to replace "due to their energy and momentum driving" with "because they transfer energy and momentum", or merge the sentence with the following one, since I find "driving" a bit too vague in this context.

3. L23: "[...] short- and long-term variations of the waves' characteristics as well as short-term variations [...]". The repetition of "short" appears to be a typo.

4. L30: "Also, the computation treatment applied to the data can limit the data [...]". I might have somehow missed the connection with the rest of the paragraph, but, as it is, this sentence is both rather vague and rather obvious. I would suggest to remove it.

5. L67: (d) image subtraction; (e) keogram construction

6. L76: the Milky Way

7. L80: A word may be missing here, as the adjective "Observational" seems to lack a corresponding noun and does not fit into the rest of the sentence.

8. L81: The Milky Way

9. L83: What does "disposing" mean here? Consider rephrasing.

10. L123: The exact meaning of the phrase "lines were unwrapped" was not clear to me.

11. The term "angular coefficient" is not very standard, I had to guess what was meant by that.

12. Section 2.3 title: I would suggest "Amplitude estimations"

13. L157: How are these standard deviations calculated?

14. Section 3 title: I would suggest to write "Synthetic image simulations" or "Simulations with

synthetic images".

15. L181: "The red line passes through the 0 point to the 15% in variance equals one, [...]". I cannot understand what the authors mean here.