

Review of:
**Winds and temperatures of the Arctic middle
atmosphere during January measured by Doppler lidar**
by Hildebrand et al.

The authors present middle atmospheric wind and temperature observations of a lidar system in northern Norway during three Januaries. These observations are compared to the ECMWF and the HWM07 model. Besides the thermal and dynamical mean state, the authors also examine the variability caused by gravity waves and large-scale waves in the observations and the model data.

In a previous review I wrote to the authors “While the collocated middle atmospheric wind and temperature measurements of the Alomar RMR lidar are unique and unprecedented in their temporal and vertical resolution, I find it hard to learn something new from the paper. As it stands right now, the paper is mainly a comparison of different profiles, but no substantial conclusions are drawn from this.” This is still the case. Thus, I can only recommend publication of the article after substantial revisions.

Please find my detailed comments below.

Major comments:

1) As said before, the paper currently lacks scientific significance. This becomes especially clear when reading the introduction: 50% of the introduction are a mere review of different techniques to observe wind speeds in the middle atmosphere. The only hint for the importance of wind observations is given in the beginning when the authors state that “together with temperature observations, they [wind observations] also offer more sophisticated studies of gravity waves”. Why is this not done in this paper? Showing different profiles of potential and kinetic energy densities does not qualify the paper as a “sophisticated study”. To put it short: the paper lacks a scientific question which is investigated and answered in the end. Without a clear scientific question the paper remains unacceptable. A mere publication of the wind and temperature observations is unjustified in my eyes, despite the fact that it is the currently most extensive data set.

2) Most of the very few conclusions drawn by the authors remain rather simple statements which purely describe the observations but the effects which lead to the observations remain in the dark. A few examples:

P. 4, ll. 26–29: the conclusion that the northern hemispheric polar middle atmosphere is highly variable can certainly be considered as textbook knowledge and is therefore redundant.

P. 5, ll. 21–29: the minor SSW and the following elevated stratopause event in 2012 have been well documented by previous studies. Also, as stated correctly by the authors, the mechanism for the formation of an elevated stratopause is known. Hence, I do not see the additional insights which are gained in this study from the combination of wind and temperature observations.

P. 8, l. 33 – p. 9, l. 2: The authors merely speculate on the effects which could cause the different gravity wave propagation conditions. Here, a thorough analysis is needed which investigates the propagation conditions in great detail.

P. 10, l. 9: Why is the E_{kin}/E_{pot} ratio larger for the ECMWF data compared to the lidar data? What does this imply?

3) P. 8, ll. 25–26: the “approach using energy ratios has the advantage that an (energy weighted) intrinsic period for the ensemble of waves is calculated”. This statement is wrong! *Geller and Gong* (2010) derive their formula from the polarization relations which are fulfilled

only for one set of wave parameters $(k, l, m, \hat{\omega})$. If a superposition of waves is to be examined you have to take the sum over the squared wave perturbations in their equations 7) and 8). If you do so and insert the summed polarization relations, you will not end up with a formula, which you can solve for the average frequency. In fact *Geller and Gong* (2010) note in their appendix A1, that their approach always results in larger values of $\hat{\omega}$ than the mean value derived by the hodograph analysis.

Furthermore, it should be noted that according to *Lane et al.* (2003) one can only see long-period inertial gravity waves in the horizontal wind speed fluctuations. Short period gravity waves exhibit more pronounced vertical wind perturbations. Thus the here applied methodology is already biased towards the large period gravity waves.

If the authors want to infer gravity wave periods from their observations they have to use the hodograph approach instead of the energy approach. The energy approach can certainly be taken in the case of a quasi-monochromatic gravity wave field as shown by *Baumgarten et al.* (2015) but for an ensemble of waves it is not applicable.

4) I still think that the comparison of the lidar measurements to the HWM07 model is not appropriate. HWM07 is a climatology and thus one cannot derive a meaningful mean profile from three years of observations in a highly variable surrounding (northern hemispheric polar middle atmosphere) which can be compared to this climatology. As a result the authors cannot differ whether the HWM07 takes too little observations into account (cf. p. 6, ll. 12–13) or whether their observations are simply too few for the comparison. Thus, I recommend removing the paragraph on the HWM07 comparison (p. 6, ll. 6–13) and instead focus the paper more on other aspects.

5) It seems to me that the ECMWF model does not contain any gravity waves above 40–50 km altitude. Here a detailed investigation of the reasons for this behavior is needed. At the moment I do not see any physical reason why the gravity waves should not propagate to higher altitudes than 40–50 km.

6) Regarding the methodology of extracting gravity waves from their observations: The authors state that they do not see any significant differences between their methodology and the Butterworth filter suggested by *Ehard et al.* (2015). If this is not the case, I wonder why the authors do not adopt the Butterworth filter? One of the reasons for using the Butterworth filter is that it ensures a comparability of different studies since the same part of the gravity wave spectrum is extracted from the observations. In fact, *Baumgarten et al.* (2017) recently showed that by applying different methods of gravity wave extraction, a different seasonal cycle of gravity wave activity can be derived.

In a response to my previous review, the authors state that a further reason for not adopting the Butterworth filter is that “When applied to ECMWF data, the Butterworth and the spline method yielded physically dubious results (see Fig. 2): E.g., altitude profiles of GWED derived with the Butterworth method always showed similar oscillating behaviour above ≈ 65 km altitude; the ratio $E_{kin}=E_{pot}$ showed values < 1 for the spline and the Butterworth method, which can’t be true for gravity waves.” This argument can be dismissed in line of my major comment 5), since if there are no gravity waves in the ECMWF model above 40–50 km altitude, the results obtained by all methods are unphysical.

Furthermore, the 10h averaging applied by the authors has a significant disadvantage when it comes to analyzing the ECMWF data. I guess (see minor comments) that the authors use data from a different ECMWF run after 00 UTC. The corresponding switch from one ECMWF run to another is very likely to introduce a sudden jump of the temperature profile, which will be detected by the authors method, but not by a vertical Butterworth filter. For example the larger E_{kin}/E_{pot} ratios by the ECMWF compared to the lidar observations (p. 10, l. 9) could

easily be an effect of the different ECMWF runs and analysis used here. In fact I think what you see in the large scale wave energy density is mostly affected by the data assimilation of the ECMWF and not the model dynamics. This has to be investigated with great care!

Minor comments:

1) In line with major comment 6): I do not know at which times the authors use analysis data and at which times they use forecast data. For example, ECMWF analysis data is available at 00, 06, 12 and 18 UTC, but one can also retrieve forecast data for these times. Also the authors do not state from which runs the data are taken (i.e. runs initialized at 00 or 12 UTC, or a combination of both). This has to be clarified.

Furthermore, I was wondering, whether you extract the lidar data really at the named position, or whether you interpolate it horizontally to your lidar position?

2) Regarding the measurement uncertainties: At which altitudes do the maximum uncertainties usually appear? How do you treat measurement profiles for which the uncertainties appear at lower altitudes, e.g. 60 km? Do you have further constraints to insure the quality of your observations?

3) P. 5, ll. 12.–13: You state the “also” (why also? what else varies?) small vertical variability of the wind profiles and in the next sentence you state “very pronounced gravity wave structures”. Aren’t both statements contradictory?

4) P. 5, l. 35: “comparison of lidar data with ECMWF (...) for the whole data set”: since you compare two different ECMWF cycles to your observations it is misleading to average both cycles like done in Fig. 4d). In fact it seems to me that by averaging both cycles the deviations between the ECMWF and the observations decrease.

Also on p. 6, l. 19, I am not astonished that the comparison is nonuniform throughout the years, since you compare different cycles to your observations. This has to be evaluated in more detail and with more care!

Also later in ll. 23–26, you should state the cycles used by the other studies.

5) P. 7, l. 4: what is the RMS, I guess the authors mean “root mean square” but of what? Please clarify and also explain the abbreviation. Maybe also give a short explanation as to why an increase of the RMS is “expected for the effect of gravity waves”.

6) Figure 4b) is unnecessary and should be removed. The information on the deviation of the different profiles from one another is already contained in the profiles and the according standard deviations (shaded area) in Figure 4a).

7) In my eyes also Figure 5 is unnecessary, since the information on gravity wave activity is already contained in Figure 6 and the paragraph (p. 6, l. 30 – p. 7, ll. 2) does not give substantial new information. Furthermore, the conclusions drawn in this paragraph again remain pure speculation.

8) A general comment regarding the Figures: most axis are rather small and difficult to read. E.g. values of the RMS profiles in Figure 5 cannot be inferred. Furthermore, all plots showing E_{pot} and E_{kin} on a log axis would definitely benefit from a larger aspect ratio so that concrete values can be inferred by the readers more easily. Furthermore, it should be avoided that plotted values are smaller than the axis values (1st panel, Fig. 3c; 3rd panel, Fig. 8a).

Technical corrections

P. 1, l. 4 and throughout the text: “month-mean” should read “monthly mean”, the same for “night-mean”.

P. 2, l. 8: “then” should read “than”

P. 2, l. 9: give the names for the models (ECMWF, HWM07) at the first appearance of the abbreviations in the text

P. 3, ll. 17–19: it might be of help for the reader to slightly change the order of the sentences: “To retrieve winds (...) The temperature retrieval relies (...) The two individually derived temperature profiles (...)” Also cite *Hauchecorne and Chanin* (1980) for the retrieval of your temperature profile.

P. 4, l. 11: the vertical resolution of the two ECMWF model cycles should be stated.

P. 4, l. 12: what is the vertical resolution of the lidar data? On p. 3, l. 27 you state that the lidar data is smoothed with a “window size of 3 km” is this the vertical resolution of the lidar data? Your profiles look way smoother than just one point every 3 km.

P. 4, l. 32: “or even split, *and* warmer air”

P. 5, l. 9: “Only *a* few days later”

P. 5, ll. 10 & 11: “some 20 K colder/warmer” – colloquial, state precise values

P. 5, ll. 11 & 12: “weak east/west/southward” should read “weakly east/west/southward”

P. 6, l. 16: “way too low” – colloquial, state precise values

P. 6, l. 20: “it is good below 60 km altitude”, please quantify. “Good” can mean anything.

P. 6, l. 26: “some deviations in the mesosphere”, please quantify.

References

- Baumgarten, G., J. Fiedler, J. Hildebrand, and F.-J. Lübken (2015), Inertia gravity wave in the stratosphere and mesosphere observed by Doppler wind and temperature lidar, *Geophys. Res. Lett.*, *42*, doi:10.1002/2015GL066991.
- Baumgarten, K., M. Gerding, and F.-J. Lübken (2017), Seasonal variation of gravity wave parameters using different filter methods with daylight lidar measurements at midlatitudes, *J. Geophys. Res.*, doi:10.1002/2016jd025916.
- Ehard, B., B. Kaifler, N. Kaifler, and M. Rapp (2015), Evaluation of methods for gravity wave extraction from middle-atmospheric lidar temperature measurements, *Atmos. Meas. Tech.*, *8*(11), 4645–4655, doi:10.5194/amt-8-4645-2015.
- Geller, M., and J. Gong (2010), Gravity wave kinetic, potential, and vertical fluctuation energies as indicators of different frequency gravity waves, *J. Geophys. Res.*, *115*, D11,111, doi:10.1029/2009JD012266.
- Hauchecorne, A., and M. Chanin (1980), Density and temperature profiles obtained by lidar between 35 and 70 km, *Geophys. Res. Lett.*, *7*, 565–568, doi:10.1029/GL007i008p00565.
- Lane, T. P., M. J. Reeder, and F. M. Guest (2003), Convectively generated gravity waves observed from radiosonde data taken during MCTEX, *Quart. J. Roy. Meteor. Soc.*, *129*(590), 1731–1740, doi:10.1256/qj.02.196.