

Response to Anonymous Referee 2

Referee comments are indicated in *italic*, followed by our reply.

General:

The authors present an algorithm to improve the accuracy of the instrumental noise of Halo Photonics Doppler lidars. Therefore, longer integration times become possible; enabling to obtain signals down to -32 dB. This is particularly useful under conditions with low aerosol load. The algorithm combines two background correction methods. The first method corrects for a variable offset in SNR at each range gate, visible as horizontal stripes in the time-height cross-sections of retrieved backscatter. The description of the source of this error in SNR as well as its correction is novel. The second method corrects for an error in scaling of the raw signal, which can result in vertical stripes in the time-height cross-sections of retrieved backscatter. The second correction method has already been published in AMT (Manninen et al., 2016). The authors show that combining both methods gives added value to the retrieval of atmospheric turbulence properties as well as aerosol observations. Additionally, the “novel post-processing” algorithm could be used in other Doppler lidars and other lidar systems. Overall, the paper is well written and the description of the algorithm is logical and well-structured. However, the authors should include a more detailed discussion of SNR statistics. Therefore, the given paper presents a valuable contribution to the field of lidar-based remote sensing. I recommend it for publication in AMT after considering the following comments.

We would like to thank the referee for the valuable comments, which have improved the manuscript.

Major comments:

1) Does applying the horizontal stripe correction prior to the Manninen correction influence the performance of the cloud screening, which included in the latter?

Applying the horizontal stripe correction makes it easier to perform cloud screening, as it is easier to discern thin clouds and aerosol layers. We have added a sentence on page 6, line 3:

“Note that typically cloud and aerosol signal is easier to discern in SNR_1 than in SNR_0 and thus cloud screening is applied after Equation 5.”

2) In Figures 6, 7 and 8 the reader can visually see the improvements in SNR looking at SNR_0 to SNR_2 , but it is essential to quantify the improvement, e.g. with a figure along the line of Figures 8 or 9 in Manninen et al. (2016). Such a statistical analysis of the SNR would present a more detailed picture of the algorithms performance. It would also help to answer comment 1).

In our opinion the major advantage of this algorithm is that it enables averaging SNR to utilise very weak signals, which is illustrated in Fig. 4d. However, to give more detailed picture of the algorithm's performance we have added histograms of SNR_0 and SNR_2 in cloud and aerosol free regime for the four cases studies as an Appendix to the manuscript. The new Fig. A1 is below.

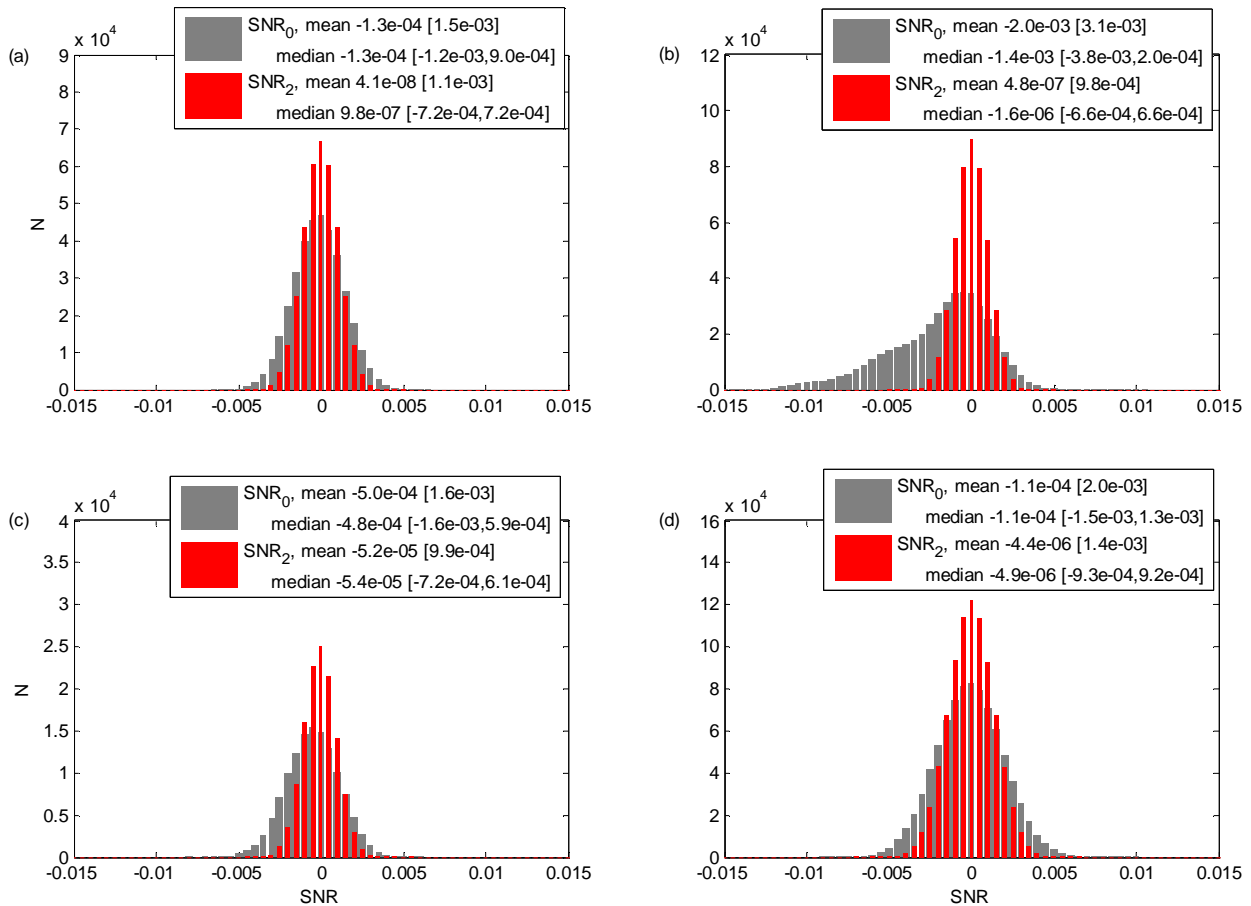


Figure A1. Histograms of SNR₀ and SNR₂ in cloud and aerosol free regime for the four case studies considered in Section 4. For each case, mean [standard deviation] and median [25th, 75th percentile] of SNR₀ and SNR₂ are included. (a) Welgegrund on 6 September 2016, 00-24 h UTC, 4800-9000 m a.g.l.. (b) Kumpula on 1 May 2018, 02-24 h UTC, 6000-12000 m a.g.l.. (c) Kumpula on 6 May 2018, 00-12 h UTC, 4000-7000 m a.g.l.. (d) Finokalia on 8 July 2014, 00-24 h UTC, 5000-96000 m a.g.l..

3) p.2 l. 18: “as uncertainty : : : wind retrievals: : :” + p.9 l.17-21: “With enhanced SNR, : : :”
 On the one hand, the introduction refers not only to turbulence but also to wind retrievals.
 On the other, the conclusion only states turbulence retrievals. Methods, such as velocity azimuth display utilize the SNR to determine reliable radial velocity measurements. Therefore, the correction algorithm increases the data availability and probably decreases the uncertainty in the retrieval.

The paper would gain more attention if this aspect is included in the discussion; either by actually showing that the novel post-processing improves the retrieved winds or by including a paragraph on the expected benefit for wind retrievals.

Thank you for bringing this up. Wind retrieval methods should be able to discard outliers (e.g. Päschke et al., 2015), so we do not expect that the retrieved wind speed and direction would be affected. However, as a lower SNR-threshold can be applied after the new post-processing, the data availability will increase. How large this effect is depends on atmospheric conditions. We have added following discussion in Section 4.1, where also vertical winds are discussed:

“Lower noise floor enables also wind retrievals with a lower SNR-threshold, which increases the data availability. The effect on data availability depends on atmospheric conditions, though. In this

case for instance (Welgeund, 6 September 2016) a 75° elevation angle velocity azimuth display (VAD) scan was utilised for horizontal wind retrieval every 15 minutes. With the new post-processing algorithm the SNR-threshold for wind retrieval could be decreased from 0.0045 to 0.0032. This decrease in SNR-threshold enabled wind retrieval for 2-13 range gates more from each VAD scan; on average, winds could be determined from 7.5 additional range gates per VAD scan. That is, vertical coverage of wind retrievals increased on average by 200 m with the new post-processing.

Wind retrievals at lower SNR will have higher uncertainty due to higher instrumental noise in radial velocity measurement, yet enhanced SNR will enable more accurate determination of the instrumental uncertainty in wind retrievals. However, as a major fraction of the uncertainty in retrieved winds arises in atmospheric turbulence (Newsom et al., 2017) the more accurate SNR will have only a limited effect on the overall uncertainty in the wind retrieval. Therefore, the uncertainty in each wind retrieval should be evaluated e.g. with the methodology of Newsom et al. (2017) before the wind retrievals are disseminated.“

We have added a note on improved data coverage for horizontal wind also in the Conclusion.

4) Manninen et al. (2016) give the code as well as an example as supporting information. This paper's algorithm could spread a lot faster, if the authors also include the code in the supporting information or point to a repository, e.g. on github.

A Matlab implementation of the algorithm presented here is available at https://github.com/manninenaj/HALO_lidar_toolbox. We have added a note on this in Conclusion and a reference to the toolbox in the references.

Minor comments:

1) Figure 1b shows the temperature dependency of the background profiles. Between 20_C and ca. 27_C the increase seems linear with only a small spread. Above 27_C the increase appears linear with a steeper slope and wider spread of values. Is it possible that part of this change (especially the wider spread) is due to the switch on of the active cooling unit of the Halo Streamline? The cooling would insert vibration and additional electronic noise into the lidar system which could alter the background measurement.

It is quite possible that this change coincides with the switching on of the fans. However, we did not find temperature dependency in the actual measurement noise (measured as the variance of SNR for cloud and aerosol free range gates). The only temperature effect we found is in P_{amp} (Fig. 2b), which becomes relevant only when integration time is several minutes or longer.

2) Figure 2 and the corresponding discussion on page 5 show that even if the technical specifications of two Doppler lidars are the same (see Table 1), their performance can be different due to instrumental characteristics. This presents an additional finding of this paper which is especially important when operating multiple Doppler lidars. The users should be aware of the possible differences in instruments of the same type. Hence, this finding deserves more attention and should be mentioned in the conclusion.

Indeed, each lidar system needs to be individually characterised by the lidar operator. We have added following paragraph in conclusion:

“Our analysis shows that even if the technical specifications of two Doppler lidar systems are identical, their instrumental noise characteristics can be quite different (Fig. 2). Therefore, the lidar operator should inspect each system individually to ensure highest data quality. Note that this algorithm or similar processing is needed to define the instrumental noise level even if raw spectra are utilised instead of the processed in data. The algorithm presented here can be applied in semi-operational use as long as at least 300 background checks (acquired in two weeks of measurements with typical configuration) are available for characterising the amplifier response to the transmitted pulse. A Matlab implementation of this algorithm is available through Github (Manninen, 2019).”

Reference:

Manninen, A: HALO lidar toolbox, GitHub, https://github.com/manninenaj/HALO_lidar_toolbox, 2019.

3) Concerning the high and low mode in the XR lidar: Do the authors know by now, why there are two modes? Halo lidars can be equipped with a depolarization channel. Is instrument number 146 such an instrument and can the authors exclude that the seemingly random shifts in modes are due to switching between co- and cross-polarization?

All instruments utilised in this study are equipped with a depolarization channel. However, the software is configured to do background check always in the co-polar mode. Our understanding is that this is a feature of the amplifier. However, whether these two modes are present in a system is easy to check by the lidar operator from the background files. Note also that some systems are configured with 10kHz pulse repetition frequency, but use the non-XR amplifier.

4) Could methods similar to the horizontal stripe correction be applied to other lidar systems such as ceilometers?

This kind of method can be applied to any lidar system that defines noise level from a finite duration background check (c.f. page 9, line 8-10). However, ceilometers do not routinely carry out such background measurement – in ceilometers the most important source of noise is solar background, which requires its own processing.

5) Some readers could be interested in using Doppler lidars semi-operationally, e.g. for a seasonal campaign or even 24/7. Which parts of the algorithm are applicable operationally? How long does the data set have to be in order to successfully apply the post-processing? What about longer data gaps? The publication would benefit from a remark to operational applicability.

The algorithm is not computationally heavy and can be utilised semi-operationally with time delay of a few minutes from measurement to post-processed data. As pointed out on page 5, line 19-20, approximately 300 background checks are needed to obtain a reliable estimate of P_{amp} . This corresponds to two weeks of measurements with hourly background checks. Such short record does not facilitate accounting for temperature dependency of P_{amp} , but that is a minor effect compared to other sources of uncertainty in the noise level.

As long as a continuous record of background checks are available, data gaps are not an issue for this post-processing algorithm.

6) *The introduction starts by mentioning “turbulent mixing” and “mixing layer height”, but the paper presents no estimate of the MLH in one of the Figures 4 through 8. Including the MLH as a line plot in Figure 4 or elsewhere would round up the discussion that started in the introduction.*

We have added MLH as a line in Fig. 5.

7) *p.5 l. 11: “t” elescope.*

Corrected on page 3, line 11.

8) *Table 1 seems to long and could be shortened by mentioning redundant information only once*

Table 1 has been condensed:

Table 1 Specifications for Halo Doppler lidars utilised in this study.

Lidar number and version	46, Stream Line 53, Stream Line Pro 146, Stream Line XR
Wavelength	1.5 μm
Pulse repetition rate	15 kHz (46 and 53) or 10 kHz (146)
Nyquist velocity	20 m s^{-1}
Sampling frequency	50 MHz
Velocity resolution	0.038 m s^{-1}
Points per range gate	10
Range resolution	30 m
Maximum range	9600 m (46 and 53) or 12000 m (146)
Pulse duration	0.2 μs
Lens diameter	8 cm
Lens divergence	33 μrad
Telescope	monostatic optic-fibre coupled
