IEA Wind Task 32 webinar summary

Approaches in filtering data from pulsed wind lidar

Authors: Leonardo Alcayaga, Rogier Floors, Pedro Santos. Editor: Andrew Clifton

Can more of a pulsed wind lidar's output data be used to get more accurate results and increased coverage?

This is a summary of the "Filtering lidar data" webinar given by Rogier Floors and Leonardo Alcayaga on 7 April 2020 in the IEA Wind Task 32 webinar series. It represents the authors' opinions.

Wind lidar and other wind measurement devices often report a range of data metrics in the output files. It is sometimes tempting to use these data as filters.

One metric often used for filtering is the Carrier-to-noise ratio (CNR). The CNR of a pulsed lidar indicates the relative strength of the signal versus background noise. It depends, among other things, on the backscatter properties of the atmosphere, the range, and the strength of the laser used in the lidar device [1]. Therefore, care must be taken to select appropriate CNR levels and not systematically exclude data.

This paper has two main threads. Firstly, the authors outline the effect of data selection on the wind climatology using ground-based profiling pulsed lidars, and how this might be mitigated. And, they present a way to achieve better spatial coverage by applying a filtering algorithm that uses several metrics.

How to determine wind climatology from profiling lidar

Wind climatologies are sensitive to data availability. Therefore the choice of CNR threshold for 10-minute mean wind speeds can impact the wind climatology.

Data collected at the FINO3 platform in the North Sea was recently used to quantify this effect [2]. When the Carrier-to-Noise-Ratio (CNR) threshold value is increased, the wind speed distribution is shifted to higher values (Fig. 1).

The factory setting of the CNR threshold for this long-range lidar was -35 dB. The data availability from this CNR threshold gives the lowest annual mean wind speed estimate using the range gate at 126 m. The mean wind speed increases to ≈13 m s^{-1} when the filtering threshold is set to -17 dB. The mean wind speed from the cup anemometer measured at 90 m shows a very similar increase, showing that the bias is not induced by the lidar, but is a consequence of **data availability**. Finally, filtering by requiring that data exceed a CNR threshold at all range gates from 126 to 626 m results in an even stronger effect, because data that fulfill this criterion at all heights are rare.

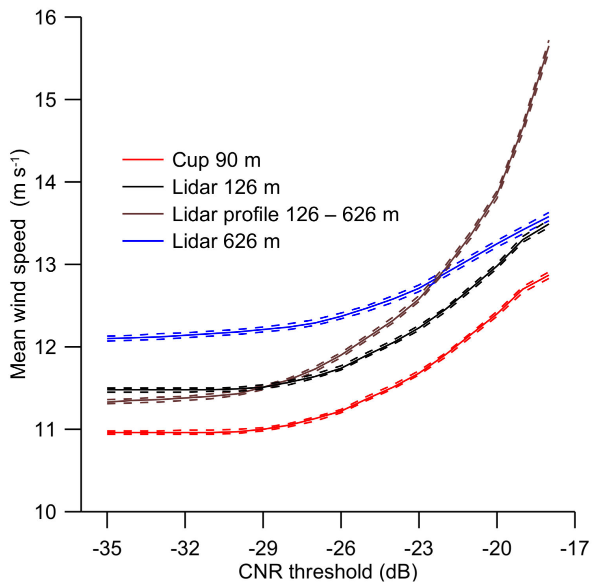


Figure 1: The effect of lidar CNR limits on the observed wind climatology at the FINO 3 platform. Data are for one year. The standard error is shown by dashed lines.

This effect of CNR threshold on data was also observed for both short- and long-range vertical profilers and for scanning lidars (see Fig. 10 in [2] or the webinar presentation [3] for more examples). The effects described here are hard to detect from e.g. a scatter plot comparison between a lidar and a met mast mounted cup anemometer: there may be a perfect correlation between the two, but the climatological wind speed at the same location can still be different when the measurements do not cover 100% of the period. This is because lower CNR values apparently correlate with lower wind speeds. The physical mechanism for this correlation was not investigated, but it was present in data collected in Germany, Denmark, and Greenland.

For more information see the webinar presentation [3] or the publication [2].

DOI 10.5281/zenodo.3908475
Version 26 June 2020: ADD DATE

DOI 10.5281/zenodo.3908475
Version 26 June 2020: ADD DATE

Increasing spatial coverage

Long-range pulsed lidars tend to show lower CNR values at longer ranges. So, using CNR as a filter would reduce availability with increasing range. This can reduce the spatial coverage of scanning lidars.

But, this filtering may be unreasonably reducing data availability. For example, the distribution of the line-of-sight wind speed (V_{LOS}) in Figure 2 shows apparently reasonable values, even for low CNR.

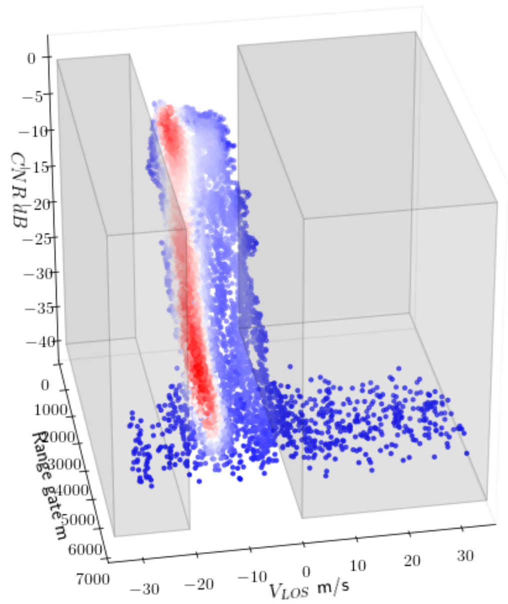


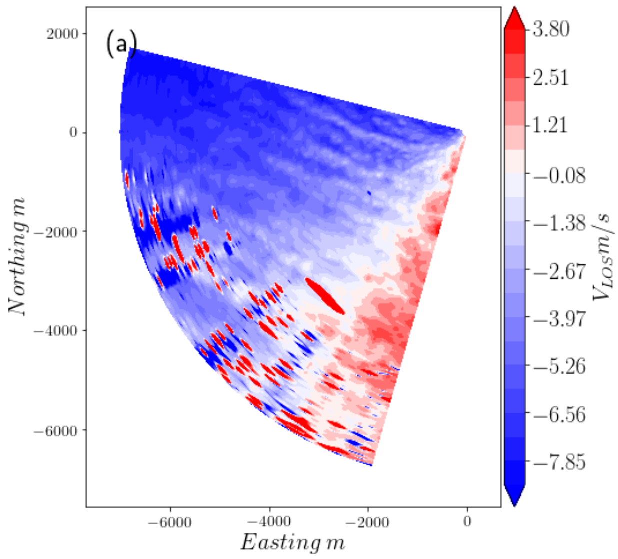
Figure 2: Sometimes low-CNR data at longer ranges still have reasonable V_{LOS} values. These data can be lost if a CNR threshold is used. (Colours indicates data density; red areas have more data than blue)

What other quality indicators could be used?

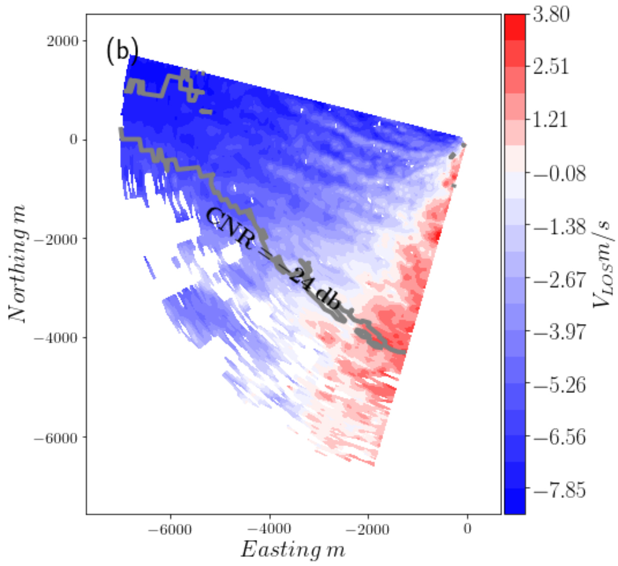
One useful indicator is the **smoothness of the line-of-sight wind speed**; a good V_{LOS} field should be continuous in the spatial domain. Therefore, a median filter can be used to detect "bad" data [4]. This is fast and efficient when the proper moving median window and filtering threshold are selected.

Another indicator is **data self-similarity**. Lidar data can be thought of as multi-dimensional data sets – described by e.g., spatial location, CNR, and wind speed – where reliable measurements cluster together [5]. These clusters become clearer as more dimensions are included, and can be detected using a clustering algorithm like DBSCAN [6]. DBSCAN detects coherent groups of data of arbitrary shape, separating them from the sparse observations associated with "noisy data". Selecting relevant data features in DBSCAN increases the data availability [7] and the spatial coverage (Fig. 3). The DBSCAN classification parameters adjust automatically and so there is no need to define a threshold for any feature.

For more information see the webinar presentation [8], or the related publication [7].



(a) The unfiltered scan data includes noisy data at farther ranges. The red and blue blobs are extreme values and indicate regions of noisy data.



(b) The choice of filter can impact coverage. The background data has been filtered using DBSCAN. The grey contour shows the furthest range achieved using a CNR threshold of -24db.

Figure 3: Filtering strategies affect spatial coverage.

DOI 10.5281/zenodo.3908475
Version 26 June 2020: ADD DATE

Summary and Implementation

To minimize the impacts of the effect of threshold filtering on the mean wind speed, one needs to achieve a data availability that is as high as possible. How this can be done depends on your goals:

- **when interested in winds at hub height** don't include measurement from heights above that, as the CNR tends to decrease with range and also the number of aerosols tends to decrease with height, thereby decreasing the data recovery rate.
- **when interested in wind shear** make sure to have a high ($>$95%) and similar availability at all heights, because the filtered mean wind speed depends on data availability which generally decreases with height (Fig. 1).
- **When availability is decreased by filtering** with a higher CNR threshold, compare the mean wind speed at the same range gate before and after the filtering to make sure they are similar.

To maximize the spatial coverage and data quality of retrievals from scanning lidars, one needs to go beyond CNR as the only quality parameter. Instead, V_{LOS} and spatial information, and filtering techniques of increasing complexity can be used. When using these techniques, consider the following:

- **First use a simpler and faster approach** based on smoothness, such as a median-like filter on V_{LOS}.
- **Check the statistics of the recovered data**. This could include comparing the distribution of V_{LOS} against high CNR data (particularly the tails).
- The moving-median window size and threshold values used in median **filters must be tuned**.
- **If the results are not satisfactory** (e.g., tails in the distribution of V_{LOS} are still heavier compared to high CNR data, there is evidence of noise, or changes in atmospheric conditions requires constant tuning of filter parameters), then a classification/clustering method based on data density may help.

References

The presentations accompanying this paper are available through Zenodo [3, 8].

[1] M. L. Aitken et al. 'Performance of a Wind-Profiling Lidar in the Region of Wind Turbine Rotor Disks'. In: *Journal of Atmospheric and Oceanic Technology* 29.3 (2012). DOI: 10.1175/JTECH-D-11-00033.1.
[2] S.-E. Gryning and R. Floors. 'Carrier-to-Noise-Threshold Filtering on Off-Shore Wind Lidar Measurements'. In: *Sensors* 19.3 (2019). DOI: 10.3390/s19030592.
[3] R. Floors and S.-E. Gryning. *Carrier-to-Noise-Threshold Filtering on Offshore Wind Lidar Measurements*. Apr. 2020. DOI: 10.5281/zenodo.3743076.
[4] R. Menke et al. 'Multi-lidar wind resource mapping in complex terrain'. In: *Wind Energy Science Discussions* 2019 (2019). DOI: 10.5194/wes-2019-85.
[5] H. Beck and M. Kühn. 'Dynamic Data Filtering of Long-Range Doppler LiDAR Wind Speed Measurement.' In: *Remote Sensing* 9(6) (2017). DOI: 10.3390/rs9060561.
[6] M. Ester et al. 'A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise'. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. KDD'96. Portland, Oregon, 1996.
[7] L. Alcayaga. 'Filtering of pulsed lidars data using spatial information and a clustering algorithm'. In: *Atmospheric Measurement Techniques Discussions* 2020 (2020). DOI: 10.5194/amt-2019-450.
[8] L. Alcayaga. *Filtering of pulsed lidar's data using spatial information and a clustering algorithm*. Apr. 2020. DOI: 10.5281/zenodo.3742961.

This webinar summary was published by IEA Wind Task 32 to accompany a webinar that took place on 7 April 2020. The information presented here is the opinion of the authors.

The International Energy Agency is an autonomous organisation which works to ensure reliable, affordable and clean energy for its 30 member countries and beyond. The IEA Wind Technology Collaboration Programme supports the work of 38 independent, international groups of experts that enable governments and industries from around the world to lead programmes and projects on a wide range of energy technologies and related issues.

IEA Wind Task 32 exists to identify and mitigate the barriers to the deployment of wind lidar for wind energy applications.

Authors (alphabetical by family name): Leonardo Alcayaga (DTU Wind Energy, Denmark), Rogier Floors (DTU Wind Energy, Denmark), Pedro Santos (DTU Wind Energy, Denmark). **Editor:** Andrew Clifton (U. Stuttgart, Germany).

Images: Banner, left to right: Alexandre Debiève on Unsplash, SWE U. Stuttgart, Markus Spiske on Unsplash. Figure 1 by R. Floors, Figure 2, 3 by L. Alcayaga.