

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2020-516-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2020-516

Anonymous Referee #1

Referee comment on "Wind measurement comparison of Doppler lidar with wind cup and L band sounding radar" by Zizhong Zhou and Zhichao Bu, Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2020-516-RC1>, 2021

General:

- Manuscript Summary

The manuscript "Wind measurement comparison of Doppler lidar (DL) with wind cup and L band sounding radar" by Zhou and Bu verifies the performance of various Doppler wind lidars manufactured in China and one Vaisala Windcube for comparison. The measurements are carried out at two different facilities, namely at the Shenzhen National Climate Observatory from November 2019 to January 2020 and at the Zhangjiakou Meteorological Bureau from 23 December 2019 to 6 January 2020. At Shenzhen, twelve lidars gathered data in two batches (first batch from 23 November 2019 until 8 December 2019, second batch 1 January 2020 until 24 January 2020) and the authors use wind cup/vane measurements from the 356 m Zhangjiakou Meteorological Tower (ZMT) as a reference. At Zhangjiakou, the authors assess the lidars' performance with the help of routine radiosonde launches, whose wind information is measurement with an L-band radar.

As verification scores, the authors use the system deviation, root mean square deviation and the correlation coefficient for the retrieved quantities wind speed and direction. The manuscript concludes that most of the tested DLs are within the WMO OSCAR requirements for near-term forecast and in addition that the L-band radar is not reliable as a reference.

- Comment

The present study follows the general outline of a journal article, but the description of individual sections lacks vital information and clarity. Further, the language makes it sometimes difficult to comprehend.

Starting with the introduction, the authors fail to give a convincing motivation of why such a comparison should be published as a journal article and not just as technical report. The literature review in this study confuses the reader, e.g. operational radar wind profilers can give extremely reliable data in the range from 500 m to 15000 m, but they take up a lot of space and are very costly compared with DLs. Some of the confusion is possibly the result of plainly copying passages from the cited works (see specific comments). This also leads to a mix-up of conclusions drawn from ground-based DLs used in Qin et al. (2019) with those based on nacelle-based DLs used on wind power plants (Schlipf et al. 2020).

In the methods section, the manuscript withholds essential information about the measurement location and the experimental setup. The instrument specifications are missing in this section and since the performance of lidar systems is dependent on the laser and beam specifications: pulsed or cw system, power per pulse, pulse duration, aperture dimension, pulse repetition frequency etc. These parameters should be stated somewhere if available. In Table 3 some of the specifications are given, but this table belongs to the results section. Here, a description of the ZMT is also necessary. The one given in He et al. (2020) shows that the ZMT has 13 wind cup/vane levels and 4 additional sonic anemometers. Why does this study not use these other instruments? In addition, the authors miss to describe the scanning mode or modes used. I would assume the Vaisala windcube used Doppler beam swinging (DBS), but this is only a speculation. A description of the wind precision of the L-band radar is also missing. Including all this missing information would help the manuscript's quality and aid the discussion of the results. As a last point, the study lacks a meteorological overview of the measurement period, since lidar performance also depends on atmospheric conditions.

The first part of results section relies solely on scatter plots of wind speed and direction and their summary statistics. Even though all tested lidars appear in the summary statistics, only four scatter plots of wind speed from different lidar systems are shown and each of which corresponding to a different tower height while at the same time the direction scatter plots are all from the same system. I think time-height cross-sections of all instruments and profiles of their mean deviation statistic would enable a clearer, height-dependent comparison. Further, it is not clear to the reader why only four measurement heights are used while the tower itself has 13 available (He et al. 2020) and the reader will surely want an uncertainty estimates of the lidars which can be obtained from the signal-to-noise ratio (SNR) as described in Pearson et al. (2009). Another important quantity is the data availability. The authors state that the lidars ran 24/7, but there is no word about data quality or data availability, preferably as a function of height.

The comparison with the L-band radiosondes in the second part of the results section is not sound for two reasons. Firstly, there is no discussion on why other studies manage to compare radiosondes and DLs successfully and what are the challenges (see Mariani et al. (2020) and Kumer et al. (2014)). Secondly, the DL measurements are afterwards used as a consistency check for the L-band winds. This is the opposite of verification and not in line with previous works. The L-band radar has in the relevant altitude range an expected wind speed accuracy of 1 m/s or less and a directional accuracy of 5° or better (Wang et al. (2018)) which is comparable to the standard GPS radiosondes. Hence, large discrepancies in some of the lidar systems has likely other reasons and the conclusion that the L-band radar should not be used as a standard is not justified by the presented results. More so, these large deviations should raise attention to potential mistakes in the analysis or in the experimental setup.

All in all, the manuscript at hand has major drawbacks and should not be considered for publication in AMT. Additionally, I agree with the authors that testing the accuracy of DLs in detail is significant for improving their quality and wide spread usage. But here, important aspect like the data availability and measurement uncertainty are neglected by the authors. Therefore, this study is not very helpful and potential readers will not be interested in it. Further, adding up all major concerns, the manuscript does not fulfill the journal the standards and does not present sound results. Hence, I advise to reject the given manuscript, but I encourage the authors to address my concerns and redo the analysis, because a study comparing twelve different DLs is of interest for scientist as well as people working the renewable energy sector.

Specific comments:

- Major comments:
 - p2 – l.2-12: The background motivation and literature review is not sufficient. Doppler wind profiler or radar wind profilers can measure between 500 m up to around 14000 m. They are a very reliable reference (see AEOLUS satellite mission), but they take up a lot of space and are very costly. This is where DLs shine. They are comparatively cheap and mainly focused on the ABL, i.e. 50 m up to around 3000 m.
 - p2 – l.18-23: This is not just a citation paraphrasing Päscke et al. (2015) but a literal quote from the paper's abstract.
 - p2 – l.23-27: Also a literal quote from Quin et al. (2019).
 - p2 – l.27-31: Again, an almost literal quote of Schlipf et al (2020), but more severely this paper is concerned with nacelle-based lidars used on wind turbines. Therefore, it is not a suitable comparison for ground-based lidars.
 - p2 – l.29: "measurement errors" and uncertainty are not addressed in later sections of the manuscript, but they are extremely important for research and more so for weather forecasting.
 - p2-3 – l.34-4: Another literal quote from a cited reference, here Lundquist et al. (2015).
 - With all those literal quotes the manuscript appears hastily put together. Even though, the original studies are cited accordingly it is at the border to plagiarism.
 - p3 – l.17-19: What is "four beam scanning synthesis"? Is this a DBS scan or a type of 4 beam PPI (point position indicator) scan. The manuscript must give a clear description of the scanning mode or modes each instrument uses, i.e. number of beams, zenith and azimuth angle etc. Otherwise the presented results are not reproducible.
 - p3 – l.28-30: How do the authors assure the measurements have not been influenced by turbulence or else? Was there some kind of filter method applied? Please explain.
 - p4 – l.1: This section leaves out important aspect: instrument specification, i.e. wavelength, pulse duration, pulse repetition frequency, aperture dimension, detector bandwidth etc. What scanning mode/modes have been used? Also an overview of the meteorological conditions at both sides during the experiments is missing. This would greatly help the interpretation of the results.
 - p4 – l.26: What Vaisala Windcube was used? From the Tables 1 and 2 I assume it was some type of Windcube V2. This information is crucial for comparisons with other studies.
 - p4 – l.27: How are instruments divided up into batches?
 - p5-15 – Results section: These are lidar systems. The uncertainty of Doppler lidars can be estimated via the Cramer-Rao lower bound. The authors should show a comparison of the lidars' SNR side by side and relate the estimated RMSE also to the calculated uncertainties. Further, what data availability and quality?
 - Additionally, 290 m max range on the Windcube v2 should be possible to extrapolate to 300 m with the right methods!
 - p5 – Figs. 1, 2, 3 and 4: The scatter plots show clearly an overestimation for high winds and a small underestimation for low wind speeds. This is not discussed in the results section. Could it be an effect of the wind cup that has a threshold for spinning to start and can suffer from overspinning. Please note that such issues need to be discussed in a scientifically sound study. Further, the scatter plots are not able to capture the atmospheric variability and its influence on the instruments. Here, a meteorological overview comes in handy. Additionally, why are the other SMT levels not used in the comparison? (e.g. in a time-height cross-section).
 - p7 – Table 1: Why is there no uncertainty estimate. This could be related to the standard deviation.
 - p8 – l.5-8: The Vaisala Windcube has clearly a system deviation which is an order of magnitude better than the other systems. Therefore, this conclusion is wrong.

- p8 – l.7: Please state the WMO OSCAR requirements explicitly.
- p11 – l.2-8: The Windcube has the smallest bias, except lidar #5. But #5's rmse is too large. The windcube shows the most consistent directions, besides an obvious bias. But the bias is still the smallest. The reader would be interested in a discussion of these differences.
Also the authors should seriously look into the possibility that the lidar system were misaligned towards true north. If this issue cannot be worked out, the directional deviation should not be used for comparison. Plus, why was there no comparison with the sonic anemometers?
- p12 – Fig.11: These figures need more explaining: Why is the RS data continuous. RS are only launched twice a day. This looks like all RS winds time stamps below a certain altitude have been matched with their corresponding DL data; and then this is stitched into a time series. This display is very unusual. A time-height cross-section would be the better choice, paired with profiles depicting the MAE, RMS and Median and Mean.
Because the interesting question is how large are the deviations in what altitude ranges. The reader learns little except a vague average over all altitudes. This is not the scope of WMO OSCAR requirements.
It also seems, that there are two types of deviation. One type with smaller speed and direction deviation occurring together and another type with large speed deviation but hardly any directional offsets. Here, a look at the SNR data and the balloon drift could help. Additionally, an investigation of precipitation and a cloud screening would be beneficial.
- p14 – l.4-7: Other studies show that the L-band radar has nominal accuracy comparable to GPS radiosondes like the Vaisala RS41 (Wang et al. (2018)). Hence, the directional RMSE should be in the range of the comparison presented in Päschrke et al. 2015, i.e. less than 20° and the speed RMSE should not be much greater than 2 m/s for an instrument to be called reliable.
- p14 – l.9-10: This is not how verification works. There are many factors at play when comparing balloon data with stationary remote sensing devices, such as balloon drift and synoptic weather conditions.
- p15 – Fig.13: The two spikes in the plot look suspicious. This analysis must include the synoptic conditions.
- p15 – l.3-6: One cannot use the instrument to be verified (DL) as a reference for verification. This simply does not work and is not in line with previous works (Kumer et al. (2014), Päschrke et al. (2015) and Mariani et al. (2020)).
- p15 – l.10-11: In the present form, this work is not significant, because it lacks key aspects touched in most verification studies of DLs.
- p15 – l.18: The numbers presented in Tables 1 and 2 show that the Windcube has better accuracy and precision. This conclusion is incorrect.
- p16 – l.9-11: Unfortunately, I was not able to download the data for verification, because one needs a Baidu account to access the download and for Baidu account one needs a Chinese mobile phone number.
- Minor comments:
 - p2 – l.29: "longitudinal wind" is not logical in the context of the manuscript. This is a result of one to one copying the passage from Schlipf et al. (2020).
 - p2 – l.28 and 30: The authors name is David Schlipf, hence the citation must read Schlipf et al. (2020).
 - p3 – l.6-9: The time periods should be listed for each location individually. Otherwise the question arises, if the identical instruments were used in both comparison or just the same of instrument.
 - p3 – l.10: The wind cup alone cannot measure direction, here the wind vane is to be mentioned as well.
 - p3 – l.23: The equation is technically only valid for monostatic DLs. Please include it in the system description.
 - p4 – l.2-4: How much data were excluded?

- p4 – l.5-6: What do the authors mean by second interval data? The archiving method is not obvious to me.
- p4 – l.15-23: It is not necessary to explain to the reader how to mean, standard deviation and correlation are calculated. Please delete this part.
- p5 – l.1-5: The description of the analysis reads like the instruction in a manual or a lab class. This should be rephrased.
- p6 – Fig.1 and 2: There seem to be artefacts of interpolation in the scatter plots. Please give an explanation of the interpolation used.
- p6/7 – l.6-15/1-5: Performing a linear fit is widely known and needs no explanation. Please delete this part.
- p7 – Table 1: Why is the floating point precision of the numbers down to three leading digits? System 2 and 3 have the exact same numbers. Is this the 1 and 10 min standard deviation of the measured wind speeds or the root mean square deviation to the reference?
- p10 – l.1: It is not possible to measure direction with the wind cup, a wind vane is needed here. What are its specifications?
- p12 – l.5: One can clearly see outliers around 9 m/s. Please discuss.
- p14 – l.3-4: A relative comparison is not helpful. Most of the time a researcher has only a single DL on site. Please remove or rephrase this point.
- p14 – l.11: When does the “night” end?

Technical comments:

- p3 – l.20: English syntax is wrong.
- p4 – l.24: Format error, this should be a subsection
- p4 – l.25: “domestic” should be changed to Chinese, because this is an international journal.
- p8 – Fig.5: Figure title does not match the caption, title 100 m/caption 50 m. Directions should be between 0° and 360°.
- p9 – Fig.6: Figure title does not match the caption, title 250 m/caption 100 m. Directions should be between 0° and 360°.
- p9 – Fig.7: Figure title does not match the caption, title 100 m/caption 250 m. Directions should be between 0° and 360°.
- p9 – Fig.8: Figure title does not match the caption, title 250 m/caption 300 m. Directions should be between 0° and 360°.
- p12 – Fig.10: Directions should be between 0° and 360°.
- p14 – l.11: “analysing”

References:

He, Y., Li, Q., Chan, P., Zhang, L., Yang, H., & Li, L. (2020). Observational study of wind characteristics from 356-meter-high Shenzhen Meteorological Tower during a severe typhoon. *Wind and Structures*, 30(6), 575–595.
<https://doi.org/10.12989/WAS.2020.30.6.575>.

Pearson, G., Davies, F., & Collier, C. (2009). An analysis of the performance of the ufaM pulsed doppler lidar for observing the boundary layer. *Journal of Atmospheric and Oceanic Technology*, 26(2):240–250. <https://doi.org/10.1175/2008JTECHA1128.1>.

Wang, Y., Xiangde, X. Zhao, Y., & Minzhong, W. (2018). Variation characteristics of the planetary boundary layer height and its relationship with pm2.5 concentration over China. *Journal of Tropical Meteorology*. 24. 385-394. 10.16555/j.1006-8775.2018.03.011.

Mariani, Z., Crawford, R., Casati, B., & Lemay, F. (2020). A Multi-Year Evaluation of Doppler Lidar Wind-Profile Observations in the Arctic. *Remote Sens.* 12, no. 2: 323. <https://doi.org/10.3390/rs12020323>

Kumer, V.-M., Reuder, J., & Furevik, B.. (2014). A Comparison of LiDAR and Radiosonde Wind Measurements. *Energy Procedia*. 53. 214-220. 10.1016/j.egypro.2014.07.230.