

Interactive comment on “Filtering of pulsed lidars data using spatial information and a clustering algorithm” by Leonardo Alcayaga

Anonymous Referee #1

Received and published: 20 March 2020

This paper presents alternative wind lidar data recovery methods, over the traditional carrier-to-noise ratio. The paper presents both a clustering technique and a median-like filter, and evaluates results on both synthetic and real lidar data.

While the paper includes some important results, the presentation is a little clumsy, and I feel the paper could be greatly improved. There needs to be general improvements to the usage of English throughout, examples of which I have highlighted below. The paper overall reads as if several authors composed different sections, there is a lot of repetition of the discussion, and the figures do not flow nicely. While some scrolling/page turning is expected, referring to figure 10 on page 7 requires the reader to turn to page 19. Perhaps there is an alternative way to make your point on page 7? Figure 7 also does not seem to be referred to in the text?

The point I would like to make most clearly is your conclusion states the clustering filter performs best in both synthetic and real data, and increases data availability between 22% and 38%, while also reducing erroneous measurements between 70% and 80%. This is a significant result, and I feel you could make more of this in the paper. There is a lot of discussion on methods used, sometimes repeated several times, but I feel comparatively little on your major results.

Improving the flow of the paper, and removing some of the repeated discussion to focus more on results will greatly enhance your paper.

Minor comments:

Title should read “lidar” rather than “lidars”

Page 1 Line 13 – replace “its adoption” with “their adoption” or similar Line 14/15/16 – the meaning of the sentence beginning “Their capability to measure...” is unclear. Do you mean a single lidar can scan a spatial domain of comparable size to a wind farm? If so, it would be helpful to include an indication on the actual size of a windfarm. By “their increasing accuracy” do you mean increased accuracy over meteorological masts? Line 17 – please be more specific with “traditional wind measurement techniques”, for example wind profiling radars can also be used, and are also susceptible to atmospheric conditions. What is “traditional”? Line 18 – please define “lack of references”, do you mean a second instrument to compare wind values to? Line 25 – please define VLOS the first time you use it, rather than the second

Page 2 Line 26 – remove the “of” in “between of line-of-sight...” Line 39 – you don’t need both “like” and “e.g.” together Line 39 – please consider rephrasing the sentence beginning “Complementing all these features...”. The sentence is very long and difficult to follow. Line 45 - “...which are capable of classify large data sets...” needs to be reworded for correct English Line 54 – swap the order of “defines” and “always” to read “which always defines a unique...” Line 56 – please define/introduce DBSCAN here, rather than on page 12 Line 58 - “...capable of identify clusters...” should read

[Printer-friendly version](#)[Discussion paper](#)

“...capable of identifying clusters...”

Page 3 Line 72 – what do you mean by “the wind speed data covers a large horizontal area”? Do you mean you wish to measure winds across a large area? Line 88 – I’m not sure I follow what a “wrong observation” is, as compared to an outlier?

Page 5 Line 99 – change “generate” to “generates” Line 102 – change “make” to “mean” or similar

Page 7 Figure 2 caption – line 3, I believe should read “next” not “nest” Line 149 - “radial” is miss-spelled Line 158 - “en” should be “in”

Page 9 Line 184 - “2” should read “section 2” as done previously Line 189 – the sentence beginning “The noisy areas show...” is very long and hard to follow. Please consider rewording.

Page 10 Line 200 to 203 – these 2 sentences seem to be a repeat of the introduction?

Page 11 Line 229 - “non” should read “not”

Page 12 Line 240 – similar to the comment above, page 10 lines 200 – 203, this section appears to be a repeat of earlier discussions

Page 15 Line 298 – I think you mean “noisy” not “nosy”

Referral to figure 7?

Page 16 Lines 315 to 320 – sentence beginning “This allows us to define...” is very long and difficult to follow Line 320 - “this metrics” should read “these metrics”

Page 18 Line 344 – I think you are missing “are” in “...that two realizations from the same distribution...” Line 365 – should read “...on the other hand...” rather than “in”

Is there a reason why you can’t do the same tests to the synthetic data as you are for the real data?

Page 19 Line 372 – remove the second “then” from “...then becomes relevant then...”

Page 20 Line 387 – remove the comma after “both” to read “...in both noisy and reliable...” Line 390 – reverse the order of “be then” to read “then be” Line 391 – replace “its” with “their” to read “...distant from their previous location...” Line 401 – remove “be” and change “benefited” to “benefit” to read “...filter will benefit by...” Line 403 – add “to” to read “...dimensions to the data description.”

Page 21 Line 406 – remove “a” to read “...of good measurements...”

I don't get the comparison to synthetic data. You site the advantages of using synthetic data are you know where the noise is, yet you don't have plots showing a comparison to the known noise is?

Page 27 Line 483 – replace “This” with “These” to read “These possible deviations...”

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-450, 2020.

Printer-friendly version

Discussion paper

