Atmospheric
Measurement
Techniques

Open Access

EGU

Interactive comment on "Detecting turbulent structures on single Doppler lidar large datasets: an automated classification method for horizontal scans" *by* Ioannis Cheliotis et al.

Anonymous Referee #1

Received and published: 21 June 2020

General Comment:

This manuscript is presenting a method to identify coherent turbulent structures using scanning wind lidar data. The method is based on the application of a machine learning algorithm to detect different patterns of spatial variations of the wind speed. The usage of this methodology to categorise data acquired by a scanning wind lidar instrument is innovative and of interest for the remote sensing research community. Its application can contribute to the analysis of large data sets by minimising the necessary amount of processing time that is required today to categorise scanning wind lidar data.

1. In the manuscript a cosine fit is applied to wind lidar data acquired during VAD scans,

in order to estimate the mean wind speed and direction. Such a data analysis method requires the assumption of horizontal homogeneity of the wind vector, as also stated by the authors. However, I am wondering to which extent this homogeneity is expected over the urban landscape of Paris. I think that a discussion on the terrain heterogeneity is missing from the article. In this direction, I think that it would be constructive to add an elevation map of the area over which the scanning wind lidar was acquiring measurements. This would provide an insight to which extent the observed spatial variations of the wind are related to temporal fluctuations of the wind and/or due to changes of the terrain elevation.

In this context and regarding the data analysis:

a. A statistical parameter is required to specify the representativeness of the fit used in the VAD scan. In the manuscript it is stated that the RMSE values of the estimated fit have been calculated but they are not stated in the document.

b. Did the authors perform a quality check of the acquired data? Do they apply any SNR filtering to the acquired radial wind speed prior the application of the data analysis?

2. The subtraction of the mean wind speed from the radial wind speeds does not compensate the fact that the individual measurements along the scanning pattern are the result of the projection of the instantaneous wind vector to the line-of-sight of the wind lidar. Therefore, the term turbulent wind field could lead to a misinterpretation. The authors should clearly state that they measure the high frequency fluctuations of the radial wind speed.

Specific Comments:

P2 Line 28. The reference of Roth 2007 is a review of the atmospheric turbulence studies over urban landscapes and it doesn't directly discuss how the pollution concentration in urban environment is dependent on the weather and on the turbulence. A reformulation of this sentence is suggested for a clarification.

P3 Line74. Why is it relevant if the campaigns are short-term or long-term? And what is the time scale of these two types of campaigns that it is relevant for the topic of this study?

P3 Line95. I am not sure that the two references of Kumer 2014 and Veselovskii et al. 2016 are the appropriate to be used here. In none of them there is an analytical explanation of the wind lidar instrument as it is stated. I suggest replacing them with more relevant references. An example could be the work of:

Cariou, J. P., R. Parmentier, M. Valla, L. Sauvage, I. Antoniou, and M. Courtney. "An innovative and autonomous 1.5 μm Coherent lidar for PBL wind profiling." In Proceedings of 14th Coherent Laser Radar Conference. 2007.

P3 Line 101. It is stated that "The duration of each scan was 3 minutes which is sufficiently short for the observation of structures". Could the authors elaborate more on this statement?

P4 Line 102. The authors state that the maximum range of the scans reached 5 km. However, in none of the figures that they have included in the manuscript the range ever reaches 5 km. The most common range in their data is between 2 – 2.5 km. Can the authors explain why is this happening?

P4 Line111. Table 2. A list of scanning patterns is included here that are not used in this study. In addition, the purpose of each scanning pattern is included without explaining the reasoning for their selection. I would suggest to either remove this table. If the authors wish to keep it then I suggest that they should elaborate more in the text about it.

P4 Line120. The "offset" refers to the vertical wind speed component? What does it mean that α is much smaller than b? Is this a common observation over the whole scanning plan? And what kind of α values does the application of the model result in?

P6. Line156-157. It is not clear how the unaligned thermals are dependent on the

C3

increased solar radiation measurements.

P6. Line 163. How do the authors estimate the value of the wind shear? Also, shouldn't the units of the wind shear be seconds to the power of -1?

P6. Line 166. The authors in P5. Line 133 state that the VAD method could not be applied in the data acquired during the night, especially at those occasions where the mean wind speed was less than 2 m/s. Does this mean that only cases with mean wind speed higher than 2 m/s were selected?

P6. Line 167. How well was the mean direction and speed estimated through the VAD method in the selected cases? I suggest adding a statistical parameter that describes the representativeness of the applied fit (equation 1) to the measured data (e.g. RMSE).

P7. Line 170. It is difficult to understand the scale of the Modis colour images. Would it be possible to either add a scale or mark on the images the scanning area?

P7. Line 176. I recommend describing very shortly the texture analysis especially in the context of remote sensing. It would be useful to add any references in the introduction regarding the previous applications of this type of analysis to remote sensing data.

P8. Line 181-182. What is the logic behind the selection of this particular values for defining the contrast?

P10. Line 243. The authors state that 60 cases of "other" patterns are used during the supervised machine learning step. They argue that this is necessary because "it is expected to be the dominant category in the classification". I am not sure that I understand what it is meant with this statement. I suggest having this part a bit more explained. Furthermore, how is the mean wind speed and direction estimated in these cases? As it is stated the VAD was not successfully applied to this data.

P10. Line 225. How do the authors explain the change in the slope of the homogeneity curve that is observed for absolute azimuth angles larger than 45 degrees?

C4

P10. Line 240. Could the authors add a reference to the literature describing the "supervised machine learning methodology"?

P11. Line 265. How do the authors physically explain this result? A low RMSE in the cosine fit, couldn't also mean that the mean wind speed and direction are not estimated correctly?

P12. Line 291. It is stated that "there were scarcely any rolls cases observed at night". However, in the sentence 288 it is written that for the classification of this study the authors "consider only thermals and rolls during daytime". To my understanding there is an inconsistency between the two statements. Can this be explained better?

P13. Line 312. In the conclusions the author state that time, wind speed and the cosine fit RMSE of the VAD method were not selected by the algorithm for the classification. However, in the results presented in figure 9 there is a time dependency in the detection of certain patterns (e.g. thermals and rolls). Could the authors comment why the inclusion of the time as a classification parameter would not improve further their results?

P13. Line 318. Given the fact that one PPI scan lasts for 3 minutes and occurs every 18 means, can the authors explain how does the acquired data set contribute to the comprehension of the development of coherent structures?

Technical Corrections:

General: There is a small inconsistency on the way that figures are referred to. Sometimes they are used parenthesis after the number of the Figure to denote a subfigure and sometimes are not.

P2 Line 34. I think that the statement "Futhermore, …." is a self-evident. I would suggest removing it. Also, I would suggest to add the reference of (Hussain, 1983) at the end of the previous sentence.

P2 Line 37. "and the lower"

P3 Line 72. I suggest reformulating this sentence. It is not clear to the reader how is the two-dimensional autocorrelation function was used. Also, I suggest changing the sentence "the observation of the scans by eye" to "visual observation of the scans".

P5 Line132. The authors state that due to the surface heterogeneity the VAD method can be applied in some cases. A surface heterogeneity will introduce an error in the VAD method regardless the wind speed.

P3 Line76. Change "month" to "months"

P3 Line83. The "Section 0" should change to "Section 3"

P3 Line86. The word "two" is used as a noun modifier and therefore the word month should be in singular form.

P3 Line96. I suggest to change the text "The lidar is sensitive only to the" to "A wind lidar is measuring the"

P3 Line100. Change the "for a" to "with a".

P4 Line103. I think that it is more grammatically correct to either use the past tense or the passive form of the "rise" verb.

P5 Line132. I suggest that the "Jussie site" is changed to "The Jussie site".

P5 Line138. Figure 3 caption. Change the "a case" to "A case", also a add a tab space between (a) and "Radial".

P6. Line149. "Sec" should be replaced by "Section".

P6. Line 164. I suggest to re-write the sentence "For many cases, the wind shear was accompanied by turbulent streaks pattern" and specify for which particular wind shear values were streaks detected.

P6. Line165. The part of the sentence "so or the training ensemble" should be rewritten.

P10. Line 230. "fore" should be changed to "for".

P11. Line251 – 253. The authors state the higher number of dimensions relative to the number of patterns lead to the "curse of the dimensionality problem". I suggest to re-write it by using appropriate scientific statements.

P11. Line 259. Correct the "Section 0".

P12. Table 5 caption: I suggest changing the "eye-made" to "visual".

P12. Line 290. The coherent structures don't have a preference. They are formed under favourable atmospheric conditions. I suggest commenting the result of Figure 9 on that basis.

P15. Line 386. "Lemone" should be changed to "LeMone". Also, necessary information of the article (e.g. journal) is missing.