Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-82-RC2, 2020

Atmospheric Measurement Techniques Discussions

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 #2

Received and published: 23 June 2020

1 General comments

The authors present the implementation of an innovative approach for identification and classification of coherent structures in atmospheric flow over urban canopy. The approach is based on measurements from a single Doppler lidar system installed within the surface layer, covering a significantly large spatial domain, and uses a machine learning technique to classify spatial patterns in horizontal wind fields under different stability conditions. The approach shows advantages over classic methodologies for coherent structure identification. It reduces data processing time, and it might allow a more automatic structure detection from large databases.

Printer-friendly version

Discussion paper

As I understand form the first paragraph in the introduction, comprehension of the flow physics it is important for monitoring atmospheric pollution. However, the physics identified here, in the form of coherent structures, are not related with bad pollution conditions, since the latter might fall in the “Other” category, and their physics are not clear from the study. The motivation of the study should be stressed form the beginning. Only when we arrive to the conclusions we can read some of the potential application of this approach. (line 316 of the text).

There are many previous studies on coherent structures, as well as lidar technology, than could be included. The description of the coherent structures that this study aims to identify is a bit vague and needs improvement.

In general, the term turbulence and turbulence fields are used frequently in the text, but the range gate resolution of the lidar scanner is 50m at a height of 75m above ground level. Turbulence and its most energetic eddies might fall within this length scale. Almost all turbulence fluctuations are filtered out by the lidar due to spatial averaging. What we can clearly see from lidar observations is medium-to-large scale fluctuations and coherent structures rather than turbulence.

The methodology behind data process could be better explained. I miss a paragraph describing how data quality was assured. Did you use a CNR/SNR threshold for filtering? What was the data recovery rate during the two months of measurements?

It is not clear from the text how the radial wind speed fluctuations are calculated. What does “stronger radial wind speed” mean? A larger absolute value?. It seems to me that the sign might come from the combination of u and $\cos(\theta)$. One suggestion is to put this definition as equations. Since the wind direction is obtained from equation (1) it would be possible to work with the streamwise component u instead of the radial wind speed u_r.

It is not clear from the results, what is the relative importance of non-distinguishable structures, bad fittings, and bad data (low CNR/SNR signals) in the “Others” category.

The authors give some information about what it seems to be the reason of one group of cases to belong to this category (bad fitting of cosine function), but no threshold on this fitting error is given.

I miss more elaboration in the description of the texture parameters used for classification, namely, why they might be relevant and if they were relevant in the end. The feature selection process is also not very clear. Cross validation is well known and well explained, but the text explaining the outcome as well as the figure used in that regard are confusing.

A brief description of the machine learning technique used could be useful for clarity. Conclusions section. In my opinion, this section should be read in a positive rather than negative way. Example: it should focus on the relevant parameters discovered (which need a bit more explanation in the corresponding section) rather than the ones excluded by the study. The sections describing the methodology used-which need some improvement-are already clear, and no repetition is needed. Same with the results highlighted.

2 Specific comments

Page 1 Line 12. Change ”manually” to “visually”.

Page 1 Line 15. Change “and installed” to “installed”.

Page 1 Line 16. It would be better to reword this sentence, maybe “The turbulent component of radial wind speed is estimated using. . .over 4577 scans.

Page 1 Line 18-21. I am not sure what the sentence describing the training set adds to the abstract if not combined with the next one. It is better to state directly the unsupervised algorithm used instead of using parenthesis. It might be better to rephrase this in a more concise way.

Page 1 Line 23. What are the remaining 20

Page 2 Line 26. Change “step for” to “step towards”.

Page 2 Line 32. A coherent structure is defined according to its phase-averaged rather than its instantaneous vorticity. I also suggest moving the Hussain 1983 reference to this sentence. A coherent structure needs to maintain its phase-averaged vorticity rather than its time-averaged vorticity or form.

Page 2 Line 35. Please specify that this is the case of atmospheric flow. Other structures are observed at laboratory scale (also in the atmosphere but not so relevant for momentum or scalar fluxes), like hairpins, or hairpin trains. Include a reference to Hutchins and Marusic (2006) and Adrian (2007).

Page 2 lines 37-44. Consider reordering the sentences here for a more fluent reading. Maybe starting the paragraph from sentence in line 41?.

Page 2 line 45. How is it that you identify rolls in the mixed layer, with sizes from few to dozen kilometers, with scans at surface layer height (75m) with spatial coverage of less than 2 kilometers?. Is this description coherent with what you are identifying?.

Page 2 lines 57-65. Since you are using lidar instead of radars it would be better to shorten the scanning pattern description using some of the given references, since they have to do only with the history on scanning patterns. There are more recent references of this regarding lidars. Cariou (2007) and Vasiljevic (2016).

Page 3 line 72. Replace “by eye” by “visual inspection” or similar.

Page 3 line 73. More than time-consuming, it might be non-systematic.

Page 3 line 74. You meant “A less time-consuming” approach? What height was the met mast?

Page 3 line 78. “This study aims to identify turbulent coherent structures from single Doppler lidar horizontal scans”. Also, please introduce here what is texture analysis

(roughly maybe) and what what machine learning technique you are using.

Page 3 line 83. Section 0 must read section 3, here and in the rest of the text.

Page 3 line 86-91. It should read “measurement campaign”. Move “in Paris” to the end of the sentence, modified to “in the urban area of Paris”. Remove the url of leosphere to the reference section maybe. More than only sensitive to the radial component, the lidar does measure and it is intended to measure only the radial component. Lidars technology and its operation principle is well known, use references (Cariou, maybe write ts paragraph in amore concise way, being specific in the corresponding table than in the text here, which is a bit confusing.

Page 3 line 101. I would say fast instead of short. Also, what type of structures and why this time window?

Page 4 line 121. What is the reason behind a is small for your case?.

Page 5 Figure 2. Why the reach of the lidar is 2 km and no 5 km?.

Page 5 line 134. Is it possible to specify the fraction of cases with low winds, and its relative importance to the number of bad fittings?. I miss an analysis of stability conditions, since it seems that stable conditions affect the most.

Page 6 line 143. Actually, for rolls, it is the opposite. Ascending motions bring low momentum to higher levels, reducing the speed, and vice versa.

Page 6 line 146. Since rolls and streaks both present areas of alternating low/high momentum with elongated shape, their main difference is their extent. What is the criteria to differentiate between them? The clouds formation shape from MODIS was used, as I understand, only for a fraction of the cases included in the training dataset.

Page 6 line 161-165. Wind shear is defined as du/dz with 1/s units, could you clarify what definition you are using here? Additionally, streaks are present in turbulent flow as well, beyond stable conditions, why do you focus in cases with low turbulence energy

(stable conditions)? It seems that high shear due to jets is only one among several mechanisms.

Page 6 line 166. How many cases did you use for the “Other” category? From table 5 seems that they are around 60, the double. What is the reason for such big number?. Can this influence the final classification output? This is explained in section 4, but it should be clear from here.

Page 7 Figure 4. What is the scale of map in (d) and (e)?

Page 7 line 176. Could you introduce what texture analysis is first? Additionally, since “Others” had a poor fitting and then uncertain wind direction, how did you align them with 0 degrees?

Page 7 line 180. Eight bins were chosen for increased contrast. Why eight?, could you develop more on this?. What is the effect of the number of bins in the output?

Page 8 line 185. The procedure for the construction of the CM matrix is a bit confusing. Could you write it in a more concise way?.

Page 9 line 212. Is it possible to elaborate more on the 4 parameters described?. It is not clear only from the equations what their characteristics are.

Page 10 Figure 6. The notation of the azimuth angle is different form the text. Why does homogeneity grow after 45 degrees for all categories? The definition of homogeneity says that CMs with large values in the diagonal might result in larger values of this parameter. The diagonal from table 3 to 4 decrease because of azimuth angle. Should homogeneity decrease monotonically from 0 to +/- 90 degrees?. Can you elaborate more on this?. How many cases are represented for each category in the figure? Only one scan? An average from many cases?

Page 10 line 231. Notation is a bit weird here.

Page 10 line 241. The description of the training set might be better place in section

3. Why is it expected that “other” category should double the rest? Please elaborate more on this.

Page 11 Figure 7. This figure is very confusing and not self-explanatory at all. Please give more information in its caption, relative to the number in parenthesis (neighbor order I suppose), state that they are all or a few of the final parameters used.

Page 12 Table 5. Change “eye-made” to a better term, like visual classification or similar.

Page 12 line 286. It is not clear if streaks were also detected during daytime, since the previous definition of the training set (line 162) says only night-time, but figure 9 says the opposite. Same for rolls and thermals. In summary, the constraint you talk about (day-time rolls and thermals, night-time streaks) does concern only the training set definition?.

Page 12 line 292. If I am correct, you tried to explain thermals during night, not others during days. Only the last word in the sentence, “reverse”, explains this. Moreover, can you elaborate more on what is the reason behind the erroneous classification of thermal as “others”? During stable conditions turbulent eddies are smaller, structures also show smaller length-scales. However, mean wind can show slight differences with no directional preference, and they can look like thermals (see Shah and Bou-Zeid, 2014).

Page 13 line 295. Stable cases during night show buoyant forces opposing vertical momentum flux and turbulence generation. Mechanical turbulence does die out under stable conditions. Mechanical turbulence destruction by buoyancy is the dominant mechanism, not the opposite.

Page 13 line 314. So thermals are not turbulent?. Why do you separate rolls and steaks form thermals? Does it has to do with pollution transport or something similar?.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-82, 2020.