Review: rev. 2 of WES-2020-96 by Siefert, *et al.*

Mark Kelly

The draft has been improved with revision; however, some issues still remain. These are pointed out below, with author comments (AC) addressed first, and later point-wise comments with line numbers referring to the file that showed revision/changes ("...ATC1.pdf").

There are also numerous linguistic errors which need to be corrected (most of them new). As in the previous review, I again suggest asking someone with native-level English proficiency to proof-read the (updated) draft; they are too numerous to note individually here.

E.g., on l.6: "in free field" is not proper English; on l.8 the word 'of' is missing after "eight".

Before getting to specific points line-by-line, I'll respond to the author comments (ACs) that addressed my previous reviewer comments (RCs), and then I add some general comments.

Replies to author comments/responses (ACs)

AC11: It is an improvement to mention the "reconstructed" aspect of wind speed, but to be open/clear, why not include your AC statement "details on the reconstruction are not available" in the text? This and/or the unknown transfer function should be mentioned; e.g. the latter can affect the direction as well as the speed.

AC12: Yaw misalignment isn't only due to "false...measurements, calibrations, or sensor installation"–especially if there is a transfer function used for nacelle-mounted 2d-anemometers. Your addition on l.106-8 helps to allay this issue.

AC15: Your response about yaw misalignment threshold and rate per 10-minutes is reasonable, but you have not included this in the revised text.

AC17: How is your approach "similar to" Taylor's, but not actually simply assuming it? Again, it appears you've assumed it over the entire range of τ (and τ_{norm}); this can become problematic for small enough U (large lags).

AC18 and l.189-205: you have used τ_{norm} before it is defined in (4); this can be quite confusing for the reader, particularly if they have not read this before. Also, why is the second 'normalization' (4) done? If $U_{max} =$ 13m/s always, and $x_{AB,mean}$ is also a simple constant for all cases, then why normalize again? If you have done this to force the peaks closer to 1, then this

should be stated. Also, how was U_{max} chosen–doesn't this just arbitrarily squeeze/stretch the correlation curves (as you wrote for $\tau_{norm,intv}$)? There appears to be no physical justification for (3) and (4) together, unless perhaps you could explain what is meant by "at least equal to the maximum possible wind speed to fit all normalised curves".

AC20: The response statement "the temporal autocorrelation of a wind turbine decorrelates in the considered time intervals of 300 s" does not make sense. Do you perhaps mean that the the correlation decreases to effectively 0 as lag (τ) approaches 300s?

AC23: If you are to insist on using the term "filtering" in place of 'data selection' or similar—knowing that WES is not a data science journal, but a wind energy journal where you are also mentioning turbulence—then you should at least include the word 'data' before it. Further, I strongly recommend section 2.1 to be renamed "Data selection and filtering" or similar—again, spectral filtering is commonly used when dealing with this kind of data in wind energy (particularly turbulence), especially when mentioning different intervals (e.g. 600s).

General comments

The labelling of peak correlations (between power fluctuations for turbine pairs) as "correlation states" is contentious, since 'states' implies different physical scenarios or flow/operational-regimes—particularly if you have not described anything like the latter. If the 'states' are basically different groups of peak correlations (magnitudes), then why not call them that? This is safer, because for different flow regimes having larger/smaller turbulence length scales (and/or other farms having different spacing, surfaces, or even hub height), then the magnitudes or groups could be quite different. Also, in the conclusion it should be mentioned how/why seemingly insignificant peak correlations (e.g. 0.2 or less) are meaningful (compared to commonly understood statistical significance being $\rho_{AB} > 0.5$); i.e. the relative values are significant given the 'noisy' turbulent flow, in addition to values consistent with others' LES results.

Some detailed comments

l.5, 7: the challenge is *spatially* variable flow, not just "highly" variable flow.

l.10-11: "decrease towards spanwise pairs" doesn't quite make sense. If the correlations decrease with angle between mean wind direction and pair separation vector, why not write that?

l.13-17: "the correlation of streamwise aligned wind turbine pairs" should be 'power correlations between streamwise-aligned wind turbine pairs'.

l.18: sorting is accomplished by a "k-means clustering algorithm", not "clustering algorithm k-means" (e.g. Likas, Vlassis, & Verbeek 2003).

l.19-20: the sentence "These groups are here referred to as correlation states." is not needed in an abstract.

l.20 repeats l.7-8.

l.18-22: "these parameters" is repeated three times; the final point is also somewhat of a repetition of l.16-17...The abstract can be cleaned up (there are more English errors in it as well).

l.71: Use of "correlation curves" here to mean "states" is ambiguous and confusing; however, in l.194 and after it is used reasonably, to refer to the actual $R(\tau)$ curves. Here in this context of "states" you are really referring to the peak correlation, as seen later in e.g. Figs.7-8.

l.151: sentence is not finished.

l.196: "noted as" –> 'denoted by'

l.212 and elsewhere: "dependency" should be 'dependence'

l.242: how does $\tau_{norm} > 1$ have any meaning, if τ_{norm} is arbitrary due to its definition via the artificial Umax?

l.381: exactly which standard deviation?